INTERACTIVE VISUALIZATION TOOLS FOR SPATIAL DATA & METADATA

by

ALISSA N. ANTLE

B.A., The University of Waterloo, 1990
B.A.Sc., The University of Waterloo, 1990

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES
(Department of Geography)

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

July, 2000

ABSTRACT

In recent years, the focus of cartographic research has shifted from the cartographic communication paradigm to the scientific visualization paradigm. With this, there has been a resurgence of cognitive research that is invaluable in guiding the design and evaluation of effective cartographic visualization tools. The design of new tools that allow effective visual exploration of spatial data and data quality information in a resource management setting is critical if decision-makers and policy setters are to make accurate and confident decisions that will have a positive long-term impact on the environment.

The research presented in this dissertation integrates the results of previous research in spatial cognition, visualization of spatial information and on-line map use in order to explore the design, development and experimental testing of four interactive visualization tools that can be used to simultaneously explore spatial data and data quality. Two are traditional online tools (side-by-side and sequenced maps) and two are newly developed tools (an interactive "merger" bivariate map and a hybrid of the merger map and the hypermap).

The key research question is: Are interactive visualization tools, such as interactive bivariate maps and hypermaps, more effective for communicating spatial information than less interactive tools such as sequenced maps? A methodology was developed in which subjects used the visualization tools to explore a forest species composition and associated data quality map in order to perform a range of map-use tasks. Tasks focused on an imaginary land-use conflict for a small region of mixed boreal forest in Northern Alberta. Subject responses in terms of performance (accuracy and confidence) and preference are recorded and analyzed. Results show that theory-based, well-designed interactive tools facilitate improved performance across all tasks, but there is an optimal matching between specific tasks and tools. The results are generalized into practical guidelines for software developers. The use of confidence as a measure of map-use effectiveness is verified. In this experimental setting, individual differences (in terms of preference, ability, gender etc.) did not significantly affect performance.

TABLE OF CONTENTS

ABSTRACT ii

TABLES OF CONTENTS iii

LIST OF TABLES v

LIST OF FIGURES vi

ACKNOWLEDGEMENTS vii

CHAPTER 1 – INTRODUCTION
- 1.1 Overview 1
- 1.2 Special Notes 5
- 1.3 Thesis Guide 6

CHAPTER 2 – THEORETICAL BACKGROUND
- 2.1 Overview 8
- 2.2 What is Cartographic Visualization? 8
- 2.3 Recent Studies in Cartographic Visualization 11
- 2.4 Recent Trends: Cognitive Cartography 14
- 2.5 An Attack and Defense of Cognitive Cartography 15
- 2.6 Research Design 16
- 2.7 Research Questions 21

CHAPTER 3 – EXPERIMENTAL DESIGN – FRAMEWORK
- 3.1 Overview 23
- 3.2 Research Hypotheses 24
- 3.3 The Maps 27
- 3.4 The Visualization Tools 29
- 3.5 The Task Set 42
- 3.6 Measuring Map Use Effectiveness 46

CHAPTER 4 – EXPERIMENTAL DESIGN – METHODOLOGY
- 4.1 Overview 49
- 4.2 Subjects 49
- 4.3 Software and Equipment 50
- 4.4 The Test Environment 51
- 4.5 Assumptions 52
- 4.6 Pilot Study One 55
- 4.7 Pilot Study Two 57
- 4.8 The Main Experiment 58
- 4.9 Statistical Analysis 60

CHAPTER 5 – RESULTS

5.1 Overview 64
5.2 Hypothesis One: Accuracy 64
5.3 Hypothesis Two: Confidence 72
5.4 Hypothesis Three: Preference 77
5.5 Demographic Variables 79
5.6 Timeouts 80

CHAPTER 6 – DISCUSSION

6.1 Overview 83
6.2 Research Hypotheses: Overview 85
6.3 Discussion: Accuracy 87
6.4 Discussion: Confidence 92
6.5 Discussion: Preference 94
6.6 Discussion: Demographics 96
6.7 Summary 97

