

RECALL OF LANDMARKS IN INFORMATION SPACE

by

Molly Emmons Sorrows

BA, Earlham College, 1987

MSIS, University of Pittsburgh, 1992

Submitted to the Graduate Faculty of

School of Information Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2004

UNIVERSITY OF PITTSBURGH
SCHOOL OF INFORMATION SCIENCES

This dissertation was presented

by

Molly Emmons Sorrows

It was defended on

April 19, 2004

and approved by

Dr. Peter Brusilovsky

Dr. Michael Lewis

Dr. Edie Rasmussen

Dr. Michael Spring

Dr. Stephen Hirtle
Dissertation Advisor

Copyright by Molly E. Sorrows

2004

RECALL OF LANDMARKS IN INFORMATION SPACE

Molly Emmons Sorrows, Ph.D.

University of Pittsburgh, 2004

Research on navigation and landmarks in physical space, information space and virtual reality environments indicates that landmarks play an important role in all types of navigation. This dissertation tackles the problem of defining and evaluating the characteristics of landmarks in information space. This work validates a recent theory that three types of characteristics, structural, visual and semantic, are important for effective landmarks.

This dissertation applies concepts and techniques from the extensive body of research on physical space navigation to the investigation of landmarks on a web site in the World Wide Web. Data was collected in two experiments to examine characteristics of web pages on the University of Pittsburgh web site. In addition, objective measurements were made to examine the characteristics of web pages with relation to the experimental data. The two experiments examined subjects' knowledge, use and evaluation of web pages. This research is unique in research on web navigation in its use of experimental techniques that ask subjects to recall from memory possible navigation paths and URLs.

Two measures of landmark quality were used to examine the characteristics of landmarks; one, an algorithm that incorporated objective measures of the structural, visual and semantic characteristics of each web page, and the second, a measure based on the experimental data regarding subjects' knowledge and evaluation of the page.

Analysis of this data from a web space confirms the tri-partite theory of characteristics of landmarks. Significant positive correlations were found between the objective and subjective landmark measures, indicating that this work is an important step toward the ability to objectively evaluate web pages and web site design in terms of landmarks. This dissertation further suggests that researchers can utilize the characteristics to analyze and improve the design of information spaces, leading to more effective navigation.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS.....	xi
1. INTRODUCTION	1
1.1. Background	1
1.2. Problem Statement.....	2
1.3. Significance of Work.....	3
1.4. Outline of the Dissertation.....	4
2. LITERATURE REVIEW	5
2.1. Overview of Navigation Research.....	5
2.2. General Theories and Frameworks of Navigation.....	6
2.2.1. Types of Knowledge.....	6
2.2.2. Types of Environments.....	7
2.2.3. Components of Navigation.....	7
2.3. Navigation Research Domains	8
2.3.1. Architecture	8
2.3.2. Biology	9
2.3.3. Neural Networks Models.....	11
2.3.4. Graph Theory.....	13
2.3.5. Robotics	14
2.3.6. Virtual Reality Environments.....	16
2.4. Navigation in Hypertext	17
2.4.1. Hypertext	17
2.4.2. Navigation.....	18
2.4.3. Maintaining Orientation.....	20
2.4.4. Navigational Aids	21
2.5. Overview of Research on Landmarks	23
2.6. Landmarks in the Physical World	23
2.6.1. Landmarks as Elements of Spatial Representation.....	24
2.6.2. Development of Spatial Representations.....	25
2.6.3. Use of Landmarks in Navigation.....	27
2.6.3.1. Landmarks as organizing concepts.....	27
2.6.3.2. Landmarks as navigational aids.....	28
2.7. Landmarks in Electronic Environments	29
2.7.1. General Descriptions	30
2.7.2. Computational Approaches to Landmark Definition	31
2.8. Summary of Landmarks in Physical and Electronic Spaces	33
2.9. A Theory of Landmarks	34
2.10. Current Research Directions.....	36

3.	RESEARCH METHODOLOGY	38
3.1.	Overview	38
3.2.	Hypotheses	39
3.3.	Measures of Landmark Characteristics	40
3.3.1.	Objective Landmark Value	40
3.3.1.1.	Measure of structural characteristics of landmarks	41
3.3.1.2.	Measure of visual landmark characteristics	43
3.3.1.3.	Measure of semantic landmark characteristics	44
3.3.2.	Subjective Landmark Value	45
3.4.	Methodology	46
3.4.1.	Experiment 1 – Navigation strategies and site knowledge	47
3.4.2.	Experiment 2 - Characteristics of landmarks.....	51
3.5.	Summary	52
4.	RESULTS	54
4.1.	Objective Landmark Values	54
4.2.	Experiment 1	58
4.2.1.	Demographics	58
4.2.2.	Path Descriptions (open-ended questions).....	59
4.2.3.	Bookmarks and URL Knowledge.....	73
4.2.4.	Starting Points (multiple-choice questions).....	74
4.2.5.	Summary	79
4.3.	Experiment 2	79
4.3.1.	Demographics	80
4.3.2.	Web Page Characteristics	80
4.3.3.	Analysis of Subjective Landmark Values.....	87
4.3.4.	Summary	90
4.4.	Chapter Summary	90
5.	ANALYSIS	92
5.1.	Evaluation of Easily Recalled Web Pages.....	92
5.1.1.	Recall of Navigation Paths	93
5.1.2.	Objective Landmark Value Ranges	93
5.1.3.	Evaluation of Hypothesis 1	94
5.2.	Evaluation of easily accessed web pages	97
5.2.1.	Evaluation of Hypothesis 2.....	97
5.2.2.	Evaluation of preferred starting points	99
5.3.	Relationship between the objective and subjective landmark measures	100
5.4.	Effect of weights on the correlation of objective and subjective measures	104
5.5.	Limitations of the analysis.....	106
5.6.	Summary	107
6.	DISCUSSION AND FUTURE WORK.....	108
6.1.	Major findings	108
6.2.	Implications for web design and use	109
6.3.	Methodology	110
6.4.	Alternate Computation of Landmark Saliency.....	116
6.5.	Future Work	117

APPENDICES	119
APPENDIX A.....	120
Web Pages.....	120
APPENDIX B.....	121
Experiment 1 Questionnaire	121
APPENDIX C.....	140
Experiment 2 Questions.....	140
APPENDIX D.....	142
Experiment 2 Web Pages	142
APPENDIX E.....	166
Student Life at Pitt Web Page.....	166
APPENDIX F	169
Semantic, Visual and Structural Components of the Objective Landmark Value.....	169
APPENDIX G.....	170
Semantic, Visual and Structural Components of the Subjective Landmark Value	170
APPENDIX H.....	171
Correlations of web page features to Subjective Landmark Value	171
APPENDIX I	172
Correlation of Subjective and Objective Landmark Value.....	172
BIBLIOGRAPHY	173
INDEX 181	

LIST OF TABLES

Table 4-1.	Objective landmark value for web pages used.	55
Table 4-2.	Overall web page references for Experiment 1.....	60
Table 4-3.	References to each web page for question 1 of Experiment 1.....	62
Table 4-4.	References to each web page for question 2 of Experiment 1.....	64
Table 4-5.	References to each web page for question 3 of Experiment 1.....	67
Table 4-6.	Number of subjects referring to each web page for question 4 of Experiment 1. ...	70
Table 4-7.	Number of subjects referring to each web page for question 5 of Experiment 1. ...	71
Table 4-8.	Starting points and successful access schemes (Exp. 1).	75
Table 4-9.	Starting points for each multiple-choice question.	76
Table 4-10.	Subjective landmark value for web pages in Experiment 2.....	88
Table 5-1.	Objective landmark value of pages from path recall questions.	95
Table 5-2.	Objective landmark value of pages easily accessed by URL or bookmarks.	97
Table 5-3.	Objective landmark value of starting points in multiple-choice questions.....	99
Table 5-4.	Comparison of subjective and objective landmark values.....	101
Table 5-5.	Correlation of subjective landmark value and the three objective components. ...	102
Table 5-6.	Correlation of subjective landmark value and features of the objective landmark value.....	103
Table 5-7.	Effect of different on correlation of subjective and objective landmark values....	105
Table 5-8.	Regression analysis of objective components.....	105
Table 6-1.	Examples of landmark significance tests for features of web pages.	117
Table A-1.	The URL and assigned short title for each web page used.....	120
Table H-1.	Subjective landmark value statistical correlations.....	171
Table I-1.	Effect of weight strategies on objective and subjective landmark value correlation.	172

LIST OF FIGURES

Figure 2-1.	Connectivity into node A and out of node B.	32
Figure 3-1.	Equation to compute the objective landmark value.	42
Figure 3-2.	Equation to compute the subjective landmark value	46
Figure 3-3.	Example of an open-ended path description question for Experiment 1.	48
Figure 3-4.	Sample multiple choice question for Experiment 1.	50
Figure 4-1.	Objective landmark value components for a selection of web pages.	57
Figure 4-2.	Path diagram for finding library hours (Exp. 1, Q.1).	63
Figure 4-3.	Paths to find campus movie times (Exp. 1, Q. 2).	65
Figure 4-4.	Paths to locate BSIS class lecture notes (Exp.1, Q. 3).	68
Figure 4-5.	Paths to find time of next home Panther football game (Exp. 1, Q. 4).	69
Figure 4-6.	Paths to find a map of the University of Pittsburgh Oakland campus (Exp. 1, Q. 5).	72
Figure 4-7.	URL knowledge and bookmark use (Exp. 1).	73
Figure 4-8.	Average answers for “How often have you been to this page?” (Exp. 2, Q.1).....	81
Figure 4-9.	Average answers for “Would you remember this page if you saw it again?” (Exp. 2, Q.2).	82
Figure 4-10.	Average answers for “How visually distinctive is this page?” (Exp. 2, Q. 3).	83
Figure 4-11.	Average answers for “How easily could you get to this page without using a search engine?” (Exp. 2, Q. 4).	84
Figure 4-12.	Average answers for “Do you know the exact URL for this page?” (Exp. 2, Q. 5).	85
Figure 4-13.	Average answers for “How important is the content of this page in general?” (Exp. 2, Q. 6).	86
Figure 4-14.	Average answers for “How important is the content of this page to you personally?” (Exp. 2, Q. 7).	87
Figure 4-15.	Subjective landmark value components for selected web pages.	89
Figure 5-1.	Objective landmark values of pages recalled in path descriptions.	96
Figure 5-2.	Objective landmark values of pages accessible by bookmark or URL.	98
Figure 6-1.	Standard deviation from the average based on four visual features.	113
Figure 6-2.	Proportion of visual features on web pages and the average across pages.	115
Figure D-1.	Pitt Home web page.	143
Figure D- 2.	Financial Aid Home web page.	144
Figure D- 3.	Capital Campaign web page.	145
Figure D- 4.	Fact Book web page.	146
Figure D- 5.	Libraries (ULS) web page.	147
Figure D- 6.	Faculty/Staff web page.	148
Figure D- 7.	Human Resources Home web page.	149
Figure D- 8.	HR Employment web page.	150

Figure D- 9.	Proposal Components web page.	151
Figure D- 10.	Events Calendar web page.	152
Figure D- 11.	Students web page.....	153
Figure D- 12.	Provost Learning web page.....	154
Figure D- 13.	Student Information Online web page.	155
Figure D- 14.	Panther Central web page.	156
Figure D- 15.	Digital Library web page.	157
Figure D- 16.	SIS Home web page.....	158
Figure D- 17.	Dental Medicine Bulletin web page.....	159
Figure D- 18.	Financial Aid Parents Info web page.....	160
Figure D- 19.	Technology Home web page.	161
Figure D- 20.	Computer Accounts web page.	162
Figure D- 21.	Cyber Career Counselor web page.	163
Figure D- 22.	Jurist web page.....	164
Figure D- 23.	Nationality Rooms web page.	165
Figure E-1.	Student Life at Pitt web page.....	166
Figure F-1.	Objective Landmark Value Components for All Pages.....	169
Figure G-1.	Subjective Landmark Value Components.	170

ACKNOWLEDGEMENTS

This dissertation would not have been possible without the support of others. I want to thank my advisor, Dr. Stephen Hirtle, who sparked my interest in navigation and cognitive science and encouraged me in my goal to pursue excellence in teaching, my dissertation committee for their guidance and encouragement in my research process, and my family and friends who have supported me in so many ways.

1. INTRODUCTION

1.1. Background

The ability to navigate is a fundamental process in all large-scale space (Golledge, 1999a). To function in any environment, a person must develop a cognitive representation of the environment. This cognitive map represents an understanding of the organization of the space and enables one to make navigational choices and plan routes. Inherent in navigation is the use of landmarks. Landmarks are prominent features of the environment, and are used in orientation and navigation. While the use of landmarks in navigation is an obvious part of one's daily life, it is also significant in animal navigation in physical space (PastergueRuiz, Beugnon & Lachaud, 1995) as well as in human navigation in virtual spaces such as hypertext and virtual reality environments (Darken & Sibert, 1996).

The World Wide Web (WWW) or a web site in the WWW is often considered an information space. In this conceptual space, web pages are conceived of as locations within the space, and movement between them is accomplished by selecting previously defined hyperlink connections to other web pages. Choices made in the information space are based on an understanding of the connections that allow movement between different sections, parts or places in the space, in a manner similar to choices along a physical route. As in physical space navigation, choices about path and direction may be made using previous knowledge, landmarks and goals.

Two research areas motivate the examination of landmarks in WWW navigation: (1) the importance of landmarks in navigation of various types of physical, electronic and conceptual spaces, and (2) the multi-dimensional complexity of the WWW. A hypertext web such as the WWW is an intricate network of nodes connected by links. The complexity of a web structure comes in part from the great flexibility in how the system is used and navigated. Users navigate by selecting links to follow or use search tools to identify potential links. The burden falls on the user to develop an understanding of the space in order to locate the desired nodes and information.

Research on complex information spaces such as the WWW indicates that the complexity of the space may hinder and disorient the user, unless design considerations related to navigation are taken into account. Work in the rapidly growing fields of information architecture and web design show many possible approaches to site architecture, with a concentration on content organization, navigational tool design and understanding the user audience (Rosenfeld & Morville, 1998; Fleming, 1998).

While the concern over disorientation in hypertext is an old question (Conklin, 1987), there is current interest in examining issues of spatial cognition both in information space and in the design of WWW sites to improve ease of navigation (Dodge & Kitchin, 2001; Neveitt, 2000; Neerincx, Lindenberg & Pemberton, 2001; Chi et al., 2003). Addressing navigation issues from the perspective of spatial cognition means applying our understanding of how a cognitive representation of physical space is developed. By examining the WWW as an information space, knowledge of the development of a cognitive map can be used to understand the characteristics and features of the space that facilitate navigation and orientation.

There are several studies that have examined cognitive issues in navigation of electronic spaces, showing the importance of design features or tools that aid in the understanding of the complexity of the space. For example, previous work found that visual cues such as background color or texture could help users differentiate regions or neighborhoods within an information space in such a way as to aid navigation (Hirtle, Sorrows & Cai, 1998). The research of Ark et al. (1998), based on the principle that people use landmarks to orient themselves and organize information, demonstrated that the addition of landmarks to a 2D iconic interface and a 3D representation of information objects resulted in faster times for locating objects in an information space. Additional work is described in sections 2.4 (Navigation in hypertext) and 3.3 (Landmarks in electronic environments).

1.2. Problem Statement

Complex spaces present challenges for navigation that may be alleviated in part by landmarks. There is a clear understanding of the role of landmarks for navigation in physical space, but this has not been easily applied to information space. In addition, research indicates that landmarks can exist in information space, but there is a need to clarify how to define and identify them. This dissertation addresses the problem of how to define the characteristics of

landmarks in information space, and how to evaluate elements as landmarks based on those characteristics. This research is an initial attempt to validate, in information space, the model of Sorrows and Hirtle (1999) that identified three categories of landmark characteristics as cognitive (later called semantic), visual and structural.

This dissertation addresses these issues in examination of the University of Pittsburgh website. It applies concepts and techniques from research on navigation in physical space to the investigation of characteristics of landmarks in the World Wide Web. For example, experimental tasks in which subjects recall navigation paths or draw maps of a particular space are common in physical space navigation research; however in web research it is much more common to examine actual user navigation paths using web server log files. This research uses both objective measurements of elements in the information space and collection of experimental data from subjects.

1.3. Significance of Work

This research analyzes a new framework of the characteristics of landmarks and applies concepts and techniques from research on physical space navigation to the investigation of landmarks in electronic information space. Research on landmarks in physical space has analyzed a broad range of individual features of landmarks such as size, visual physical qualities, cultural importance and familiarity. Similar research in electronic spaces has been limited to structural aspects such as link structure and access frequency. The presentation of a theory of three categories of landmark characteristics provides an excellent framework in which to examine a range of features in electronic information space. If support is found for the tri-partite theory of visual, structural and semantic landmarks, then there is the potential to use this theory to re-examine the range of landmark characteristics identified in physical space by Sadalla, Burroughs and Staplin (1980) and others. A theory that applies to landmarks in both physical and conceptual spaces can provide a significant step in how navigation and design issues are addressed in a variety of environments.

Success in finding correlation between the objective and subjective measures of landmark quality may be used as a basis for the ability to objectively evaluate web pages for landmark characteristics. In addition, an understanding of the characteristics of landmarks resulting from this work may be used to develop guidelines to create web sites that are easier to navigate.

Consideration of the multiple dimensions of landmark characteristics may enable more effective web page design as well as web site designs based on more effective landmarks.

1.4. Outline of the Dissertation

This dissertation presents a method for evaluating potential landmarks in an electronic information space. It investigates the key characteristics of landmarks in an electronic information space and examines the use of landmarks during navigation of that space. This chapter mentioned some of the areas that provide support for this work. Chapter 2 describes research on navigation in a number of domains as well as key research specifically on landmarks in physical space and in electronic environments. The last three sections (2.8 to 2.10) present an analysis of the essence of landmarks across environments, describing the tri-partite theory of Sorrows and Hirtle (1999). Following this summary of related work, Chapter 3 describes the research methodology used in two experiments examining landmark characteristics and use, and defines the hypotheses examined in the present work. It includes an examination of how landmark quality may be examined by both objective and subjective measures and provides a definition of these measures. Chapter 4 presents the results of the experimental work and the analysis is provided in Chapter 5. The final chapter discusses broader issues in this research area and describes potential future work.

2. LITERATURE REVIEW

2.1. Overview of Navigation Research

One's ability to live and work in any environment is dependent on forming an understanding of that environment and on knowing the consequences of actions in that space. There are two basic requirements to being able to function in a space: to know where you want to go and how to get there. The first part is addressed by concepts such as idea generation. The second part, the action of "getting there", is the core of navigation. The concept of navigation is second nature, as many forms of navigation are used in the interaction with different kinds of spaces. Navigation is basically the process of planning, describing and controlling a course through an environment. That environment might be a natural, physical space, an electronic information space, or one of many variants.

The field of Information Science is concerned with many kinds of information spaces, including databases, document spaces, telecommunication networks, and even cognitive spaces. The ability to function in these information spaces starts with an information need and the generation of an appropriate question that may, when answered, address that need. The user decides what kind of information is sought or what part of the system to go to. These issues are addressed by a variety of domains within Information Science, including decision-making, information seeking behavior, and cognitive psychology. After determining a goal, which may be as general as exploration of the space, various actions allow the user to navigate the space in different ways. For example, the ability to manipulate information in a database involves planning and describing the retrieval of certain items of information from the space, and controlling how and what information is retrieved: This planning, describing, and controlling of actions in an environment can be examined in terms of navigation.

Navigational features, such as how the user selects choices in a system, changes to a different mode, or opens a new space in the environment, can be examined in terms of both human factors issues and computational issues. Navigational features are implemented in interfaces for information systems in many different domains such as databases and other storage

and retrieval environments, document processing systems, document management, and hypertext. Hypertext structures provide a uniquely interesting environment within information science to examine navigation because of the potential complexity of the networks.

One feature important to navigation in all spaces is a landmark. Landmarks are commonly understood to be prominent features of the environment that provide one with orientation information for navigation. This basic understanding is used throughout this work, and landmarks are examined in detail beginning in Section 2.5.

This chapter presents a literature review of research on navigation and landmarks. It begins with a presentation of key concepts of navigation, starting with background on some of the general theories of navigation. Next, navigation and navigational issues from a variety of different domains are presented. These sections provide the relevant background from some of the many disciplines that discuss the topic of navigation, to create a foundation for discussing navigation in hypertext. Section 2.4 provides further background in examining the domain of hypertext, defining its structure, identifying the navigational issues in working with complex graph structures, and introducing the role of landmarks in this domain. Sections 2.5, 2.6 and 2.7 review the research on landmarks in both physical and electronic environments, and the chapter closes with two sections on the common aspects of landmarks in these two environments.

2.2. General Theories and Frameworks of Navigation

This section discusses key issues related to navigation, providing definitions and structures that will be referred to throughout the paper. These include types of knowledge, types of environments and components of navigation. This review will provide a picture of navigation issues and will lead into the topic of landmarks.

2.2.1. Types of Knowledge

One key factor in the ability to navigate through an environment is spatial knowledge of that specific environment. There are several similar definitions of spatial knowledge. Hirtle & Hudson (1991) describe place, route, and survey knowledge as a model that has received general acceptance. Place knowledge is knowledge of specific locations. Route knowledge is knowledge of a sequence of locations, for example along a route or path. Survey knowledge is knowledge of the general interrelationships between locations in a region and is said to incorporate Euclidean relationships. Most researchers agree that in a new environment people first acquire

place knowledge, and then learn specific routes between those places, and finally develop survey knowledge of the area (Hirtle and Hudson, 1991). This theory is used in section 3.1 to discuss the role of landmarks in the creation of a cognitive map.

2.2.2. Types of Environments

There are two basic environments for navigation, open terrain navigation and networked space navigation. Examples of open terrain navigation are found in orienteering, military planning, robot navigation and animal navigation when movement is restricted by the method of locomotion but not by specific pathways and can include avoidance of obstacles. The VIBE system (Korfhage, 1995) provides a type of open terrain navigation in that it restricts method of action (how to place a point of interest or “POI”), but does not restrict where the action takes place. Virtual reality interfaces to information systems may also allow open terrain style navigation, for example when the navigation mode allows the user to pass through the surfaces or walls defined in a virtual reality space. Examples of networked space navigation include any animal, human or robot navigation performed along predefined pathways such as automobiles on a road, rabbits in a burrow, or hypertext.

2.2.3. Components of Navigation

Reasons for navigation can include exploratory travel returning to a known location, trying to reach a novel goal, or navigating to a familiar goal (Allen, 1999; Kuipers, 1983). Either as a goal or as an unintended side effect, people also identify objects during navigation, including recognizing landmarks, finding categories or clusters of objects, and obtaining information about objects in the space (Benyon and Höök, 1997).

Navigation in physical space (or active navigation) has been defined as consisting of a cognitive component, often referred to as wayfinding, and a motor component which is physical locomotion (Darken, 1997). The motor component of navigation refers to the actual locomotion involved. Locomotion is behavior or movement from one point to another that is guided by one of the senses, most typically vision. The role of the perceptual systems will be discussed briefly in a later section, but the main focus of this paper is on the cognitive components of navigation.

Navigation involves the use of both environmental cues, which are inherent to the space and navigation tools, which are augmentations to the space (Darken, 1997). An example of an environmental cue is a landmark in the space, such as the Cathedral of Learning, which is a 42-

story building on the University of Pittsburgh Oakland campus. Examples of tools that augment the space are maps, route descriptions, and signs.

2.3. Navigation Research Domains

This section reviews the literature on navigation from the perspective of a variety of domains. The fields of architecture, biology, graph theory, neural networks, robotics, and virtual reality have been selected. Since much of our understanding of electronic environments is based on metaphors that help us compare those environments to ones that are more familiar, examining the nature of navigation in different domains is important background work. These areas may provide correlates or may suggest new ways to apply a spatial metaphor of navigation to understanding navigation in electronic environments more completely. The domain of hypertext is central to this work, and is examined in detail at the end of this chapter, beginning with section 2.4. Other fields, such as orienteering and cartography, are not explicitly examined here. Perspectives from navigation in the physical world are drawn instead from architecture and biology. Cultural or individual differences in navigation are set aside for the present, in order to develop a clearer understanding of the core ideas, and will be addressed where they significantly affect the findings or descriptions.

2.3.1. Architecture

Lynch (1960) in the *Image of the City*, provides an outline of the key issues in understanding a physical space from the point of view of urban design which has remained a well respected foundation for research in this area. He examined the design of cities by focusing on the physical and spatial attributes of the real world environment. He defined a set of principles that make an environment one which we can navigate physically and which enable the observer to create a mental map of the environment. Lynch (1960) presents “legibility” as an essential characteristic of a city. Legibility refers to “the ease with which its parts can be recognized and can be organized into a coherent pattern” (Lynch, 1960, p. 2-3). Lynch recognizes legibility to be particularly important in environments at the urban scale. Humans, like all animals, must be able to categorize and develop an understanding of the surrounding environment, and large, complex environments provide unique challenges to developing mental maps.

Lynch describes five building blocks of cognitive mapping of a space: landmarks, nodes, paths, edges and districts. Paths are the pathways that are potentially available in a space. Edges are boundaries or barriers that run between two regions. In an urban environment, districts are medium to large sections of a city which are identified by some common characteristic. Both nodes and landmarks are strategic points of reference, the key difference in Lynch's definition being that the user of the space can enter into nodes, but not into landmarks. Nodes are strategic points such as the convergence of multiple paths. Landmarks may be close or distant, and used to identify or locate one element in the scene of the environment. These elements are intricately related in the environment. Each is important in the unique aspect of the picture that it provides the observer. Landmarks in particular will be examined extensively in Chapter 3. Legibility of an environment can provide a sense of security and confidence, and therefore provides the visitor of that space with an opportunity to experience the environment more intensely (Lynch, 1960).

Passini is an architect who was strongly influenced by Lynch's work in urban design. Passini (1984, 1996) integrates examination of the wayfinding process with observation of the information processing performed by humans during navigation. Although some (e.g. Darken, 1997) define wayfinding as the physical locomotion in the navigational process, Passini (1996) emphasizes the cognitive component, and defines wayfinding to include plans, execution, and information processing. Wayfinding requires consistent use and organization of the sensory information perceived from the environment (Lynch, 1960).

Passini (1984) observes that a person's understanding of a space is based on the use of certain organizational principles. Starting with the basic Gestalt principles, Passini describes the use of geometric laws and geometric forms in making sense of a space. He describes how people will impose an organization in a complex environment even if it does not fit well. Passini (1996) supports the theory of selective perception as a functional necessity. That is, our environment contains many more elements and details than we can immediately or effectively comprehend, and we must use selective perception and consistency to reduce information overload. Passini's work has focused on different types of environmental meanings: functional, socio-cultural, and sensory, as they affect architectural design and the navigational processes.

2.3.2. Biology

The domain of biology provides background from neuroscience on the processes of the brain involved in navigation and from animal behavior on the use of landmarks by different

animal species. Research in biology and neuroscience has explored issues of navigation in general, and the use of landmarks in particular, in a variety of animals. Some of the documented uses match key ways in which humans use landmarks as well: 1) Landmarks are documented as keys in how animals recognize scenes when displaced to a different but familiar location, 2) they are used as “beacons” that animals aim for during navigation, and 3) they are used in a consistent manner during repeat traversal of a path or area (Collet, 1996).

Current research indicates that brain activity during navigation orientation activity is primarily centered in the hippocampus. Extensive research (particularly on rats) on the hippocampus has indicated that several categories of cells provide different functions during navigation that uses landmarks (see, for instance, Gothard, Skaggs, Moore and McNaughton, 1996). Some hippocampal cells encode the location in a “fixed spatial frame”, and others encode location “with respect to different reference frames associated with the task-relevant, mobile objects” (Gothard et al., 1996). Different cells of the hippocampus fire based on specific locations, on proximity to a goal or landmark, and on locations at special places such as a start or end. For example, place-field cells fire at their maximum rate based on the animal’s relative location to a set of distal landmarks (Zipser, 1986). The place-field cells are unable to function in this way if too many landmarks are removed, however they can properly compute location in an artificial/imaginary world if all landmarks are presented as scaled equivalently to a different size (Muller, Kubie, & Ranck, 1983 as cited in Zipser, 1986).

Most animals use visual landmarks, and auditory, tactile and odiferous landmarks are possible as well. The perceptual systems in many animals studied are very different from the human perceptual system: for example, wasps and honeybees are believed to use a very basic image matching process. This process requires the landmark to be in a standard retinal position for comparison, and related in size by some parameter (Collet & Rees, 1997).

A variety of this research appears to correlate with what is known about human navigation, though there are some significant differences. This area of research contributes to the understanding of the use of landmarks in navigation in that the strategies of (1) recognizing scenes, (2) biased detours, and (3) aiming at beacons (Collet, 1996) are all ways that humans also use landmarks. Also, it is likely that there is also a distribution of the navigation functions and the use of landmarks in different sets of (hippocampal) cells in the human, just as in rats and other animals (Touretzky and Redish, 1995). Despite the similarities, the process of basic image

mapping in insects described by Collet and Rees (1997) is primitive with respect to our current understanding of the human visual perceptual processes.

Experimentation with rats shows generally accepted conclusions that the number, identity and geometric configuration of landmarks are important to the animal's search performance. The geometric representation created by the rat contains information about each of these aspects of landmarks in the search space (Prescott, 1997). The development and contents of this representation are similar in some basic ways to the development of cognitive maps of space by people, which is described in section 3.2.

Some animals such as ants have been observed to use an understanding of global landmark-landmark relationships in addition to the basic isolated landmark-goal relationship (PastergueRuiz, Beugnon & Lachaud, 1995), which illustrates a complex understanding of the environment. Landmark-goal relationships provide information about the distance and direction of the goal from the landmark and may be specific to a particular path or goal. Landmark-landmark relationships, on the other hand, provide more global information about the location of objects in the environment and their orientation to one another. This information provides a more complete understanding of the space, such as is acquired in survey knowledge. Some studies have been performed on both humans and animals, which provides more direct data for comparison. For example, pigeons and humans show the same pattern of knowledge in learning about sets of landmarks on a path to a goal: Within a set of landmarks, subjects (animal and human) had the most extensive knowledge about the landmark that was closest to the goal (Spetch, 1995).

The biology behind navigation provides significant background information about processes that are modeled in neural networks and robotics, and provides the biological basis for understanding the cognitive processes involved in navigation.

2.3.3. Neural Networks Models

As described in the previous section, there are place-field cells in the hippocampus of the rat that fire at a maximum level to indicate a location at a particular position relative to landmarks. The ability to determine this location within an environment is essential in creating an ability to navigate there. Zipser (1986, p. 435) developed a model for navigation that examines the configuration of distal landmarks and relates this visual, structural information to specific characteristics of the place fields. The Beta model is a neural network model is based on

the geometrical constraints of computing a location based on visible features of the distal landmarks (Zipser, 1986). This model requires the use of a minimum of three landmarks, so that the estimation of distance to each landmark provides a single point of intersection where the location would match the correct distance for all three landmarks.

In a typical neural network implementation, the model has a two-layer architecture, in which the output of the first layer provides the input to the second layer. In Zipser's Beta model, each unit in the first layer deals with one landmark and one place field. The neural network measures the difference between the values of each location parameter and the value of the same parameter that is stored in the unit (Zipser, 1986). In the real world that this is modeling, the sensory system would provide detection and description of the landmarks, which is represented by the input to the first layer. As the descriptive information becomes available and the landmarks are detected in the scene, the output of this first layer would increase. The second layer in the model produces output similar to the place-field neuron, providing the data necessary to define the location and shape of the place fields. Zipser (1986) notes that an interesting indication of the flexibility of this model, or class of models, is that it can be implemented using either a step function or a Gaussian function for the first layer computation, each with success, depending on what features are measured. However, the model is dependent on the accuracy of the distance estimates to compute the location.

Zipser's (1986) analysis of the Kubi, Muller & Ranck (1983) study indicates that rats do not use absolute distance as a measure of landmark location. This conclusion is based on data of the rat's correct output of the place-field neurons during a set of dilation experiments that would cause the animal to misjudge the distance of a set of landmarks. Zipser (1986) proposed examining this problem by using either the visual angle of observation or the retinal area of the image as alternatives to absolute distance that would not be affected by incorrect scaling due to dilation.

Two alternatives to Zipser's Beta model of navigation are a direction-averaging model and an extension to the Beta model. McLaren (1995) proposed one modification of the Beta model that allows for orientation information to be deduced from only two landmarks, when a third landmark is too far away to accurately estimate the distance. The direction-averaging model deals with the conditions of being presented with conflicting cues presented by an array of

landmarks (Cheng, 1994). The model is based on data from experimentation on landmark-based search by pigeons that found distance and direction to be determined independently.

In a different neural network approach to examining how environmental information is gained, Ghiselli-Crippa and Munro (1994) examined the emergence of global structure from local associations. They implemented a neural network to examine certain characteristics of the internal representations, or hidden units, in a system based on the encoder architecture. The application task was to map locations to sets of neighboring locations. The purpose of the network design was to show that a network trained to learn local spatial associations could result in a system in which the hidden units exhibit global spatial properties of the environment. One method implemented in this system used a subset of patterns in one phase of training until these landmark locations were learned, and then including all training patterns in a second phase of training, with some restrictions. This process was combined with the introduction of noise into the system. Landmark learning was one of three methods discussed that had a positive influence on how global spatial structure can emerge.

2.3.4. Graph Theory

Graph theory provides a theoretical basis for examining structural factors of the environment such as the complexity of spaces, the organization of node clusters, and the placement of landmarks. It can be applied to research in both real and electronic environments, such as path configuration and hypertext design. A brief example of research in this area is discussed here.

Navigation in a networked space, such as that by a robot between particular node locations, can be examined from the perspective of graph theory. If certain nodes are made to be distinctive “landmarks,” then the robot can locate itself within the space by measuring its distance to different nodes. In physical spaces with visual landmarks, this location would be done visually. Two key differences of a graph space are that there is really no concept of direction and there is no visual detection of landmarks (Khuller, Raghavachari and Rosenfeld, 1996). Khuller, et al. (1996) examined issues surrounding the number and location of landmarks required in a graph space to enable a robot to determine its location within the space and to navigate. In analyzing graphs, it is often useful to look at the subset of the graph that consists of the landmark nodes, called the “metric basis” of the graph. The minimum number of landmarks

required for uniquely determining a position in the graph is called the “metric dimension” of the graph.

2.3.5. Robotics

One of the key features of autonomous systems (natural and artificial) is the ability to develop a model of the environment through exploration and direct experience. Such a model is task-independent, and provides a structure for understanding and using the environment during subsequent tasks. It is easily adaptable to environmental changes or additional information, and serves as a reference for recognizing meaningful states or situations. In humans and animals, this model is often referred to as a cognitive model. In artificial systems, this model is part of the database with which the system reasons.

Artificial navigation systems such as mobile robots require a knowledge base of information about their physical environment in order to navigate, just as humans and other animals do. For mobile robots using machine vision, physical and visual landmarks are often central in training in new environments. The perception of sensory-motor stimulants during use or exploration of the environment provides an autonomous system (e.g. a human being or a mobile robot) “the ability to learn and adapt internal task-independent models of environments incrementally” (Bachelder and Waxman, 1995, p. 267).

Mobile robots frequently need to re-calibrate their position within an environment because they may lose their position due to environmental factors such as wheel slippage (Lin and Tummala, 1997). Most robotic systems, whether designed either for indoor navigation or for outdoor navigation, use some type of vision system along with map or landmark information to enable re-calibration of the robot. This type of navigation by autonomous or semiautonomous mobile robots (or robotic vehicles) is referred to as perception-based navigation (Bhanu, 1994).

Traditional approaches to robots learning about environments are based on mainstream artificial intelligence and machine vision techniques. These systems typically use a reconstructive approach to perception (Bachelder & Waxman, 1995). Reconstruction is based on perception using stereopsis or ultrasonic range-finding and then using time-intensive triangulation calculations to construct 3-D representations including grid models and convex polygons of the environment (Ellis, 1993; Bachelder & Waxman, 1995). This approach has a number of problems, including 1) the mechanical problems of time intensity and high error rates of stereopsis and ultrasound, and 2) the fact that the resulting representations require a search

through a huge number of possibilities for valid paths, and 3) that the representations do not hold up to variance in the environment (Ellis, 1993).

An alternative to the traditional approach is a relational map-making approach. This general theory is based on learning the adjacency relations between distinctive places. This learning requires long-term, global map-making strategies to be separated from short-term, local map-making. The long-term planning provides for map-making of the large-scale space, made up of the distinctive places. The local planning provides the robot with knowledge of its immediate surrounding, so it can plan routes and avoid obstacles within distinctive places. These methods can be image-based, in which case they rely on images and require clear viewing conditions, or they can be landmark-based, which rely on the ability to identify specific distinguishable, unoccluded structures in the environment (Bachelder and Waxman, 1995). Both methods have trouble in dynamic environments (Bachelder and Waxman, 1995).

Relational map-making is based on work on learning state transitions for task-independent models (Bachelder and Waxman, 1995). Task-independent models can be developed using symbolic, statistical or neural network approaches. A symbolic approach might use classical planning techniques from artificial intelligence to construct a simple graph model of a deterministic environment using a finite state machine. Statistical methods often use a variant of first order Markov models (e.g. hidden Markov model). Each variant of the model has its strengths, but also its limitations. Recurrent neural networks bridge at least one of the gaps in the other systems: they can be used to implement a finite state machine, and can also be used to “approximate complex dynamical systems.” (Bachelder and Waxman, 1995, p270), but they present additional challenges in design, analysis and user training.

Bachelder and Waxman (1995) have proposed that the neural system for map-making or learning an environment consists of two neurocomputational architectures, one for place learning and recognition, and one for action consequence learning. This system recognizes the importance of a variety of inputs. Landmark features provide the vision system with cues to identify what objects and features exist in the environment. The eye and head positions provide egocentric data of where the landmarks are located in relation to the body. The place learning architecture component uses these two types of information, *what* landmarks are in the environment and *where* they are located, in a process of unsupervised place learning. The action consequence learning component incorporates knowledge of *when* actions occur as well as the

consequences of those actions, through measurements of rotation, translation and heading from the locomotive system.

Similar to the Beta model extension discussed in section 2.3.3 (McLaren, 1995), Lin and Tummala (1997) described a unique landmark-based navigation technique for a mobile robot that required fewer landmarks (one or two rather than the traditional three) for determining position and orientation. The system estimated robot position using a camera and a set of navigational landmarks made up of specially designed geometrical patterns and was implemented in a semi-structured indoor environment (Lin & Tummala, 1997). Open terrain environments are often considered ‘semi-structured’ because although they don’t include specific road paths, they do contain landmarks that help to structure the space. In this particular system, each landmark consisted of two concentric circles. The main, large circular disk was black on a white background, and the second, smaller disk was white, and was placed a short distance in front of the large black disk (Lin & Tummala, 1997). Comparing the center axis of the second disk to the center axis of the main landmark disk provided data for determining the angle of direction between the robot and the landmark. That is, when the robot camera was aimed directly at the landmark, the vertical center axes would match.

The circular disk pattern was chosen for the landmarks for three reasons: 1) a circular disk projection can be easily approximated by an ellipse using the elliptical Hough transformation technique, 2) the circular disk pattern is not easily confused with objects in a typical indoor environment consisting of rectangular and polygonal shapes, and 3) circular patterns are more robust than polygons to noise and occlusion (Lin & Tummala, 1997).

The choice of patterns that are clear in form and can be readily recognized reflects some of the important characteristics of landmarks that will be examined in chapter 3. Robotics is an important area of research with significant research departments dedicated to it. This section has reviewed some of the key elements in navigation and landmark research in robotics, but has barely touched on the breadth or depth of this field.

2.3.6. Virtual Reality Environments

Navigation in a virtual world can follow many of the features and requirements of navigation in physical spaces. However, in a virtual world, several different perspectives on a scene can be used, providing different types of information about where the user is located. In

addition, the methods and rules of navigation may be different. Just as it is important in a physical space, the importance of knowing where you are in the space, or the ability to “self-locate,” is viewed as one of the central concepts in navigating in a virtual world (Steuer, 1992). Some systems create the sensation of the user moving through the environment, while others are created such that the user manipulates an object, such as a figure or a car through the space.

Other types of virtual reality environments such as the Virtual Brain (1999) require the user to manipulate an extension of the body, such as a probe, in a virtual environment using clues in both the virtual and physical spaces. In this situation, the user manipulates one or more objects in the virtual space based on the landmarks corresponding to the physical space, and this manipulation causes a change in the physical environment.

Darken & Sibert (1996) examined wayfinding in large virtual worlds by applying key concepts from physical space navigation. Their work is the basis for a great deal of navigation research in virtual environments. Recent work by Modjeska & Waterworth (2000) used these ideas to create a series of prototype virtual reality worlds with designs that specifically addressed navigation needs and concerns. They concluded that with training and careful design, virtual reality worlds could provide excellent navigational environments for information visualization and information retrieval tasks in which users could easily apply skills from wayfinding in physical space. Extensive research in this rapidly changing area can be examined from these references.

2.4. Navigation in Hypertext

2.4.1. Hypertext

Hypertext is an information technology that provides for the creation of a multi-linear, network structure of information. In hypertext, links are used to connect blocks of words, images, or sounds. These building blocks are self-contained units; they have been termed “lexias” by Landow (1997), and are similar to the concept of the node. Multiple paths link the different lexias in the collection, resulting in a web structure. Since readers of a hypertext or users of an information system implemented with a hypertext structure make individual choices about which links to follow, each person’s path may be different, and the hypertext structure is therefore appropriately called multi-linear, although it is frequently referred to also as nonlinear.

Hypertexts are commonly dynamic spaces; that is, they are subject to frequent changes in the structure and content provided.

The original vision for hypertext came from Vannevar Bush (1945) who described a system called Memex, based on microfilm and mechanical projection equipment. The term hypertext was coined by Ted Nelson in 1965. Nelson developed one of the early electronic hypermedia systems, Xanadu (Nelson, 1990). Hypertext can be implemented using electronic media or conventional media (e.g. books, video, and microform). Hypertext structures are used in a variety of application areas; the most commonly considered hypertext now is the World Wide Web. Other applications include help and documentation systems, electronic reference works with cross-referencing, CASE tools for software engineering, tourist information systems, interactive fiction and poetry, and translation of previously written literature. For further background, Begoray (1990) provides an overview of hypertext issues and some of the early hypertext information systems, and Landow (1997) is prominent in the area of hypertext literature, fiction, and translation system development.

The domain of hypertext has unique characteristics that make the study of navigation particularly interesting. The domain of hypertext is an example of the environment that this research is focused on: a dynamic, electronic networked information space such as the World Wide Web. These sections provide additional background material, with a focus on the type of environment that will be the basis for study. Issues of navigation, maintaining orientation in an environment, and specific types of navigational tools are discussed.

2.4.2. Navigation

Using hypertext is different from reading a traditional text such as a book. Conventional representations of information are made up of pieces of information which are presented in a specific structure, typically a linear structure for texts such as fiction, or hierarchical one for texts such as technical manuals (Begoray, 1990). References to other sections within a text or citations to separate texts create network connections (Begoray, 1990). A typical book is traditionally navigated sequentially, from beginning to end, or used by selecting from a hierarchical table of contents to locate a beginning point for linear reading. Hypertext, on the other hand, is comprised of many interlinked blocks of self-contained information that compose a complex network or graph structure. Begoray (1990) describes this complex network structure as one of the defining characteristics of hypertext.

Hypertext requires that the user (or reader) actively select links to follow to examine different lexias. This active selection results in a type of navigation, as the user examines portions of the graph structure and then at each decision point makes a choice as to which portion of information to examine or which link to follow next. The structure of hypertext creates a natural environment for associative browsing.

Readers of a traditional text typically have few choices in navigating the text: to read in a sequential fashion, to search forwards or backwards, or possibly to use an index or table of contents. Hypertext users on the other hand, have many choices available to them at each decision point (within or at the end of each block of information), and may choose among the links leading to new lexias or may choose to backtrack to the previously explored lexia. Just as in reading a traditional text or using another type of information system, there are many possible user goals. For example, the user's goal may be to find an answer to a specific question (whether the existence of the answer is known or unknown), to look for unknown items which might be interesting, or to understand what information is available. The user may seek to achieve this goal in a directed or an undirected way. There are numerous combinations of goals and navigational styles employed by users in a complex information space. Studies such as those by Canter, Rivers and Storrs (1985) and McAleese (1989) have shown that these navigation strategies include:

1. Scanning – skimming to cover a broad area without much depth
2. Browsing – following paths by association, selecting links by personal interest
3. Searching – seeking an answer to a specific question or an explicit goal
4. Exploring – determining the extent of the information space
5. Wandering – purposeless, unstructured exploration of the space

The WWW differs from classic hypertext, which is a closed corpus with predefined link structure. An evolving hypertext space such as the WWW is made more complex for navigation because of its size and its dynamic nature. Units can be modified or added to the space; links can be changed, added or deleted. The complexity of a hypertext space and the flexibility of how the user of a hypertext utilizes the space cause some challenges for the user in maintaining an egocentric orientation.

Chen et al. (2001) developed a model of the elements in a hypermedia system and their functions that they have used to examine World Wide Web pages. In this model, the basic elements on a web page are examined as having one or more function: informative, decorative,

navigational, or interactive (Chen et al., 2001). Informative elements convey content information, while decorative elements are graphics or color elements for purely decorative purposes. Navigational elements are items such as links, or navigation bars or lists. Examples of interactive elements are a search field, a button for a display option, or a submission button. This model may be helpful in examining user perception of different web pages (see section 4.3.1.2).

2.4.3. Maintaining Orientation

Orientation is important in hypertext because of the types of navigation used. The intricate network of links defined between nodes creates a tremendous flexibility in how users browse and use the system. This flexibility puts a burden on the user to understand clearly not only the task or goal, but also to have or to develop an understanding of the hypertext space in order to locate the desired nodes and information (Conklin, 1987). In addition, during browsing a user may often come across multiple interesting paths, and must then keep track of those link locations to return to and follow later. At that later time, the user must recall a sufficient amount about the node and its location to be successful. Shum (1990), in work on spatial cognition, noted the importance of the human desire and need to impose structure on an environment. Ark, Dryer, Selker, and Zhai (1998) emphasized that people must impose structure to help simplify the vast amounts of information received from the world.

The need to maintain one's orientation in a hypertext system has been widely discussed in terms of the problem of disorientation (Conklin, 1987; Foss, 1989; Utting & Yankelovich, 1989). Disorientation is described as including a number of different aspects. It refers to the problem of the user not knowing where he/she is within the space, confusion about where to go next, or not knowing how to locate something that is believed to exist in the space (Conklin, 1987; Utting & Yankelovich, 1989). Definitions also include the sense of not knowing the "boundaries of the information space" (Utting & Yankelovich, 1989, p. 61). Kim and Hirtle (1995) and Smith and Wilson (1993) provide extensive summaries of the literature on navigation problems and the cognitive constraints hypertext navigation places on the user.

The focus of work on disorientation is on the design of the information technology, and not on the content of the hypertext system. That is, a reader may become disoriented because he or she does not understand the content presented, or because the defined links do not perform as expected. From the point of view of hypertext authoring guidelines, Landow (1997) points out that it is important to recall that "we all know that readers often experience confusion and

disorientation simply because they fail to grasp the logic or even meaning of a particular argument” (Landow, 1997, p. 116). Design features such as navigational aids and clues to orientation cannot help to correct confusion caused by the content.

Landow (1997) discusses whether disorientation is really a “crippling and dis-enabling” problem. Landow argues that disorientation may be positive in the humanities. Disorientation of the reader in literature, such as Dante’s *Divine Comedy* “has one important parallel to that encountered in some forms of hypertext: in each case the neophyte or inexperienced reader finds unpleasantly confusing materials that more expert readers find a source of pleasure” (Landow, 1997, p. 117). Landow also describes the reactions of students using a hypertext system for browsing, searching, and exploring, as defined above, which indicate that experienced users often enjoy the surprises of exploring an unknown space and not always maintaining a sense of bearing or a clear direction. Some users even criticized a system because they were unable to become disoriented, with links always leading to overview style nodes rather than to nodes with more content and new links to explore (Landow, 1997). One user commented “Orientation devices such as these explained and categorized links rather than allowing me to make my own connections and categories” (Landow, 1997, p. 121-122). There is little other research supporting Landow’s recent perspective on positive aspects of disorientation, but the observations are worthy of consideration. These comments indicate the importance of allowing the user to discover the structure of the space individually, and of providing a rich enough space that various structures or landmarks can emerge, as they are relevant.

2.4.4. Navigational Aids

Many hypertext systems provide additional navigational aids to help users orient themselves in the information space or help them find the desired information. This section briefly describes possible search facilities, guided tours, bookmarks, and various graphical overview maps. These aids are particularly helpful in very large hypertext systems. Even if these aids do not locate the specific goal for the user, they may place the user in an appropriate “region” of the hypertext space, which the user can then explore by following links.

The most basic navigational aid provided by most systems is a method of returning to the previous node after following a link. One common navigational style in exploring a hypertext space is to follow several links in one direction, and then use the “back” facility to return to a known space before venturing again along another path.

One type of navigation is the use of search facilities to locate potential target information. Search engines typically allow the user to enter a specific word or phrase and the system locates matches for different lexias in the hypertext space. Rather than following individual links from the current place in the hypertext to the desired lexia, a search engine produces a list of lexia sites that the user can navigate to directly. There are many variations in search engine structure and algorithms, including keyword and full-text search, and the application of fuzzy matching and different similarity measures (Salton, 1989). Any method can be implemented to provide direct navigation to a result from a search request.

Guided tours also depart from the basic browsing strategy of following hypertext links. Guided tours are designed by the author of the content, and present a pre-defined pathway through the hypertext space. In guided tours, navigation is basically limited to “next” and “previous” commands, but the user can also follow links which take them out of the tour, and then return to the tour structure later if they wish.

Many hypertext systems allow the user to specify bookmarks or favorite locations within the hypertext system. The bookmark list allows the user to select an item and return to that place within the hypertext space with a single selection, rather than needing to recall and follow the individual links to access the element.

Another category of navigational aid in hypertext is graphical overview maps. Main topics in a hierarchy can be listed in a table of contents display or arranged in a graphical site map. Fisheye views show the area of interest in detail and surrounding areas in less detail (Furnas, 1986). Focus+context views specifically show the immediate neighborhood of a node of interest, and then show only landmark nodes in the surrounding area (Mukherjea and Hara, 1997). Various filtering and analysis methods have been used to select the subset of nodes to represent.

Ingram and Benford (1995) applied the features defined in Lynch's analysis of urban environments (see section 3.2) to three dimensional information visualizations of the WWW. Their prototype, LEADS, included implementation of landmarks, districts, paths, nodes and edges, adapted from the urban analysis. The LEADS prototype was also implemented in a virtual reality document database system called VR-VIBE. While not all of the features seemed applicable initially, it is an interesting attempt to apply physical space navigation research to an information space.

2.5. Overview of Research on Landmarks

Differentiation has long been recognized as an important issue in navigation. While people learn about an environment, they develop place knowledge, route knowledge, and survey knowledge about the space (see section 2.2.1). The development of each of these types of knowledge is dependent on the ability to differentiate between specific locations and objects in the environment, for example to recognize one element as distinct from its surroundings or to recognize what is or is not part of a particular route. This knowledge contributes to the creation of a cognitive map. The ability to navigate along specific routes or to varied locations in an environment is enhanced by the formation of a spatial representation of that environment. Landmarks play a key role in the creation of such a cognitive map. Landmarks are fundamental to navigation in both open terrain and networked spaces (see section 2.2.2), and this chapter discusses how they are used in both real and electronic environments.

For the purpose of this work, the possible distinction between the terms *landmark*, *reference point* and *anchor point*, is not significant. In critical related works, these terms are often used similarly (Golledge, 1999). I will consistently use the term landmark for this concept. To examine the nuances in depth, refer to Sadalla et al. (1980) and Rosch (1975) for discussion on reference points, and see Couclelis et al. (1987) for discussion of the anchor point theory.

The next section presents a discussion of landmarks in the real world, examining first the characteristics of landmarks described by Lynch in architecture and urban design, which is the basis for much of the more recent work. Section 2.7 presents an extensive look at definitions of landmarks in the electronic world. The chapter ends with some conclusions about the relations between landmarks in these two environments.

2.6. Landmarks in the Physical World

This section discusses landmarks in the physical world. The characteristics defined by Lynch are presented first, because his research is widely respected and was used as the basis of other research that will be presented. Other research on the characteristics of landmarks as elements of a spatial representation is presented. The second section describes the work of Siegel and White on the development of spatial representations, and discusses additional research on wayfinding and cognitive map development. Section 3.2.3 presents several perspectives on how landmarks are used, as well as problems encountered in situations that lack

distinctive landmarks. Finally, the section summary describes one perspective on the difference between focusing on landmarks as objects or landmarks as they are used in navigation.

2.6.1. Landmarks as Elements of Spatial Representation

Landmarks are physical elements in the environment that serve as point references for an observer (Lynch, 1960). These elements can be singled out from among other elements in the environment based on particular characteristics that are “unique and memorable in the context” (Lynch, 1960, p. 78). According to Lynch, the following characteristics make a landmark easier to identify (Lynch, 1960, p.78-79):

- Clear form
- Contrast with its background or surroundings
- Prominence of spatial location

Clear form refers to an understandable shape or outline of the structure that helps distinguish it from its surroundings. Lynch found that the concept of contrast with the background is the most significant factor, and is a very broad concept. The background might be something immediate such as a large building behind a smaller object, or the background might be something distant, such as the sky, or it might be a conglomerate of objects, such as an entire city providing a backdrop to one particular, unique building. The distinction between an object and its surroundings might also be based on features like the age or cleanliness of the objects in view, or differences in size, shape or position of a building. Prominence of spatial location refers to the position of a landmark in a location that is significantly noticeable, or that stands out in its visibility. These characteristics can be related to basic information theory. Shannon’s measure of information (Shannon and Weaver, 1962) is based upon the differentiation or unexpectedness of the data received. Lynch’s characteristics of a good landmark relate to Shannon’s characteristics of (high) information content.

Lynch also found that buildings or objects could be landmarks based on the meaning or use of the object, even if it lacks the visual characteristics described above. A building may be a landmark because of an activity that takes place there, or from an understanding of its historical significance or because of a unique name.

Appleyard (1969) applied Lynch’s concept of imageability to analyze the distinctive form of frequently remembered buildings in a city in Venezuela. He outlined three categories of attributes that might contribute to a building’s imageability: attributes of physical form, of

visibility, and of use and significance. Physical form attributes included the size, shape, surface color and texture, the amount of movement in the area, the quality of construction, the presence and clarity of signs, and the amount of separation between buildings or between a building and its surrounding vegetation. The visibility of different buildings was measured based on three attributes: the number of people viewing it, its location at a major decision point, and its centrality to the main line of view. The final category, use and significance, included attributes such as the building's frequency of use, the uniqueness of its function, and its symbolism, such as its economic or political importance. He found that form, visibility and significance attributes correlated with map and trip recall.

In work based on Appleyard (1969), Lynch (1960) and Rosch (1975), Sadalla, Burroughs and Staplin (1980) tested several operational definitions of reference points and landmarks. They examined potential characteristics of landmarks in physical space and found the results clustered into six categories: size, familiarity, cultural importance, physical quality and complexity, label quality and finally visual salience. Rating scales completed by subjects were used to examine the importance of these features. A score of referentiality was developed based on the theory that locations are understood in terms of their relation to reference locations or landmarks. They found that referentiality was predicted by familiarity, visibility from a distance, domination of nearby places, and cultural importance.

Lynch defined landmarks primarily by their characteristics, and others follow similar descriptions with slight modifications. Still others focus on how landmarks are used. Landmarks are typically thought of as elements that are commonly recognized by many people; but some landmarks may be salient for an individual in a way that is equivalent to these more widely recognized features. In a survey of landmarks in Columbus, Ohio, Golledge and Spector (1978) found that half of the places rated as most familiar were significant on an individual level and half were commonly identified throughout the group. Whether relevant to individuals or large communities, landmarks are clearly elements that are used to develop an understanding of the organization of space through their use as reference points. The development of this spatial representation is discussed in the following section, and then the use of landmarks is taken up.

2.6.2. Development of Spatial Representations

Siegel and White (1975) proposed a sequence of development of knowledge about an environment. The basic theory described the acquisition of landmark, route and survey

knowledge in the development of a spatial representation of the space, and the process was divided into a series of stages that are similar for children and adults.

According to Siegel and White, the first stage of environmental learning for children is noticing and remembering landmarks. In the second stage, children can act using the landmarks as reference points in their decision making process. At this stage, children develop knowledge of sequences and use landmarks to follow a route. The third stage is the development of “mini-maps” which consist of clusters of landmarks, although Siegel and White (1975) admitted that this might not indicate a significant change in the child’s understanding. The fourth stage is the child’s formation of an objective frame of reference. Finally, the fifth developmental stage is characterized by the use of both routes and an objective frame of reference to develop survey maps or survey knowledge of the environment.

Siegel and White emphasized the similarity between the development of children’s environmental knowledge and the way an adult develops knowledge about a new environment. One difference is that adults are more facile at understanding different points of reference, and therefore the development of an objective frame of reference is not as significant.

Blades (1991) analyzed a number of studies of the development of wayfinding abilities. He summarized the findings of several studies of children learning routes in unfamiliar environments. One common treatment in a number of experiments was the amount of information in the environment pointed out by the experimenter. In each of these cases, subjects performed better when they were advised to look around at each decision point, or when characteristics of the environment were mentioned along the route, such as during normal conversation. Both of these resulted in better observation of landmark features in the environment. Blades concluded that the studies of both children and adults show that individuals recognize and learn landmarks and routes before developing more complete configurational or survey knowledge of the environment.

Heft (1979) examined adults’ wayfinding in a novel area, and compared the strategies used in the natural environment to the same space with the addition of distinctive artificial landmarks. The route was through a Biological Reserve, and the two conditions were the natural environment with no changes and the addition of artificial landmarks such as plastic lawn ornaments and flags. The task was to re-trace a route including over twenty intersections and describe reasons for each decision point choice, after having been shown the route once.

Subjects were successful in performance in both conditions, but the results showed that different types of landmarks were used during navigation. Artificial landmarks placed in the environment were used as landmarks whenever present, but when the artificial landmarks were not present subjects used natural features as landmarks and used other strategies such as remembering specific turns to re-trace the route.

2.6.3. Use of Landmarks in Navigation

Landmarks serve multiple purposes in wayfinding. Sadalla, Burroughs and Staplin (1980) summarized three uses of the term landmark. “The term has been used to denote discriminable features of a route, which signal navigational decisions, discriminable features of a region, which allow a subject to maintain a general geographical orientation, and salient information in a memory task” (Sadalla et al., 1980, p. 516). The distinction between the use of landmarks at decision points and the significance of landmarks in defining regions is echoed in the work of others. There is wide support for two main uses of landmarks: as an organizing concept for a space or as a navigational aid (see Golledge, 1999; Presson and Montello, 1988).

Anooshian (1988) suggested that focusing on either component could yield different approaches to distinguishing landmarks from ordinary places. Anooshian argued that the spatial cognition literature emphasizes the usefulness of *memory as an object*, which she terms the ‘place perspective’. In contrast, a ‘procedural perspective’ focuses on navigational procedures, and *examines memory as a tool*. In one experiment subjects were encouraged to use one or the other approach in learning about an environment. The experimental results showed several differences in memory for features and route decisions. Anooshian (1988) concluded that the two aspects of the definition of landmark might reflect different ways of learning and remembering environments. These two main uses of landmarks are examined next.

2.6.3.1. Landmarks as organizing concepts

Landmarks can represent a cluster of objects at a higher level of abstraction or scale (Golledge, 1999). In this way, a landmark can become an anchor for understanding local spatial relations (Couclelis et al., 1987; Golledge, 1999). Stephens and Coupe (1978) found that people made systematic errors in judging the relationship between two locations because the judgment was based on the relationship of a superordinate category (e.g. the classic example: 'Is Reno, Nevada east or west of San Diego, California?'). The key conclusion of the study was that there is a close association between a landmark and the area in which the landmark is located. Related

work has shown that landmarks can affect the organization of one's cognitive map both through direct use of an environment and through learning a map of the environment (Hirtle and Mascolo, 1986; Hirtle and Jonides, 1985).

Presson and Montello (1988) gave examples of two types of organizing concepts. Symbolic landmarks may symbolize a whole region, as the Eiffel Tower in Paris has come to be a symbol of the city. One of the key features of landmarks as spatial reference points (Sadalla et al., 1980) is that there is an asymmetric relationship between landmarks and other objects. Non-reference points are judged or remembered with relation to the landmark, but landmarks are not recalled in relation to non-landmarks.

2.6.3.2. Landmarks as navigational aids

Heth et al. (1997) described two ways that landmarks are used in navigation. Landmarks are the memorable cues that are selected along a path, particularly in learning and recalling turning points along the path. Landmarks also enable one to encode spatial relations between various environmental objects and features, enhancing the development of a cognitive map of a region. This distinction can also be described as two types of relationships: landmark-goal relationships, where landmarks are cues along a path to a goal, and landmark-landmark relationships, which provide a global understanding of the environment (PastergueRuiz, Beugnon & Lachaud, 1995). Landmark-goal knowledge may be particularly used in active navigation, and landmark-landmark knowledge may be most essential in orientation activities.

Lynch described the use of sequences of landmarks to navigate in a city. In this situation, seeing one landmark may trigger the person to anticipate another detail or landmark that should be seen next. Stronger associations are created for landmarks that are located at the intersection of paths or at other points that require decisions. Additional details are usually remembered about landmarks located at decision points, and at the final destination. This begins to show some of the interaction between the elements that Lynch defines: Paths, edges, districts, nodes and landmarks. Lynch emphasizes that each is important, both individually and together with the others, in creating an understandable environment (Lynch, 1960, p.84).

While some landmarks are prominent and visible from a distance in a large area, many are used more locally, and are recognizable or useful only in a close context. The familiarity with an area seems to affect whether a person uses distant landmarks much or whether they tend to use more local ones (Lynch, 1960). Tall, distant landmarks were found to be most

recognizable by their top from a distance, with the actual location of the base of the landmark often uncertain even to people familiar with the area.

The essential nature of landmarks is discussed in the field of architecture by Passini (1996), who points to the lack of distinctive units in labyrinths as the central reason people do not understand the spatial layout of labyrinth environments. Although many people enjoy solving paper maze puzzles, there is a certain level of discomfort at being placed in a physical maze and attempting to navigate out. One can seek to envision the layout of turns during navigation through a maze, but the uniform appearance of the labyrinth's walls or tunnels make each intersection look like a variation of the last. This shows the importance of differentiation in composing a cognitive understanding of an environment. People's sense of disorientation in mazes is essentially caused by the lack of landmarks (Passini, 1996). The problem of navigating in an environment that does not contain distinguishing landmarks is also illustrated in observations of the King Saud University in Saudi Arabia. Abu-Ghazze (1996) described the unusual appearance of the University, which contains a set of buildings with nearly identical internal floor plans and external architecture and connected by a system of standardized covered walkways. His research showed that the uniform visual nature of the space made the environment extremely difficult and frustrating to learn and navigate (Abu-Ghazze, 1996). Passini (1996) describes how people depend on landmarks, and will impose organization in a complex environment, even if the imposed structure does not match the environment well.

Landmarks provide key information about the relationships of locations, objects and paths, and are used in developing spatial representations of the environment and in active navigation and orientation tasks. The distinction of what a landmark looks like and how it is used is also discussed in the next section on electronic environments.

2.7. Landmarks in Electronic Environments

Many of the same cognitive principles described above in terms of navigation of physical space are also involved in navigation in electronic environments. The main focus in this section will be on hypertext. The importance of research on differentiation of nodes in hypertext has been recognized but not fully explored in applications of hypertext research, although it may be one of the keys in identifying or creating landmarks in hypertext systems. In hypertext, differentiation can be achieved through use of a variety of visual cues such as background color

and texture, and may provide essential cues to the user of his location in a local area (Kim and Hirtle, 1995, Hirtle et al., 1998). Cues that help the user differentiate regions or neighborhoods within the information space will help the user navigate in the space. Rich electronic environments must contain landmarks for effective navigation, just as real worlds do.

Most research uses a basic definition of landmarks as “prominent nodes.” Some researchers agree that “prominent nodes can be provided which can always be accessed from anywhere in the system” (Balasubramanian, 1994), and some separate the definition of local and global landmarks (Glenn and Chignell, 1992, Nielsen, 1995). Although there are many ways the term landmark is used in electronic spaces, and many of them are compatible, agreement on an algorithmic process of determining landmarks would make their definition and use much clearer.

2.7.1. General Descriptions

The “home” page or introductory screen in a hypertext system is frequently implemented to be a landmark, which is accessible from many parts of the web. The home page may contain certain characteristics such as an overview of the environment, a particular layout, or a unique graphic design. In addition to the home page, an author may design local landmarks within the system to provide a useful and usable structure for the user. All of these landmarks may be prominent in an overview diagram or other navigational aid.

Nielsen (1995) describes the opening screen of a particular hypertext system as the “landmark” for that system. The screen is accessible from any page within the system through the use of an icon displayed on each page, and it provides the user with basic features such as help information and the ability to reset options (Nielsen, 1995). This feature of a “home” page that is directly accessible from all other pages in a particular system provides one of the most common landmark designs in hypertext systems.

Particular graphic patterns, color, or other design features may also uniquely identify pages or sections. Nielsen (1995, p.24) describes a page which is framed with a unique graphic design “to emphasize its landmark status and to differentiate it from the text of the other nodes.” Again, the landmark node is available from all other screens through an overview diagram.

Glenn and Chignell (1992) suggest that landmarks may be selected for various reasons. Nodes might be selected as landmarks because they serve as basic categories (Rosch et al., 1976). Rosch found that people name particular items, or prototypes, most often as examples for a category. These prototypes are easier for people to learn, recognize and categorize. Glenn and

Chignell (1992) recognized that these characteristics are effective attributes for landmarks as well. Landmark nodes may be ones that are particularly well connected and visually memorable, as indicated by Lynch (1960) and described extensively in section 3.2.1. Nodes may also be selected as landmarks because they have high salience. Salience refers to a striking point or significant feature. The node may be conspicuous in its different use of graphical design or layout as described above, or it may be prominent in relation to other nodes.

Glenn and Chignell (1992) described landmarks as part of a symbol system which is both visual and cognitive, and in which the visual and cognitive functions are intricately tied. They described visual landmarks as aspects which are particularly visual in nature, such as icons which people remember as pictures and as well as by their particular location on the screen. Cognitive landmarks have two origins that may or may not coincide: the landmarks may be inherent to the information based on the relationships between the topics, or the author may design the landmarks through the design of the hypertext structure. Glenn and Chignell (1992) gave examples of visual and cognitive landmarks in the physical world: a medieval church steeple provided a visual marker, and the church itself served as a cognitive marker for events that occur there or for the beliefs of the church. While the descriptions are useful, the methods for selecting landmarks in electronic space are not specified.

The question of how to select which nodes should be indicated as landmarks is a complex one, and is at the core of this research. Nielsen (1995) simply stated, for example, that an author might start by looking at the connectivity levels of each candidate node. He indicated a preference for author-selected landmarks over automatically selected landmarks based on connectivity. The next section presents approaches to computational analysis of the structure of web sites and the selection of landmarks.

2.7.2. Computational Approaches to Landmark Definition

People have attempted to measure aspects of both document content and hypertext structure as a way to determine what nodes are landmark nodes in a hypertext. Glenn and Chignell (1992) developed an algorithm that involved measuring the co-occurrence of index terms within the documents of the hypertext, and calculating the second-order connectivity for each term as a measure of landmark quality. First-order connectivity for a term was defined as the number of terms that were directly related to it. Second-order connectivity for a term was defined as the number of terms that could be reached by following two relational links. They

used this process to construct a landmark view for any node, allowing the user to indicate the desired distance from the selected node.

In the context of the World Wide Web, Mukherjea & Hara (1997) defined a landmark as a node that is important to the user because it helps to provide an understanding of both the organization and the content of that part of the information space. They identified three features that determine the importance of a node:

- 1) connectivity, based on the number of links it contains and the number of nodes within the site that link to it,
- 2) frequency with which the node, site or page is accessed, and
- 3) depth, based on the URL.

Mukherjea and Foley (1995) defined one method of showing the user certain contextual information about the location of the current node in relationship to important or landmark nodes. The algorithm they developed to identify landmarks relies on structural analysis of the web network based on four factors: the node's outdegree, the node's indegree, its second-order connectedness, and its back second-order connectedness. See Figure 2-1 for a graphical illustration of these measures.

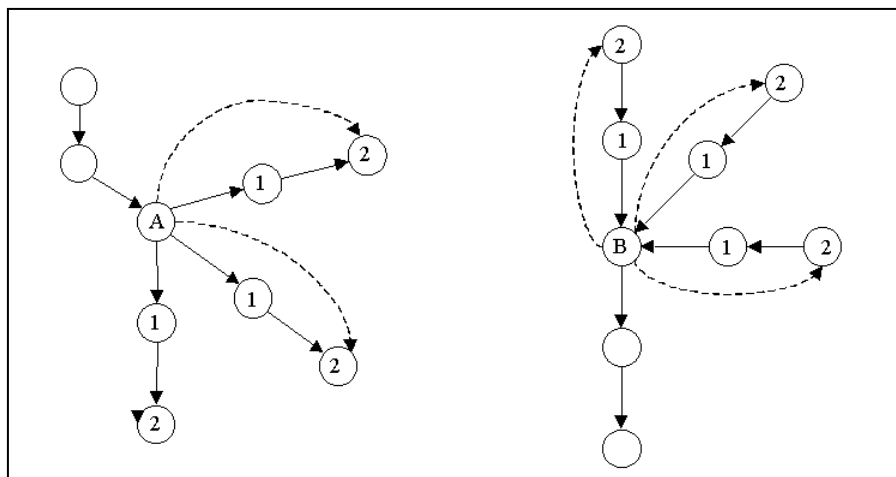


Figure 2-1. Connectivity into node A and out of node B.

First-order (outdegree) and second-order connectivity to node A and back first-order (indegree) and back second-order connectivity to node B. (based on Glenn and Chignell, 1992 and Mukherjea and Hara, 1997)

The outdegree of a node is the number of nodes that can be reached from the node by following only one link. The indegree of the node is the number of nodes that can reach that

node by following only one link. The second-order connectedness is essentially the same as described above, the number of nodes that can be reached by a node when following at most two links. And the back second-order connectedness is the number of nodes that can reach the specified node by following two links. The close landmark nodes are identified for a given node (e.g. the current location of the user), and then the procedure is called recursively to identify more distant landmarks related to those found in the first round. Mukherjea and Foley (1995) conclude that contextual analysis such as the use of access counts is also needed to determine the importance of a node.

2.8. Summary of Landmarks in Physical and Electronic Spaces

This review of the role of landmarks in real and electronic environments concludes with a summary of how the characteristics described in physical space can be applied to electronic spaces, and vice versa. A number of these characteristics are summarized briefly here to present how they can be applied to landmarks in both physical and electronic spaces. For more extensive descriptions and discussion of other characteristics, see Sorrows and Hirtle (1999).

Clear form refers to a clearly identifiable shape or outline of a structure (Lynch, 1960). In a document, an example of a landmark with clear form is a table of contents section. In hypertext, an example of clear form is a navigation bar at the top or bottom of a page.

Singularity or sharp contrast with its surroundings means a landmark visually conspicuous in its environment (Lynch, 1960). Examples of contrast in physical environments include distinction of a building from its immediate surroundings due to difference in size, shape, position, age, or cleanliness. In hypertext, a landmark node might contain a unique graphical design or layout that is not used in other parts of the system.

Prominence: A building may be *prominent* that is visible from many locations (Lynch, 1960). In a hypertext environment, prominence could correspond to the inverse of the length of the URL. For example, the main University of Pittsburgh site, <http://www.pitt.edu> is a logical place to begin for a wide variety of information about the University.

Accessibility is a characteristic defined in research in electronic space and indicates that a variety of paths are available to access a given page (Nielsen, 1995). In hypertext, a page such as a home page may be designed to be accessible from other pages in the system by placing an icon or text link on each page or may be made accessible through an overview diagram. In the

physical world, a building located at an intersection of multiple roads (e.g. near a rotary) would have this characteristic.

Content or unique meaning or use can also render an object a landmark (Appleyard, 1969). A building or site may be a landmark because of its historical significance. In electronic spaces, a landmark based on meaning might be a FAQ page, which provides an expert's answers to "frequently asked questions" and gives the expert's recommendation on links to follow for more information or assistance on a given topic. Another example of a landmark that is based on meaning, not visual characteristics, is an official company site, such as the official Volkswagen site, <http://www.vw.com>, as opposed to the many other sites that also offer information on VWs.

Prototypicality: An object may also be characterized a landmark because of its *prototypicality*, or how typically it represents a category (Sorrows and Hirtle, 1999; Glenn and Chignell, 1992). This concept is based on the research of Rosch et al. (1976) who found that people name particular items, or prototypes, most often as examples for a category, and that these prototypes are easier for people to learn, recognize, and categorize. In the same way that a single prototype often represents an entire category (Rosch et al., 1976), prototypical landmarks can gain a new meaning as well.

2.9. A Theory of Landmarks

The research on navigation and landmarks presented in this paper shows how research developed in one domain may be applicable to another. The previous discussion provides the background to discuss landmarks and to judge various characteristics. Sorrows and Hirtle (1999) proposed that it would be helpful to structure a theory that defines different types of landmarks. Further characterization will enable evaluation of a landmark's effectiveness. As described in section 3.3.1, Glenn and Chignell (1992) working on hypermedia design, suggested that landmarks may be particularly visual or cognitive, and that these functions are intricately tied as parts of a symbol system. Sorrows and Hirtle (1999) proposed a theory of landmarks that defines three types of landmarks. The characterization of landmarks by type provides a structure to evaluate why and when a specific landmark is effective and what purpose it fulfills. This section examines their expanded notion of landmarks and its relevance across domains.

Sorrows and Hirtle (1999) proposed three categories of landmarks: visual, cognitive (or semantic), and structural. Each of these types of landmark characteristics plays a different role in navigation.

A *visual landmark* is an object that is a landmark primarily because of its visual characteristics. As described above (sections 3.2.1 and 3.3.1), these may include the features of contrast with surroundings, prominence of spatial location, and visual characteristics that make the landmark particularly memorable. An example of a visual landmark in physical space is the Washington Monument in Washington, D.C. It is a visual landmark because visual characteristics of the monument such as its unique size, shape, and location are what make it memorable. An example in electronic space could be a typical university home page, where photographs and layout are unique when compared to all the subordinate pages. Visual landmarks have remained prominent in the study of landmarks in physical space, and have been similarly recognized in virtual reality environments, but have not received as much attention in other electronic environments such as the World Wide Web.

A *cognitive or semantic landmark* is one in which the meaning stands out. A feature or object may be a semantic landmark because it has typical meaning in the environment. Semantic landmarks may have unusual or important content, and the content might contrast with the surrounding content. Cultural or historical importance may also identify a semantic landmark. An example given by Sorrows and Hirtle (1999) is a resident advisor's room in a dormitory, which although structurally and visually identical to the other rooms, may form a landmark for the dormitory residents based on its unique status. Semantic landmarks are often personal, are learned through interaction with the space, and can be missed by those not familiar with the environment, unless they have some distinguishing markings or signage.

A *structural landmark* is one whose status comes from its role or location in the structure of the space. This class of landmarks is likely to be highly accessible, and may have a prominent location in the environment. One example of a structural landmark in hypertext is a navigation bar. The designer of a hypertext document or space makes certain decisions about the links between documents and at what points visitors are expected to enter the site or its sub-sections; these decisions can result in structural landmarks. In addition, designers create certain spaces, intersections or aspects in the environment that can be considered structural landmarks, for example a downtown plaza.

These three categories, visual, semantic, and structural, can be used to describe the variety of landmarks in both real spaces and electronic information spaces. However, these categories are not discrete. An object in the environment may have landmark characteristics in more than one category. For example, a home page in a hypertext environment could be a structural landmark because of the importance of its location within the network, and a visual landmark because of its unique graphic layout. Sorrows and Hirtle (1999) asserted that the strongest landmarks in an environment would be landmarks in terms of all three elements: visual, semantic and structural. They also note that research on landmarks in physical space has included very little about the structural dimension, and that research in electronic spaces has focused around the structural dimension and has minimized interest in the visual dimension. Examining each of the three categories more completely in both environments may provide insights into the design and use of landmarks.

2.10. Current Research Directions

Several recent papers have examined the role of landmarks in greater detail. Raubal and Winter (2002) apply the Sorrows and Hirtle (1999) definition of landmark components to physical space to evaluate landmarks for selection in wayfinding instructions. Features for the visual, semantic and structural components included similar or corresponding features as in the present work, as well as a few different ones. Size and color were evaluated in contrast to surrounding elements, and building shape was examined for its deviation from a basic rectangle shape. Cultural or historical importance and the presence of signage were used as semantic features. Structural features examined were based on Lynch's (1960) elements of nodes, boundaries and regions. While features were evaluated primarily on a local level, Winter (2003) continues that work in further examining the effect of incorporating urban features at a distance into the landmark computation.

The proposal that landmarks in both physical space and electronic space have similar components is one aspect of spatial information that has been examined in both environments. Hochmair & Raubal (2002) look at the how the nature of cognitive maps influences decisions made during wayfinding in both physical and Web space. The issues related to local and global landmarks in navigation in virtual environments also continue to be an active area of research (Steck and Mallot, 2000).

Berendt and Brenstein (2001) examine methods of analyzing individual navigation patterns on the Web and creating visualizations from navigation paths. Pilgrim and Leung (2001) use links and page attributes to identify potential landmarks and present them in an expanding outline format. Pearson and Schaik (2003) examine how presentation issues within a page affect the usability of that page and they present an analysis of concerns with regard to web page design guidelines.

Elias (2003) uses data mining to select appropriate landmarks from spatial databases based on landmark components in order to enhance car navigation system instructions. The initial work shows potential for two methods in landmark selection, one using decision trees and the other a clustering algorithm. The stored characteristics allow the selection of landmarks based on relative uniqueness in the environment.

There is also a variety of work in more theoretical aspects of landmarks and navigation. Worboys (2001) presents a formal analysis of distance relations such as ‘nearness’ in physical space navigation using three different logic approaches. The ability to understand the placement of objects in relation to others in an environmental space affects one’s ability to navigate.

3. RESEARCH METHODOLOGY

3.1. Overview

This dissertation addresses the problems of (1) how to identify web pages that may be used as landmarks within a web site and (2) what characteristics of a web page make it a likely candidate to be used as a landmark. Data was collected regarding characteristics of a variety of web pages, and two experiments were performed to gather subjective data related to landmark characteristics and use. The two experiments took the form of questionnaires. The first experiment collected data from subjects on how they would find certain information on the University of Pittsburgh web site and on their knowledge of the web site. The second experiment gathered subjective data of landmark characteristics by asking subject to rate a number of web pages with respect to essential visual, structural and semantic landmark characteristics.

The experiments are based on research methods that have been used to examine physical space navigation and landmarks. Appleyard (1969) examined the characteristics of buildings and structures in a city and found that attributes of physical form, visibility and significance correlated with map recall and trip recall by subjects (see section 2.2.1). Sadalla, Burroughs and Staplin (1980) had subjects rate eighteen potential landmark characteristics, which were found to cluster into six categories: size, familiarity, cultural importance, physical quality and complexity, label quality and finally visual salience. A score of referentiality or the potential of a location to be used as a landmark was predicted by familiarity, visibility from a distance, domination of nearby places, and cultural importance (see section 3.2.1).

This dissertation uses two measures to evaluate characteristics that may identify potential web pages as landmarks within the University of Pittsburgh web site. These measures are the objective landmark quality value and the subjective landmark quality value. The objective landmark value of a web page is computed using direct measures of the page, including the number of links to and from the web page, the percentage of text and graphics present, and a lexical analysis of similarity to the subject profile. The subjective landmark value of a web page

is computed from the responses of subjects regarding its visual characteristics, its importance, and their familiarity with the location of the page. These measures are described in detail in section 3.3.

The remainder of this chapter is structured as follows. The hypotheses are presented next, in section 3.2, followed by the detailed definition of the objective and subjective landmark values in section 3.3. The methods that were used for the two experiments are described in section 3.4.

3.2. Hypotheses

Four hypotheses regarding the dependent variables of subjective and objective measures of landmark quality were tested. Data to examine hypotheses 1 and 2 was collected in the first experiment on recall and use of web pages. Data to address hypotheses 3 and 4 was collected in the second experiment on subjective web page characteristics.

Hypothesis 1: The pages easily recalled will have stronger objective landmark characteristics.

A higher objective value is expected for the pages that are remembered and listed by the subjects in path descriptions than for pages in general on the web site. The null hypothesis is that the objective value will be less or that there will be no difference in the objective landmark quality value between the web pages that are indicated by the subject and other pages on the web site.

Hypothesis 2: The pages easily accessed will have stronger objective landmark characteristics.

Knowledge of the exact URL or having a bookmark for a page provides easy access to the page. It is expected that the subject will indicate knowing the URL or having a bookmark more frequently for pages that have a high landmark quality rating. The null hypothesis is that there is no difference or that the objective value will be lower for pages that subjects would access by knowing the URL, using bookmarks or lists of links, using a search engine, or accessing in another way. Data is collected in the first experiment by asking subjects to write the URL or indicate if they have a bookmark for each of ten pages within the University of Pittsburgh web site.

Hypothesis 3: The subjective landmark quality value and the objective landmark quality value will be positively correlated.

It is expected that web pages that have a high objective value will also have a high subjective value of landmark quality. The null hypothesis is that there is not a significant positive correlation between the two values (it is either negatively correlated or not correlated at all). The subjective value is computed from the subject rating data collected in the second experiment. The objective value is computed from objective measures of the page characteristics as described below.

Hypothesis 4: The degree of correlation can be shown to be dependent on the weights selected for combining the visual, semantic and structural characteristics in the objective measure.

The objective landmark value is defined as a combination of the structural, visual and semantic characteristics combined with equal weights of 1/3, 1/3 and 1/3. This hypothesis examines the degree of correlation between the objective landmark value and the subjective landmark value when the weights used to combine the components of the objective value are changed. A variety of weights are used to analyze the effect on the correlation.

3.3. Measures of Landmark Characteristics

This section defines in detail the computation of the objective and subjective landmark values which are analyzed in this dissertation. The objective value is made up of separate measures of structural, visual and semantic characteristics that can be made of any web page, independent of experimental data. The subjective value is a set of rating questions designed as a questionnaire for subjects, and it also encompasses these three types of characteristics. The measures for both the objective and subjective values are based on concepts from the literature, though not all of the measures have been used previously. Development of each of the measures is described in the following sections.