CHAPTER 7 – CONCLUSION

7.1 Overview 99
7.2 Limitations 102
7.3 Future Research 104
7.4 Last Words 104

BIBLIOGRAPHY 106

APPENDIX A – TASKS 115

APPENDIX B – TRAINING SCRIPT 120

APPENDIX C – SAMPLE TOOL EVALUATION WORKSHEET 124

APPENDIX D – SAMPLE LAB HANDOUTS 126

LIST OF TABLES

4.1 Pilot one: median task times. 56

5.1 Mean accuracy (%) by task complexity. 66

5.2 Mean accuracy (%) by task complexity and tool type. 66

5.3 ANOVA results: accuracy differences across task complexity types for each tool. 69

5.4 Tukey's HSD for merger tool. 69

5.5 Tukey's HSD for hypermap-merger tool. 70

5.6 One-way ANOVA results: accuracy differences across tools for each task complexity type. 71

5.7 Tukey's HSD for specific tasks. 71

5.8 Correlation between accuracy and confidence across tool types. 74

5.9 Correlation between accuracy and confidence across task types. 74

5.10 Mean confidence (5 – high) by task complexity. 75

5.11 Tukey's HSD for all tools. 75

5.12 Mean confidence (5 – high) by tool. 76

5.13 Mean accuracy and confidence for preferred and not-preferred tools. 78

5.14 Mean accuracy and confidence per gender. 79

5.15 Mean accuracy per native language. 80

5.16 Frequency of timeouts per task type. 81

5.17 Frequency of timeouts on specific tasks per tool group. 81

5.18 Frequency of timeouts on general tasks per tool group. 82

5.19 Frequency of timeouts on integrated tasks per tool group. 82

LIST OF FIGURES

2.1	The realm of cartographic visualization	11
3.1	Hypothesis one: accuracy, task complexity and tool.	25
3.2	Hypothesis two: accuracy and confidence.	26
3.3	Hypothesis two: confidence, task and tool.	26
3.4	Base map: species composition.	28
3.5	Certainty map.	29
3.6	Side-by-side map display.	31
3.7	Toggle display tool.	32
3.8	Desaturation of green.	34
3.9	Sample of desaturated map.	34
3.10	Full colour: thematic map.	36
3.11	Full desaturation: certainty map.	37
3.12	Colour desaturation: thematic and certainty map.	38
3.13	Merger tool.	39
3.14	Hypermap-merger tool: query.	41
3.15	Hypermap-merger tool.	42
4.1	Subject profile information screen.	53
4.2	Basic screen layout.	54
5.1	Mean accuracy by task complexity and tool.	67

ACKNOWLEDGEMENTS

My doctoral dissertation research was funded in part through the Natural Sciences and Engineering Research Council of Canada (NSERC), the B.C. Science Council GREAT program, and the University of British Columbia Graduate Fellowships. I wish to thank my supervisor, Brian, for his endless support and unwavering confidence that I would actually, some day, finish this document.

Special thanks to all the many people who helped and supported me along the way including: Peter Marshall, Peter Keller, Jack Snoeyink, Tom Rutledge, Trevor Davis, Darren Ham, Catherine Griffith, Vincent Kujala, Jim Mintha, Steve Cummings, Mike Joy, Rose Klinkenberg and David Vogt. I wish to thank all my friends and family, especially Andra Smith, who never let me forget that when I wasn't working full time, training, buying my first home, or raising a puppy, that I still had Ph.D. to finish. And to Maxwell, now no longer a puppy, who is still as at home in the corridors of Geography, as romping through the University Endowment Lands. Maxwell has shown me the true meaning of patience. And to Jen who will be proud of me simply for persevering.

CHAPTER 1 – INTRODUCTION

1.1 Overview

The map is one of the simplest data visualization tools there is. The development of complex digital environments has opened the door for new methods of data visualization and for the adaptation of traditional data display methods. Exploring, analyzing and displaying spatial data and metadata are central activities in a geographical information systems (GIS) environment. Two innovations from the field of computer science have been recently exploited in experimental GIS. Techniques from visualization in scientific computing and interactive multimedia provide dynamic interactive tools for the visualization of spatial data and metadata.

Since the 1987 Commission on Visualization in Scientific Computing, new visualization tools have emerged rapidly and are being incorporated into fields as diverse as cosmology and nuclear biology. Although many scientific breakthroughs have been achieved by visualizing solutions to problems, either mentally or using graphical displays, traditional science has favored mathematical or statistical modeling over visual approaches. Recently, however, the ability to store, process and display large volumes of data on modern graphics workstations is making scientific visualization an effective strategy for analysis. The concurrent emergence of Jacques Bertin's theories of graphics (Bertin 1981, 1983) has lead to the application of computerized visualization techniques to the field of cartography. The resulting set of techniques is rapidly evolving to form the field known as cartographic visualization. In Hall's popular book *Mapping the Next Millennium: the Discovery of New Geographies* (1992), the author stresses the importance of using maps in fundamental science to establish where things are. The author does not restrict this map use to geographic scales, but rather advocates the use of maps to examine data at all scales (from the sub-atomic to the galactic). While Hall does not use the terms cartographic visualization, the book is perhaps the best manifesto for both scientific and cartographic visualization that has yet been written.

The digital online map is a visualization tool found in most disciplines where spatial location is relevant. In GIS, the online map communicates geographical information and facilitates map-use tasks. Numerous static cartographic methods exist for visualizing multivariate geographic information. Bivariate choropleth maps, side-by-side maps and cartograms are used routinely (e.g., Tufte 1990, Dorling 1995). Interactive methods developed by computer scientists and geographers working in tandem are quickly replacing these static methods. Sequenced maps and animated maps are two of the interactive tools created from this collaboration (e.g., Slocum et al. 1990, Fisher 1994) although neither of these tools appears in standard GIS packages.