3.3.1. Objective Landmark Value

This research tested a proposed set of objective measures of landmark quality in order to provide an evaluation of the characteristics of a web page that make it likely to be used as a landmark. These characteristics are aspects that may make the web page easy for users to remember or to access, or that provide the means to accomplish certain tasks such as the ability to go to the relevant next choices. For the objective value, research indicates that the measure of

structural characteristics alone may not be a sufficient measure of landmark quality in a web site (Mukherjea & Foley, 1997). Lynch's (1960) analysis of environmental images at the city level describes three components: identity (identification), structure, and meaning. Lynch (1960, p.8) observed, "It is useful to abstract these for analysis, if it is remembered that in reality they always appear together."

As described in Sorrows & Hirtle (1999) and discussed in the review of literature on landmarks (see section 2.5), three categories of landmark characteristics have been defined: structural, visual, and semantic. Specific measures are presented for each category, and these are combined to give the *objective landmark value* used in this research. Equal weights of 1/3 each are used to combine the components, with the expectation that they are equally important in this measure. The equation used to compute the objective landmark value is given in Figure 3-1. The measures of structural, visual and semantic characteristics that are combined to generate the objective landmark value in this equation are described in the next three sections.

3.3.1.1. Measure of structural characteristics of landmarks

Structural landmark characteristics refer to importance related to the role or location in the structure of the space. Structural characteristics can include the connectivity of the node, meaning the number of links into and out of the node, the relative frequency with which the node is accessed, and the position of the node in the web space of the particular web site. In work on interpreting web pages for wireless applications, Chen et al. (2001) report high precision at identifying index pages versus content pages using the number of links into and out of the page. Mukherjea & Foley (1995) also developed an algorithm to compute landmark quality based on the link connections, which was revised by Mukherjea & Hara (1997) to include not only the links into and out of the page, but also the access frequency and position of the page within the web site. For this study, structural landmark characteristics were measured using a variation of the algorithm developed by Mukherjea & Hara (1997), which was presented in the literature review (section 2.7.2).

The computation of the structural component of landmark quality used the length of the URL (depth), the number of links going into the node and the number of navigational links leading out of the node. Each of these parts of the structural component was assigned weights based on previous literature.

Objective Landmark Value =

$$\begin{aligned} &= \text{StructuralComponent} / \text{MAX}[\text{StructuralComponent}] \cdot \text{STRUCTURALwt} \\ &\quad + \text{VisualComponent} / \text{MAX}[\text{VisualComponent}] \cdot \text{VISUALwt} \\ &\quad + \text{SemanticComponent} / \text{MAX}[\text{SemanticComponent}] \cdot \text{SEMANTICwt} \end{aligned}$$

where

StructuralComponent

$$\begin{aligned} &= (\text{LinksOut} / \text{MaxLinksOut}) \cdot \text{LinkOutWt} \\ &\quad + (\text{LinksIn} / \text{MaxLinksIn}) \cdot \text{LinkInWt} \\ &\quad + (\text{DepthWt} / \text{depth}) \end{aligned}$$

Visual Component

$$= (\text{GraphicsArea} / \text{TotalArea})$$

SemanticComponent

$$= (\text{WordMatch} / \text{TotalWordsOnPage})$$

Variables

LinksOut = number of links out of the web page

MaxLinksOut = maximum LinksOut on web site

LinksIn = number of links into the web page

MaxLinksIn = maximum LinksIn on web site

Depth = depth of the URL

GraphicsArea = area of web page covered by graphical elements

TotalArea = area of web page

WordMatch = number of words on the web page matching subject profile

TotalWordsOnPage = total number of words on the page

Weights

LinkOutWt = 0.4

LinkInWt = 0.4

DepthWt = 0.2

STRUCTURALwt = 1/3

VISUALwt = 1/3

Figure 3-1. Equation to compute the objective landmark value.

The measurement of each of the components of the structural components is described here. The depth of the URL is computed based on the top page of the domain. For the University of Pittsburgh web site, the top page is at www.pitt.edu, which is assigned a depth of 1. Other page depths are computed by adding one to the depth for each level added to the URL. The addition of a unit to either the host name or to the directory structure is an additional level. For example, the admissions page (www.pitt.edu/admissions.html) and the library page (www.library.pitt.edu) are at depth 2, and the Department of Information Science and Telecommunications page is at a depth of 3 (www.sis.pitt.edu/~dist).

The number of links going into the web page was determined using the Google search engine (www.google.com). The Google search was first restricted to the University of Pittsburgh domain, www.pitt.edu. Then the advanced search feature for links was used to look for the links into a specific page. For example, from the restricted site search, the command “`link:http://www.sis.pitt.edu/~dist`” was entered to return the number of pages within the University of Pittsburgh site that refer to that page. The number of links out of the web page was found by examining the source document for the web page and counting the number of “`href`” tags.

3.3.1.2. Measure of visual landmark characteristics

The second component of the objective landmark value is the visual component. Visual landmark characteristics refer to visual features or presentation. Lynch (1960) examined the visual quality in the design of cities, a quality that he denoted “legibility,” meaning how easily one can recognize the parts of what is seen and organize that view into a pattern. Visual presentation for a web page includes the presence, layout and design of text, graphics and white space. For the present study, the visual landmark characteristics examined are the amount of space used on a web page by text, graphics, links and white space.

Chen et al. (2001) have classified elements of hypermedia systems according to four functions: informative, decorative, navigational and interactive, as described in section 2.4.2. Using this model, the key visual elements can be examined as (1) the informative elements that convey content, which will generally be text on a web page, and (2) the decorative elements, which will be graphics, logos, interactive elements such as search fields, or other graphic elements.

Based on the visual characteristics described by Chen et al. (2001), a simple measure was developed to compare the visual qualities of different web pages. The measure of visual landmark characteristics is a ratio of the area of the page that was covered by decorative elements to the total area of the page. For this study, the visual component was defined as the percentage of page area covered by graphics. Graphics included decorative elements such as photographs, drawings, symbols, logos or other graphic designs. Interactive elements were measured as graphics, but their unique qualities were not reflected in any other manner by this measure. The areas covered by each type of element were measured on a printed copy of the entire web page. The total area of the page was measured similarly.

3.3.1.3. Measure of semantic landmark characteristics

The third component of the objective landmark value, the semantic component measures landmark characteristics related to the content of the web page and the importance of the content. As described by Sorrows and Hirtle (1999), semantic landmarks may be more personal, for example they may be significant only for a particular group.

For this study, semantic landmark characteristics were measured using lexical analysis to compare a web page to a typical subject profile. Lexical analysis is used extensively in information retrieval research to analyze the similarity between two documents, between a document and a query, or between a user profile and a document. Subjects for the experiments were undergraduate students at the University of Pittsburgh, so the content of the web page for "Student Life at Pitt" (Appendix E, formerly <http://www.pitt.edu/~osaweb/lap>) was used as the subject profile.

The "Student Life at Pitt" page contained text links to areas of significance to the University of Pittsburgh undergraduate student population, including registration, financial aid, academic and library resources, health, wellness and fitness information, social life and sports, and other campus and related information. The subject profile consisted of the text of this web page. The computation of the semantic component involved counting the number of words on the web page that matched words on the "Student Life at Pitt" page. The number of word matches was divided by the total number of words on the page provided a ranking of the pages in terms of importance of the content for this population. This provided a measure of the semantic landmark value, or the importance of the page on a semantic basis for the population examined.

In general, student affairs and admissions offices at a university analyze the needs and interests of students. At the University of Pittsburgh, this information is reflected in the "Student Life at Pitt" page compiled by the Student Affairs Department. Profiles could be developed for other populations of users by collecting data of interests and needs. For the University of Pittsburgh site, faculty, staff, prospective students and alumni are obvious examples of other populations for whom profiles could be developed.

The equation that was used to combine the semantic, visual and structural components into the objective landmark value was given above, in Figure 3-1. The next section describes the computation of the other measure of landmark quality, the subjective landmark value.

3.3.2. Subjective Landmark Value

The second measure of landmark quality used in this research is the *subjective landmark value*. Like the objective value described above, this provides an evaluation of the characteristics of a web page that make it likely to be used as a landmark. The major difference is that this measure is derived from ratings made by subjects on specific web pages. The rating items are designed to gather data on the three categories of landmark characteristics, visual, structural and semantic characteristics. This measure is new in the literature, so a list of potential questions was generated and a set of them were tested in a pilot study. There is some work in hypertext that similarly asks about knowledge of URLs, and research in navigation asks related questions about recognition and distinctness of visual aspects of a physical environment. The questions were narrowed down and classified to reflect the three categories of characteristics. The questions for each category are listed here.

The questions regarding visual characteristics are:

- How often have you been to this page?
- Would you remember this page if you saw it again?
- How visually distinct is this page in its design/layout?

The questions regarding structural characteristics are:

- How easily could you get to this page without using a search engine?
- Do you know the exact URL for this page?

The questions regarding semantic characteristics are:

- How important is the content of this page in general?
- How important is the content of this page to you personally?

The scores of the 5-point ratings for these questions were combined as follows to generate the subjective value. In each case, the high positive response was given a value of 5, and the low or negative response was given a value of 1. The scores for each question were summed across the subject pool and then manipulated to get an average answer for that category. The product of the number of questions in the category and the number of subjects gave the divisor for this manipulation. Thus, for the visual characteristics, the sum of all the answers for the three questions was then divided by 90 (3 questions x 30 subjects). This result was scaled by dividing by 5 to provide a subjective landmark value between 0 and 1. The equation for this computation is given in Figure 3-2.

<p>Subjective Landmark Value =</p> $\left[\sum (Q1 + Q2 + Q3) / 90 + \sum (Q4 + Q5) / 60 + \sum (Q6 + Q7) / 60 \right] / 5$ <p>where</p> <p>Q1, Q2, and Q3 are the questions related to the visual characteristics</p> <p>Q4 and Q5 are the questions related to the structural characteristics</p> <p>Q6 and Q7 are the questions related to the semantic characteristics</p>

Figure 3-2. Equation to compute the subjective landmark value

This completes the definition of the objective and subjective landmark values, which are the dependent variables in this dissertation. The next section describes the methods used for the two experiments.

3.4. Methodology

Two experiments were used to collect data to analyze the objective and subjective landmark characteristics just described and to address the four hypotheses described in section 3.2. The first experiment had three parts. It collected path descriptions, data of URL knowledge and bookmark use as well as typical starting points for tasks on the University of Pittsburgh web

site. The first experiment is described in the next section. The methodology for the second experiment, a questionnaire regarding web page characteristics, is presented in section 3.4.2.

3.4.1. Experiment 1 – Navigation strategies and site knowledge

Procedure

Subjects were recruited for Experiment 1 through announcements in undergraduate Information Science classes and posters near the classrooms. The 30 subjects recruited were undergraduate students taking courses in the Department of Information Science and Telecommunications. All subjects identified themselves prior to the experiment as strong in English and familiar with the University of Pittsburgh web site. The experiment was administered in classrooms on the University's Oakland campus. Verbal and written instructions introduced the subject to the several types of questions and asked the subject to complete all sections as fully as possible. The experiment was implemented as a paper questionnaire, and took subjects between 20 and 45 minutes to complete. Subjects were paid for their participation.

Task design

The Experiment 1 questionnaire consisted of 5 path description questions, 10 URL and bookmark identification questions and 10 multiple-choice questions about navigation starting points, as well as basic demographic information (Appendix B). The sections were always asked in this order, but the order of presentation of questions within a section was random. The descriptive and multiple-choice questions were both of the type "How would you use the University of Pittsburgh web site to find _____?"

The path description questions had an open-ended format to elicit path descriptions. This section asked the subjects to give the first five steps they would use in answering that question using the University web site, as shown in the example in Figure 3-3. The instructions for the open-ended questions asked the subject to list actions he/she would take to find the answer (up to five actions) using the University web site. The actions might be URLs of web pages or a title of a page. The instruction at each question included a reminder to start with a URL or to indicate how the first page would be accessed.

An example of a path description question is:

How would you find the starting time of a movie playing on campus?

Give up to five steps:

_____ (5 blank lines given) _____

If the first step was not a URL, please indicate how you would get to that first page.

Figure 3-3. Example of an open-ended path description question for Experiment 1.

The second section consisted of asking about the subject's ability to access 10 web pages on the University web site. A descriptive title of each web page was listed, and the subject was asked to write the URL for the page if they knew it, and to indicate whether they had a bookmark or favorites list item for that page. Finally, subjects were asked to list the complete URLs for any other pages at the University of Pittsburgh web site that they knew them for.

The third section presented multiple-choice questions about the starting point that would be used to answer each of ten navigation questions related to the University web site. The instructions asked the subject to select one of the starting points listed or to choose "other" and indicate with a URL or a page title where they would start. The starting points offered in the multiple-choice list included the ten web pages that were asked about in section 2 of the experiment. The additional multiple-choice options included the search field on the University home page, and an Internet search page such as Google or Yahoo. An example of a multiple-choice question is given in Figure 3-4.

These three sets of questions provided a variety of data about the subjects' familiarity with and ability to navigate the University of Pittsburgh web site.

Data

The data collected from the first section, the open-ended path description questions, included URLs and titles of web pages as well as description of initial strategies for starting that navigation, such as knowledge of the URL, or indication of use of a bookmark or search tool. The subject may know the specific web page with the answer, or may give a series of steps to achieve the goal.

The URL and Bookmark questions provided data as to which of the listed web pages were easy for the subjects to access in these ways. The data from the multiple-choice questions identified the web pages that subjects would use to start a variety of navigation tasks on the University web site.

Demographic data was also collected, including age, sex, year in college, self-assessment of familiarity of use the World Wide Web in general and of the University of Pittsburgh web site specifically and self-assessment of English language ability.

Data from this experiment address hypotheses 1 and 2, that the pages easily recalled or accessed will have stronger objective landmark values. Results for this experiment are given in section 4.2 and analyzed in sections 5.1 and 5.2.

How would you find the office hours for the student health center?

I would start at:

- a. *The University homepage and follow links*
- b. *The University homepage search field and enter the search phrase _____*
- c. *The SIS home page*
- d. *The DIST home page*
- e. *The University Library System*
- f. *The Student Life at Pitt page*
- g. *CourseWeb*
- h. *Pitt News*
- i. *The Pittsburgh Panthers official site*
- j. *The Information Technology home page*
- k. *Student Information Online*
- l. *A web search engine and enter the search phrase _____
(Google, Yahoo, Altavista, Lycos, etc.)*
- m. *Other _____*

If other, you would get to this page by:

- ☐ *typing the URL which is
http://_____*
- ☐ *using my bookmark/favorites list*
- ☐ *Other:
_____*

Figure 3-4. Sample multiple choice question for Experiment 1.

3.4.2. Experiment 2 - Characteristics of landmarks

Procedure

Subjects were recruited for Experiment 2 through announcements in undergraduate Information Science classes and posters near the classrooms. The 30 subjects recruited were undergraduate students taking courses in the Department of Information Science and Telecommunications. All subjects identified themselves prior to the experiment as strong in English and familiar with the University of Pittsburgh web site. The experiment was administered in classrooms on the University's Oakland campus. Verbal and written instructions introduced the subject to the several types of questions and asked the subject to complete all sections as fully as possible. The experiment was implemented as a paper questionnaire in color, and took subjects between 30 and 50 minutes to complete. Subjects were paid for their participation. Subjects were permitted to have previously participated in Experiment 1. Demographic data was collected as in the first experiment, including age, sex, year in college, self-assessment of familiarity of use the World Wide Web in general and of the University of Pittsburgh web site specifically and self-assessment of English language ability.

Task design

Experiment 2 asked subjects to rate a selection of University of Pittsburgh web pages using questions regarding their familiarity with the page and their observations looking at the page. Twenty-four (24) web pages were selected from the University of Pittsburgh web domain (<http://www.pitt.edu>) for examination in this study. Potential web pages were analyzed to calculate the objective landmark value as described earlier in section 3.2.1. The final web pages selected included ones with a range of objective landmark values from low to high, and included seven that were indicated frequently by subjects in the course of Experiment 1.

Images of the web pages were captured using consistent web browser settings and window size to display the portion of the web page that would typically be viewed in a moderate size window on a 16" monitor. The URL and browser buttons were hidden from the display. The images were printed in color, one per page, with a set of questions below them. Each question included a five-point response scale, and the same questions were asked about each web page. The questions were listed in section 3.3.2, and the full questionnaire is included as Appendix C.

Data

These data provided the subject's impression of the importance of visual, structural and semantic characteristics of the web page, and were used to compute a score of the subjective landmark value of each web page. The first three questions relate to visual aspects, the next two to structural aspects, and the final two to semantic content. The ratings were combined using the method described above in section 3.3.2 to determine the subjective landmark value. The results of Experiment 2 are presented in section 4.4. The analysis of this data in relation to hypotheses 3 and 4 regarding the correlation of the objective and subjective landmark values is done in section 5.3.

3.5. Summary

The two experiments gathered data with respect to the dependent variables of objective and subjective landmark values, which were defined at the beginning of this chapter. Experiment 1 compiled data on navigation paths described by the subjects, and on knowledge of URLs and use of bookmarks. Experiment 2 gathered subjects' judgment of several visual, structural and semantic characteristics that might indicate the value of that web page as a landmark.

Hypothesis 1 is that the pages easily recalled will have stronger objective landmark characteristics. This hypothesis will be examined using the web pages listed in the path description section of Experiment 1. For each web page the objective landmark quality value will be calculated. To address Hypothesis 2, that the pages easily accessed will have stronger objective landmark characteristics, the URL and bookmark access data from Experiment 1 will be analyzed.

Hypothesis 3 is that the subjective and objective landmark values are positively correlated. This will be analyzed for the web pages in Experiment 2. For each web page, the two values will be computed as described in section 4.3. The objective landmark value is computed using objectively measured data about the page, and the subjective landmark value is computed from the data gathered in Experiment 2.

Finally, Hypothesis 4 is that the degree of correlation between the objective and subjective values is dependent on the weights selected for combining the characteristics within

the objective measure. The predicted weights for combining the structural, visual and semantic components of the objective landmark value are $1/3$, $1/3$, $1/3$. The data collected from the experiments is presented in Chapter 4 and analyzed with relation to the hypotheses in Chapter 5.

4. RESULTS

Two experiments were performed to provide data for this research. The experiments were performed at the University of Pittsburgh and, for each, 30 subjects were recruited to complete the experiment questionnaires. The first experiment required subjects to recall navigation paths for the tested web site. The second experiment examined individual evaluation of characteristics of web pages on the tested site. Both of the experiments examined web pages at the University of Pittsburgh web site.

The research methodology defined two variables to examine web pages for potential usefulness as landmarks: the objective landmark value and the subjective landmark value. The computation of the objective landmark value was independent of the experimental data, and so the objective landmark value data is presented in section 4.1. Presentation of the data for Experiments 1 and 2 follow in sections 4.2 and 4.3. The subjective landmark value is computed for the web pages used in experiment 2, and therefore is presented in section 4.3.3 as part of Experiment 2.

4.1. Objective Landmark Values

The objective landmark value was computed for a set of web pages at the University of Pittsburgh web site, using the methods defined in the research methodology. To summarize, the measures of the structural, visual and semantic characteristics of the page were combined in equal parts to give the objective landmark value. For the visual component, the areas covered by text, graphics, links and white space were measured on printed copies of the web pages. The structural component used counts of the number of links into the web page and the number of links out of the web page, combined with the depth of the URL. The semantic component used word matching to determine the relevance of the web page. The complete algorithms for these computations are described in section 3.3.1. The objective landmark value was computed for 39 web pages on the University of Pittsburgh web site. The results are presented in Table 4-1. The data are grouped into categories ranging from very high to very low. The URLs for these web pages are listed in Appendix A.

Table 4-1. Objective landmark value for web pages used.

PageName	Objective Landmark Value	Range
Pitt Home	0.540	High >= 0.29
Student Life at Pitt	0.490	
Student Activities	0.429	
Student Information Online	0.425	
Pitt Program Council	0.383	
Pitt Site Index	0.379	
Nationality Rooms	0.353	
Campus Map	0.322	
Virtual Tour	0.317	
Events Calendar	0.309	
SIS Home	0.300	
Panther Central	0.292	
Future Students	0.282	Medium 0.20 – 0.29
Cyber Counselor	0.281	
Digital Library	0.280	
Course Web	0.265	
Athletics	0.263	
Financial Aid Home	0.263	
Technology Home	0.246	
Pitt News	0.229	
Admissions	0.227	
Students	0.225	
DIST Home	0.220	
Academics	0.211	
About Pitt	0.206	
SIS Library	0.202	
Libraries (ULS)	0.189	Low < 0.20
Budget and Controller	0.182	
Fact Book	0.179	
Financial Aid Parents Info	0.170	
Jurist	0.146	
Human Resources Home	0.144	
Faculty/Staff	0.130	
Capital Campaign	0.117	
Provost Learning	0.102	
Computer Accounts	0.090	
HR Employment	0.087	
Dental Medicine Bulletin	0.059	
Proposal Components	0.027	

The pages with the highest objective landmark values include the home page for the University of Pittsburgh web site, called “Pitt Home”, and two pages related to the Student Affairs Department, the Student Life at Pitt page and the Student Activities page. Also in the high category were the Student Information Online page of the Registrar’s office, the Pitt Program Council, the Site Index for the University web site, and the School of Information Sciences Home page (SIS Home). The mid-level range includes the entry page for an electronic blackboard course materials system called Course Web, the Students page accessible from the Pitt Home page, and the Information Technology Home page. Several pages in the ‘low’ range were the Faculty/Staff page accessible from the Pitt Home page and the Information Technology Department’s Computer Accounts page for students.

One unusual situation is that there are two pages that could reasonably be used as starting points for the University of Pittsburgh libraries, and the objective landmark analysis for them resulted in very different values. The Libraries (ULS) page is accessed from the Pitt Home page by clicking on the Libraries link or by the URL <http://www.pitt.edu/libraries.html>, and it had an objective landmark value of 0.189, in the low category. This page serves as a home page linking all of the main University libraries. The Digital Library page is accessible from the Libraries (ULS) page or with the URL <http://www.library.pitt.edu> and is often considered the main library page. It has several graphical elements and well connected in the web site. It had a medium level objective landmark value of 0.280. The Library (ULS) page has more text, only a small photo, and fewer links to it from other pages in the web site.

The components of the objective landmark value are now examined more closely. As stated before, the objective landmark value is computed based on three compound components, the semantic component, the visual component and the structural component, each of which is based on direct measures of the web page. The semantic component is based on a semantic analysis of the page for relevance to undergraduate students, the visual component quantifies the amount of graphics such as photographs or logos on the page, and the structural component is computed based on the number of links into and out of the page and the depth of the URL. Figure 4-1 shows the individual component values for a sample of pages. Data for the full reference set of pages is given in Appendix F.

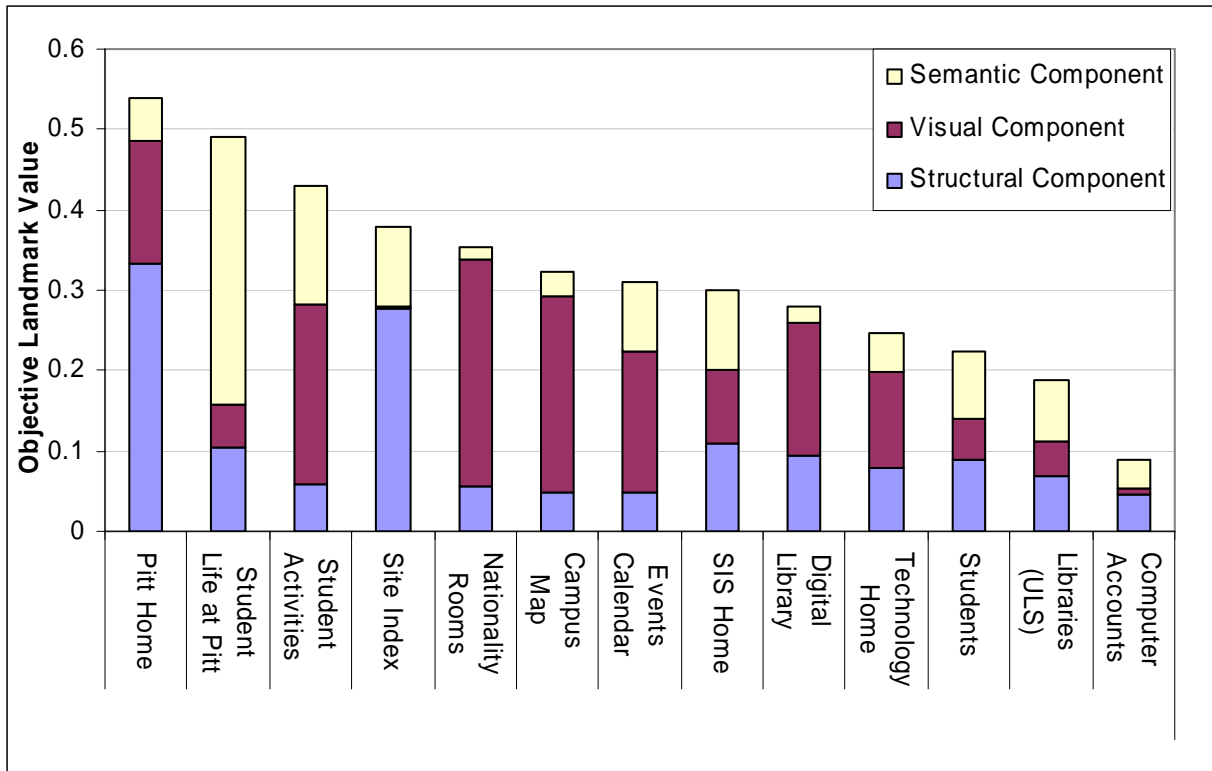


Figure 4-1. Objective landmark value components for a selection of web pages.

The data presented in Figure 4-1 show the variety of proportions of the different components found in the web pages. The Site Index has moderate or high values on both the semantic and structural components and a value of zero on the visual component, whereas the Pitt Home page has high or moderate structural and visual components and a low semantic component. The objective landmark value for the Nationality Rooms and the Campus Map come mostly from the visual component.

The Site Index on the University web site is a very long list of text links including over 450 words or phrases. It has no graphical features, and thus has essentially a zero value for the visual component. The very high structural component is due primarily to its very high number of outgoing links, but also its high number of links into it from many different pages on the web site, and fairly short URL. The semantic component is moderately high due to the presence of so many terms that are pertinent to the student population.

The Pitt Home page has the highest structural component in the sample, due in large part to having a short URL and an extremely high number of links into the page from other pages in

the site. This page has a large photograph in the center, and a variety of small photographs and logos on the page, which when compared to the total page area gave a moderately high visual component value. The presence of only a few specific terms relevant to students gave the low but measurable semantic component.

The Nationality Rooms page has one large graphic and several smaller icons, creating a very high value on the visual component. The structural component for this page came primarily from having a fairly short URL, but also a moderate number of links to it. The small amount of text had a very few words matching the subject profile, providing a very low value on the semantic component.

The computation of the semantic component needs to be mentioned with respect to the objective landmark value of one page in particular, the Life at Pitt page. As described in the research methodology, this web page was used as the source for relevant terms for the semantic analysis comparison. Thus, this page had a maximum value for the semantic component.

The objective landmark values will be examined in relation to the results of the two experiments, and will be examined further in the discussion chapter. The next section presents the results for Experiment 1.

4.2. Experiment 1

Experiment 1 consisted of basic demographic information, five descriptive questions related to navigation paths, 15 multiple-choice questions about starting points for using the University of Pittsburgh website, and ten URL identification questions. The entire questionnaire is included as Appendix B.

4.2.1. Demographics

The 30 subjects recruited for Experiment 1 were undergraduate students taking courses in the Department of Information Science and Telecommunications. There were 24 men and six women, ranging in age from 18 to 31. They were seven freshmen (first year students), five sophomores (second year), seven juniors (third year) and eleven seniors (fourth year). The group included 28 native English speakers, one with “excellent” English skills and one with “some difficulty,” according to the self-report.

The subjects rated themselves on five-point scales in terms of familiarity with the WWW and familiarity with the University of Pittsburgh web site as well as their frequency of use of the

University web site and of the WWW in general. Self-reporting of familiarity with the World Wide Web showed 25 subjects were “very familiar,” four were “moderately familiar,” one had “some familiarity,” and none were in the bottom two categories of “slightly familiar” and “not at all familiar.” The report on familiarity with the University of Pittsburgh web site was approximately one step lower than for the web in general. Four were “very familiar,” 18 were “moderately familiar,” seven had “some familiarity,” and one was “slightly familiar.” The self-reporting on frequency of use showed that 26 use the World Wide Web “frequently” and four use it “moderately often,” with no one reporting less frequent use. For frequency of use of the University of Pittsburgh web site, seven indicated “frequently,” thirteen “moderately often,” eight reported “some” use, and two “rarely,” with no one reporting “not at all.”

4.2.2. Path Descriptions (open-ended questions)

Subjects were asked to identify the steps they would use to answer five questions using the University of Pittsburgh web site, such as “How would you find the hours of operation of the SIS Library for the upcoming vacation period?” Subjects considered each task, and then, using their knowledge of the web site, recalled from memory the selections they would make to select and follow links to find the required information. The web site was not available to them, and they were asked to write their path suggestions on paper. Subjects listed up to five steps for the navigation to find each answer.

In the course of answering these five questions, all 30 subjects used the Pitt Home page at least once, and more than 50% of the subjects referred to six other pages at least once: the Campus Map, Athletics, SIS Home, CourseWeb, and the Students page. Table 4-2 lists all of the pages that were mentioned by at least two subjects overall. In the web page names, the abbreviation ULS refers to the University Library System. The Library (ULS) page is the main library page accessible from the Pitt Home page. Some of the distinctions between it and the Digital Library were described above in section 4.1. Other pages relevant to the whole University Library System are labeled with ULS as part of their name in this work, which, for example, makes it easy to distinguish between the SIS Library Hours page and the page that lists the operating hours for all of the libraries within the University system.

Table 4-2. Overall web page references for Experiment 1.

Web page name	Number of Individuals
Pitt Home	30
Campus Map	23
Athletics	21
SIS Home	20
Movies	18
CourseWeb	16
Students	15
Student Activities	15
SIS Library Hours	13
SIS Library	11
Search field	10
Academics	9
Events Calendar	8
Admissions	8
Virtual Tour	8
Student Life at Pitt	7
ULS Library Hours	6
Digital Library	5
Library (ULS)	5
Pitt Program Council	4
SIS BSIS Current Course Schedule	4
Future Students	3
Pitt News	2
About Pitt	2
Pitt Site Index	2
ULS List of Libraries	2

This section presents the data question by question. For each of the five questions, the data is presented in two ways, in tables and in path diagrams. Presented first each time is the table showing how many subjects referred to each web page. The tables provide the total references to each page and also show a breakdown of how many subjects listed the page at each step in the navigation path. Each table includes only web pages and search fields listed by the subjects; it does not include steps such as “look for a relevant link” or “look for” certain information. The second presentation tool for each question is a path diagram. These figures show the sequence that the web pages were indicated in the subjects’ answers and also reflect the objective landmark value category for the available pages. The path diagrams were created using the software program Inspiration, version 7.5. The tables and figures are introduced extensively in the presentation of the first question. Questions were presented in random order in the experiment, but are given numbers here for ease of discussion.

Question 1: How would you find the hours of operation of the SIS library for the upcoming vacation period?

The first data column of Table 4-3 shows the overall references to different web pages. Twenty-one subjects listed the Pitt Home page, and seventeen subjects listed the SIS Home page. Eleven subjects referenced the SIS Library page and six referenced the Academics page. The table also shows the breakdown of page references by each step given. For the first step in answering question 1, 21 subjects listed the Pitt Home page, seven indicated the SIS Home page, and one listed the Digital Library. In addition, one person indicated starting at “Student Services,” but did not give a URL. This may refer to the page by this name that is available from the University web portal, <http://my.pitt.edu>. The remaining columns in this table show the number of subjects who listed each of these pages at subsequent steps in the navigation path for their answer. For example, as the second step in the navigation path four subjects listed a selection from the choice box on the Pitt Home page that lists Departments, two indicated search terms for a search, and seven listed the SIS Home page.

A sketch of the paths taken by each individual provides a view of the number and sequence of references to different pages (Figure 4-2). The lines are labeled with the number of subjects who noted that path, and the most commonly referenced web pages are labeled with a circle showing the total number of subjects who listed that page as part of their solution. The

objective landmark value for the web page is reflected in the color of the web page symbol. Pages with a high objective landmark value such as the Pitt Home page are shown with a dark green symbol, pages with a medium value such as the Academics page are shown in medium green, and pages with a low objective landmark value such as the Library (ULS) page have a light green symbol.

Table 4-3. References to each web page for question 1 of Experiment 1.

Location	Total references	Step 1	Step 2	Step 3	Step 4	Step 5
Pitt Home	21	21				
Pitt Home Departments choice box	4		4			
Pitt Home Search field	2		2			
SIS Home	17	7	4	6		
SIS Library Hours	13			3	8	
SIS Library	11		3	5	3	2
Academics	6		6			
ULS Library Hours	6			6		
Library (ULS)	5		4	1		
Digital Library	5	1	4			
Pitt Site Index	2		2			
Other	1	1				

The diagram shows that subjects suggested four different routes to get from the Pitt Home page to the SIS Home page. A variety of possible pages were also listed to get from the SIS Home page to the SIS Library page or to a listing of the SIS Library hours.

There are two types of recall errors that are accommodated in the diagrams. One is that a step in the navigation path may be left out. Where it was possible to identify one missing page that would make the path work, the page is shown in a light blue symbol, and the path is shown by a dashed line. This situation indicates an understanding that there was a way to get from one page to the other, but the subjects did not correctly recall or know the name of the needed intermediate page. The second type of recall error is that a page may be named by the subject but no corresponding page was identified on the web site by the researcher. The unidentified pages are indicated with a white symbol and the path is a dashed line.

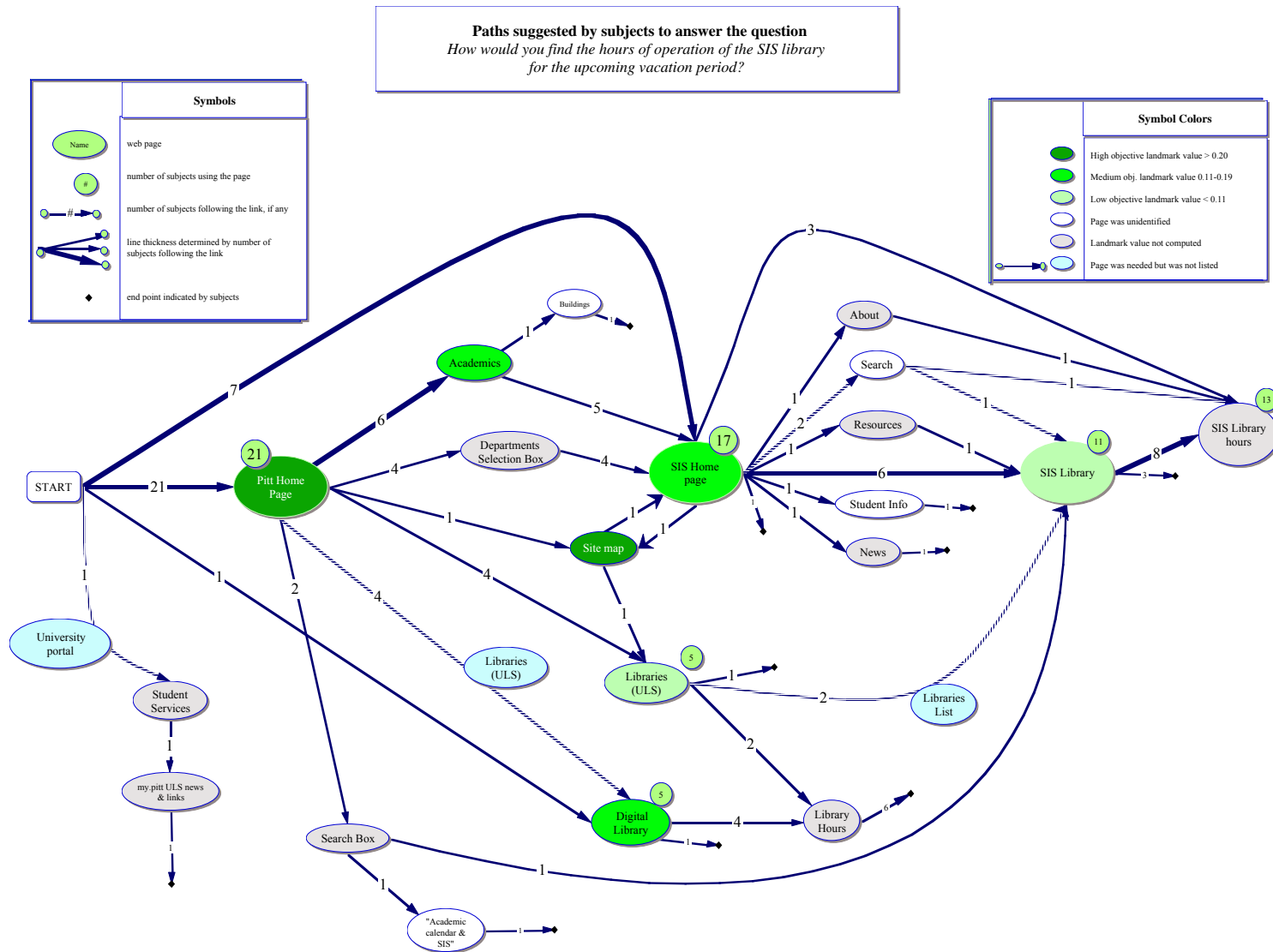


Figure 4-2. Path diagram for finding library hours (Exp. 1, Q.1).

Question 2: How would you find the starting time of the movie showing on campus this weekend?

Table 4-4 shows the web pages referenced by subjects in the answers for question 2. Twenty-eight subjects indicated the Pitt Home page as the starting point for finding campus movie times. One person indicated the Pitt Program Council as their starting point, and one listed an external site. In other steps of their answers to this question, 15 subjects included the Student Activities page, 12 listed the Students page and seven listed the Student Life at Pitt page.

Table 4-4. References to each web page for question 2 of Experiment 1.

Location	Total References	Step 1	Step 2	Step 3	Step 4	Step 5
Pitt Home	28	28				
Pitt Home Search field	3		3			
Student Activities	15		3	12		
Students	12		12			
Events	8		2	3	3	
Student Life at Pitt	7		7			
Pitt Program Council	4	1		2	1	
Pitt News	2		1	1		
External Site	1	1				
Movies	18		1	5	9	3

The path diagram to find the show times for a campus movie, Figure 4-3, shows a variety of potential paths to get to both the Student Activities page and the Events Calendar, though not all of the paths were valid. While many of the paths suggested for answering this question are not perfectly correct, they are close, and if actually navigating, subjects might often have seen the intermediate or alternate page they needed although they did not express the correct one from memory.

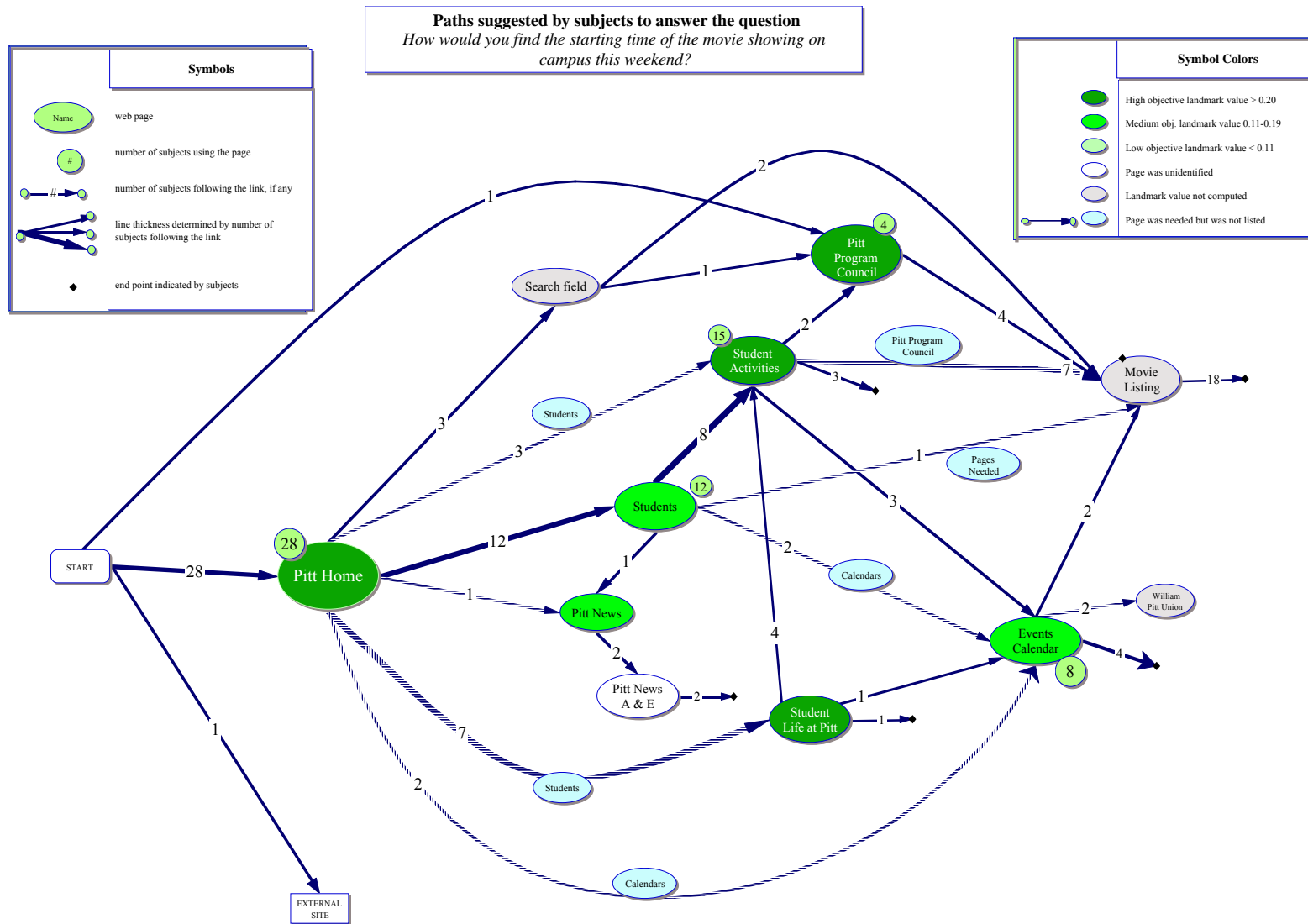


Figure 4-3. Paths to find campus movie times (Exp. 1, Q. 2).

The portion of the path diagram that includes the Pitt Home, Students, Student Activities and the Student Life at Pitt pages is interesting to examine in depth. Neither the Student Activities page nor the Student Life at Pitt page is accessible directly from the Pitt Home page, though seven subjects listed one of these paths and three listed the other. These paths are shown with dashed lines to indicate that they are not directly successful paths. Both pages are accessible from the Students page and the Student Activities page is also accessible from the Student Life at Pitt page. The pages that were missing from the answer but were needed in the navigation path are shown in light blue along the dashed lines. Where multiple pages were left out, a single light blue symbol is used and the missing pages are not specified.

One of the key aspects of the definition of a landmark examined in this study is that a landmark is more memorable than other elements around it in the environment. The Students page, which was left out of ten individual navigation answers has an objective landmark value in the medium range compared to the other pages around it that were mentioned in the paths which all have high objective landmark values. This will be examined further in the analysis of hypothesis 1 in section 5.1.

Question 3: How would you find the lecture notes for a BSIS class you missed?

The web pages listed by subjects in their answers to question 3 are listed in Table 4-5. The most frequently mentioned pages were the CourseWeb, Pitt Home, and SIS Home pages. The CourseWeb is a system for making course materials and discussion forums available to students registered for a course. At the time of this research it was primarily available by its URL and was not available from pages such as the Students or Academics pages. Courses that use the CourseWeb system may also have their own web page for the course that may contain a link to the CourseWeb.

There was more variation in the starting points listed for this question than for any of the other questions. Twelve subjects listed starting at the CourseWeb, eleven at Pitt Home and five at the SIS Home page, and one subject listed a URL to access a SIS faculty page as the first step. One subject listed the “Student Services” page as the start. This was the same subject that listed this page in question 1, and again it may refer to a page off the University portal <http://my.pitt.edu>. The various paths are shown in Figure 4-4. The two main methods described by subjects to find BSIS class lecture notes were using the CourseWeb or, alternatively, locating

class or instructor information off the SIS Home page. Since not all classes use the CourseWeb, there were two clear potential goals to obtain this information, either the CourseWeb or the home page for the course. Web pages that would be used after signing on to the CourseWeb are not shown in these results.

Table 4-5. References to each web page for question 3 of Experiment 1.

Location	Total References	Step 1	Step 2	Step 3	Step 4	Step 5
CourseWeb	16	12		4		
Pitt Home	11	11				
Pitt Home Search field	1		1			
SIS Home	9	5		4		
Individual Course Home page	7		1	2	3	1
Individual Faculty Home page	6	1		4		1
Academics	4		4			
SIS BSIS Current Course Schedule	4		1	1	2	
Students	3		3			
SIS Academics	1		1			
Pitt Site Index	1		1			
SIS Faculty	1				1	
Other	1	1				

Question 4: How would you find the date and time of the next home Panther game?

Table 4-6 shows the web pages listed by subjects in the paths they would use to find the requested game time information, and Figure 4-5 shows the related path diagram. As starting points to answer this question, 25 subjects indicated the Pitt Home page, two listed the Athletics page and one listed their personal web page. Of the two subjects whose start is listed as “other,” one listed an incorrect URL that appeared to be a guess at the Athletics page URL, and the other declined to answer the question, writing just “I don’t know.” Despite this, there were fewer incorrect paths or web pages references in answers to this question than for other questions.

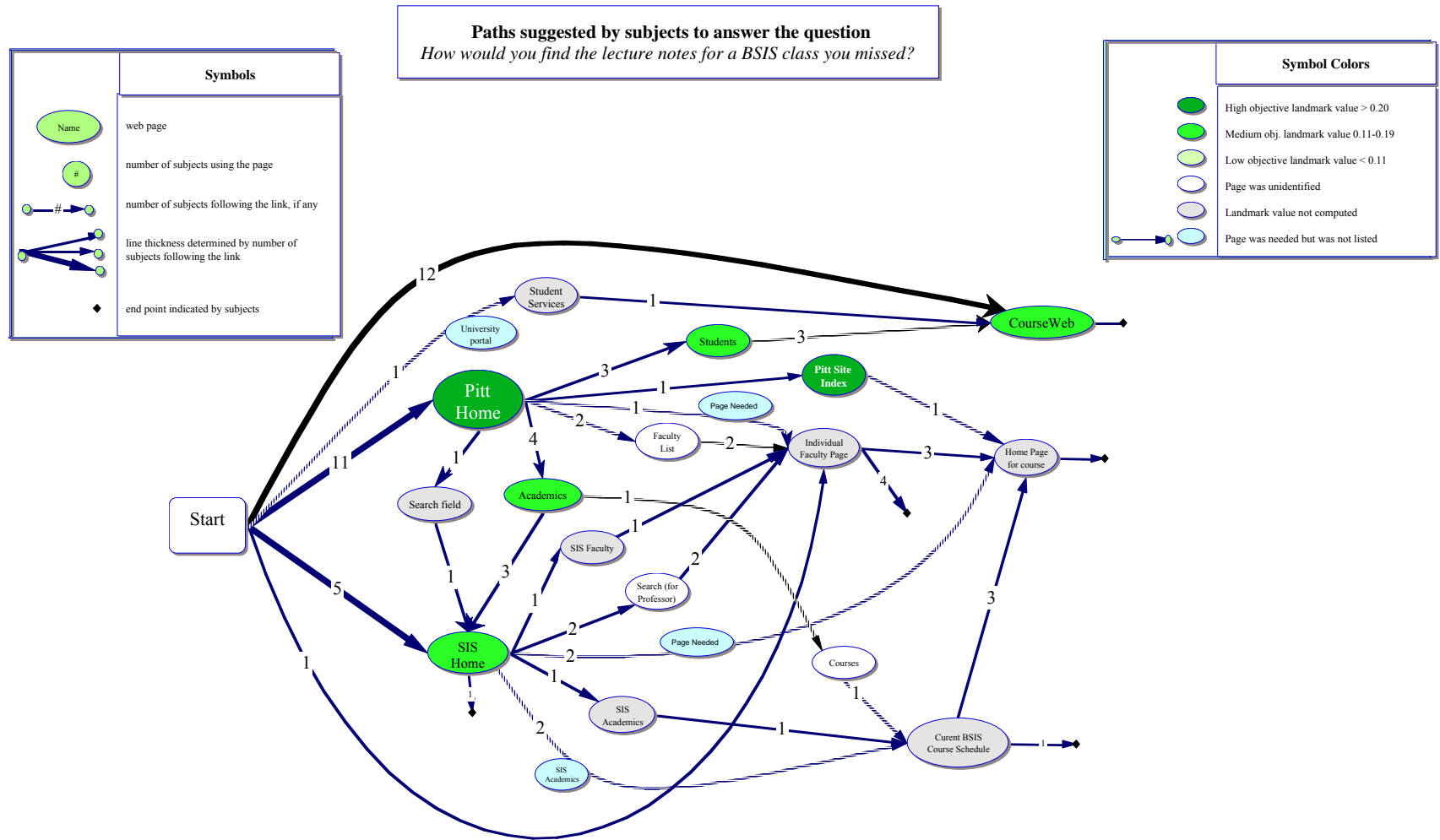


Figure 4-4. Paths to locate BSIS class lecture notes (Exp.1, Q. 3).

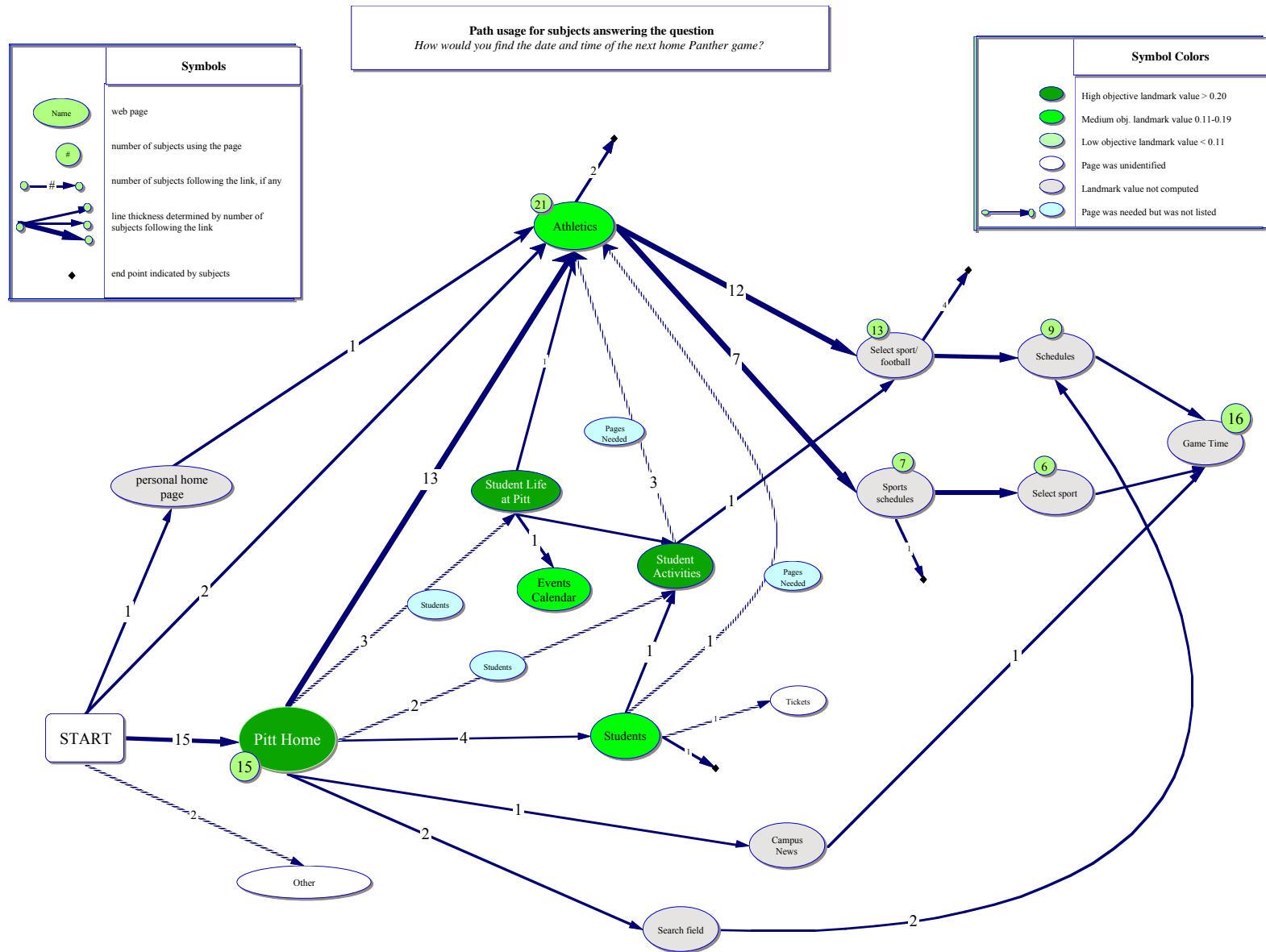


Figure 4-5. Paths to find time of next home Panther football game (Exp. 1, Q. 4).

Many of the subjects were familiar with the Athletics link from the Pitt Home page and used that to answer this question. In all, 21 subjects referred to the Athletics page in some step of their answer. As shown in Figure 4-5, from the Athletics page there are two parallel paths to find sports game information. The sport can be selected first and then schedule information requested, or schedules can be selected first and then there is a choice of which sport to display information for. While more subjects indicated the sport first, both paths were used by this subject pool.

Table 4-6. Number of subjects referring to each web page for question 4 of Experiment 1.

Location	Total References	Step 1	Step 2	Step 3	Step 4
Pitt Home	25	25			
Pitt Home Search field	2		2		
Athletics	21	2	14	3	2
Students	3		3		
Student Life at Pitt	3		3		
Student Activities	4		2	2	
Events Calendar	1			1	
Pitt Campus News	1		1		
Personal Home page	1	1			
Other	2	2			

The path diagram shows dashed lines indicating missing steps in the navigation path descriptions. Two of these originate at the Pitt Home page, with one going to the Student Life at Pitt page and the other to the Student Activities page. In both cases, including the Students page as an intermediate step would make these paths work. The Pitt Home page, Student Life at Pitt page and Student Activities page which were mentioned in the paths in question all have high objective landmark values, while the Students page that was left out of the path answers has a medium objective value. Appropriate intermediate pages were not yet identified by the researcher for the other paths requiring additional steps.

Question 5: How would you find a map of the University of Pittsburgh Oakland campus?

To locate a map of the University of Pittsburgh Oakland campus, 28 subjects listed the Pitt Home page as their starting point, one listed the SIS Home page, and one referred to an external site. In the steps of their answers, fifteen indicated the Campus Map, eight indicated the Admissions page, five used the Virtual Tour, and three listed to the About Pitt page in their solutions. Table 4-7 and Figure 4-6 show the results for this question. Only four subjects listed more than three steps in their answer for this question.

Table 4-7. Number of subjects referring to each web page for question 5 of Experiment 1.

Web Page	Total References	Step 1	Step 2	Step 3	Step 4
Pitt Home	28	28			
Pitt Home Search field	4		4		
Campus Map	23		2	17	4
Admissions	8		8		
Virtual Tour	8		3	5	
Future Students	3		3		
Students	2		2		
About Pitt	2		2		
SIS Home	1	1			
About SIS	1		1		
Student Life at Pitt	1			1	
External site	1	1			

It is interesting to notice that many more subjects used the Admissions and Future Students pages to look for this information than the Students page, and in fact the Students page is a more direct way to access the campus map than either of the other two. It is also noteworthy that while the Virtual Tour is available on the main Pitt Home page, only three subjects used that path, and it would have been one of the shortest methods of getting to the Campus Map page.

The path information for these five questions provides one look at the pages that subjects remember within the University of Pittsburgh web site. The paths and objective landmark values for these pages will be discussed further in Chapter 5, in conjunction with other data presented in the following sections.

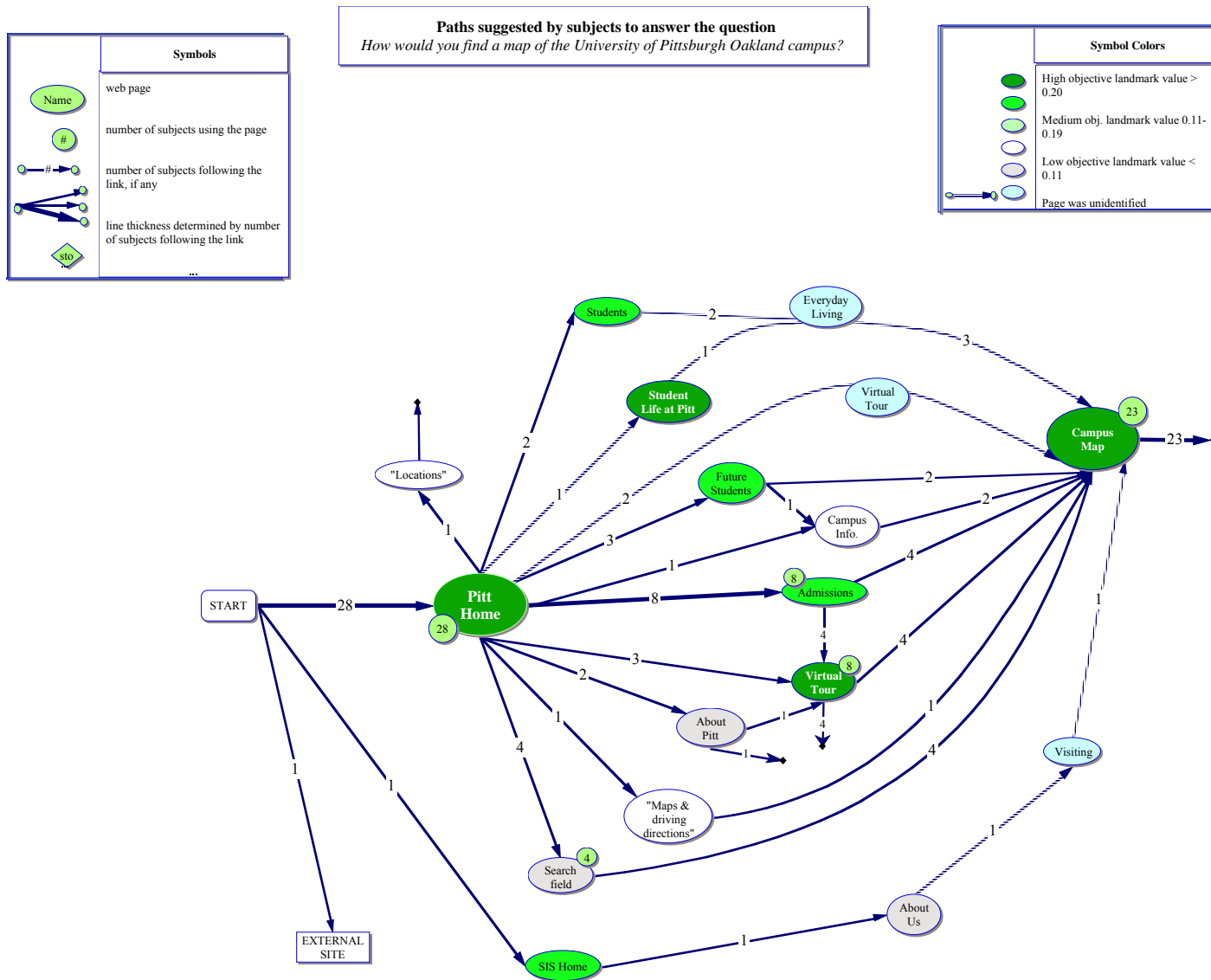


Figure 4-6. Paths to find a map of the University of Pittsburgh Oakland campus (Exp. 1, Q. 5).

4.2.3. Bookmarks and URL Knowledge

One aspect of this study was to examine whether the study population had knowledge of different URLs within the University of Pittsburgh web site as well as whether they keep any of those web pages on a list of bookmarks or favorite sites for easy access. The 30 subjects examined a list of ten web pages from the University of Pittsburgh web site, and were asked to state the URL if they knew it, and to indicate whether they had that page on a list of bookmarks. Figure 4-7 shows the knowledge of the URLs and the availability of bookmarks for the indicated pages. In cases where more than one URL is available for a page, any of the valid URLs was accepted. Similarly, the University Library (ULS) Home page, the Digital Library page, and Pittcat were all considered valid even though the question stated the “University Library System.”

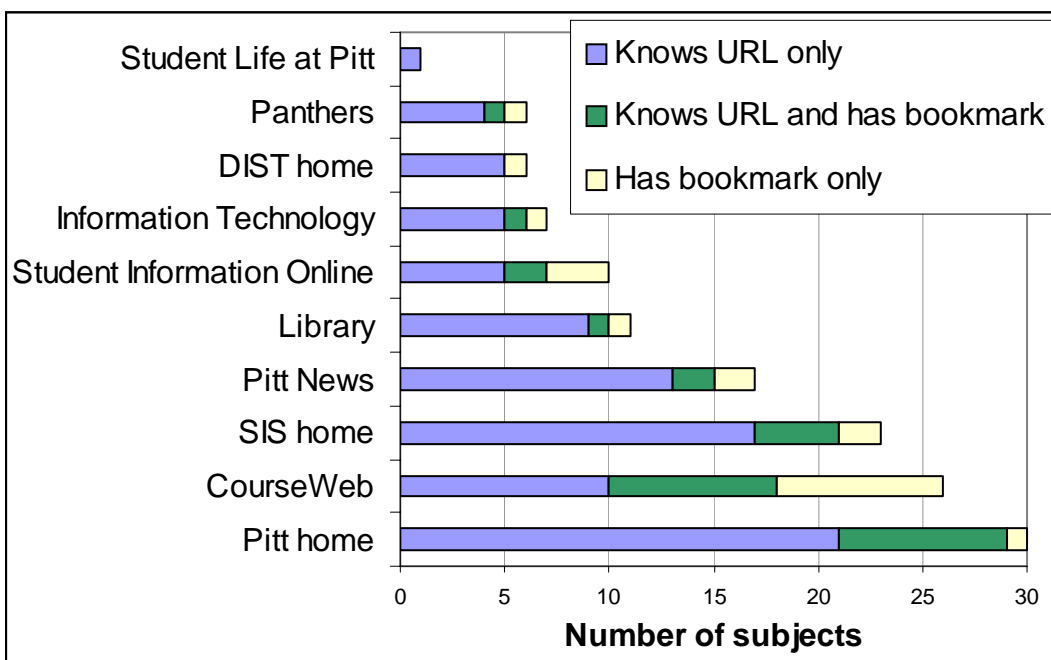


Figure 4-7. URL knowledge and bookmark use (Exp. 1).

It is expected that people develop methods for easily accessing web pages that are useful to them as landmarks. Many of the pages proposed as starting points in the experiment were ones that subjects had easy access to, either by URL or by bookmark. With the combination of methods, six pages were found to be easily accessible by one third or more of the subjects (the

Pitt Home page, CourseWeb, SIS Home, Pitt News, and the University Library). More than 75% of the subjects correctly identified the URLs or had bookmarks for the Pitt Home page, the CourseWeb page and the SIS Home page. It is interesting to note that while more subjects knew the URL for the SIS Home page than for the CourseWeb, when the use of bookmarks was included, more subjects could easily access the CourseWeb. More than 15% had bookmarks for the Pitt Home page, the CourseWeb page, the SIS Home page and Student Information Online.

In addition to identifying use of bookmarks and knowledge of URLs for these pages that could be used as starting places for particular tasks, subjects were also asked to list any other pages at the University of Pittsburgh that they typically use to start their work. These pages included personal web pages, various department pages (math, computer science, engineering, Russian literature, geology), administration pages (accounts, registrar, placement, academics, technology), and specific course or project pages. The subjects mentioning each page listed valid URLs for nearly all of these pages, and indicated use of bookmarks for several. None of these pages was listed by more than one individual, so the objective landmark values were not computed.

4.2.4. Starting Points (multiple-choice questions)

The next section of Experiment 1 posed similar questions about finding information using the University of Pittsburgh web site, but asked specifically for the first step the subject would take. The multiple choice answers included ten pages at the University web site, an option for a web search engine or use of the search field on the Pitt Home page, and a place to fill in another answer if none of the offered choices were the preferred starting point. All thirty subjects answered the ten questions, numbered 6-15 in this analysis. The summary data is presented first, followed by results for each question.

Five of the web pages examined were chosen as starting points by more than half of the individuals in this experiment. The commonly referenced pages were the Pitt Home page, the SIS Home page, the University Library System, the Student Life at Pitt page, and the Student Information Online, which are labeled a, c, e, f, and k in Table 4-8. In addition, the search field on the Pitt Home page was also indicated as a starting point for more than half of the subjects.

Table 4-8. Starting points and successful access schemes (Exp. 1).

Answer Choices	Number of subjects selecting page as starting point	Number who had bookmark or URL knowledge
a. The University Home page	29	29
b. University Home page search field	20	20
c. The SIS Home page	23	20
d. The DIST Home page	2	1
e. The University Library System	23	
Digital Library		10
Pittcat		2
f. The Student Life at Pitt page	17	1
g. CourseWeb	6	4
h. Pitt News	0	n/a
i. The Pittsburgh Panthers official site	0	n/a
j. The Information Technology page	10	3
k. Student Information Online	16	6
l. A web search engine	3	3
m. Other	12	12

It is interesting to examine the choices for starting points in combination with the data of URL knowledge and bookmark use for each individual described in Section 4.1.3. The relevant data is shown in the right most column of Table 4-8. All or nearly all of the subjects who chose the Pitt Home page and the SIS Home page as starting points had successful strategies for accessing them. More than 50% of the subjects also selected three other pages (the Library, Student Information Online, and Student Life at Pitt) as starting points, but most did not have a bookmark or knowledge of the URL. It is interesting that the population envisions these pages as effective and useful places to begin their work, but could not directly access them. The data for the ability to access the library is given according to the access point listed by the subjects. Ten subjects listed access methods for the Digital Library page and two for the Pittcat page, while none of the subjects listed the correct URL for the Library (ULS) Home page, <http://www.pitt.edu/libraries.html>.

The breakdown of answers used for each question is shown in Table 4-9. The questions are listed across the top of the table by number, and are described in the following discussion (see Appendix B for the full questionnaire). For each question, numbered 6-15, the number of

subjects who indicated each of the listed web pages as the starting point is given. Blank cells indicate that no subjects referenced the page as a starting point. The Pitt News and the Pittsburgh Panthers site were not indicated by any of the subjects in this portion of the study.

Table 4-9. Starting points for each multiple-choice question.

Answer Choices	Question										Total
	6	7	8	9	10	11	12	13	14	15	
a. The Pitt Home page, to select links	4	5	4	8	8	11	20	14	5	10	89
b. The Pitt Home page search field		8	2	1	8		6	10	5	7	47
c. The SIS Home page	6				2	13			15		36
d. The DIST Home page									2		2
e. The University Library System	13		20								33
f. The Student Life at Pitt page		14	1		1		1	3		12	32
g. CourseWeb	5			4		2					11
h. Pitt News											
i. The Pittsburgh Panthers official site											
j. The Information Technology page					8				2		10
k. Student Information Online	2			13	2	1	2	2			22
l. A web search engine		2						1			3
m. Other		1	3	4	1	3	1		1	1	15

The first two of the multiple-choice options both used the University of Pittsburgh Home page, and asked the subjects to distinguish whether they would follow links from that page or would use the search field. The Pitt Home page was used by at least some subjects for every question.

Question 6: How would you find out what items are held on reserve for a current IS class?

Thirteen subjects indicated the University Library System as the starting point for this task. Six subjects marked the SIS Home page, five indicated the CourseWeb page, four the Pitt Home page, and two indicated Student Information Online. For this question, eighteen of the thirty subjects would have been successful with their choice, and the others would have had to find another way to navigate to the answer.

Question 7. How would you find whether Pitt has a chess club for undergraduate students?

Fourteen subjects indicated they would start at the Student Life at Pitt page, eight would use the search field on the Pitt Home page and five would begin by selecting links from the Pitt Home page. None of the fourteen had bookmarks or URL knowledge for the Life at Pitt page, but the thirteen starting at the Pitt Home page would have been successful.

Question 8. How would you find whether the journal *Communications of the ACM* is received in electronic edition by the University library?

Twenty subjects indicated they would start directly at the University Library page, and six would begin at the Pitt Home page. The subjects have the impression of easy access to the Library, even if only seven of them could actually get directly there according to the data they provided. The library is directly accessible from the Pitt Home page, so in reality the subjects could find a quick alternative in this case.

Question 9. How would you check your grades for the previous semester?

To access their student records, thirteen subjects chose Student Information Online as their starting point, nine indicated the Pitt Home page, and four went to CourseWeb. Six of those choosing the Student Information Online would have had success in getting to their starting point, as would the subjects who indicated the Pitt Home page and CourseWeb.

Question 10. How would you find the phone number for the help desk at the University computing center?

Three starting methods were chosen by eight subjects each: the Information Technology page, the Pitt Home page search field, and the Pitt Home page to select a link to follow. Nineteen out of 24 would be able to get directly to that starting point.

Question 11. How would you check the BSIS course offerings for next semester?

Thirteen subjects selected the SIS Home page as the best place to start for this question, and twelve of those could get there using a bookmark or URL entry. Eleven subjects chose to start at the Pitt Home page.

Question 12. How would you find a description of the University financial aid scholarships?

Nearly all of the subjects (26) chose to start at the Pitt Home page to find financial aid information. Six indicated that they would use the search field, and 20 would have followed links from there.

Question 13. How would you find the office hours for the student health center?

Twenty-four of the subjects chose to start at the Pitt Home page, and, of those, ten would have made use of the search field while fourteen planned to follow navigational links. The three subjects choosing the Student Life at Pitt page, and the two indicating Student Information Online, did not have bookmarks or know URLs for those pages.

Question 14. How would you find the complete mailing address of the Department of Information Science and Telecommunications?

Twenty-six of the thirty subjects chose starting places for this question that they could get to directly. Fifteen subjects selected the SIS Home page, ten selected the Pitt Home page, two selected the DIST Home page and two the Information Technology page.

Question 15. How would you get information on how to become a resident advisor in the dorm?

Twelve subjects thought the best place to start finding about resident advisor positions was the Student Life at Pitt page, but none of them indicated knowing the URL or having a bookmark for this page. Seventeen subjects would have started at the Pitt Home page.

Throughout the questions, the most successfully used page, besides the Pitt Home page, was the SIS Home page. It could be directly accessed by nearly all of the subjects who chose it. It provided a successful start strategy for fourteen out of fifteen subjects for question 14 (for the Department of Information Science and Telecommunications mailing address), and for twelve out of thirteen subjects for question 11 (for BSIS course offerings). The SIS Home page was also indicated successfully for question 6 (IS class library reserves) by five out of six subjects, and for question 10 (the University Help desk line) by two out of two subjects.

The results for the Student Life at Pitt page pose some questions. For two different questions this was seen as the best starting point by many of the subjects, but none of those subjects had a bookmark or knew the URL. For question 7, to find out about chess clubs, fourteen subjects wanted to start at this page, and for question 15, to look for information about resident advisors in a dorm, twelve subjects indicated it. One possibility is that subjects knew easily how to get to that page, for example by going to the Students page off the Pitt Home page, and therefore envisioned that as the start to the challenging part of the navigation task. Data could be collected in different ways to explore this kind of issue further.

4.2.5. Summary

Several web pages were indicated in the three parts of Experiment 1 as pages the subjects would use more frequently and access more easily than other pages. The Pitt Home page and the SIS Home page were among the highest in each of the three sections of this experiment, suggesting that they are used often as landmarks. The Pitt Home page was used by subjects in each of the path description questions, and the Students page was used in 4 out of 5 of these questions. Another indication of familiarity with the pages was that many subjects knew the URLs or had bookmarks for the Pitt Home page, the Course Web, the SIS Home page, and the Pitt News.

Experiment 1 captured a record of the web pages that the subjects could recall easily and that they use often. In chapter 5, this data will be analyzed with respect to the objective landmark values. The second experiment gathered data on characteristics of web pages.

4.3. Experiment 2

Experiment 2 consisted of seven questions that were asked to 30 about each of 24 web pages from the University of Pittsburgh web site. The questionnaire is included as Appendix C. Throughout the results and discussion, each web page is referred to by a short title rather than the more cumbersome URL. The URLs and page names are given in Appendix A. At the time of the experiment, two different web pages for the University library were accessible through typical means, and both were included in the experiment. Screen shots of a sample of the 24 web pages used are given in Appendix D. The demographics of the subject population are presented in section 4.2.1, and the results for each of the seven questions are given in section 4.2.2.

4.3.1. Demographics

The 30 subjects recruited were undergraduate students taking courses in the Department of Information Science and Telecommunications. There were 7 women and 23 men, ranging in age from 19 to 31. They were 19 seniors (fourth year students), 9 juniors (third year students) and one sophomore (second year student). The group included 25 native English speakers, 4 subjects with “excellent” English skills, and 1 with “moderate” English skills, according to their self-reports.

Using the same five-point scales that were used in Experiment 1, the subjects rated themselves in terms of familiarity with the World Wide Web and familiarity with the University of Pittsburgh web site as well as their frequency of use of the University web site and of the Web in general. Twenty-four subjects were “very familiar” with the World Wide Web, five were “moderately familiar,” and one had “some familiarity” with the Web in general. In terms of familiarity with the University of Pittsburgh web site, 8 were “very familiar,” 14 were “moderately familiar” and 8 had “some familiarity.” Frequency of use showed similar distributions, with 25 subjects indicating use of the World Wide Web “frequently,” 2 “moderately often,” and 3 “some.” Frequency of use of the University of Pittsburgh website included 7 subjects who use it “frequently,” 11 “moderately often,” 9 “some” and 3 “rarely.”

4.3.2. Web Page Characteristics

The seven questions in Experiment 2 addressed different web page characteristics. Subjects rated each web page for each of the characteristics using a five-point scale. The results are presented for each question.

Question 1: How often have you been to this page?

Several distinct groups are clear in the results for this question, as shown in Figure 4-8. The most frequently visited pages were CourseWeb, Pitt Home, Student Information Online, and the SIS Home page. The second tier of pages was in the mid-range for frequency and included the two Library pages, the Students page, the Accounts page and the Information Technology pages. The figure shows a gap of 0.8 points between the results for the SIS Home page and the University Library page, and a gap of 0.6 points between the Information Technology page and the Events calendar pages.

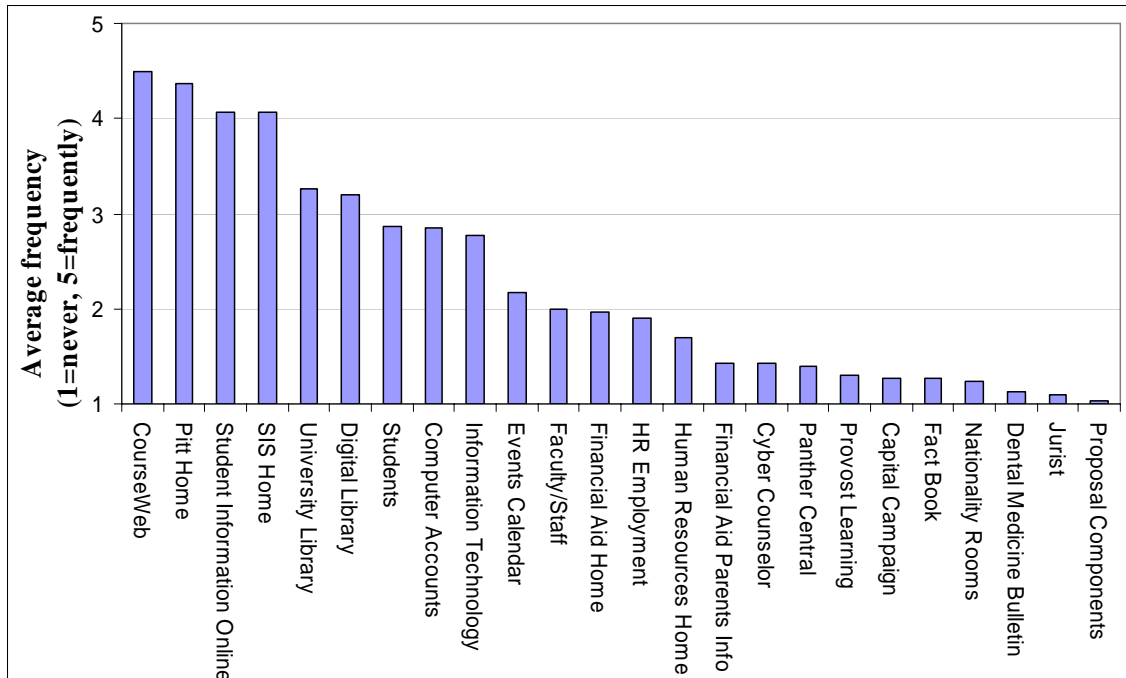


Figure 4-8. Average answers for “How often have you been to this page?” (Exp. 2, Q.1).

Responses were rankings of frequently (5), moderately often (4), a few times (3), rarely(2), or never (1). Average response was 2.3, median response was 1.9, and standard deviation 1.14.

The most frequently used pages may fall into two groups. Some, such as the CourseWeb, Student Information Online, and the Library pages enable access to specific, key information for students. Other pages, such as the Pitt Home and SIS Home pages, in addition to enabling access to known types of information, may be starting points for browsing or searching.

Question 2: Would you remember this page if you saw it again?

The pages most likely to be remembered according to question 2 were the CourseWeb, Pitt Home and SIS Home, followed closely by the Student Information Online and the Digital Library page, as shown in Figure 4-9. The University Library page, the Information Technology page, the Cyber Counselor and the Events Calendar are all clustered around the second highest level of “probably would remember.” All of the pages except two had average ranks of mid-level or higher. The two not likely to be remembered were the pages with Dental School Information and Proposal Components.

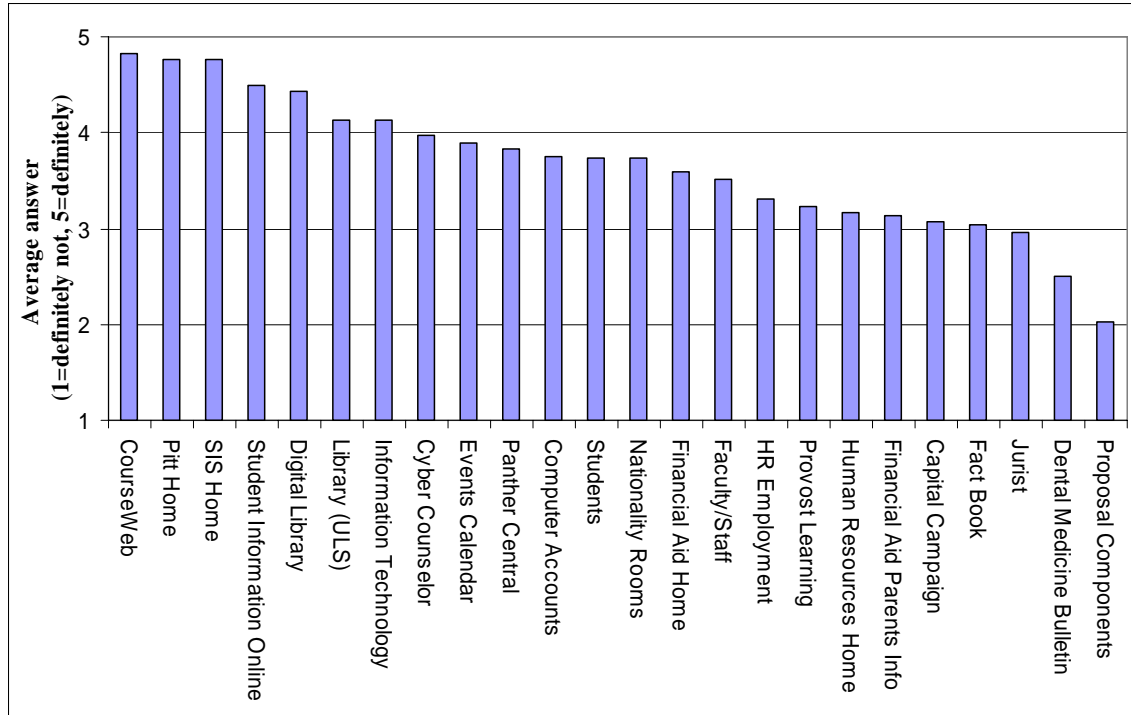


Figure 4-9. Average answers for “Would you remember this page if you saw it again?” (Exp. 2, Q.2).

Response choices were definitely (5), probably(4), maybe(3), probably not (2), or definitely not (1). Average response was 3.7, median response was 3.7, and standard deviation 0.72.

A variety of factors may affect whether a person is likely to remember the page, including prior familiarity with the page and visual uniqueness. Comparing these results to those for question 1 in Figure 4-8, the same pages are found within the top six for each question, as being memorable and having been visited frequently, and the same two pages have the bottom two average ranks as not memorable, and almost never visited by this subject population.

Question 3: How visually distinctive is this page?

Fourteen pages were ranked at the mid-level or higher in terms of being visually distinctive, but the Digital Library page, the Cyber Counselor and the Pitt Home page were the only pages with an average response of 4 (quite distinctive) or higher. Figure 4-10 shows the average answer for each page.

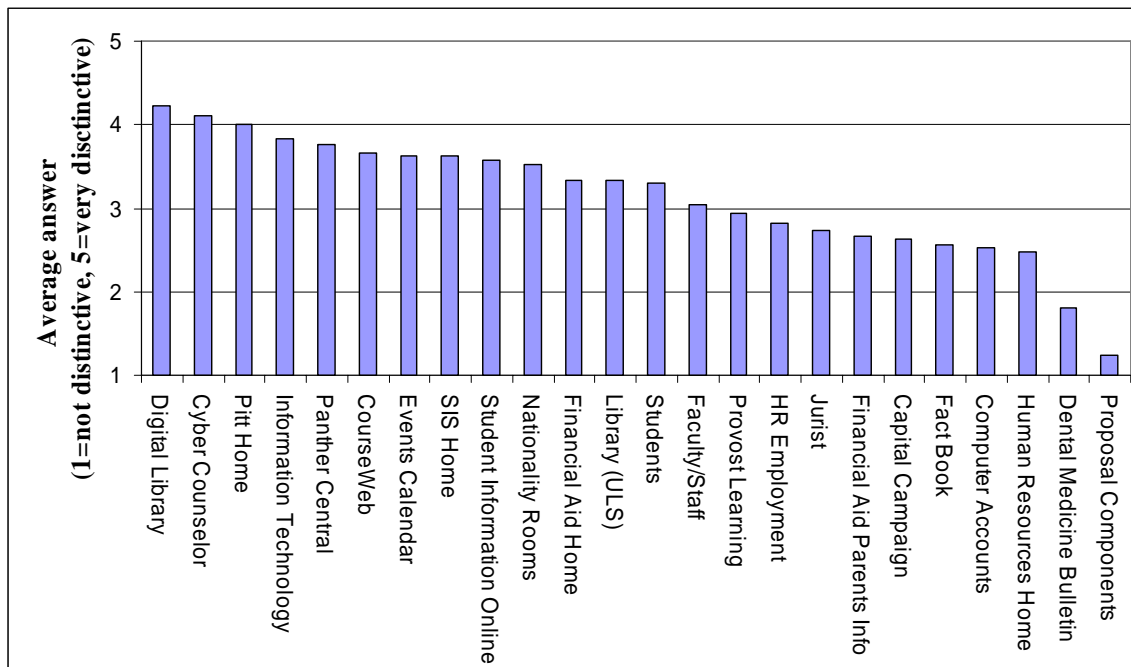


Figure 4-10. Average answers for “How visually distinctive is this page?” (Exp. 2, Q. 3).

The answer choices were very(5), quite(4), some distinctive features (3), slightly distinctive (2), or not distinctive (1). Average response, was 3.2, median response was 3.3, and standard deviation 0.73.

The pages that were ranked the highest all use a variety of photographs and graphics, while the two pages with the lowest ratings on visual response both contained basically just text with minimal variation of font or formatting. For example, the Digital Library page has a dark background, with a large photograph of the main library (Hillman Library) lit up against the night sky. The menu on the left uses color and white lettering against dark background to display options and a contrasting color box at the bottom of the page displays library news. The Cyber Counselor page has an eye-catching green robot or alien figure, with a black header and sidebar.

Question 4: How easily could you get to this page without using a search engine?"

Six pages were rated very highly by subjects in considering whether they could access the page without using a search engine, as shown in Figure 4-11. On average, subjects said it would be at least “moderately easy” to get to the Pitt Home page, the SIS Home page, the CourseWeb, Student Information Online, and the two Library pages.

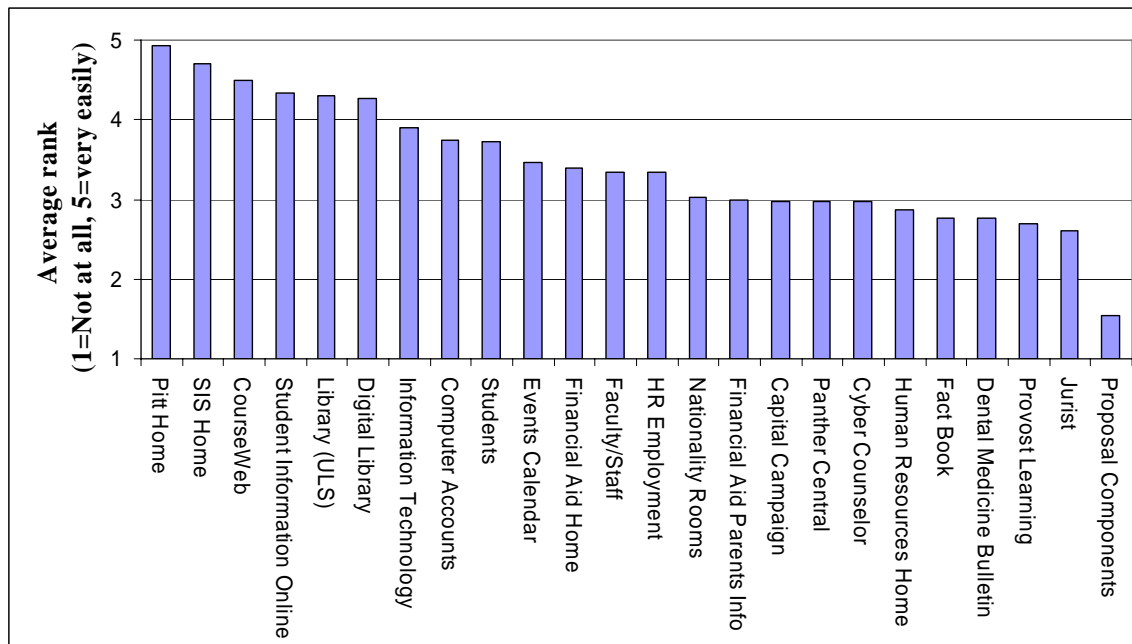


Figure 4-11. Average answers for “How easily could you get to this page without using a search engine?” (Exp. 2, Q. 4).

The response choices were very easily (5), moderately (4), could possibly (3), probably not (2), or not at all (1). The average response was 3.4, median response was 3.3, standard deviation 0.80.

This question required subjects to consider whether they knew the URL for the page, or whether they could tell enough about the page to navigate to it using links from a familiar web page such as the Pitt Home page. While six pages were ranked below the mid-level, only the Proposal Writing page was ranked much below the median.

Question 5: Do you know the exact URL for this page?

Subjects indicated a high level of confidence in knowing the URL for the Pitt Home page, the CourseWeb and the SIS Home page, and showed an average between “maybe” and “probably” knowing the URL for the Student Information Online, the Library pages, and the Information Technology Home page, as shown in Figure 4-12.

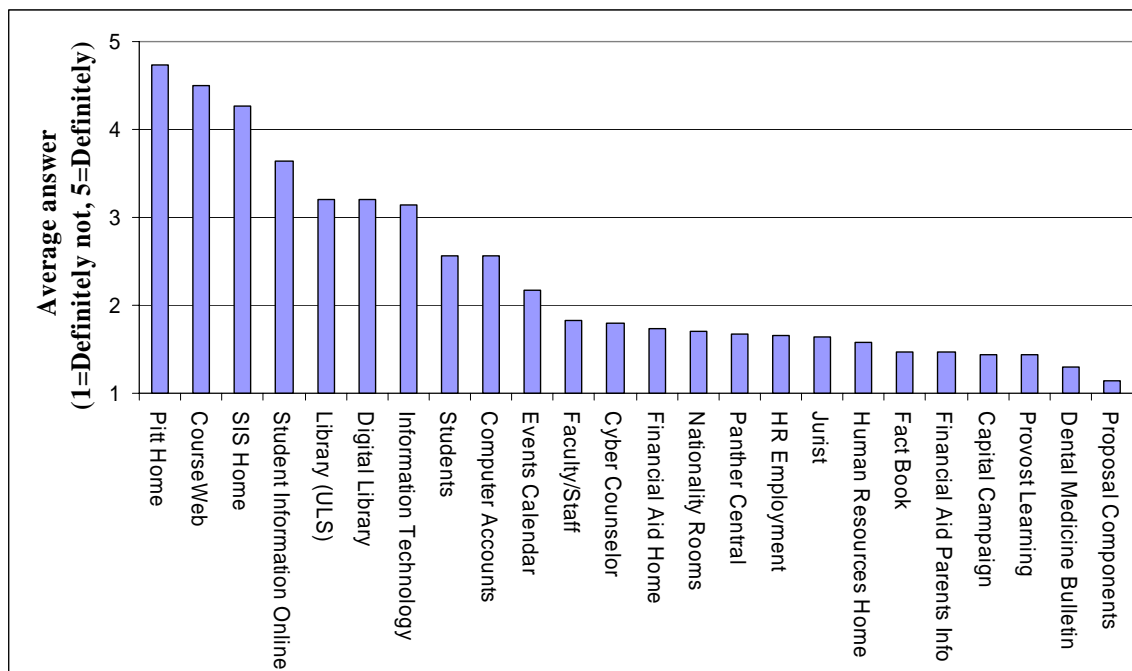


Figure 4-12. Average answers for “Do you know the exact URL for this page?” (Exp. 2, Q. 5).

The answer choices were definitely (5), probably (4), maybe (3), probably not (2), or definitely not (1). The average response was 2.3, median response was 1.8, and standard deviation 1.08.

The average data for the top four pages of this question corresponds to the number of subjects who indicated the top category for each page. The breakdown by individual shows that out of 30 subjects, 26 reported “definitely” knowing the URL for the Pitt Home page, 24 subjects indicated they “definitely” knew the CourseWeb URL, 20 subjects “definitely” knew the SIS URL, and 15 subjects “definitely” knew the URL for the Student Information Online page.

Question 6: How important is the content of this page in general?

The results of this question show 19 of 24 pages with an average importance of the mid-level response or higher. The Student Information Online page had the highest average rank of 4.5, nearly 0.4 points higher than the next highest. As shown in Figure 4-13, six pages were clustered around the 4.0 mark, with five more pages following closely ranked near 3.7.

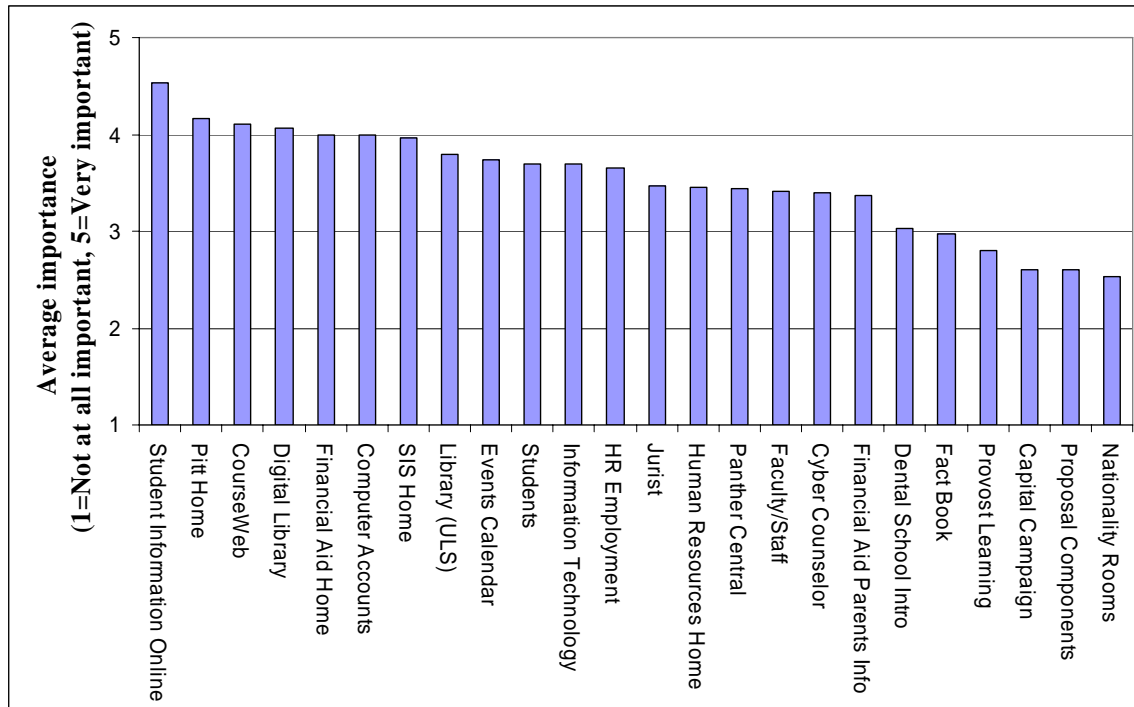


Figure 4-13. Average answers for “How important is the content of this page in general?” (Exp. 2, Q. 6).

The response choices were very important (5), moderately (4), neutral (3), not very (2), or not at all important (1). The average response was 3.5, median response was 3.6, and standard deviation 0.54.

Question 7: How important is the content of this page to you personally?

The Student Information Online and CourseWeb pages topped the list in importance for this subject population. Also of high importance to them were the SIS Home page, Pitt Home page and Computer Accounts, as shown in Figure 4-14.

The CourseWeb and SIS Home pages were more important to the subjects of this population than they believed they were in general. For question 6 above, the average answer was 4.1 for the CourseWeb and 3.97 for the SIS Home page. For question 7, the average answer for the CourseWeb was 4.43 and for the SIS Home page was 4.2.

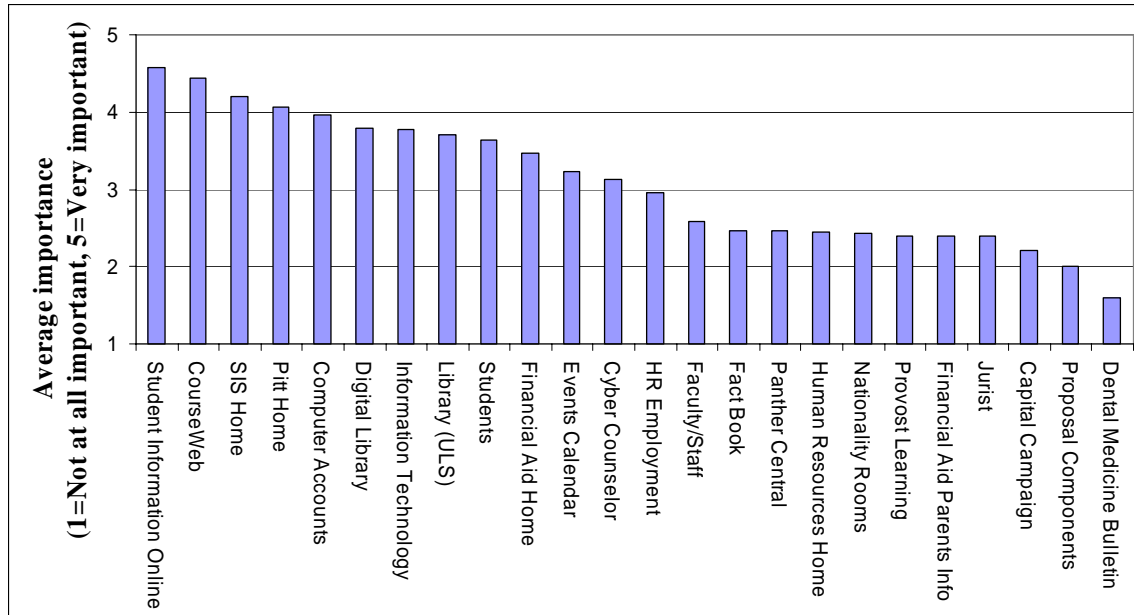


Figure 4-14. Average answers for “How important is the content of this page to you personally?” (Exp. 2, Q. 7).

The response choices were very important (5), moderately (4), neutral (3), not very (2), or not at all important (1). The average response was 3.1, median response was 3.1, and standard deviation 0.84.

4.3.3. Analysis of Subjective Landmark Values

The seven questions in this experiment touched on visual, structural and semantic aspects of the subject’s reactions to each web page. These results are combined to compute the subjective landmark value by summing the responses for each question for the page and combining them proportionally, as described fully in the research methodology (section 3.3.2). The results are presented in Table 4-10. As was done for the objective landmark values, the data are grouped into categories of high, medium and low.

The Pitt Home page, Course Web page, SIS Home page and the Student Information Online page had the highest subjective landmark values of those measured. The Students page, the Events Calendar, and the Financial Aid Home page were among the mid-level values. Those in the very lowest category included the Nationality Rooms page, the Human Resources Home page, and the Pitt Fact book.

Table 4-10. Subjective landmark value for web pages in Experiment 2.

Page name	Subjective Landmark Value	Range
Pitt Home	0.886	High > 0.70
Course Web	0.873	
SIS Home	0.848	
Student Information Online	0.839	
Digital Library	0.775	
Libraries (ULS)	0.739	
Technology Home	0.722	
Students	0.674	Medium 0.53 – 0.70
Events calendar	0.636	
Computer Accounts	0.626	
Financial Aid Home	0.618	
Cyber Career Counselor	0.588	
Panther Central	0.551	
HR employment	0.547	
Faculty/Staff	0.544	
Nationality Rooms	0.509	Low < 0.53
Financial Aid parents info	0.502	
Human Resources Home	0.489	
Jurist	0.488	
Provost Learning	0.477	
Fact book	0.475	
Capital Campaign	0.461	
Dental Medicine bulletin	0.411	
Proposal Components	0.338	

The data that makes up the subjective landmark value can also be examined with relation to the three different components. Questions 1, 2, and 3 related to visual characteristics, questions 4 and 5 to structural aspects, and questions 6 and 7 to semantic issues. The results for each question for the component were summed and then normalized in order to combine them in equal proportions. The breakdown of these components is shown in Figure 4-15 for a representative sample of the pages, with the sum giving the subjective landmark value.

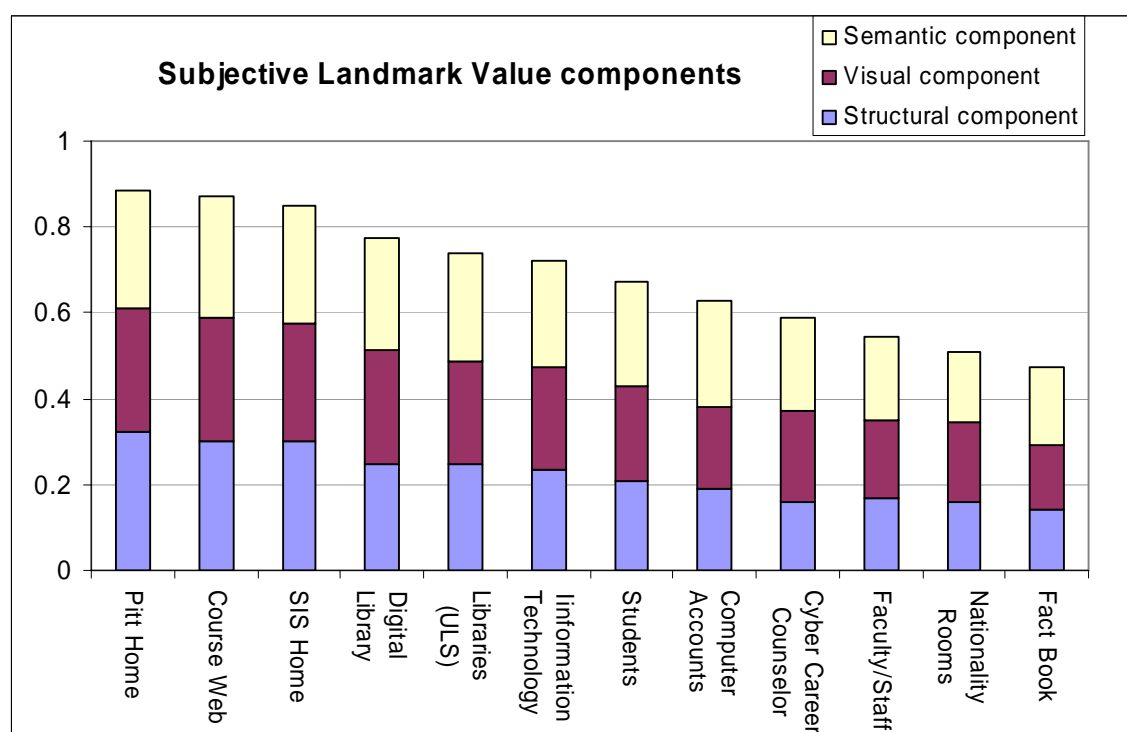


Figure 4-15. Subjective landmark value components for selected web pages.

It is interesting to note that there is not a lot of variability in the proportion of each component to the whole subjective landmark value. A corresponding table with the subjective landmark value components for all twenty-four web pages is included in Appendix G.

4.3.4. Summary

The analysis of web page characteristics using the data from Experiment 2 quantifies how individuals perceive each of the pages and allows for the computation of a value of the potential for the page to be used as a landmark based on the subjective data. The user ratings suggested that there is a continuum of landmark quality, with four pages notably higher than the rest. In addition, all three components of semantic, visual and structural distinctiveness were important criteria in determining what page is considered a landmark. The subjective landmark values are further analyzed in the next chapter.

4.4. Chapter Summary

The goal of this research was to examine two different ways of evaluating web page characteristics that reflect on the quality of the page as a landmark. The research reported here provided different ways of examining web pages, evaluating the relevance of the page for the subject population as well as the ease with which subjects use, recall or recognize different pages.

The objective landmark value of a page, computed from measures such as the link structure of the web surrounding a page, the relevance to the subject population and the level of visual characteristics, provides one way to examine the possible significance of the page as a landmark. The objective landmark values computed for a reference set of pages at the University of Pittsburgh web site have several widely spaced very high values and a very wide middle range. The subjective landmark value was computed from the subjective ratings of visual, semantic and structural features in Experiment 2. While the subjective landmark values showed uniformity in the importance of each of the components, with each web page having fairly equal proportions of the three types of characteristics, in the objective landmark value, the importance of the components in the computation varied widely. For example, for some pages a high objective landmark value came almost entirely from one of the components.

The path descriptions collected in Experiment 1 show which pages the subjects remembered and were able to recall within the web site. Four pages were mentioned by more than half the subjects in the course of the experiment: the Pitt Home page, the SIS Home page, the Athletics page, and the Campus Map. More than half the subjects also knew URLs or had

bookmarks for four pages; the Pitt Home page, the CourseWeb, the SIS Home page and the Pitt News page. The next chapter examines the pages that were recalled and the pages that were listed as easily accessible through the use of a URL or a bookmark, and evaluates those pages with respect to the objective landmark values. Integrating the data and examining the relationship of the objective and subjective landmark values is the task of the next chapter.

5. ANALYSIS

The two experiments in this dissertation provide data about how web pages are used, how well they are known, and how they are perceived. Examining two measures of landmark quality in relation to these data and to one another provide two key issues to be analyzed in this section. The objective landmark value was computed from direct measures taken of the web pages. The subjective landmark value was derived from the experimental data. The analysis focuses first on the relationship between the objective landmark value and the use and recall of web pages by the subjects, and second, on the relationship between the objective and subjective values, to determine whether objective measures could appropriately predict one's impression of a page as a landmark. Statistical analyses were performed using SPSS version 12. Correlation routines and stepwise linear regression analysis routines were used, and all correlation results presented are Pearson R correlations.

In this analysis, the first issue presented focuses on use and recall of the web pages, and examines those in relation to the objective landmark value for each page. The findings on ease of recall and hypothesis 1 are discussed in section 5.1, and on ease of access and hypothesis 2 are presented in section 5.2. The second main issue is to evaluate the relationship between the objective landmark value and the subjective landmark value for individual pages. A significant correlation between them is found and analyzed, and this evaluation is presented in section 5.3 and 5.4 with respect to hypotheses 3 and 4. Section 5.5 addresses limitations of the analysis and section 5.6 includes a brief summary.

5.1. Evaluation of Easily Recalled Web Pages

This section examines the classification of web pages as easily recalled, and looks at the objective landmark values computed for those pages. In the first part of Experiment 1, subjects were asked to describe how they would accomplish certain tasks using the University web site. The navigation paths provided by the subjects were generated without prompting, and thus

provide the recall data for this analysis. Several aspects of this data are considered, and then the objective landmark values of the web pages are examined with relationship to the research hypotheses.

5.1.1. Recall of Navigation Paths

The five open-ended path questions in Experiment 1 asked subjects to respond to task requests by providing several steps of a navigation path. Multiple paths were possible to accomplish the tasks, as shown in the figures in section 4.2.2, and the subjects each indicated one path, which was the path that they recalled most easily.

Because of the web structure, subjects listed different paths to navigate to specific pages. For example, a large proportion of the subjects might mention a particular page at the third or fourth step in their solutions, and because they indicated different paths, there was not an individual precursor page that was listed as many times. Data such as this may illuminate possible landmarks within the web structure, not just at the starting points. The later page that is arrived at using different navigation paths may be a potential landmark.

This availability of multiple navigation paths to solve each of the tasks is significant in understanding these results. A subject may be aware of multiple ways to complete the task, but only the one path is recorded in this analysis. For example, a given task might be equally well completed starting at either the Pitt Home page or the SIS Home page or by entering the valid URL for the target page. If the subject knows and provides the URL for the target page, they may also know how to access the Pitt and SIS Home pages, but that information is not recorded during the task. In addition, if the desired page can be reached from the SIS Home page, subjects starting at the Pitt Home page can use one of several methods of navigating from there to the SIS Home page. The sequence that is recorded is the set of steps that the individual envisioned actually using to complete the task; the path that is the most natural for them.

5.1.2. Objective Landmark Value Ranges

The objective landmark values were computed for over 30 pages in the University of Pittsburgh web site and the results ranged from 0.027 to 0.540, as presented in Table 4-1. Also identified in that table are three levels objective landmark values: high (over .29), medium (.20-.29), and low (below .20). The research methodology proposed conceptual descriptions for these categories. A high objective landmark value indicates a web page with very strong characteristics of a landmark. A medium value indicates a web page that has some landmark

qualities, and it is proposed that such a page is likely to be used as a landmark in certain topics or types of work, or by a particular sub-population. A low objective landmark quality value indicates that the web page has minimal features that have been identified as corresponding to landmark quality.

5.1.3. Evaluation of Hypothesis 1

Hypothesis 1 states that the pages easily recalled will have stronger objective landmark characteristics. The main idea behind this hypothesis is that people are more likely to remember for later use those web pages that have higher than average objective landmark characteristics. To examine this, the objective landmark value of the pages that were recalled in the path descriptions are compared to the reference set of pages in this analysis.

The pertinent results are presented in Table 5-1, combining the path description data from Table 4-2 and the objective landmark values given in Table 4-1. The objective landmark values are color coded as they were in the path diagrams, making it easy to see that the pages recalled in the path descriptions have predominantly high and medium objective landmark values, colored dark and medium green respectively. The average objective landmark value for the pages that were recalled by subjects in the path descriptions was 0.30 compared to an average objective landmark value of 0.25 for the full set of web pages examined. The average objective landmark value for the pages not recalled was 0.19.

The presence of medium and high objective landmark values for 18 out of 19 of these pages indicates they do have significant landmark qualities. This finding along with the difference in the average value for these pages compared to the reference set provides support for the hypothesis.

The objective landmark value for the Library (ULS) page was the only one that fell in the low category. The differences between the Library (ULS) page, the Digital Library page, and their objective landmark values were discussed briefly in section 4.1. The Digital Library page can be accessed from the Library ULS page, and services such as a list of libraries or the hours for all the libraries is in fact accessible from either Library page, though these options could be considered more obvious selections on the Digital Library page. In the path listings if the subject indicated the “library” from the Pitt Home page and then listed other links they would follow, it was assumed they would have used the links on the Library (ULS) page that is presented from the Pitt Home page. It is possible that some would have passed through the ULS page by

selecting the Digital Library page and then made their selections from there, but that they neglected to mention both steps.

Table 5-1. Objective landmark value of pages from path recall questions.

Page name	Number of Subjects who mentioned page in path descriptions	Objective Landmark Value
Pitt Home	30	0.540
SIS Home	17	0.300
Campus Map	15	0.322
Student Activities	15	0.429
Students	14	0.225
Course Web	13	0.265
SIS Library	13	0.202
Athletics	11	0.263
Events Calendar	9	0.309
Academics	9	0.211
Admissions	8	0.227
Life at Pitt	7	0.490
Digital Library	5	0.280
Library ULS	5	0.189
Virtual Tour	4	0.317
About Pitt	4	0.206
Pitt Program Council	4	0.383
Pitt Site Index	2	0.379
Pitt News	2	0.229

The SIS Library page was also recalled by many subjects and has an objective landmark value at the very low end of the medium range. Like the Library (ULS) page, this page was recalled with respect to the question in Experiment 1 about the hours of the SIS Library. Of the ten subjects who accessed the SIS Library page, four ended their path listings there, and six indicated following a link for the SIS Library hours, which was then the end point of those paths. For these ten subjects, either the SIS Library or the link to hours from it was considered the goal of this question, so while the SIS Library page may be on the low side in terms of objective landmark characteristics, the page was essential to the task and was well known by many of these subjects.

Figure 5-1 shows how the objective landmark values of the pages recalled in the path descriptions compare to the objective landmark values of the reference set of pages. The pages that were recalled are marked in blue, with the shade indicating the number of subjects who listed the page. Darker shades indicate a higher number of subjects who mentioned the page in the path descriptions.

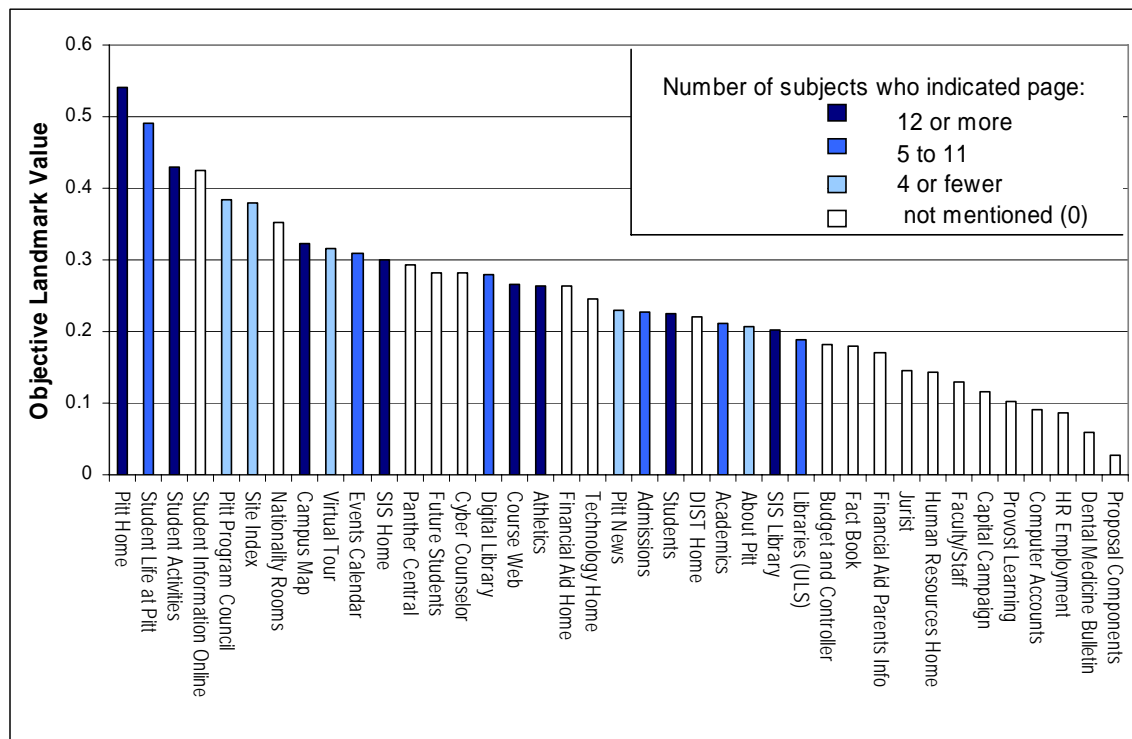


Figure 5-1. Objective landmark values of pages recalled in path descriptions.

The histogram shows that in comparison to the reference set of web pages, the recalled pages lie in the range of high and medium landmark values. Hypothesis 1 states that the pages easily recalled will have stronger objective landmark characteristics. The data supports this hypothesis with all but one of the pages recalled having a medium or high objective landmark value, and the set of pages having a higher objective landmark value compared to the reference set presented in Table 4-1.

5.2. Evaluation of easily accessed web pages

It is expected that people develop methods for easily accessing web pages that are useful to them as landmarks. One aspect of this study was to determine whether the study population had knowledge of a list of URLs at the University of Pittsburgh web site as well as whether they keep any of those web pages on a list of bookmarks or favorite sites for easy access. This data was presented in section 4.2.3. This section examines the classification of web pages as easily accessed, and looks at the objective landmark values for those pages. This data is used to evaluate the second research hypothesis.

5.2.1. Evaluation of Hypothesis 2

Many of the pages proposed in the experiment were ones that subjects had easy access to, either by URL or by bookmark. The results of the URL and bookmark knowledge are presented together with the objective landmark values in Table 5-2. Examining the set of pages with respect to the objective landmark value, it is found that four of the pages have high objective landmark values and six have mid-level values. The average objective landmark value for these pages is 0.33 compared to the average objective landmark value of all pages, which is 0.25.

Table 5-2. Objective landmark value of pages easily accessed by URL or bookmarks.

Page name	Number of Subjects who could access page by bookmark or URL	Objective Landmark Value
Pitt Home	30	0.540
CourseWeb	26	0.265
SIS Home	23	0.300
Pitt News	17	0.229
Student Information Online	10	0.425
Digital Library	9	0.280
Technology Home	7	0.246
DIST Home page	6	0.220
Athletics	3	0.263
Life at Pitt	1	0.490

With the combination of methods, nine pages were easily accessible by three or more subjects; and the remaining page, the Life at Pitt page, was accessible by one subject. The medium and high objective landmark values correspond to the expectation of the hypothesis that

the pages which subjects can access easily would be ones with higher than average objective landmark values. Figure 5-2 shows the objective landmark values for these pages in blue superimposed on the full reference set.

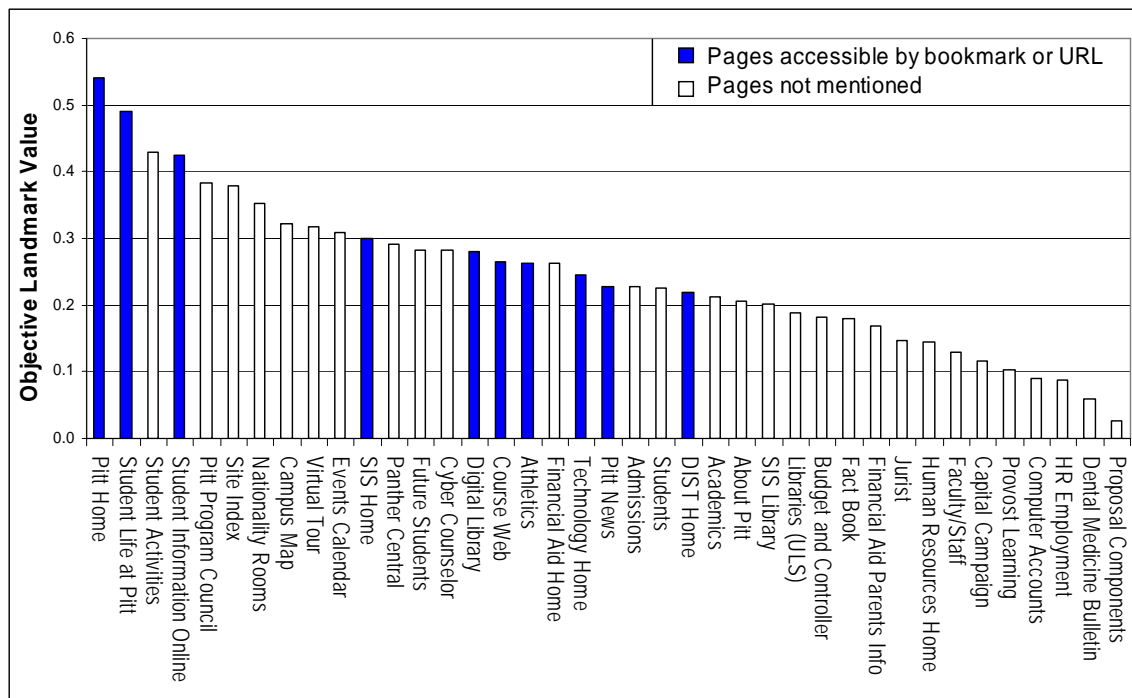


Figure 5-2. Objective landmark values of pages accessible by bookmark or URL.

The data show some support for hypothesis 2, that the pages easily accessed will have stronger objective landmark characteristics. The pages that were easily accessed by this subject population using URLs or bookmarks had high or medium objective landmark values. The easily accessed pages have stronger objective landmark characteristics than the norm, as evaluated by the objective landmark value.

5.2.2. Evaluation of preferred starting points

A second aspect of the research study is related to the issue of different pages that are easily accessed. The multiple choice style questions presented in section 4.2.4 asked subjects to determine the starting point to use for various web tasks, such as “how would you check your grades for the previous semester?” The information of which pages an individual can easily access, as described above, and the pages they choose as starting points can be examined in relation to the objective landmark values to create a picture of one use of the pages as landmarks. Table 5-3 shows the eight pages that were indicated as starting points in the experiment, how many subjects selected the page as a starting point, and the objective landmark value of the page.

Table 5-3. Objective landmark value of starting points in multiple-choice questions.

Page name	Number of subjects who indicated page as a starting point	Objective Landmark Value
Pitt Home	29	0.540
SIS Home	23	0.300
Digital Library	23	0.280
Life at Pitt	17	0.490
Student Information Online	16	0.425
Technology Home	10	0.246
Course Web	6	0.265
DIST Home page	2	0.220
Pitt News	0	0.229
Athletics	0	0.263

The pages that were used as starting points have medium or high objective landmark values, with an average of 0.35 compared to the overall average of 0.25. This result is important because it shows that, for a variety of tasks with different content areas, a large proportion of the subjects chose as a starting point a page with a moderate or high objective landmark value.

These pages account for 94% of the chosen starting places for the questions asked. The breakdown of how many other pages were used as starting points was shown in Table 4-8. For

five of the questions, only one person used another method, and for one question no one did. For the remainder of the questions, at most four subjects chose a different method of starting than one of these pages.

5.3. Relationship between the objective and subjective landmark measures

This section examines the relationship between the objective and subjective landmark values for a set of pages. The objective landmark value is computed from aspects that are measurable from the web page itself, while the subjective landmark value is computed from the responses of subjects regarding the web pages. Hypothesis 3 of this work is that the subjective landmark quality value and the objective landmark quality value will be positively correlated. Table 5-4 presents the two landmark values, with the high, medium and low ranges for each represented by the dark green, medium green and light green colors.

The results presented in Table 5-4 are significantly correlated with a correlation coefficient of .729 ($p=.01$). This important result indicates that the objective landmark value could be used to predict which web pages a user would find useful as a landmark. This research supports hypothesis 3, that the subjective and objective landmark values will be positively correlated.

The colors on the table give a visual answer that the two values are related. The top eight pages in the table have a combination of high and medium values of the two measures. On the subjective side, this means that subjects envisioned the pages as accessible, memorable and significant in content. On the objective side, the pages had characteristics of being well-linked and thus accessible, having visual variety in terms of presence of graphics, and containing content relevant to this subject population. Three entries in the table have medium values on both measures. Medium values were defined as indicating pages that have some landmark qualities and might be used as a landmark in certain conditions. Two pages have the extremes of a high value on one scale and low on the other, and three pages have medium subjective values and low objective ones. Eight pages have low measures on both scales. The low measures are also important, because this part of the research explicitly examines subject evaluation of pages that have low objective landmark values. In the parts of this research that required subjects to recall information such as URLs for web pages, there was no explicit way to select and analyze those pages that were not recalled.

Table 5-4. Comparison of subjective and objective landmark values.

Page Name	Objective Landmark Value	Subjective Landmark Value
Pitt Home	0.540	0.886
Student Information Online	0.425	0.839
SIS Home	0.300	0.848
Course Web	0.265	0.873
Digital Library	0.280	0.775
Technology Home	0.246	0.722
Events Calendar	0.309	0.636
Panther Central	0.292	0.551
Libraries (ULS)	0.189	0.739
Students	0.225	0.674
Financial Aid Home	0.263	0.618
Cyber Counselor	0.281	0.588
Nationality Rooms	0.353	0.509
Computer Accounts	0.090	0.626
Faculty/Staff	0.130	0.544
HR Employment	0.087	0.547
Financial Aid Parents Info	0.170	0.502
Fact Book	0.179	0.475
Jurist	0.146	0.488
Human Resources Home	0.144	0.489
Provost Learning	0.102	0.477
Capital Campaign	0.117	0.461
Dental Medicine Bulletin	0.059	0.411
Proposal Components	0.027	0.338

From the point of view of a designer who might want to make use of the potential for an item to be a landmark, the correlation of the objective and subjective landmark values is important. This indicates that objective measures may be able to indicate the use of an element in the environment as a landmark. To be effective, however, further work will be needed to examine the components of the objective value in more detail. The relationship of the objective and subjective values can also be examined in terms of the individual objective features. An initial analysis is provided here.

The objective landmark value is made up of three components; visual, semantic and structural. The subjective landmark value has a significant correlation to each of these components, as shown in Table 5-5, and none of the components are significantly correlated with each other.

Table 5-5. Correlation of subjective landmark value and the three objective components.

	Subjective Landmark Value	Visual Component	Semantic Component	Structural Component
Subjective Landmark Value	1	.414(*)	.670(**)	.511(**)
Visual Component	.414(*)	1	.115	.047
Semantic Component	.670(**)	.115	1	.227
Structural Component	.511(**)	.047	.227	1

* Correlation is significant at the 0.05 level (1-tailed).

** Correlation is significant at the 0.01 level (1-tailed).

Each of the three components was based on several direct measures of features of the web page. The structural component was computed using three direct measures, the number of links into the web page, the number of links available on the web page, and the depth of the URL. The visual component used the amount of the page covered by graphics and the total area of the page to evaluate the visual aspects of the page. Finally, the semantic component was the percent of words on the page that matched the subject profile. These measures are fully described in section 3.3.1. An examination of the relationship of the individual objective features to one another and to the subjective landmark value shows a number of significant correlations that are notable. The correlations are given in Table 5-6, showing that the subjective landmark value is significantly correlated with four of the five features used in the objective measure.

The features and a selection of the correlations presented in Table 5-6 will be discussed in turn. The subjective landmark value had highly significant correlations ($p = .01$) with the depth of the URL and the percent of words present on the page that matched the subject profile. It was also significantly correlated ($p = .05$) with the number of links to a page and the percentage of the page covered by graphics. Thus, a high subjective landmark value is correlated with a short URL, a high number of links into the page, a high percentage of related terms, and a high percentage of graphics on the page. The number of links out of the page is a count of how many

links are present, enabling navigation from that page. The number of links out of the page was not found to be significantly correlated to the subjective landmark value as measured in these experiments.

Table 5-6. Correlation of subjective landmark value and features of the objective landmark value.

	Subjective Landmark Value	# Links Out from Page	# Links to Page	Depth of URL	Percent Graphics Area	Percent Matching Words
Subjective Landmark Value	1	.133	.413(*)	-.689(**)	.399(*)	.650(**)
# Links Out from Page	.133	1	.267	-.320	-.453(*)	.276
# Links to Page	.413(*)	.267	1	-.397(*)	.069	.169
Depth of URL	-.689(**)	-.320	-.397(*)	1	-.399(*)	-.502(**)
Percent Graphics Area	.399(*)	-.453(*)	.069	-.399(*)	1	.098
Percent Matching Words	.650(**)	.276	.169	-.502(**)	.098	1

* Correlation is significant at the 0.05 level (1-tailed).

** Correlation is significant at the 0.01 level (1-tailed).

Pages with shorter URLs had a high number of links to the page as shown by the significant negative correlation. This is easily envisioned by an example of the home page, <http://www.pitt.edu>, which is accessible from links on many if not nearly all pages in the domain of the University web site. Depth of URL was also found to have a significant negative correlation with the percent of words matching the subject profile. This means that pages with a short URL had a high percentage of related terms on the page.

Additional features beyond those used to compute the objective landmark value were measured for each of the web pages. In addition to the measurement of the amount of graphics on the page and the total area of the page, the areas of three other types of elements on the page were measured: the area covered by links, by plain text and by blocks of contiguous unused or blank space. In addition, data for the frequency of access was available for several of the web pages examined, and estimates of access were made for the other pages. A table of the bivariate

correlations for all of these features is included in Appendix H, and provides an opportunity for further work in the improvement of the efficiency of the objective landmark value algorithm.

In summary, the analysis shows support for hypothesis 3 in indicating a correlation between these two computations of landmark quality. A significant positive correlation between these values indicates that the direct measurement of objective characteristics may provide a reasonable estimation of what pages are likely to be useful as landmarks for navigation by this population. Detailed analysis of the web page features mentioned in this section is saved for future work.

5.4. Effect of weights on the correlation of objective and subjective measures

This research used values of equal weights for the combination of the structural, visual and semantic factors in the objective landmark value computation (Section 3.3.1). As shown in the previous section, the objective landmark value computed with the equally weighted components had a positive correlation with the subjective landmark value. This section examines the change in the relationship between the objective and subjective landmark values based on the weights used for combining the objective factors, in order to examine hypothesis 4. Hypothesis 4 states that the degree of correlation can be shown to be dependent on the weights selected for combining the visual, semantic and structural characteristics in the objective measure.

A variety of weights were tested to examine the effect of proportions of the objective landmark components on the correlation between the objective and subjective landmark values. Each component was positively correlated with the subjective landmark value; however none was as good a predictor alone as combined with the other components. As stated, the equally weighted components of one third gave a correlation coefficient of 0.729 ($p=.01$). Omitting any one or two components gave much lower correlations, ranging from 0.414 to 0.692 ($p=.05$ or $p=.01$).

While the correlation is strong with the equal weights, tests that included all of the components with one or two weighted more heavily than the others showed that a stronger correlation could be found. Combinations that included a high weight on the semantic component gave higher correlations than other combinations tested, and are shown in Table 5-7. Table 5-7 shows that assigning a high weight of 3/5 to the semantic component and lower weights of 1/5 each to the structural and visual components gave the highest correlation of these

options, though at the lower significance level (0.817, $p=.05$). For comparison, the other combinations of weights tested that were less significant are shown in Appendix I.

Table 5-7. Effect of different on correlation of subjective and objective landmark values.

Weights for each Objective Measure Component			Correlation Coefficient	
Semantic	Structural	Visual		
3/5	1/5	1/5	0.817	*
3/7	3/7	1/7	0.774	**
1/3	1/3	1/3	0.729	**
1/5	3/5	1/5	0.693	**
1/7	3/7	3/7	0.661	**
3/7	1/7	3/7	0.660	**

* Correlation is significant at the 0.05 level

** Correlation is significant at the 0.01 level

A stepwise linear regression analysis was used to find the optimal combination of the three objective measure components in comparison to the subjective landmark value. The results of the regression analysis are shown in Table 5-8. The model shows a very strong relationship of the weighted components to the subjective landmark value ($R=.833$), with weights of 0.547 for the semantic component, 0.371 for the structural component, and 0.333 for the visual component. It accounts for more than two-thirds of the variation of the subjective landmark value, with an R Square of .694.

Table 5-8. Regression analysis of objective components.

	Objective Landmark Components			Model Summary		
	Semantic	Structural	Visual	R	R Square	Adjusted R Square
Standardized Coefficients (Beta)	0.547	0.371	0.333	.833	.694	.648

This analysis supports hypothesis 4, that the degree of correlation between the objective and subjective measures is dependent on the weights used to combine the objective components. It demonstrates that the semantic component is the most important in predicting the degree of correlation, but that the structural and visual components are also important.

5.5. Limitations of the analysis

There are ten pages shown in Table 5-4 that are ranked at different levels by the subjective and objective landmark measures. Discussion of several examples will permit some examination of the reasons for the discrepancies. The CourseWeb page had a high subjective landmark value and medium objective landmark value. The CourseWeb page is the password protected entry point into the system for online course materials. The text on the page is very sparse, not reflected on the subject profile and it is not well linked to other pages. These factors led to low values on the semantic and structural components. The Information Technology Home page was also ranked high on the subjective measure and medium on the objective measure. This page is important to the University student population because it provides access to managing one's computer accounts as well as providing information about the software and computer labs available. In terms of content, however, it contains a lot of current news and reflects current issues, which do not necessarily provide a good semantic match with the subject profile. This is similar to what was seen for the student newspaper, Pitt News, which also had a low semantic component on the objective measure.

Pages such as the CourseWeb and Information Technology Home page are well-used, so incorporating the access data in the objective measure would improve the match with the subjective landmark values. One example described in section 2.7 is the algorithm of Mukherjea and colleagues which incorporated access frequency with a stronger weight than link structure and URL depth. In addition, improvements in the subject profile would address additional differences. At the time of this research, the Student Life at Pitt page that was selected as the subject profile did not mention several key topics and departments important in student academic life.

5.6. Summary

The data from these experiments provide some support for each of the hypotheses. This analysis shows moderate support for the first two hypotheses and strong support for the Sorrows and Hirtle (1999) tri-partite theory of landmark characteristics examined in the second two hypotheses. The objective landmark value and its three components have a significant correlation to the subjective landmark value, and the analysis produced an optimal model for generation of the objective landmark value. There was also evidence that web pages that are remembered well and accessed easily have characteristics that are identified with important landmark qualities.

A strong correlation was found between the objective and subjective measures of landmark quality in hypothesis 3. The subjective landmark value was computed from responses to questions regarding the visual, semantic and structural significance of the page. The objective landmark value used measurements of the web page that could be made directly. The aspects evaluated on the web pages using each of the measures came from literature on the characteristics of landmarks in both physical and electronic space, and provide an indication of which web pages might be best recalled and used as landmarks in the space. The significant positive correlation between the two landmark measures is exciting because it means that the ability to predict which pages will be envisioned as landmarks can be measured with objective characteristics.

The analysis of hypothesis 4 provided a potential modification to the algorithm used to compute the objective landmark value. Three components make up the objective measure, structural, visual and semantic. Regression analysis showed that the optimal weights of 0.547 semantic, 0.371 structural and 0.333 visual gave the highest correlation between the objective and subjective landmark values. Strong support for this model, shown by the high R value, emphasizes the importance of each of the three objective components.

Some support was found for the first two hypotheses as well. Although there was not a strong statistical tool to examine the ease of recall and ease of access, it was shown that the web pages that were recalled in the path descriptions and those that were easily accessed using bookmarks or URL knowledge did have higher average objective landmark values than the reference set of pages.

6. DISCUSSION AND FUTURE WORK

6.1. Major findings

Essential to effective and efficient navigation in a large-scale space is the development of a cognitive map of the environment. This cognitive representation includes an understanding of the organization of the space and enables one to make essential navigational choices and to plan routes. Landmarks are prominent elements or features in an environment that are more easily recognized and recalled than typical elements. A large body of research has shown that landmarks are key elements in a cognitive map of an environment and are used in orientation and navigation in physical space. For example, in studies asking subjects to recall details about a space, landmarks are more likely to be recalled than other elements in the environment. Landmarks in real space are defined by characteristics such as a clear and distinguishing shape, contrast between features of one structure and those that surround it, and prominence of spatial location.

This dissertation looked at landmarks on the World Wide Web, and examined how research on landmarks and navigation in physical space apply to an electronic information space. Strong evidence was found that landmarks do exist on the Web. They follow the nature of landmarks in real space: they are used as key elements in navigation and they are recalled easily. A variety of visual, semantic and structural characteristics are associated with landmarks in both the physical and electronic domains.

Landmark characteristics of web pages were evaluated with a measure called the objective landmark value for which a high value indicates an element that would make a higher quality landmark. The objective landmark value is based on measures related to three categories of characteristics: visual, semantic and structural. This dissertation found that the web pages easily recalled by subjects for navigation paths had higher objective landmark values than the average page. These pages had a greater presence of landmark characteristics that made them more visually memorable, more semantically meaningful, and/or more structurally useful. This research also found that web pages that could be easily accessed using bookmarks or URL

knowledge also had higher objective landmark values. These are navigational shortcuts that give people easy and efficient ways to get to pages that are used frequently, such as for starting points for navigation.

The research presented here also found that there is a significant correlation between the objective landmark value and one's impression of a web page as a landmark. A second measure, the subjective landmark value was calculated from subject rankings on questions related to whether the page was recognized or was likely to be remembered, whether it looked meaningful and useful, or how well the subject might be able to access the page. Together the responses created an indication of the page as a landmark according to each subject.

6.2. Implications for web design and use

This research has implications for two audiences. Web designers may make use of these results in initial design or re-design work. Users of web spaces may find navigation enhanced through the use of landmarks in navigation tools or improved design. While the activity of these two groups is interconnected, it is useful to examine the implications of this research from each perspective.

There are several ways for web site designers to make use of the findings presented here. The evidence that landmarks exist in the World Wide Web may encourage a designer to design specific landmarks in the web space or to use landmark quality to represent pages differently in overview diagrams or index pages.

Within a web site, web pages could be designed as landmarks by creating pages that have high values on each of the components of the objective landmark value: visual, semantic and structural. All of the characteristics of these objective measures can be easily computed in an iterative design cycle with the exception of the links into the node, which may increase with links from externally created pages thus also increasing the landmark quality of the page. Analysis of the target interest groups for the web site would give relevant information to create landmarks pertinent to individual groups or to the entire population.

The initial design of an entire web site with planned landmarks for an intended population may sound ideal, but the significance of landmarks and the methods to design them might also be applied to the re-design of individual pages within a web site with effective results. In this situation, data of access frequency can be used to identify web pages that are frequently

used. Evaluating the relationship of the frequently used pages to the remainder of the web site and then modifying the appropriate pages by enhancing characteristics related to landmark quality would result in more effective landmarks. It would also enhance the significance of landmarks in a site to be sure that pages that are not landmarks do not have the characteristics that would make one try to use them as such.

The findings also have interesting implications for users of web spaces. First, using the objective measure of landmark quality described in this research, landmarks could also be indicated effectively on navigational aids such as overview maps, fisheye views, graphical site maps or text site indices. A visually distinctive landmark will be noticeable even in a thumbnail image. The significance of a page in terms of landmark quality could be shown using various features such as icon size, color or font attributes. When features such as these are implemented that help identify potential landmarks, users of the space can incorporate this information to implement more effective navigation techniques.

From the point of view of an individual navigating a web site, the design of landmarks in the site or the use of a navigational aid with landmark information encoded should facilitate navigation and use of the site. The availability of access to the home page of a site has become widespread in web site design, and most users understand the meaning and effect of the “home” button when they see it. In addition to considering what characteristics and features should be present on a page to increase the likelihood that it is noticed and remembered as a landmark, research could examine whether there are features one could train users to notice or ways to include standard elements in the design that become learned as landmarks similarly to the “home” feature. Designs can incorporate visual, semantic and structural characteristics to complement one another creating effective landmarks in the space.

6.3. Methodology

The experimental work of this dissertation examined landmarks in two ways, through recall of elements in a web space, and by examining characteristics of elements in the space. This research was unique in that the first experiment looked at memory of URLs and memory for web pages that would be used for navigation tasks. Examining recall of information is common in research on navigation in physical environments, however most web space navigation research has made use of web server logs to analyze actual navigation. Research on navigation in

physical space has shown that recall methods are useful in examining elements that may serve as landmarks (Golledge, 1999). The second aspect of the present research examined landmarks through an analysis of web page characteristics. Objective and subjective measures were used to evaluate various web page characteristics.

Two issues noticed in this research are related to the data collection. First, not all of the URLs listed in answers were valid. And second, not all of the steps listed in the paths of the open-ended questions could be identified as valid web pages. In future research, it might be valuable to record the level of confidence of the subject has with regard to the items they have recalled. This method is widely used in memory studies.

This research provides the basis for several suggestions of modifications to the computation of the objective landmark value components. For the structural component, it makes sense to examine each data point in relation to its closeness to an ideal value, though this is not necessarily straightforward. The structural component was computed from the number of links into a node, the number of outgoing links on the page and the depth of the URL. The ideal for the number of links into a node seems infinite; such links make the node easily accessible without any detrimental effect. The ideal number of outgoing links on a page may be limited however, because when evaluating whether a page is truly useful as a landmark, usability comes into play. A very large number of links may decrease usability because the user may not effectively find the links that are significant for the task at hand. The ideal depth of a URL for the computation is also more complex than expected. A long URL is not necessarily harder to remember, so it might be that semantic analysis could be used to evaluate both relevance and ease of memorability of the elements in the URL.

The semantic component of the objective landmark value could be improved by use of more complex and complete text analysis algorithms that take into account not just the presence of the same words, but also levels of importance of different terms and other factors. There is extensive work in the field of information retrieval on the methods to compute relative importance of documents or text passages.

One approach that has been used to determine the semantic importance of a web page is the notion of “Information Scent.” This is an evaluation of whether it will be worthwhile to follow a particular link when looking for specific information (Chi, Pirolli and Pitkow, 2000). The method involves judging how well the links and surrounding information indicate the target

information. This takes into account that users make navigation choices based on the clues they have about whether the link will lead to something that addresses their current task. Chi et al. (2003) have automated aspects of determining Information Scent on a web site using a simulator of user activity. Aspects of Information Scent research might integrate well with the semantic component described in the present research.

To focus on visual distinction of a page compared to its surroundings, computation of the visual component of the objective landmark value could be modified to look at deviation either from a perfectly balanced page with equal amounts of graphics, text and links, or at deviation from the average values for those elements on the pages within the site. This suggestion is examined in more depth in the remainder of this section. The intensity of color and use of background color and texture could be incorporated more directly in the visual portion of the algorithm as well.

The reference set of 39 web pages had an average of 18.5% graphics, 23.4% text, 26.8% links, and 21.0% blocks of white space with no other elements. The standard deviation from the average was used for an initial exploration of the idea that visual difference from surrounding pages is an important characteristic for a landmark. Two figures present the presence of the four types of visual features and the deviation of each page from the average. Both figures show the pages in a range from those the greatest deviation from the average at the left side to those with the least deviation from average at the right.

The standard deviation of the presence of visual features on each page from the average is shown in Figure 6-1. This computation included four features: the percentages of graphics, text, links and blocks of white space. Zero on this scale means that the page does not differ from the average values for the features examined. The pages which have a very large or very small proportion of one or more features are found on the left of this figure. They deviate the most from the average values.

The proportion of visual features for each of the web pages is shown Figure 6-2. The average for each feature is presented in the right hand column of the chart. It is essential to note that for this research, measurements of graphics and link features were not exclusive, so a photo or graphic design that was also a link was included in the computation of both features. Thus the total percentages are sometimes greater than 100. Blocks of white space were considered to have potential importance visually, and so were measured whereas small blank portions of a page such

as margins between different elements were not measured. Therefore not every portion of the page is included, so pages do not total to 100 percent. While not ideal, this data nonetheless gives the opportunity for a basic exploration of this concept.

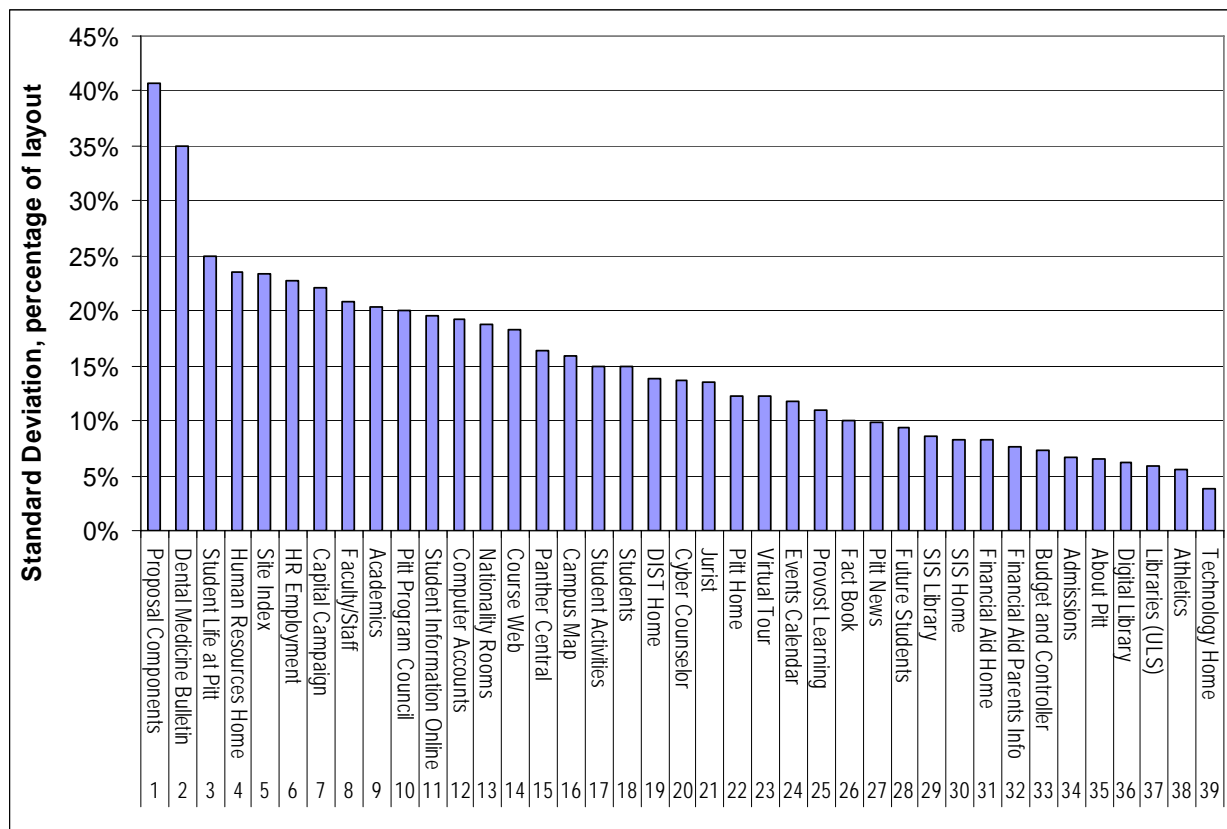


Figure 6-1. Standard deviation from the average based on four visual features.

There is a wide variety in the presence of the different features shown in Figure 6-2. The site index (#5) has a similar proportions of just two features, links and blocks of white space, while the Proposal Components and Dental Medicine Bulletin pages (#1and #2) are primarily text, and the Nationality Rooms (#13) and Student Information Online (#11) have a high proportion of graphics.

The preliminary analysis presented in these two figures indicates that while deviation may be significant in determining distinction from the surrounding elements of the environment, deviation from the average on all features may not be the best indicator of a landmark in other

senses of the definition. A page with only plain text and no outgoing links such as the Proposal Components page is unlikely to be useful as a landmark in the traditional sense for navigation and orientation, though it may be critical in importance for a given population based on its content. Deviation from a particular set of proportions rather than from the average may provide a better indication of the visual features that will make a page more memorable as a landmark.

While a number of researchers have already found the Sorrows and Hirtle (1999) theory of landmark components useful, it will continue to take some work to determine how to best implement this theory and develop appropriate, accurate and efficient measures for each of the features that make up the components important in landmarks.

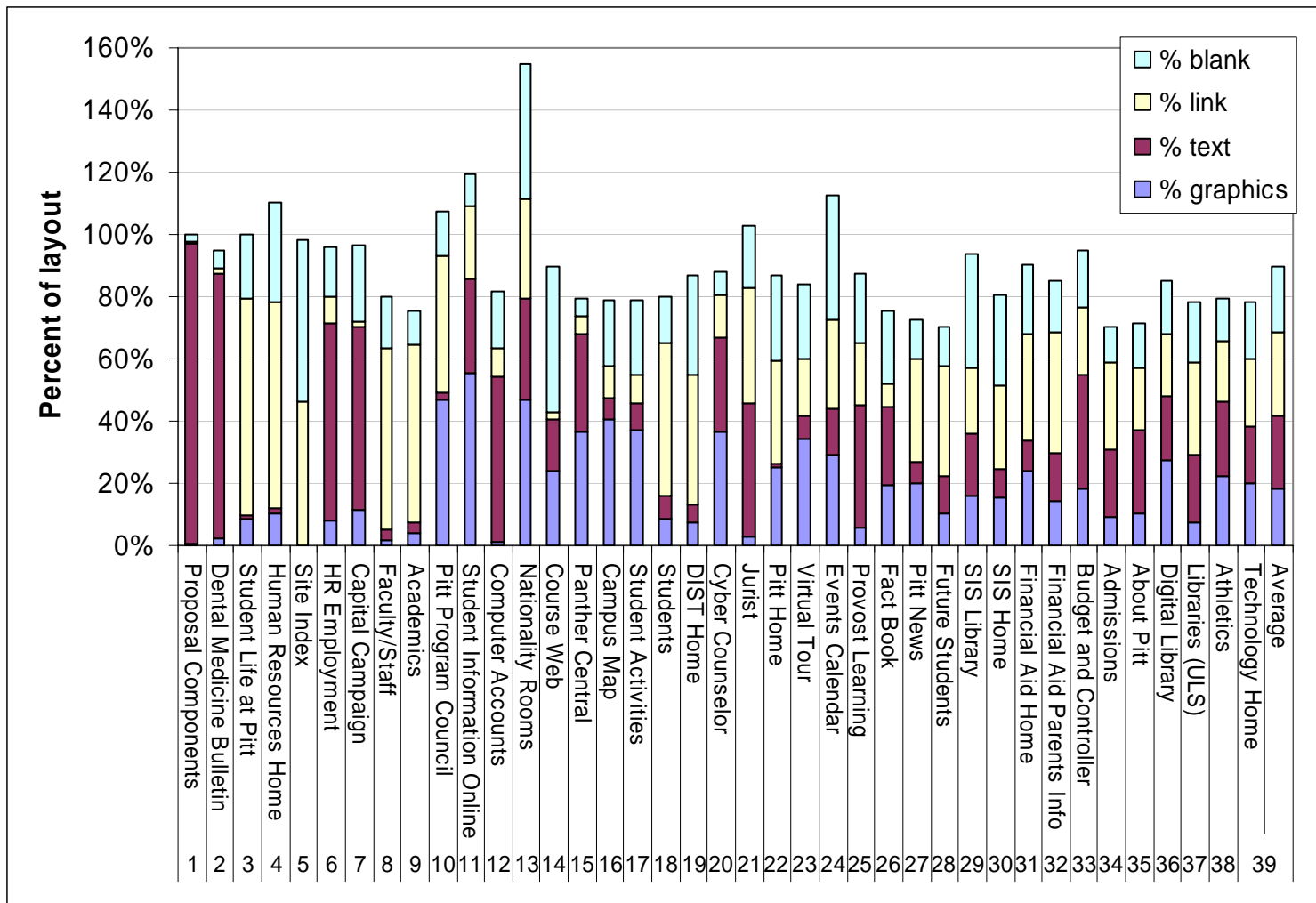


Figure 6-2. Proportion of visual features on web pages and the average across pages.

Areas of graphics and links may overlap, resulting in percentages beyond 1.0.

6.4. Alternate Computation of Landmark Saliency

In terms of fine tuning the objective landmark measure, it is worth examining the method of Raubal and Winter (2002) who implemented the tripartite theory to evaluate potential buildings as landmarks to be included in wayfinding instructions for physical space. Similar to the approach of this dissertation, their implementation of the theory involved gathering data for specific structural, visual and semantic features and then combining them in a measure of landmark saliency. Their work suggests an alternate algorithm for computing a landmark evaluation measure from this data.

The objective landmark value algorithm presented in this dissertation involved scaling the measurements of each component so that each was distributed over the same range, and then computing a landmark value from the weighted combination of these components. Raubal and Winter's (2002) algorithm evaluates each individual feature as significant or not, and then combines the significance flags in a weighted measure. For example, the structural features of URL depth, and number of links into and out of a node would each be tested to determine whether they were a significant contribution to the page as landmark. Then each of those data points of "significant" or "not significant" would be combined.

A preliminary analysis is presented to show the potential of applying their method to the data examined in this dissertation. Hypotheses tests are defined to evaluate a potential landmark on the basis of each individual feature. Values are defined for which the feature is considered to contribute to the importance as landmark. The potential landmark is evaluated with respect to each feature. Table 5-8 lists potential hypothesis tests for some of the web page features measured, in the manner presented by Raubal & Winter (2002). An examination of recent work as well as new research is needed to determine the appropriate cut off values for each test.

To complete this analysis, the results for each significance test of 1 for significant or 0 for not significant would be averaged for each of the components. The sum of these three significance results gives the final measure of landmark saliency. Raubal and Winter (2002) were successful in using this method to evaluate features of buildings and determine the relative landmark saliency of different buildings at an intersection.

Table 6-1. Examples of landmark significance tests for features of web pages.

Component	Feature	Property	Potential Test
Structural	Number of Links out of page	A	$20 \leq \alpha < 200$
	Number of links into node	B	$\beta > = 50$
	Depth of URL	Γ	$\gamma \leq 2$
Visual	Percent graphics	Δ	$\delta > .20$
	Percent text	E	$\varepsilon < .20$
Semantic	Percent of related terms	Z	$\zeta > .20$

6.5. Future Work

This dissertation took a tri-partite theory of visual, semantic and structural characteristics of landmarks based on research in physical and electronic space and examined it in an electronic information space. The theory proposed by Sorrows and Hirtle (1999) stood up to this initial examination with data providing a variety of support of the existence and importance of these three components of landmarks. A broader look at the extensive research on navigation and landmarks in physical space shows that there are additional findings that might be applied to electronic space. As described in the previous section and in section 2.10 of the literature review, the Sorrows and Hirtle (1999) theory of landmark components is already being examined in physical space as well.

As described in the previous section, refinements are still needed to the evaluation of landmarks through objective measures. Using the same theoretical base, Raubal and Winter (2002) similarly define formal measures to evaluate landmark saliency in physical space. Their conclusions reinforce the notion that such refinements are worth pursuing. They coded the semantic attribute by hand in their analysis, suggesting that this attribute is also particularly difficult to measure in physical environments. Further examination of this method will show which features can be most efficiently and cost effectively measured while still producing useful results.

There is some research and a great deal of anecdotal evidence that indicate that one of the key forms of navigation in the World Wide Web is the use of a search engine. It is interesting to note that there was indication of only minimal use of the search option in the experiment described in this work. Various site designs could be examined to determine if the presence of

well designed landmarks decreases use of a search engine while at the same time increasing effective and efficient use of a site.

Navigation in physical space has shown that people make use of identifiable and expected patterns within similar environments. For example, people make use of different expectations regarding the common features and buildings and their layout in cities than in rural towns. On a smaller scale, the layout of a shopping mall with department stores at the ends or middle of a concourse with small specialty shops between is another environmental pattern that enhances navigation by basic expectations and understanding. There are a variety of web design suggestions that outline expected content or structure for different types of sites such as academic, corporate or sales sites, but the universal understanding and expectations of these appear to be in their early stages. The development and influence of identifiable patterns in web site design is an area that may be enhanced by the present work on landmarks.

There is a great potential for difference between sites that are fully designed by one person or team and those that continue to grow in an organic fashion with many different page developers. It would be interesting to examine the differences between the presence of landmark pages in these two types of sites. One of the most significant characteristics for indicating the quality of a landmark according to the present research was the level of link connections of a web page. In a rapidly changing organically expanding site, pages that are found to be useful and significant will be linked to frequently by other pages, and therefore will see an increase in their objective landmark value. Examining this change over time would provide insight into a somewhat different arena of landmarks, the idea that landmarks even if not initially designed or intended may emerge over time.

APPENDICES

APPENDIX A

Web Pages

Table A-1. The URL and assigned short title for each web page used.

Short page title	Web page URL
Pitt Home	www.pitt.edu
Financial Aid	www.pitt.edu/~oafa/fahome.html
Capital Campaign	www.giveto.pitt.edu/capitalcampaign.html
Fact Book	www.ir.pitt.edu/factbook/general/index.html
University Library	www.pitt.edu/libraries.html
Faculty/Staff	www.pitt.edu/for_faculty.html
Human Resources	www.hr.pitt.edu/toc.htm
Temporary Jobs	www.hr.pitt.edu/employment/tempemploy/tempjobs.htm
Proposal Writing	www.pitt.edu/~offres/proposal/propwriting/components.htm
Events Calendar	www.events.pitt.edu/view.asp
Students	www.pitt.edu/students.html
Provost/Learning	www.pitt.edu/~provost/learning.html
Student Info Online	student-info.pitt.edu
Panther Central	www.pc.pitt.edu
Digital Library	www.library.pitt.edu
SIS Home	www2.sis.pitt.edu
Dental School Intro	www.univ-relations.pitt.edu/bulletins/dental/da.html
CourseWeb	courseweb.pitt.edu
Parents Financial Aid	www.pitt.edu/~oafa/parentsonly.html
Information Technology	technology.pitt.edu
Computer Accounts	technology.pitt.edu/accounts/index.html
Cyber Counselor	www.careers.pitt.edu/cybercounselor/cybercounselor.html
Jurist	jurist.law.pitt.edu
Nationality Rooms	www.pitt.edu/~natrooms

APPENDIX B

Experiment 1 Questionnaire

Demographic information

Please complete all items:

Age: _____

Sex: Male Female

Year in college: _____

Rate yourself on the following:

Familiarity with the use of the WWW in general

Very ----- Moderately ----- Some ----- Slightly ----- Not at all

Familiarity with the University of Pittsburgh web site

Very ----- Moderately ----- Some ----- Slightly ----- Not at all

Frequency with which you use the WWW

Frequently ----- Moderately often ----- Some ----- Rarely ----- Not at all

Frequency with which you use the University of Pittsburgh web site

Frequently ----- Moderately often ----- Some ----- Rarely ----- Not at all

Your English language ability

Native ----- Excellent ----- Moderate ----- Some difficulty ----- Very poor

Part 1

These questions ask about how you would find information using the University of Pittsburgh web site.

Answer the questions in this section by listing the actions you would take to find the answer. Please give up to five (5) sequential steps you would take.

If the first step in your process is not a URL, please state how you would get to that page.

A sample question and answer are given as an example:

Sample: How would you find the virtual tour of the University's Oakland campus?

If the first step is not a URL, please indicate how you would get to that first page.

Answer to Sample:

Step 1: Go to the University home page – this is the default page on my browser.

Step 2: Go to Admissions

Step 3: Find virtual tour

Turn the page to begin the questions.

Please answer:

How would you find the lecture notes for a BSIS class you missed?

If the first step is not a URL, please indicate how you would get to that first page.

1. _____

2. _____

3. _____

4. _____

5. _____

Please answer:

How would you find the date and time of the next home Panther game?

If the first step is not a URL, please indicate how you would get to that first page.

1. _____

2. _____

3. _____

4. _____

5. _____

Please answer:

How would you find the starting time of the movie showing on campus this weekend?

If the first step is not a URL, please indicate how you would get to that first page.

1. _____

2. _____

3. _____

4. _____

5. _____

Please answer:

How would you find a map of the University of Pittsburgh Oakland campus?

If the first step is not a URL, please indicate how you would get to that first page.

1. _____

2. _____

3. _____

4. _____

5. _____

Please answer:

How would you find the hours of operation of the SIS library for the upcoming vacation period?

If the first step is not a URL, please indicate how you would get to that first page.

1. _____

2. _____

3. _____

4. _____

5. _____

Part 2

Some University web pages are well-known and well-used. Please indicate for the following pages whether or not you know the URL, and if so what it is, or whether you have the page on your list of bookmarks/favorites.

The University home page http:// _____
Bookmark ☐ Yes ☐ No

The SIS home page http:// _____
Bookmark ☐ Yes ☐ No

The DIST home page http:// _____
Bookmark ☐ Yes ☐ No

The University Library System http:// _____
Bookmark ☐ Yes ☐ No

Life at Pitt http:// _____
Bookmark ☐ Yes ☐ No

CourseWeb http:// _____
Bookmark ☐ Yes ☐ No

Pitt News http:// _____
Bookmark ☐ Yes ☐ No

The Pittsburgh Panthers official site http:// _____
Bookmark ☐ Yes ☐ No

Information Technology home page http:// _____
Bookmark ☐ Yes ☐ No

Student Information Online http:// _____
Bookmark ☐ Yes ☐ No

Part 3

For the next section, we are only interested in where you start your work, not the complete process you would use. The format and answer choices are the same for each question.

Please answer what your first action would be:

How would you find whether Pitt has a chess club for undergraduate students?

I would start at:

- a. the University home page and follow links
- b. the University home page search field and enter the search phrase _____
- c. the SIS home page
- d. the DIST home page
- e. the University Library System
- f. the Student Life at Pitt page
- g. CourseWeb
- h. Pitt News
- i. the Pittsburgh Panthers official site
- j. the Information Technology home page
- k. Student Information Online
- l. A web search engine and enter the search phrase _____
(Google, Yahoo, Altavista, Lycos, etc.)
- m. Other _____

If other, you would get to this page by:

- ☐ typing the URL which is http:// _____
- ☐ my bookmark/favorites list
- ☐ Other: _____

Please answer what your first action would be:

How would you check your grades for the previous semester?

I would start at:

- a. the University home page and follow links
- b. the University home page search field and enter the search phrase _____
- c. the SIS home page
- d. the DIST home page
- e. the University Library System
- f. the Student Life at Pitt page
- g. CourseWeb
- h. Pitt News
- i. the Pittsburgh Panthers official site
- j. the Information Technology home page
- k. Student Information Online
- l. A web search engine and enter the search phrase _____
(Google, Yahoo, Altavista, Lycos, etc.)
- m. Other _____

If other, you would get to this page by:

- ☐ typing the URL which is http:// _____
- ☐ my bookmark/favorites list
- ☐ Other: _____

Please answer what your first action would be:

How would you find the phone number for the help desk at the University computing center?

I would start at:

- a. the University home page and follow links
- b. the University home page search field and enter the search phrase _____
- c. the SIS home page
- d. the DIST home page
- e. the University Library System
- f. the Student Life at Pitt page
- g. CourseWeb
- h. Pitt News
- i. the Pittsburgh Panthers official site
- j. the Information Technology home page
- k. Student Information Online
- l. A web search engine and enter the search phrase _____
(Google, Yahoo, Altavista, Lycos, etc.)

m. Other _____

If other, you would get to this page by:

- ☐ typing the URL which is http://_____
- ☐ my bookmark/favorites list
- ☐ Other: _____

Please answer what your first action would be:

How would you find whether the journal *Communications of the ACM* is received in electronic edition by the University library?

I would start at:

- a. the University home page and follow links
- b. the University home page search field and enter the search phrase _____
- c. the SIS home page
- d. the DIST home page
- e. the University Library System
- f. the Student Life at Pitt page
- g. CourseWeb
- h. Pitt News
- i. the Pittsburgh Panthers official site
- j. the Information Technology home page
- k. Student Information Online
- l. A web search engine and enter the search phrase _____
(Google, Yahoo, Altavista, Lycos, etc.)

m. Other _____

If other, you would get to this page by:

- ☐ typing the URL which is http://_____
- ☐ my bookmark/favorites list
- ☐ Other: _____

Please answer what your first action would be:

How would you check the BSIS course offerings for next semester?

I would start at:

- a. the University home page and follow links
- b. the University home page search field and enter the search phrase _____
- c. the SIS home page
- d. the DIST home page
- e. the University Library System
- f. the Student Life at Pitt page
- g. CourseWeb
- h. Pitt News
- i. the Pittsburgh Panthers official site
- j. the Information Technology home page
- k. Student Information Online
- l. A web search engine and enter the search phrase _____
(Google, Yahoo, Altavista, Lycos, etc.)
- m. Other _____

If other, you would get to this page by:

- ☐ typing the URL which is http:// _____
- ☐ my bookmark/favorites list
- ☐ Other: _____

Please answer what your first action would be:

How would you find a description of the University financial aid scholarships?

I would start at:

- a. the University home page and follow links
- b. the University home page search field and enter the search phrase _____
- c. the SIS home page
- d. the DIST home page
- e. the University Library System
- f. the Student Life at Pitt page
- g. CourseWeb
- h. Pitt News
- i. the Pittsburgh Panthers official site
- j. the Information Technology home page
- k. Student Information Online
- l. A web search engine and enter the search phrase _____
(Google, Yahoo, Altavista, Lycos, etc.)
- m. Other _____

If other, you would get to this page by:

- ☐ typing the URL which is http:// _____
- ☐ my bookmark/favorites list
- ☐ Other: _____

Please answer what your first action would be:

How would you find the office hours for the student health center?

I would start at:

- a. the University home page and follow links
- b. the University home page search field and enter the search phrase _____
- c. the SIS home page
- d. the DIST home page
- e. the University Library System
- f. the Student Life at Pitt page
- g. CourseWeb
- h. Pitt News
- i. the Pittsburgh Panthers official site
- j. the Information Technology home page
- k. Student Information Online
- l. A web search engine and enter the search phrase _____
(Google, Yahoo, Altavista, Lycos, etc.)
- m. Other _____

If other, you would get to this page by:

- ☐ typing the URL which is http:// _____
- ☐ my bookmark/favorites list
- ☐ Other: _____

Please answer what your first action would be:

How would you find the complete mailing address of the Information Science and Telecommunications department?

I would start at:

- a. the University home page and follow links
- b. the University home page search field and enter the search phrase _____
- c. the SIS home page
- d. the DIST home page
- e. the University Library System
- f. the Student Life at Pitt page
- g. CourseWeb
- h. Pitt News
- i. the Pittsburgh Panthers official site
- j. the Information Technology home page
- k. Student Information Online
- l. A web search engine and enter the search phrase _____
(Google, Yahoo, Altavista, Lycos, etc.)

m. Other _____

If other, you would get to this page by:

- ☐ typing the URL which is http://_____
- ☐ my bookmark/favorites list
- ☐ Other: _____

Please answer what your first action would be:

How would you get information on how to become a resident advisor in the dorm?

I would start at:

- a. the University home page and follow links
- b. the University home page search field and enter the search phrase _____
- c. the SIS home page
- d. the DIST home page
- e. the University Library System
- f. the Student Life at Pitt page
- g. CourseWeb
- h. Pitt News
- i. the Pittsburgh Panthers official site
- j. the Information Technology home page
- k. Student Information Online
- l. A web search engine and enter the search phrase _____
(Google, Yahoo, Altavista, Lycos, etc.)

m. Other _____

If other, you would get to this page by:

- ☐ typing the URL which is http://_____
- ☐ my bookmark/favorites list
- ☐ Other: _____

Please answer what your first action would be:

How would you find out what items are held on reserve for a current IS class?

I would start at:

- a. the University home page and follow links
- b. the University home page search field and enter the search phrase _____
- c. the SIS home page
- d. the DIST home page
- e. the University Library System
- f. the Student Life at Pitt page
- g. CourseWeb
- h. Pitt News
- i. the Pittsburgh Panthers official site
- j. the Information Technology home page
- k. Student Information Online
- l. A web search engine and enter the search phrase _____
(Google, Yahoo, Altavista, Lycos, etc.)
- m. Other _____

If other, you would get to this page by:

- ☐ typing the URL which is http:// _____
- ☐ my bookmark/favorites list
- ☐ Other: _____

APPENDIX C

Experiment 2 Questions

Demographic information

Please complete all items:

Age: _____

Sex: Male Female

Year in college: _____

Rate yourself on the following:

Familiarity with the use of the WWW in general

Very ----- Moderately ----- Some ----- Slightly ----- Not at all

Familiarity with the University of Pittsburgh web site

Very ----- Moderately ----- Some ----- Slightly ----- Not at all

Frequency with which you use the WWW

Frequently ----- Moderately often ----- Some ----- Rarely ----- Not at all

Frequency with which you use the University of Pittsburgh web site

Frequently ----- Moderately often ----- Some ----- Rarely ----- Not at all

Your English language ability

Native ----- Excellent ----- Moderate ----- Some difficulty ----- Very poor

Experiment 2 questions for each web page

1. How often have you been to this page?

Frequently ----- Moderately often ----- A few times ----- Rarely ----- Never

2. Would you remember this page if you saw it again?

Definitely ----- Probably ----- Maybe ----- Probably Not ----- Definitely not

3. How visually distinctive is this page?

Very ----- Quite ----- Some distinctive features ----- Slightly distinctive ----- Not distinctive

4. How easily could you get to this page without using a search engine?

Very easily ----- Moderately ----- Could possibly ----- Probably not ----- Not at all

5. Do you know the exact URL for this page?

Definitely ----- Probably ----- Maybe ----- Probably not ----- Definitely not

6. How important is the content of this page in general?

Very important ----- Moderately ----- Neutral ----- Not very ----- Not at all important

7. How important is the content of this page to you personally?

Very important ----- Moderately ----- Neutral ----- Not very ----- Not at all important

APPENDIX D

Experiment 2 Web Pages

A sample of the web pages used in Experiment 2.



Figure D-1. Pitt Home web page

<http://www.pitt.edu>



Figure D-2. Financial Aid Home web page

<http://www.pitt.edu/~oafa/fahome.html>

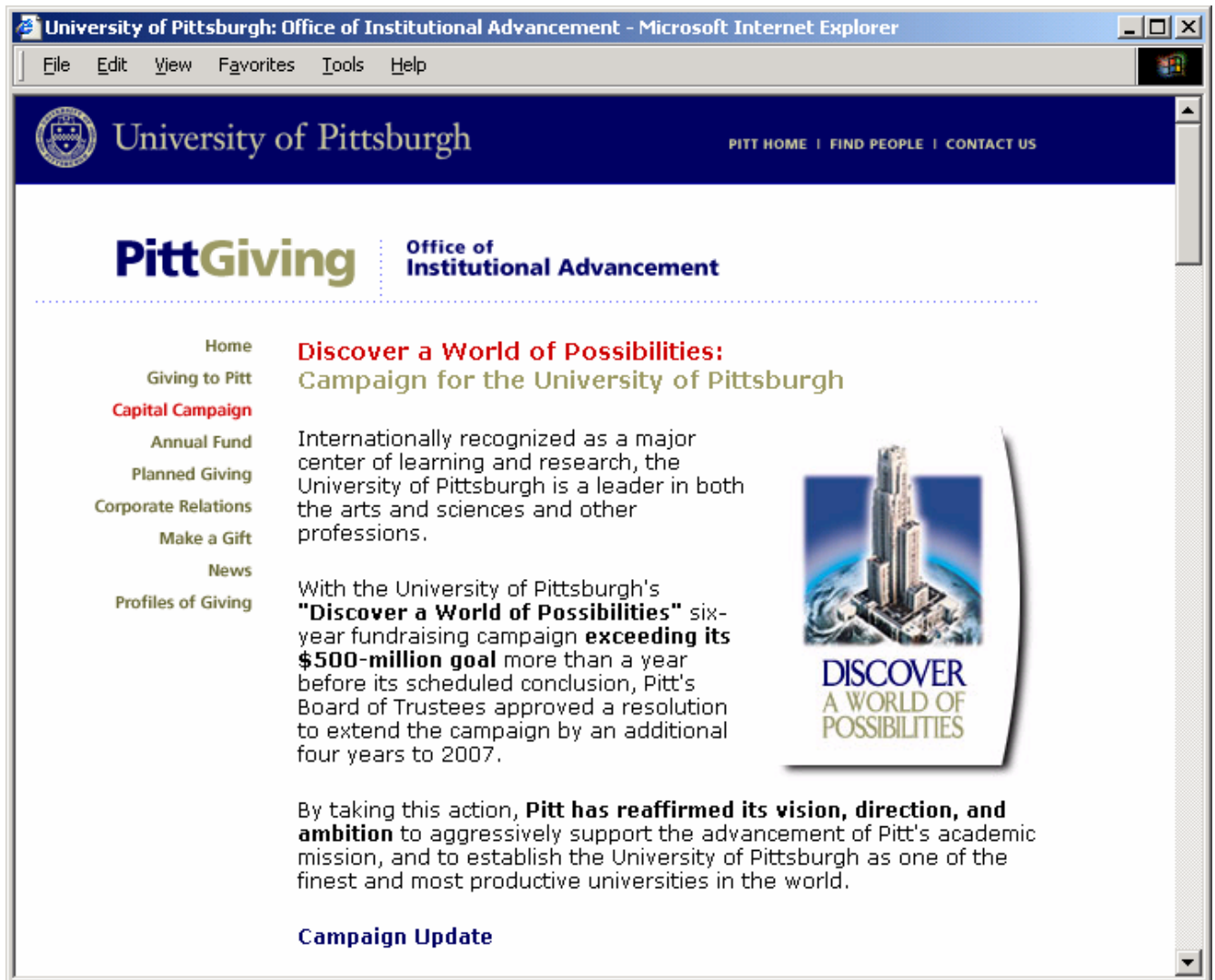


Figure D-3. Capital Campaign web page.

<http://www.giveto.pitt.edu/capitalcampaign.html>

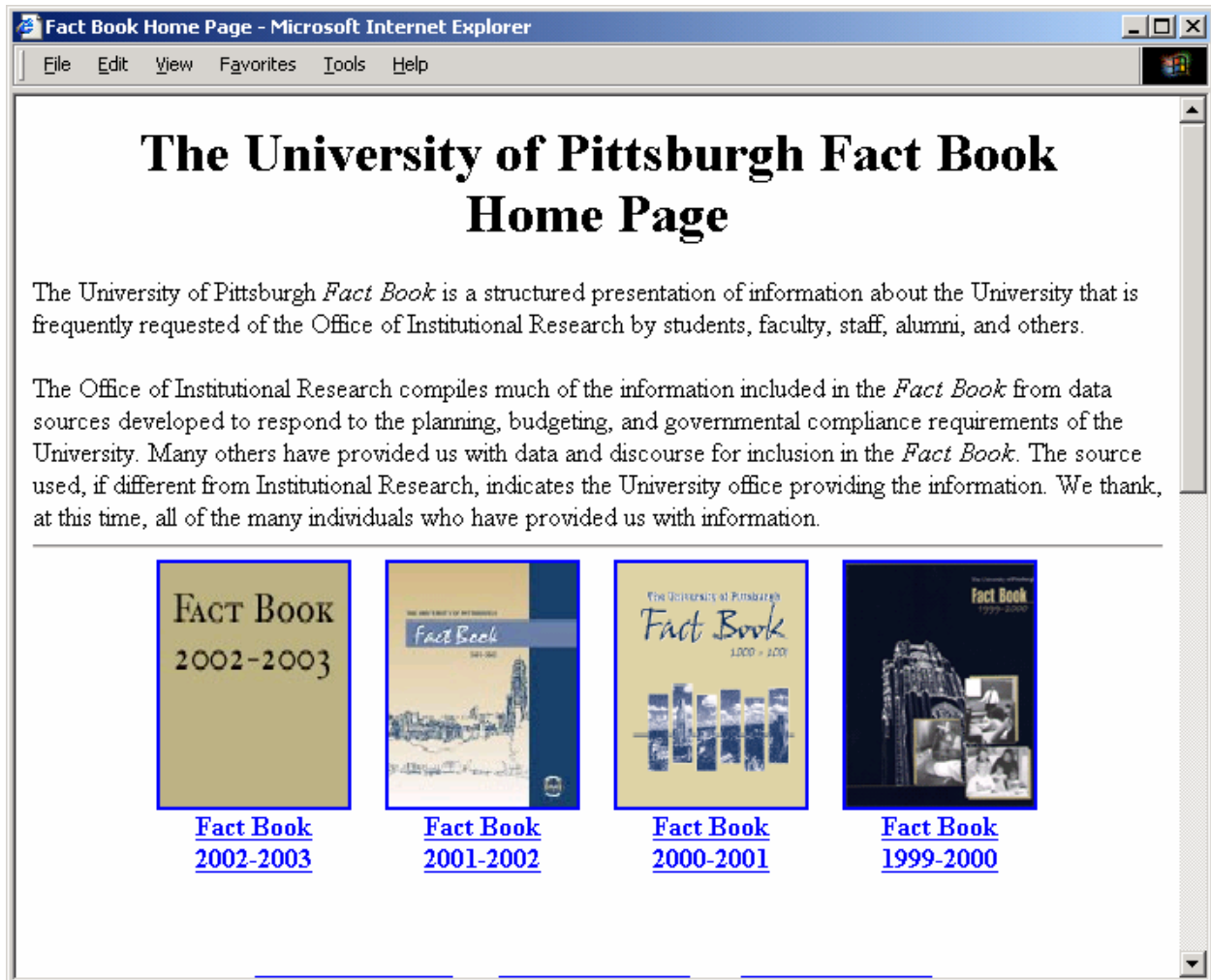


Figure D-4. Fact Book web page.

[http:// www.ir.pitt.edu/factbook/general/index.html](http://www.ir.pitt.edu/factbook/general/index.html)

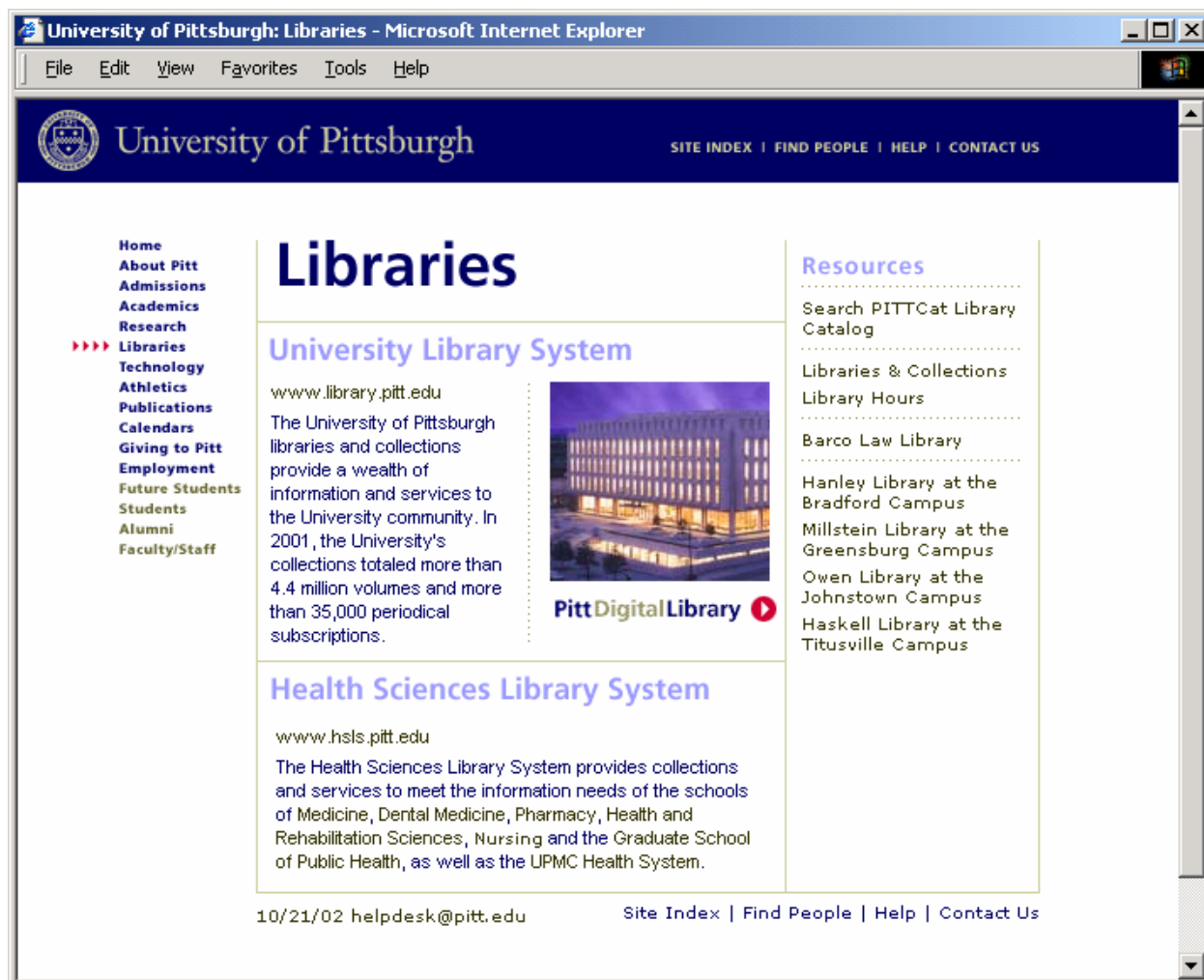


Figure D-5. Libraries (ULS) web page.

[http:// www.pitt.edu/libraries.html](http://www.pitt.edu/libraries.html)

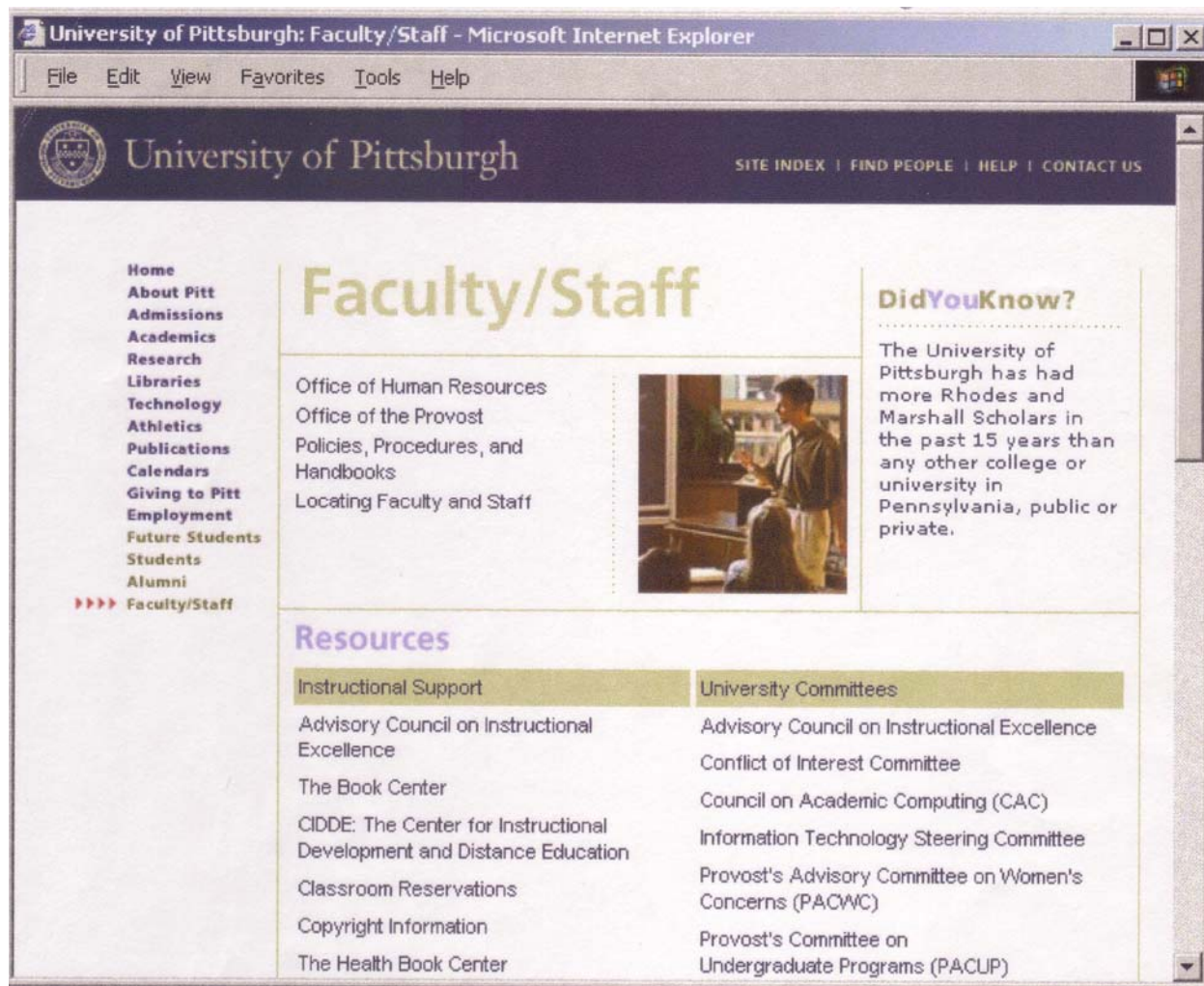


Figure D-6. Faculty/Staff web page.

http://www.pitt.edu/for_faculty.html



Figure D-7. Human Resources Home web page.

<http://www.hr.pitt.edu/toc.htm>



Figure D-8. HR Employment web page.

[http:// www.hr.pitt.edu/employment/tempemploy/tempjobs.htm](http://www.hr.pitt.edu/employment/tempemploy/tempjobs.htm)

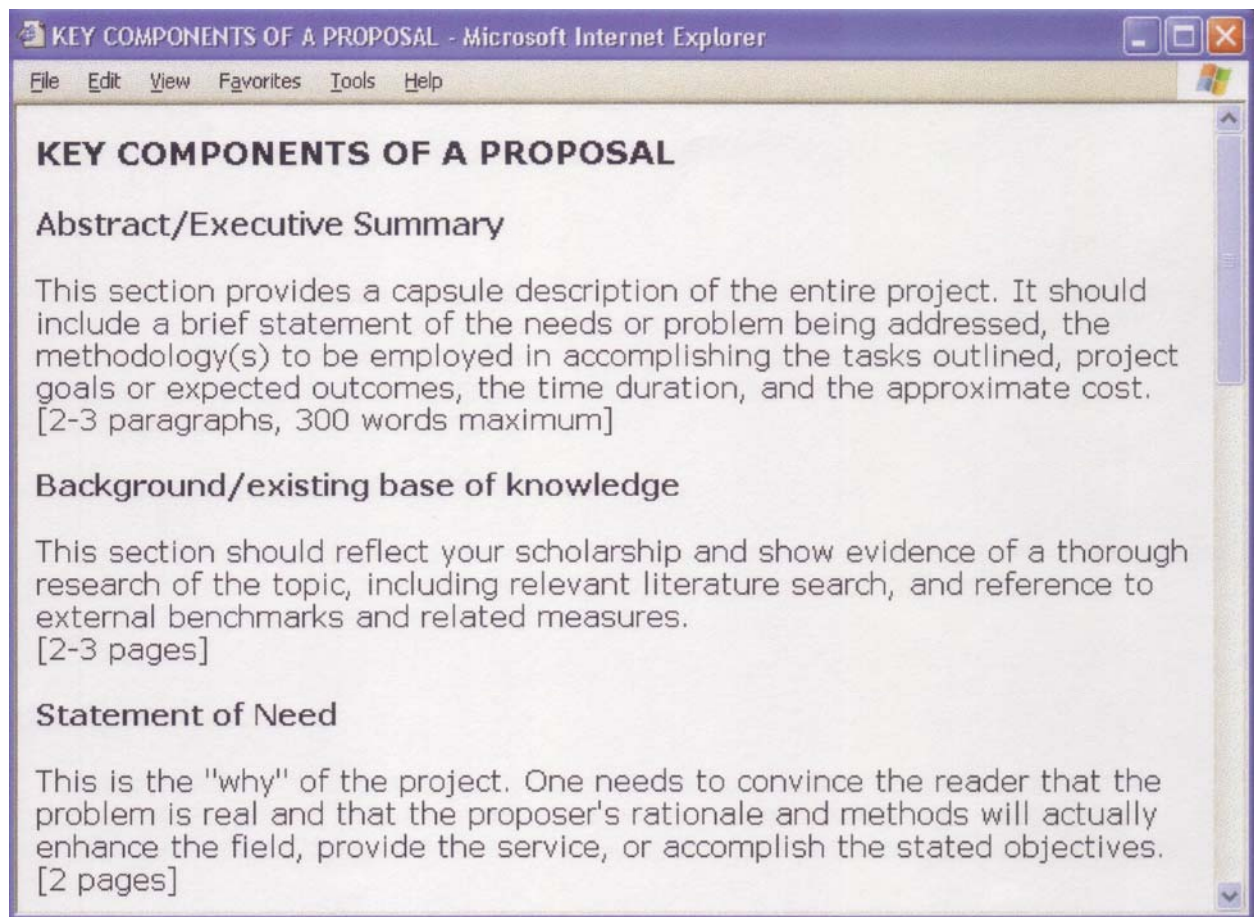


Figure D-9. Proposal Components web page.

<http://www.pitt.edu/~offres/proposal/propwriting/components.html>

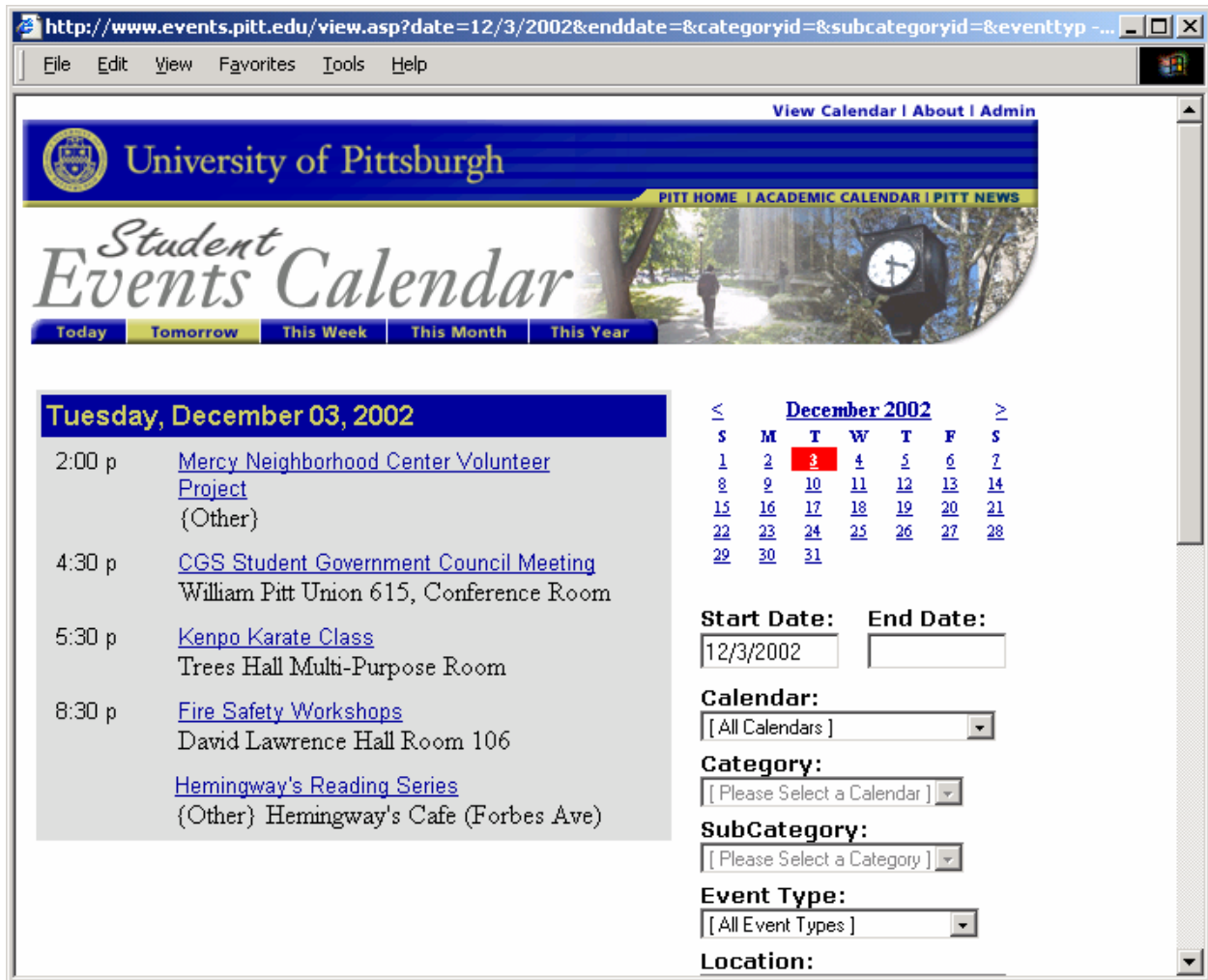


Figure D-10. Events Calendar web page.

<http://www.events.pitt.edu/view.asp>

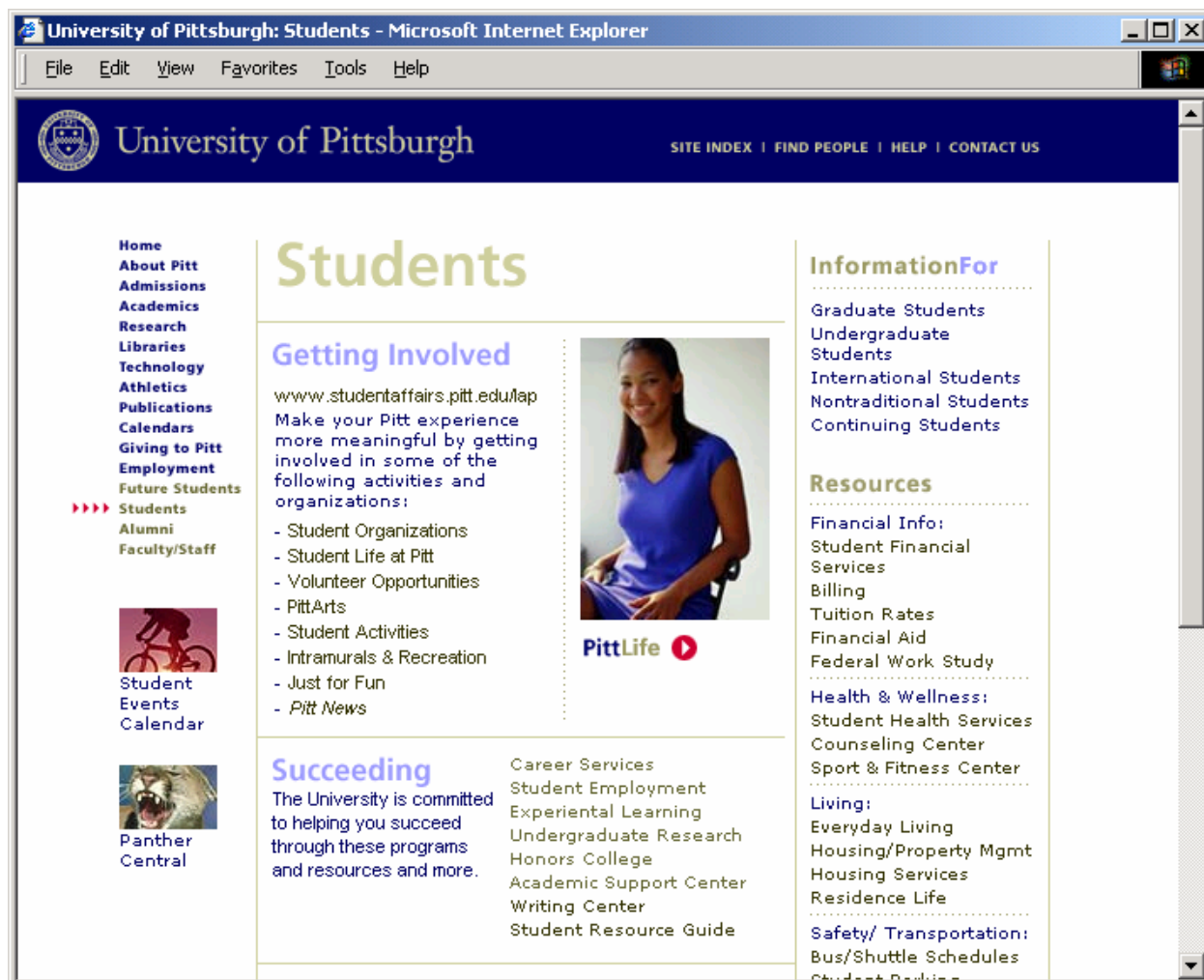


Figure D-11. Students web page.

<http://www.pitt.edu/students.html>



Figure D-12. Provost Learning web page.

<http://www.pitt.edu/~provost/learning.html>



Figure D-13. Student Information Online web page.

[http:// student-info.pitt.edu](http://student-info.pitt.edu)



Figure D-14. Panther Central web page.

[http:// www.pc.pitt.edu](http://www.pc.pitt.edu)



Figure D-15. Digital Library web page.

<http://www.library.pitt.edu>



Figure D-16. SIS Home web page.

<http://www2.sis.pitt.edu>

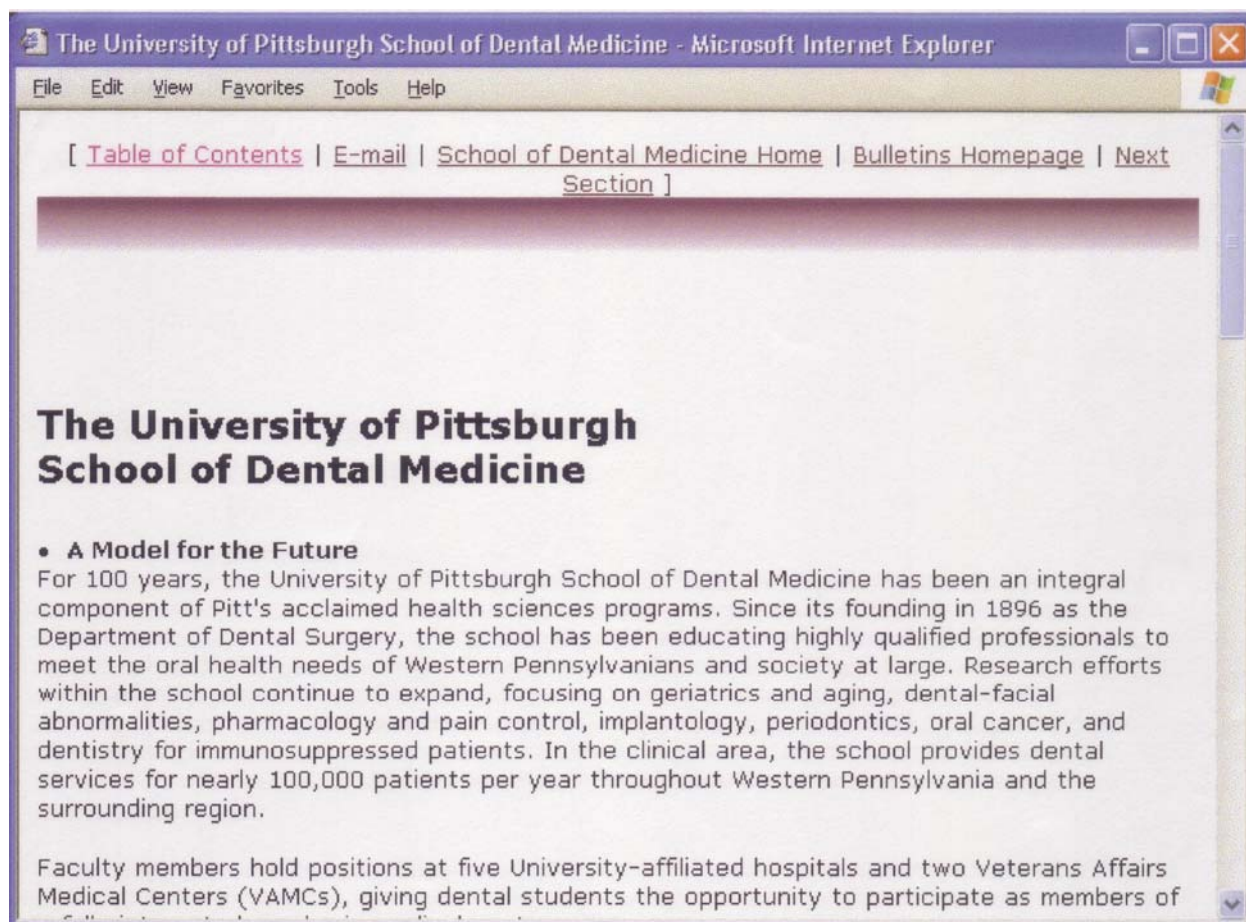


Figure D-17. Dental Medicine Bulletin web page.

[http:// www.univ-relations.pitt.edu/bulletins/dental/da.html#1](http://www.univ-relations.pitt.edu/bulletins/dental/da.html#1)

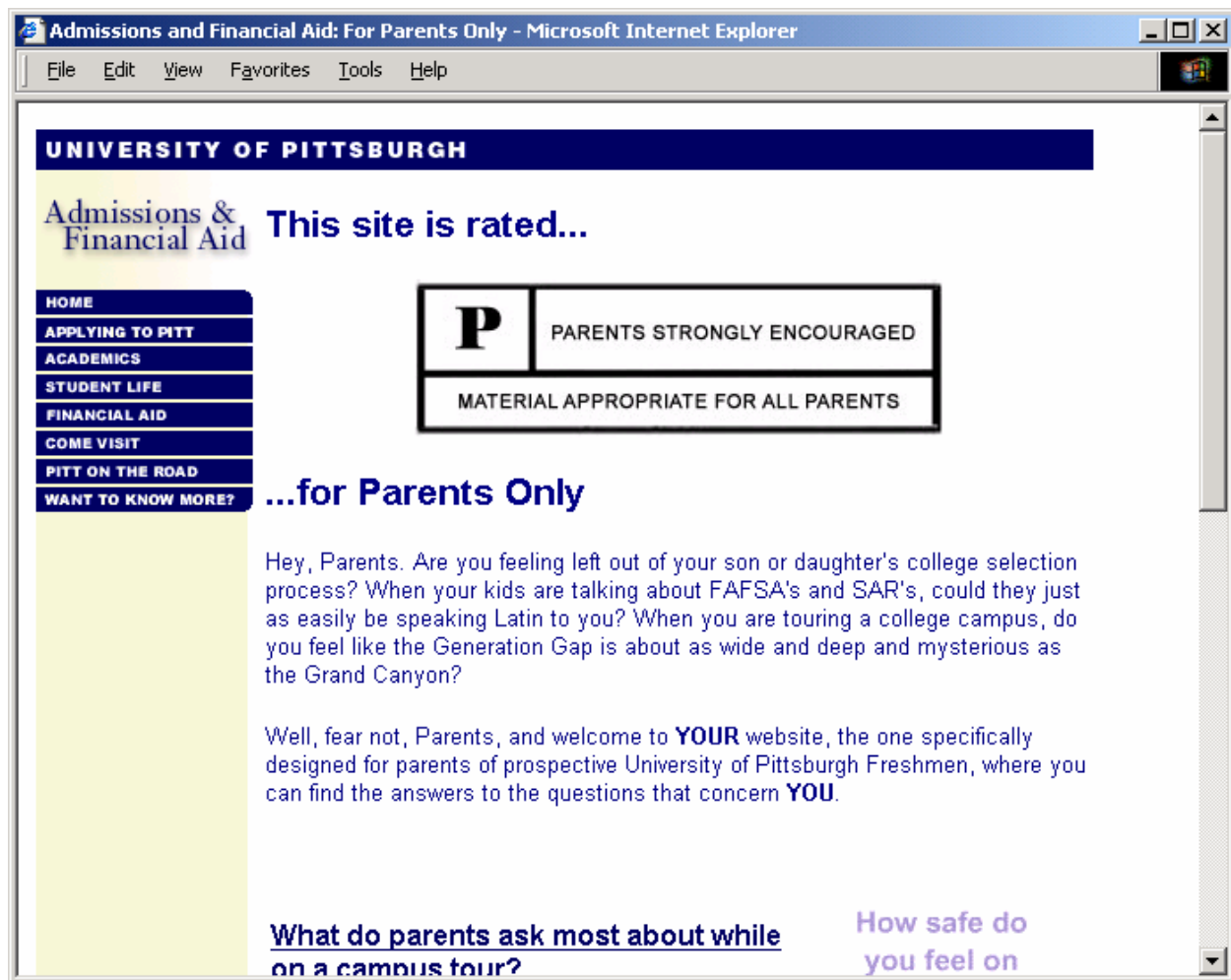


Figure D-18. Financial Aid Parents Info web page.

[http:// www.pitt.edu/~oafa/parentsonly.html](http://www.pitt.edu/~oafa/parentsonly.html)

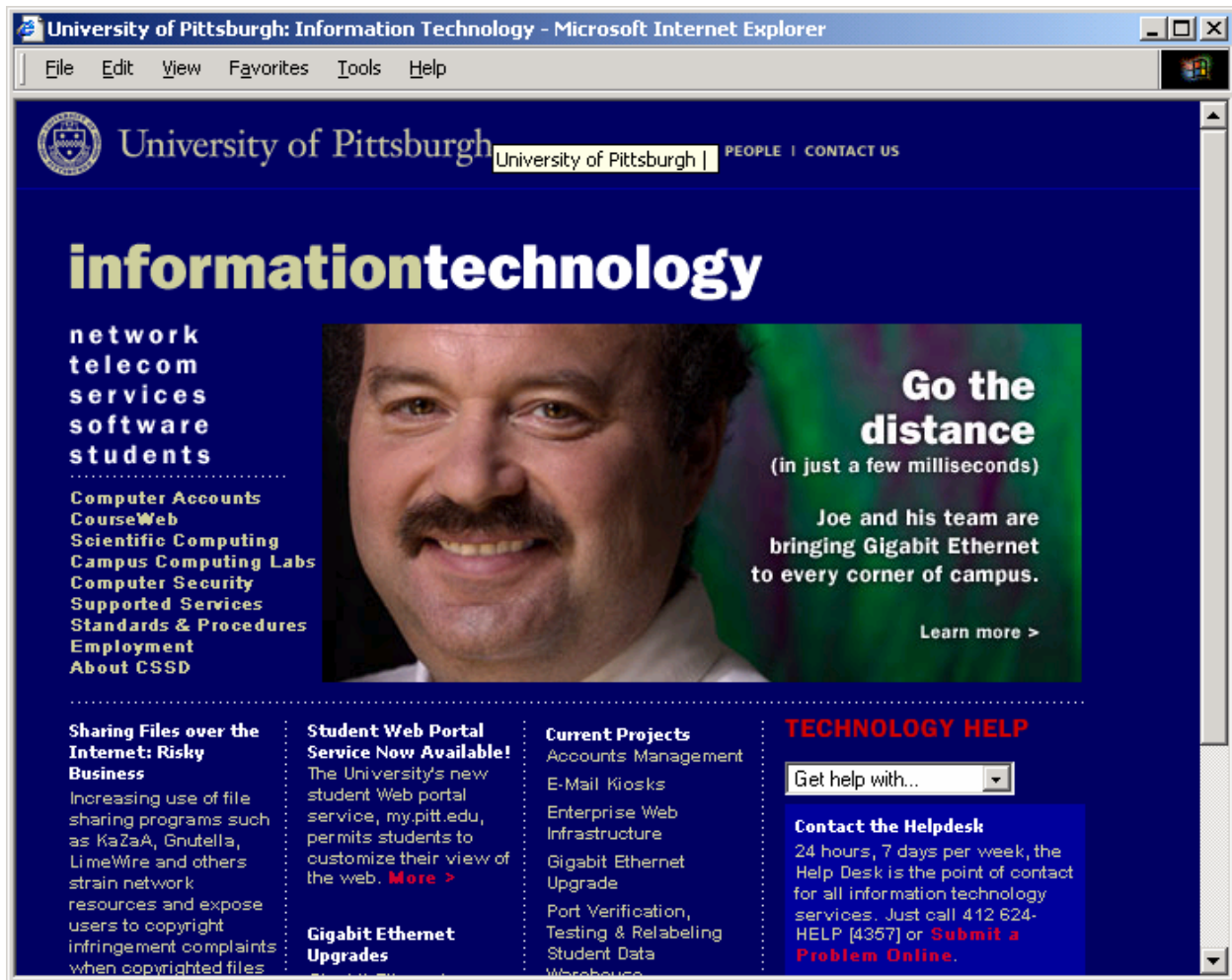


Figure D-19. Technology Home web page.

<http://technology.pitt.edu>

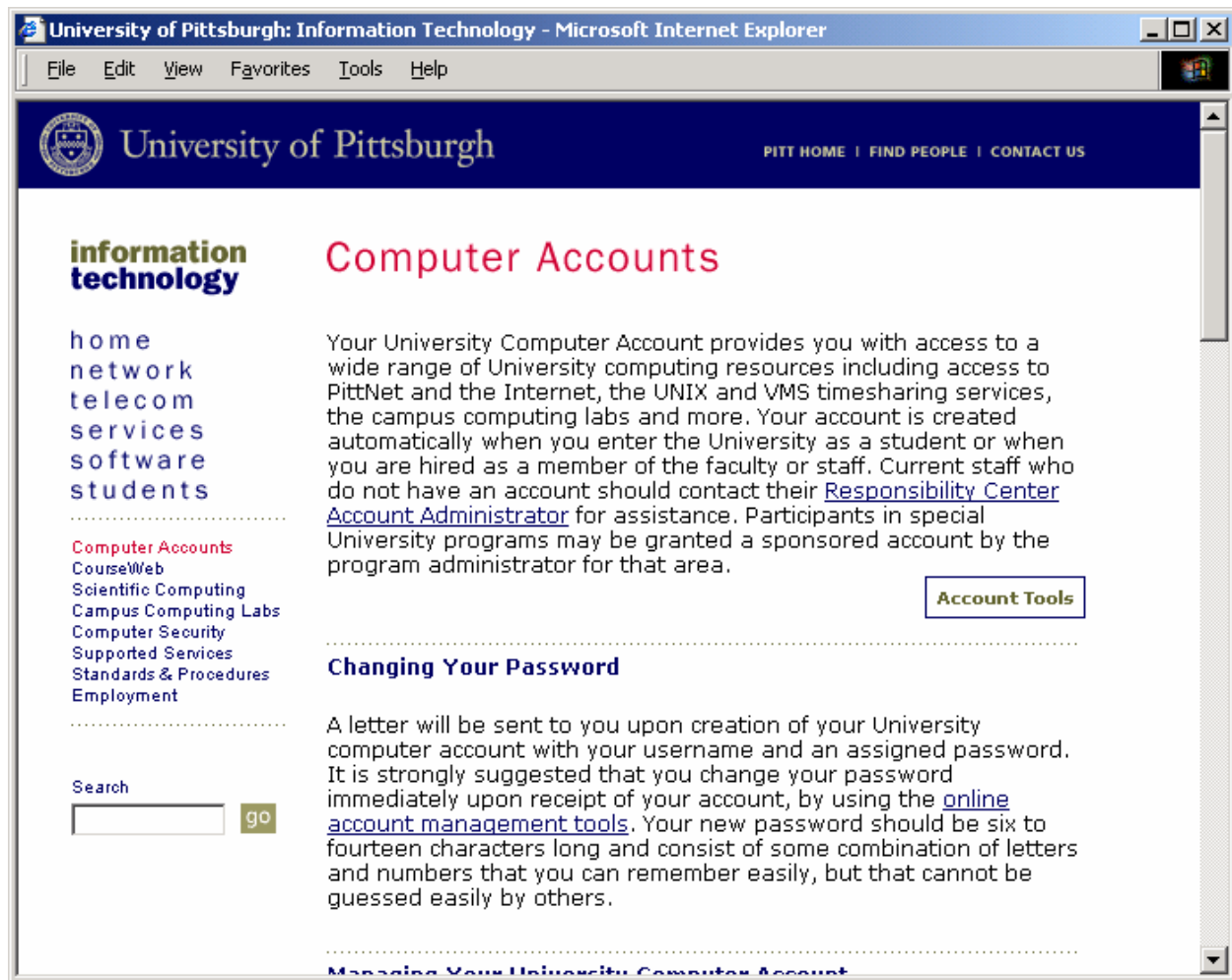


Figure D-20. Computer Accounts web page.

[http:// technology.pitt.edu/accounts/index.html](http://technology.pitt.edu/accounts/index.html)



Figure D-21. Cyber Career Counselor web page.

[http:// www.careers.pitt.edu/cybercounselor/cybercounselor.html](http://www.careers.pitt.edu/cybercounselor/cybercounselor.html)



Figure D-22. Jurist web page.

<http://jurist.law.pitt.edu>




Figure D-23. Nationality Rooms web page.

<http://www.pitt.edu/~natrooms>

APPENDIX E

Student Life at Pitt Web Page page 1



Life at Pitt
Sponsored by the Division of Student Affairs

- ✧ [Student Affairs](#)
- ✧ [Getting Involved](#)
- ✧ [Everyday Living](#)
- ✧ [Academics](#)
- ✧ [Fun Stuff](#)
- ✧ [Student Resource Guide](#)
- ✧ [Student Events Calendar](#)
- ✧ [Club Sports and Recreation](#)
- ✧ [Diversity Mission Statement](#)

Site Index

<u>Student Affairs</u>	<u>Getting Involved</u>
<ul style="list-style-type: none">• <u>Disability Resources and Services</u>• <u>First Year Experience</u>• <u>Learning Skills Center</u>• <u>Office of International Services</u>• <u>Office of Residence Life</u>• <u>Office of Student Activities</u>• <u>Placement and Career Services</u>• <u>Sexual Assault Services</u>• <u>Student Health Services</u>• <u>University Counseling Center</u>• <u>University Student Judicial System</u>• <u>Vice Provost and Dean of Students</u>	<ul style="list-style-type: none">• <u>Black Action Society</u>• <u>Club Sports and Recreation</u>• <u>College of General Studies Student Gov't</u>• <u>Graduate Professional Student Assoc. (GPSA)</u>• <u>Greeks (Fraternities and Sororities)</u>• <u>Intramurals and Recreation</u>• <u>Office of Student Activities</u>• <u>Pitt Arts</u>• <u>Pitt Program Council (PPC)</u>• <u>PittNews</u>• <u>Resident Student Association (RSA)</u>• <u>Student Events Calendar</u>• <u>Student Government Board</u>

FigureE-1. Student Life at Pitt web page.

Student Life at Pitt Web Page, continued

Everyday Living

- [Bicycle Registration](#)
- [Book Center](#)
- [Buses](#)
- [Calendar/Academic](#)
- [Computer Help Desk](#)
- [Computing Labs](#)
- [Code of Conduct--Student](#)
- [e-Store](#)
- [Food Services Housing \(Student\)](#)
- [ID Center](#)
- [Information Technology](#)
- [Intramurals and Recreation](#)
- [Maps--Pitt](#)
- [Maps--Pittsburgh](#)
- [PittTV](#)
- [Parking, Transportation & Services](#)
- [Pitt News](#)
- [Pitt Promise](#)
- [Pittsburgh Links](#)
- [Pitt Shop](#)
- [Pitt TV](#)
- [Police, University](#)
- [Ridesharing](#)
- [Student Health](#)
- [Student Parking at the University](#)
- [Student Telephone Service](#)
- [Telefact](#)
- [University Child Development Center](#)
- [University Times](#)
- [Vehicle Rentals](#)

Student Resource Guide

- [Academic Advising/Scheduling Problems](#)

- [Student Organization Resource Center \(SORC\)](#)
- [Student Volunteer Outreach](#)
- [Varsity Marching Band](#)
- [WPTS Radio Station](#)

Fun Stuff

- [Amusement Park-Kennywood](#)
- [Andy Warhol Museum](#)
- [Athletics](#)
- [Ballet](#)
- [CNN](#)
- [Carnegie Museum](#)
- [Carnegie Museum of Art](#)
- [Carnegie Science Center](#)
- [Mellon Arena](#)
- [Friday Nite Improvs](#)
- [In Pittsburgh Newsweekly](#)
- [LPGA](#)
- [Letterman Top 10](#)
- [M - Metropol](#)
- [Movies](#)
- [NBA](#)
- [NFL](#)
- [Nationality Rooms](#)
- [Opera](#)
- [PGA](#)
- [Parks-Pittsburgh](#)
- [Penguins](#)
- [Phipps Conservatory](#)
- [Pirates](#)
- [Pitt Arts](#)
- [Pitt Program Council \(PPC\)](#)
- [Pittsburgh Dance Council](#)
- [Pittsburgh Links](#)
- [Pittsburgh Post Gazette](#)
- [Pittsburgh Public Theatre](#)
- [Pittsburgh Symphony](#)
- [Pittsburgh Zoo](#)
- [Rachel Carson Homestead](#)
- [Rosebud](#)
- [Sandcastle Waterpark](#)

Student Life at Pitt Web Page, continued

- [Address/Telephone Changes](#)
 - [Adjusting to the U.S.](#)
 - [Alcohol/Drug Issues](#)
 - [Career Direction](#)
 - [Computer Assistance](#)
 - [Dental](#)
 - [Difficulty Understanding Course Material](#)
 - [Disciplinary Problems](#)
 - [Disabilities](#)
 - [Financial Problems](#)
 - [Food \(Meal Plans\)](#)
 - [Generalized Anxiety](#)
 - [Getting Involved in Campus Activities](#)
 - [Grades/Enrollment Certification](#)
 - [Home or Family Problems](#)
 - [Housing](#)
 - [Housing Roommate /Residence Hall Problems](#)
 - [Housing Off-Campus](#)
 - [Loneliness, Personal Problems](#)
 - [Medical Problems/Physical Illness or Pharmacy/Prescription/Healthy Lifestyle Questions](#)
 - [Mentoring for African American and Latino Freshmen](#)
 - [Name Changes](#)
 - [Pregnancy/Unplanned Pregnancy](#)
 - [Registration or Add/Drop](#)
 - [Relationship Problems](#)
 - [Religious Life](#)
 - [Sexual Assault](#)
 - [Social Security # Change](#)
 - [Stress/Tension](#)
 - [Poor Study Habits \(Lack of Study Skills\) Test Anxiety](#)
 - [Time Management](#)
 - [Transportation](#)
 - [Veteran's Services](#)
 - [Volunteer Opportunities](#)
 - [Senator John Heinz Pittsburgh Regional History Center](#)
 - [Steelers](#)
 - [TV](#)
 - [Weather](#)
- ### Academics
- [Book Center](#)
 - [Computing and Information Services](#)
 - [Departments and Programs](#)
 - [Libraries](#)
 - [Online Schedules and Grades](#)
 - [Registrar's Office](#)
 - [Schedule of Classes](#)
 - [Schools, College and Centers](#)

Page last revised: June 8, 2001
Send comments to: dsc+t@pitt.edu

APPENDIX F

Semantic, Visual and Structural Components of the Objective Landmark Value

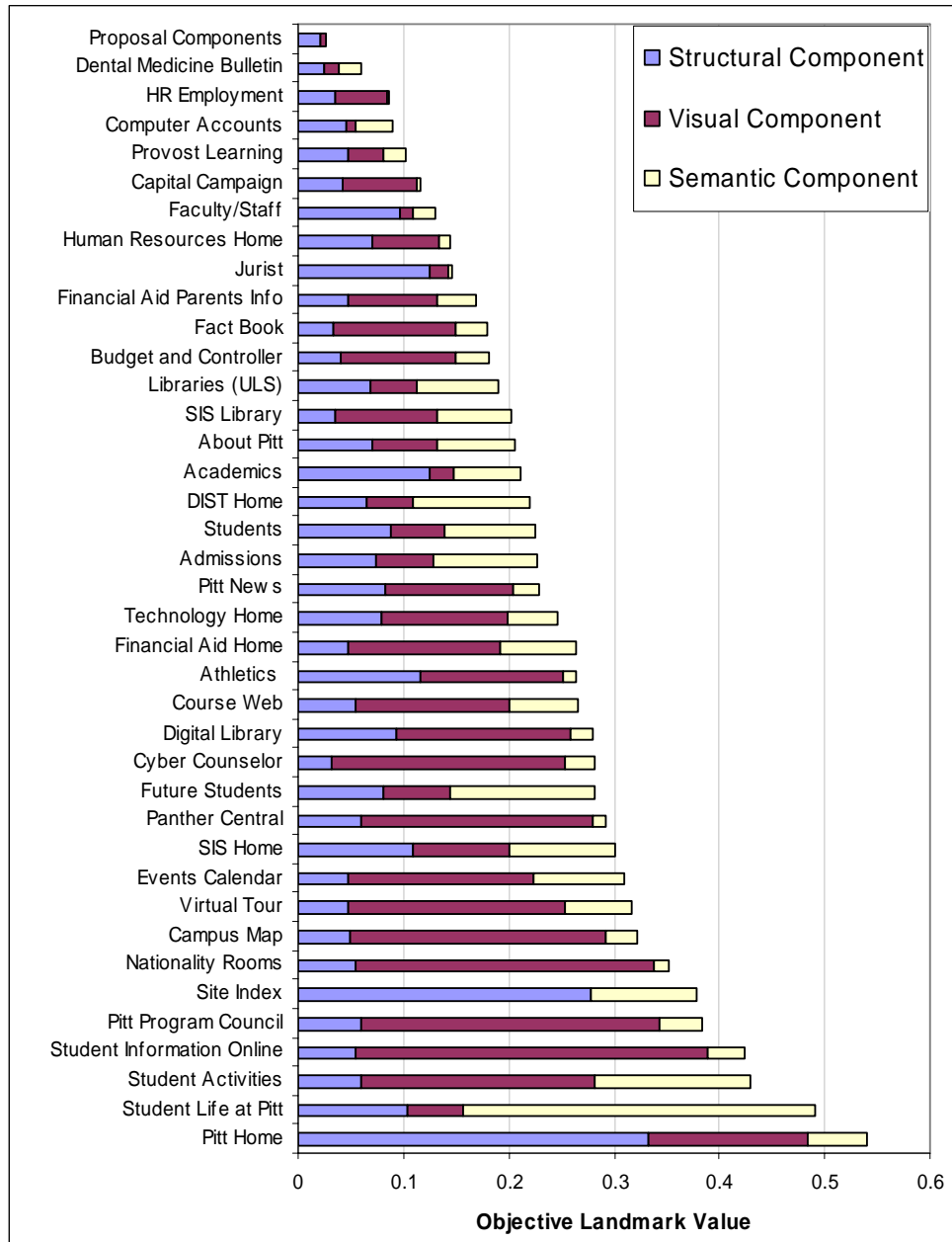


Figure F-1. Objective Landmark Value Components for All Pages.

APPENDIX G

Semantic, Visual and Structural Components of the Subjective Landmark Value

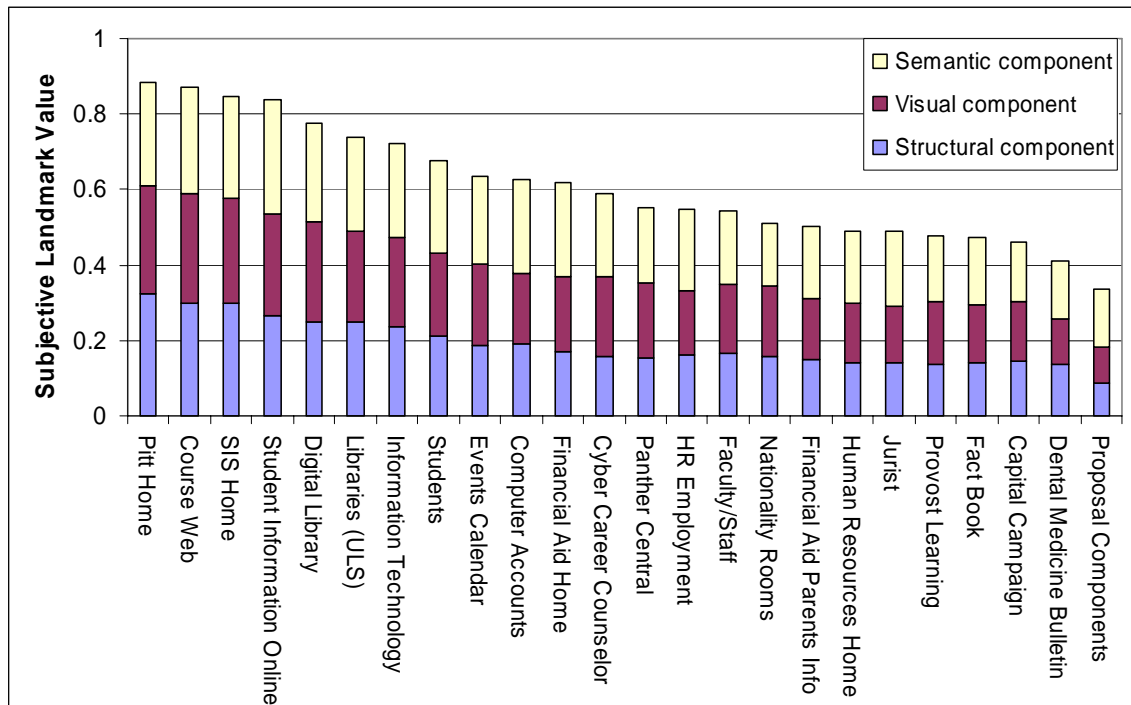


Figure G-1. Subjective Landmark Value Components.

APPENDIX H

Correlations of web page features to Subjective Landmark Value

Table H-1. Subjective Landmark Value Statistical Correlations.

	Subjective Landmark Value	# Links Out	Depth of URL	# Links Into Node	Estimated Frequency of Access	% Blank Space	% Links	% Graphics	% Plain Text	Total Page Area	% Matching Words	Total Words
Subjective Landmark Value	1	.133	-.689(**)	.413(*)	.458(*)	.327	.063	.399(*)	-.588(**)	-.489(**)	.650(**)	-.544(**)
# Outgoing links	.133	1	-.320	.267	.200	.096	.691(**)	-.453(*)	-.448(*)	.074	.276	.079
Depth of URL	-.689(**)	-.320	1	-.397(*)	-.430(*)	-.407(*)	-.423(*)	-.399(*)	.761(**)	.598(**)	-.502(**)	.588(**)
# Links Into Node	.413(*)	.267	-.397(*)	1	.974(**)	.128	.127	.069	-.257	-.144	.169	-.097
Estimated Frequency of Access	.458(*)	.200	-.430(*)	.974(**)	1	.144	.121	.091	-.296	-.215	.264	-.185
% Blank Space	.327	.096	-.407(*)	.128	.144	1	.241	.190	-.461(*)	-.150	.285	-.342
% Links	.063	.691(**)	-.423(*)	.127	.121	.241	1	-.064	-.709(**)	-.308	.232	-.347(*)
% Graphics	.399(*)	-.453(*)	-.399(*)	.069	.091	.190	-.064	1	-.316	-.535(**)	.098	-.538(**)
% Plain Text	-.588(**)	-.448(*)	.761(**)	-.257	-.296	-.461(*)	-.709(**)	-.316	1	.598(**)	-.532(**)	.728(**)
Total Page Area	-.489(**)	.074	.598(**)	-.144	-.215	-.150	-.308	-.535(**)	.598(**)	1	-.507(**)	.832(**)
% Matching Words	.650(**)	.276	-.502(**)	.169	.264	.285	.232	.098	-.532(**)	-.507(**)	1	-.485(**)
Total Words	-.544(**)	.079	.588(**)	-.097	-.185	-.342	-.347(*)	-.538(**)	.728(**)	.832(**)	-.485(**)	1

** Correlation is significant at the 0.01 level (1-tailed).

* Correlation is significant at the 0.05 level (1-tailed).

APPENDIX I

Correlation of Subjective and Objective Landmark Value

Table I-1. Effect of Weight Strategies on Objective and Subjective Landmark Value Correlation.

Weight strategy	Weights for each Component			Correlation Coefficient
	Structural	Visual	Semantic	
Equal weights	1/3	1/3	1/3	0.729**
Only one component	1	0	0	0.511**
	0	1	0	0.414*
	0	0	1	0.670**
One component heavily weighted	3/5	1/5	1/5	0.693**
	1/5	3/5	1/5	0.570
	1/5	1/5	3/5	0.817*
Two components heavily weighted	3/7	3/7	1/7	0.661**
	1/7	3/7	3/7	0.660**
	3/7	1/7	3/7	0.774**
Only two components included	1/2	1/2	0	0.617**
	0	1/2	1/2	0.582**
	1/2	0	1/2	0.692**
Regression results	0.371	0.333	0.547	0.807**

** Correlation is significant at the 0.01 level

* Correlation is significant at the 0.05 level

BIBLIOGRAPHY

- Abu-Ghazzeh, T. M. (1996). Movement and wayfinding in the King Saud University built environment: A look at freshman orientation and environmental information. *Journal of Environmental Psychology*, 16, 303-318.
- Allen, G. L. (1999). Spatial abilities, cognitive maps, and wayfinding: Bases for individual differences in spatial cognition and behavior. In R. G. Golledge (ed.), *Wayfinding Behavior: Cognitive mapping and other spatial processes* (pp. 46-80). Baltimore: Johns Hopkins University Press.
- Anooshian, L. J. (1988). Places versus procedures in spatial cognition: Alternative approaches to defining and remembering landmarks. *British Journal of Developmental Psychology*, 6, 389-390.
- Appleyard, D. (1969). Why buildings are known. *Environment and Behavior*, 1, 131-156.
- Ark, W., Dryer, D.C., Selker, T., and Zhai, S. (1998). Landmarks to aid navigation in a graphical user interface. Workshop on Personalised and Social Navigation in Information Space. Stockholm (March 16-17, 1998). (In association with IFIP working group WG 13.2). Retrieved on January 1, 2001 from the World Wide Web: <http://www.sics.se/humle/projects/persona/web/workshop/>
- Bachelder, I. A. and Waxman, A. M. (1995). A view-based neurocomputational system for relational map-making and navigation in visual environments. *Robotics & Autonomous Systems* 16, 267-289.
- Balasubramanian, V. (1994). State of the art review on hypermedia issues and applications. Unpublished manuscript. Rutgers University, Newark, New Jersey. Converted to HTML by D. Duchier. Retrieved on July 14, 2001 from the World Wide Web: http://www.isg.sfu.ca/~duchier/misc/hypertext_review/index.html
- Bhanu, B. (1994). Introduction to the special section on perception-based real-world navigation. *IEEE Transactions on Robotics and Automation* 10, 725-726.
- Begoray, J. A. (1990). An introduction to hypermedia issues, systems and application areas. *International Journal of Man-Machine Studies*, 33, 121-147.

- Benyon, D. and Höök, K. (1997). Navigation in information spaces: Supporting the individual. *HCI Interact '97*, 39-46.
- Blades, M. (1991). Wayfinding theory and research: The need for a new approach. In D. Mark and A. Frank (eds.), *Cognitive and linguistic aspects of geographic space* (pp. 137-165). The Netherlands: Kluwer Academic.
- Bush, V. (1945). As We May Think. *The Atlantic Monthly*, 176, 101-108. Retrieved on July 14, 2001 from the World Wide Web: <http://www.theatlantic.com/unbound/flashbks/computer/bushf.htm>
- Canter, D., Rivers, R. and Storrs, G. (1985). Characterizing user navigation through complex data structures. *Behavior and Information Technology*, 4, 93-102.
- Chen, J., Zhou, B., Shi, J., Zhang, H., and Wu, Q. (2001). Function-based object model towards website adaptation. *Proceedings of the 10th WWW Conference* (May 1-5, 2001, Hong Kong).
- Cheng, K. (1994). The Determination of Direction in Landmark-based spatial search in pigeons - a further test of the vector sum model. *Animal Learning and Behavior*, 22, 291-301.
- Chi, E. H., Pirolli, P. and Pitkow, J. (2000). The scent of a site: a system for analyzing and predicting information scent, usage, and usability of a web site. In *Proceedings of the CHI 2000 conference on human factors in computing systems*, pp. 161-168.
- Chi, E.H., A. Rosien, G. Suppattanasiri, A. Williams, C. Royer, C. Chow, E. Robles, B. Dalal, J.Chen and S. Cousins. (2003). The Bloodhound Project: Automating Discovery of Web Usability Issues using the InfoScentTM Simulator. In *Proceedings of the Conference on Human Factors in Computing Systems*, CHI 2003, Ft. Lauderdale, Florida, USA.
- Collet, T. S. (1996). Insect navigation en route to the goal: Multiple strategies for the use of landmarks. *Journal of Experimental Biology*, 199: 22-235.
- Collet, T. S. and Rees, J. A. (1997). View-based navigation in Hymenoptera: Multiple strategies of landmark guidance in the approach to a feeder. *Journal of Comparative Physiology: A Sensory Neural and Behavioral Physiology*, 181, 47-58.
- Conklin, E. J. (1987). Hypertext: An introduction and survey. *IEEE Computer*, 20, 17-41.

- Couclelis, H., Golledge, R. G., Gale, N.D., and Tobler, W. R. (1987). Exploring the anchor-point hypothesis of spatial cognition. *Journal of Environmental Psychology*, 7, 99-122.
- Darken, R. P. (1997). Navigating virtual worlds: Wayfinding and locomotion in real and not-so-real environments. [Presentation: Carnegie Mellon University HCI Seminar series, April, 16, 1997].
- Darken, R.P. and Sibert, J.L. (1996). Wayfinding strategies and behaviors in large virtual worlds. *Proceedings of ACM CHI '96*, pp. 142-149.
- Devlin, A. S. and Bernstein, J. (1995). Interactive wayfinding: Use of cues by men and women. *Journal of Environmental Psychology*, 15, 23-38.
- Dodge, M. and Kitchin, R. (2001). *Mapping cyberspace*. New York: Routledge.
- Elias, B. (2003). Extracting landmarks with data mining methods. In W. Kuhn, M.F. Worboys and S. Timpf (eds.), *Lecture Notes in Computer Science #2825, Spatial Information Theory: Foundations of Geographic Information Science (COSIT 2003)*. Berlin: Springer-Verlag.
- Ellis, S. R. (1993). *Pictorial communication in virtual and real environments (2nd Ed.)*. London: Taylor & Francis Ltd.
- Foltz, M. A. (1998). Designing navigable information spaces. Unpublished master's thesis. MIT Department of Electrical Engineering and Computer Science. Retrieved on July 14, 2001 from the World Wide Web: <http://www.infoarch.ai.mit.edu/publications/mfoltz-thesis/>
- Foss, C. L. (1989). Tools for reading and browsing hypertext. *Information Processing Management*, 25, 407-418.
- Furnas, G. W. (1986). Generalized fisheye views. In *Proceedings of the ACM SIG-CHI '86 Conference on Human Factors in Computing Systems*, pp. 16-23.
- Ghiselli-Crippa, T. B. and Munro, P. W. (1994). Emergence of global structure from local associations. In J. D. Cowan, G. Tesauro, & J. Alspector (eds.), *Advances in Neural Information Processing Systems*, 6. San Francisco, CA: Morgan Kaufmann.
- Glenn, B. T. and Chignell, M. H. (1992). Hypermedia: Design for browsing. In H. R. Hartson & D. Hix (eds.), *Advances in Human-Computer Interaction, Vol. 3*. Norwood, NJ: Ablex Publishing Corp.

- Golledge, R. G. (1999). *Wayfinding behavior: Cognitive mapping and other spatial processes*. Baltimore: The Johns Hopkins University Press.
- Gothard, K.M., Skaggs, W. E., Moore, K.M. and McNaughton, B.L. (1996). Binding of hippocampal CA1 neural activity to multiple reference frames in a landmark-based navigation task. *Journal of Neuroscience*, 16 (2), 823-835.
- Heft, H. (1979). The role of environmental features in route-learning: Two exploratory studies of wayfinding. *Environmental Psychology and Nonverbal Behaviour*, 3, 172-185.
- Heth, C. D., Cornell, E. H. and Alberts, D. M. (1997). Differential use of landmarks by 8-and 12- year-old children during route reversal navigation. *Journal of Environmental Psychology*, 17, 199-213.
- Hirtle, S. C. and Hudson, J. (1991). Acquisition of spatial knowledge for routes, *Journal of Experimental Psychology*, 11, 335-345.
- Hirtle, S. C. and Jonides, J. (1985). Evidence of hierarchies in cognitive maps. *Memory and Cognition*, 13, 208-217.
- Hirtle, S. C. and Mascolo, M. F. (1986). The effect of semantic clustering on the memory of spatial locations. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 12, 181-189.
- Hirtle, S. C., Sorrows, M. E., and Cai, G. (1998). Clusters on the World Wide Web: Creating neighborhoods of make-believe. *Hypertext 98*, 289-290. New York: ACM.
- Hochmair, H. and Raubal, M. (2002). Topologic and metric decision criteria for wayfinding in the real world and the WWW. In Symposium on Geospatial Theory, Processing and Applications, Ottawa, 2002.
- Ingram, R. and Benford, S. (1995). Legibility enhancement for information visualisation. *Proceedings of the IEEE Conference on Visualization (IEEE VIZ '95)*. Atlanta, GA (Nov., 1995).
- Khuller, S., Raghavachari, B., and Rosenfeld, A. (1996). Landmarks in graphs. *Discrete Applied Mathematics*, 70, 217-229.
- Kim, H. and Hirtle, S. C. (1995). Spatial metaphors and disorientation in hypertext browsing. *Behaviour & Information Technology*, 14, 239-250.

- Korfhage, R. R. (1995). VIBE: Visual Information Browsing Environment. In *Proceedings of the 18th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, Seattle.
- Kuipers, B. J. (1983). The cognitive map: Could it have been any other way? In H. Pick and L. Acredolo (eds.), *Spatial orientation: Theory, research, and applications* (pp. 345-359). New York: Plenum Press.
- Landow, G. P. (1997). *Hypertext 2.0*. Baltimore, MD: Johns Hopkins University Press.
- Lin, C.-C. and Tummala, R. L. (1997). Mobile robot navigation using artificial landmarks. *Journal of Robotic Systems*, 14 (2), 93-106. New York: John Wiley & Sons.
- Lynch, K. (1960). *The Image of the City*. Cambridge, MA: MIT Press.
- McAleese, R. (Ed.). (1989). *Hypertext: Theory into practice*. Norwood, NJ: Ablex Publ. Corp.
- McLaren, I. P. L. (1995). A better beta model of navigation. *British Journal of Mathematical and Statistical Psychology*, 48, 51-55.
- Modjeska, D. and Waterworth, J. (2000). Effects of desktop 3D-world design on user navigation and search performance. *Proceedings of IV 2000, International Conference on Information Visualization*, London, England (July, 2000), IEEE Press.
- Mukherjea, S. and Foley, J. D. (1995). Visualizing the World-Wide Web with the Navigational View Builder. GVU report, GVU-TR 95-09. Retrieved on July 14, 2001 from the World Wide Web: <ftp://ftp.gvu.gatech.edu/pub/gvu/tech-reports/95-09.ps.Z>
- Mukherjea, S. and Hara, Y. (1997). Focus+context views of world-wide web nodes. *Hypertext '97: The Eighth ACM Conference on Hypertext* (Southampton, UK). New York, NY: ACM Press.
- Neerincx, M. A., Lindenberg, J. and Pemberton, S. (2001). Support concepts for web navigation: a cognitive engineering approach. *ACM WWW10* (May, 2001, Hong Kong).
- Nelson, T. (1990, September). On the Xanadu Project. *Byte*, pp. 298-299.
- Neveitt, W. T. (2000). Spatial knowledge navigation for the World Wide Web. Unpublished PhD. dissertation. MIT Department of Electrical Engineering and Computer Science.

- Nielsen, J. (1990). *Hypertext and Hypermedia*. New York: Academic Press, Inc.
- Nielsen, J. (1995). *Multimedia and hypertext: The Internet and beyond*. New York: Academic Press, Inc.
- Passini, R. (1984). *Wayfinding in Architecture*. New York: Van Nostrand Reinhold.
- Passini, R. (1996). Wayfinding design: Logic, application and some thoughts on universality. *Design Studies*, 17, 319-331.
- PastergueRuiz, I., Beugnon, G. and Lachaud, J.P. (1995). Can the ant *Cataglyphis cursor* (Hymenoptera: Formicidae) encode global landmark-landmark relationships in addition to isolated landmark-goal relationships? *Journal of Insect Behavior*, 8, 115-132.
- Pearson, R. and van Schaik, P. (2003). The effect of spatial layout of and link colour in web pages on performance in a visual search task and an interactive search task. *Int. Journal Human-Computer Studies* 59, 327-353.
- Pilgrim, C.J. and Leung, .K. (2001). Utilising landmarks for web site navigation. In W. A. Lawrence-Fowler and J. Hasebrook (eds.), *Proceedings of WebNet 2001*, 1004-1008.
- Prescott, T.J. (1997). Spatial representation for navigation in animals. *Adaptive Behavior*, 4, 85-123.
- Presson, C. C. and Montello, D. R. (1988). Points of reference in spatial cognition: Stalking the elusive landmark. *British Journal of Developmental Psychology*, 6, 378-381.
- Raubal, M. and Winter, S. (2002). Enriching wayfinding instructions with local landmarks. In M.J. Egenhofer and D. M. Mark (eds.), *Lecture Notes in Computer Science #2478, Geographic Information Science (GIScience 2002)*. Berlin: Springer-Verlag.
- Rosch, E. (1975). Cognitive reference points. *Cognitive Psychology*, 7, 532-547. Referenced in Sadalla, Burroughs & Staplin (1980).
- Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., and Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, 8, 382-439.
- Sadalla, E. K., Burroughs, W. J. and Staplin, L. J. (1980). Reference points in spatial cognition. *Journal of Experimental Psychology: Human Learning and Memory*, 5, 516-528.
- Salton, G. (1989). *Automatic text processing*. Reading, MA: Addison-Wesley.

- Shannon, C. E. and Weaver, W. (1962). *The mathematical theory of communication*. Urbana: University of Illinois Press.
- Siegel, A. W., and White, S. H. (1975). The development of spatial representations of large-scale environments. In H. W. Reese (ed.), *Advances in child development and behavior* (Vol. 10, pp. 9-55). New York: Academic Press.
- Shum, S. (1990). Real and virtual spaces: Mapping from spatial cognition to hypertext. *Hypermedia*, 2, 133-158.
- Smith, P. A. and Wilson, J. R. (1993). Navigation in hypertext through virtual environments. *Applied Ergonomics*, 24, 271-278.
- Sorrows, M. E. and Hirtle, S. C. (1999). The nature of landmarks for real and electronic spaces. In C. Freksa and D. M. Mark (eds.), *Lecture Notes in Computer Science #1661, Spatial Information Theory: Cognitive and Computational Foundations of Geographic Information Science (COSIT '99)*. Berlin: Springer.
- Spetch, M. L. (1995). Overshadowing in landmark learning: Touch-screen studies with pigeons and humans. *Journal of Experimental Psychology: Animal Behavior Processes*, 21, 166-181.
- Steck, S. D. and Mallot, H. A. (2000). The role of global and local landmarks in virtual environment navigation. *Presence*, 9, 69-83.
- Stevens, A. and Coupe, P. (1978). Distortions in judged spatial relations. *Cognitive Psychology*, 10, 422-437.
- Steuer, J. (1992). Defining virtual reality: Dimensions determining telepresence. *Journal of Communication*, 42, 73-93.
- Touretzky, D. and Redish, D. (1995). Landmark arrays and the hippocampal cognitive map. In L.F. Niklasson, & M.B. Boden, (eds.), *Current trends in connectionism: Proceedings of the Swedish Conference on Connectionism 1995* (pp. 1-13). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Tullis, T. S. (1998). A method for evaluating web page design concepts. In *Proceedings of the ACM CHI '98* (April, 1998), 323-324.

- Utting, K. and Yankelovich, N. (1989). Context and orientation in hypermedia networks. *ACM Transactions on Information Systems*, 7, 58-84.
- Virtual Brain Project. (1999). Virtual Centre for Health Informatics, Aalborg, Denmark. Web site: <http://www.vision.auc.dk/~jl/vbp/index.htm>
- Winter, S. (2003). Route adaptive selection of salient features. In W. Kuhn, M.F. Worboys and S. Timpf (eds.), *Lecture Notes in Computer Science #2825, Spatial Information Theory: Foundations of Geographic Information Science* (COSIT 2003). Berlin: Springer-Verlag.
- Worboys, M.F. (2001). Nearness relations in environmental space. *Int. J. Geographical Information Science*, 15, 633-651.
- Yamauchi, B. and Langley, P. (1997). Place recognition in dynamic environments. *Journal of Robotic Systems*, 14, 107-120.
- Zipser, D. (1986). Biologically plausible models of place recognition and goal location. In J.L. McClelland and D.E. Rumelhart, (eds.), *Parallel Distributed Processing: Explorations in the microstructure of cognition, vol. 2: Psychological and biological models*. Cambridge, MA: MIT Press.

INDEX

A

access of web pages. *See* web page access
algorithms. *See* computational approaches
anchor points. *See* entries for landmarks
artificial intelligence, 14–16

B

bookmarks, web, 22, 40
 usage, 47–50, 74–80, 98–99
buildings as landmarks, 24–26, 117

C

Chi, E. H., 2, 112–113
cognitive landmarks. *See* semantic landmark
 characteristics
cognitive maps, 1, 2, 9, 26.
 and landmarks, 28–29, 110
computational approaches, 32–33, 43, 47, 117–118
content, text, in web pages. *See* semantic landmark
 characteristics

D

data. *see* statistical data
design, urban, 8–9
disorientation, hypertext, 20–21

E

environments, types of, 7

F

favorite sites. *See* bookmarks, web

G

Golledge, R. G., 1, 26, 28
graph theory, 13–14
graphical overview maps, 22, 34, 110–111
graphics, web. *See* visual landmark characteristics
guided tours, web, 22, 112. *See also* navigational aids

H

hierarchies, 22
Hirtle, S. C., 2–4, 28, 30, 34–37, 107, 118
home pages, 30–31, 35–36, 112
 as landmarks, 30–31
hyperlinks. *See* links
hypertext, 18
 orientation, 20–21
 node connectivity, 32–33, 42–44, 113
hypertext navigation. *See* navigation, web

I

information architecture, 2
 research implications, 111–112

K

knowledge, navigational, 6, 26

L

landmarks, physical, 6, 37
 characteristics, 24–26, 34
 as navigational aids, 28–30, 110
 types of, 35–36
landmarks, semantic web. *See* semantic landmark
 characteristics
landmarks, structural web. *See* structural landmark
 characteristics
landmarks, visual web. *See* visual landmark
 characteristics
landmarks, web, 30–33, 32, 109–111. *See also* web
 page elements
 types of, 30
 objective value
 hypotheses, 40–41
 computational approaches, 42–44
 characteristics, 44–46
 experiments, 48–50, 55–59, 91
 analysis, 93–108, 171, 174
 conclusions, 109–110

landmarks, web (continued)
 subjective value
 hypotheses, 40-41
 evaluation, 46-53
 experiments, 80-91
 analysis, 101-109, 172, 173, 174
 conclusions, 111
 statistical analysis, 56-58
 weighted analysis, 105-107, 118-119
 learning, spatial knowledge, 26-27
 links, 18-19, 42-44. *See also* structural landmark characteristics
 statistical data, 104
 Lynch, K., 8-9, 23-25, 29, 31, 37, 44

M

map-making, relational, 14-16
 map recall. *See* recall in navigation
 mental maps. *See* cognitive maps
 methodology, 39-40, 48, 52, 112-116
 algorithms, 43, 47, 118-119
 standard deviation, 114-115
 Mukherjea, S., 32-33, 42

N

navigation paths, web, 48-49, 60-73, 75-80, 94.
 See also starting points
 navigation, physical, 5, 7, 27-28
 urban, 8-9, 29
 neurological aspects, 10-13
 animal, 11
 robot, 14-16
 virtual reality, 17
 learning, 26-27
 recall, 39, 110
 and landmarks, 28-30, 119-120
 navigation, web, 2-3, 23-30, 110-111.
 See also web page access
 aids, 22-23, 112
 evaluation, 48-50, 75-80
 paths, 48-49, 60-73, 75-80
 navigational aids
 landmarks as, 28-30
 web, 22-23, 112
 navigational features, 5
 networked space navigation. *See* space, networked
 neural networks, 12-13
 neuroscience, 10-13
 nodes, 9, 13-14, 23, 30-33
 connectivity, 32-33, 42, 113

O

objective landmark value. *See* landmarks, web, objective value
 open-terrain navigation. *See* space, open-terrain

organizing concepts for landmarks, 28
 orientation, hypertext, 20-21

P

Passini, R., 9, 29
 page, web, *See entries for* web page
 paths, web navigation, 48-49, 60-73, 75-80.
 See also starting points
 recall of, 94. *See also entries for* recall in navigation
 physical landmarks. *See* landmarks, physical
 points of reference. *See entries for* landmarks

Q

questionnaires, 46-47, 48-53
 participants, 59-60, 81
 questions, 51, 77-79, 81-88

R

Raubal, M., 37, 118-119
 recall in navigation, 39, 110
 of web pages. *See* web page recall
 reference points. *See entries for* landmarks
 relational map-making, 14-16
 research implications, 111-112
 research method, 39-40, 48, 52
 robotic navigation, 14-16

S

Sadalla, E. K., 3, 25, 27, 39
 search engines, 22, 119-120
 semantic landmark characteristics, 36
 evaluation, 45-47, 86-88
 correlation with subjective landmarks, 103, 105-107
 recommendations, 113-114
 Sorrows, M. E. and Hirtle, S. C., 2-4, 34-37, 42, 45, 107, 118
 space, networked, 7, 13
 space, open-terrain, 7, 16
 spatial cognition, 2, 20, 27. *See also* cognitive maps
 spatial representation of landmarks, 24-26
 starting points, 48-49, 57, 68. *See also* paths, web navigation
 statistical data, 75-77, 100
 statistical data
 navigation paths, 60-63, 65, 68, 71, 72, 85
 objective landmark values, 56-58, 84, 102-107
 subjective landmark values, 89, 90, 102-107
 URL knowledge, 74, 86
 structural landmark characteristics, 36. *See also* links
 evaluation, 42, 44, 46-47
 correlation with subjective landmarks, 103, 105-107
 recommendations, 113
 URL length, 34, 42
 subjective landmark values. *See* landmarks, web,

subjective value
surveys of users. *See* questionnaires

T

text content in web pages. *See* semantic landmark
characteristics
tours, web guided, 22

U

urban design, 8–9
URL knowledge, 49, 74–76, 85–86, 98–99
URL length, 34, 42
usability, 113
user interface navigational features, 5

V

virtual reality, 17
visual cues, 2
visual landmark characteristics, 35
evaluation, 44–47, 83–84
correlation with subjective landmarks, 104, 105–107
recommendations, 114–117

W

wayfinding. *See* navigation, physical
web graphics. *See* visual landmark characteristics
web navigation. *See* navigation, web
web page access, 80. *See also* navigation, web
conclusions, 109
hypothesis, 40
measurability, 42–44
statistical data, 84–85, 98–99
web page elements, 20, 42–46, 105, 112, 114–116.
See also landmarks, web
web page landmarks. *See* landmarks, web
web page recall, 80
conclusions, 109, 112
evaluation, 93–97
hypothesis, 40, 53, 95
statistical data, 82–83, 96–97
web site design, 2, 44–46, 103, 120
research implications, 111–112
Winter, S., 118–119
World Wide Web, 18, 20, 110–111