The phenomenal growth of the Internet has made digital spatial data available to millions of users through the Web environment. In order to make dissemination of this data useful, both GIS and Web-based tools must exist which facilitate the visualization (including interactive exploration, interpretation and analysis) of online maps (Antle and Klinkenberg 1999). One visualization tool that has evolved for visual map analysis on the Web is the online interactive map. Using techniques from interactive multimedia, Web developers have created hypermaps that allow users to point and click on map areas in order to display auxiliary information. More sophisticated Web tools include: pan, zoom and query. For example, interactive maps provide up-to-date real estate information for communities across the United States and Canada (see http://www. mapquest.com). An interactive map of the University of British Columbia campus provides information about building names, uses and addresses in response to user queries (see http://www.math.ubc.ca/). Despite the rapid development of new interactive tools, few mainstream GISs incorporate sophisticated interactive map visualization tools or directly support Web-based interactive maps. ESRI's MapObject application builder software creates online maps from a GIS (which provides the digital map) that can then be dynamically placed in a Web page. However, the interactive toolbox for MapObject is limited to pan, zoom and query type operations. Ultimately, the development of sophisticated interactive map tools is limited only by our imagination and our common sense.

The creation of online maps and the acceptance of interactive map tools should not proceed without warnings. Users may have a naive belief that digital maps (and the interactive maps made from them) accurately reflect reality (Goodchild 1991) (especially when they are presented with fancy graphics and accompanied by sophisticated tools). Digital maps tend to imply a degree of locational and attribute accuracy not always warranted by the data (Buttenfield and Beard 1992). Poorly designed informational displays frequently miscommunicate information, even if that information is accurate (Weibel and Buttenfield 1992). Few of the digital maps available online that are used in policy formation, decision-support and environmental modeling are produced by cartographic experts, often resulting in information that is inaccurate, ambiguous and poorly documented. History has shown us that informational displays are powerful tools that can be used to intentionally distort the truth. And, lastly, there exists little research that justifies the use of these interactive visualization tools despite almost fanatic acceptance.

In spite of these warnings, Howard and MacEachren (1996) state that more online maps are produced each week than their hardcopy counterparts. Design guidelines for the hardcopy display of geographic data are well documented in the cartographic literature, although there exist opposing viewpoints. Graphic designers such as Edward Tufte (1983, 1990) have promoted informational displays with high data density in order to promote detailed examination and comparison of data. Others, such as Hocking and Keller in their 1992 study of atlases, disagree, and state that many informational displays are too detailed and cluttered to be interpreted easily. Even if there was a consensus on design guidelines, adaptation to the digital environments is not straightforward. Guidelines developed for static paper media cannot simply be extended to the digital world (Goodchild and Gopal 1989). By assuming that the same design guidelines apply, the dynamic and interactive aspects of an online environment are neglected.

MacEachren (1994a) suggests that visualization tools have changed the way we represent data and the way we interact with online maps. The former concerns symbolization issues, and the latter concerns display or interface issues. While some research exists that examines symbolization in an interactive environment, little research exists that examines the interaction among user, interface and information. The effectiveness of these interactive visualization tools must be demonstrated if we are to heed the warnings stated above and go beyond using technology simply because it exists.

This type of research is especially important in the field of GIS where spatial accuracy and its representation are a growing concern (Goodchild and Gopal 1989, Beard and Buttenfield 1991, Morrison 1995, Aspinall and Pearson 1995). Metadata, or data about data, is one type of auxiliary information that documents data accuracy. Metadata may include data quality information such as historical information, as well as locational and attribute uncertainty information. While data quality is a longstanding issue in cartography (e.g., Wright 1942), the growth of digital cartography has led to new and larger problems as data layers are merged, altered and distributed without their associated spatial data quality information, and this information (or misinformation) is used subsequently in decision-making (MacEachren 1991b). In their daily use most digital maps are soon separated from their quality reports (van der Wel, Hootsman and Ormeling 1994). Methods that can incorporate both spatial data and data quality information into cartographic materials must be developed (Beard, Buttenfield and Clapham 1991). Reliability of digital maps is particularly important in fields such as resource management and environmental planning, where the potential long terms effects of decisions based on land use and land cover information are irreversible and can be devastating (Evans 1997).

Many authors have adapted traditional methods or proposed new methods for the visualization of this spatial data quality information (e.g., Buttenfield 1993, Beard and MacKaness 1993, MacEachren 1993, Klinkenberg and Joy 1994, van der Wel, Hootsman and Ormeling 1994, Aspinall and Pearson 1995). Several of these methods have been criticized because they were not designed based on how users perceptually and cognitively interact with digital spatial data presented in virtual environments. Although few such theories on map user interaction have been validated, the design of interactive visualization tools should be based on an understanding of the how users perceive and think about spatial data (McGranaghan 1996). Two new visualization tools, based on an extension to theories of cognitive map formation to a virtual environment and on an understanding of colour theory, were created for this study. These tools are unique and subject to patent.

As well as good design, visualization tools must be evaluated in real world settings in order to determine their effectiveness and to iteratively fine-tune their design (Mersey 1990). Only a few very recent studies have examined the effectiveness of these techniques (e.g., Schweizer and Goodchild 1992, MacEachren et al. 1998, Evans 1997). Experimental studies that examine the impact of different interactive visualization tools for spatial data and metadata are needed.

The design, development and evaluation of interactive visualization tools for spatial data and metadata must address several questions. *How do we effectively represent both data and* *metadata in a digital display? How can the dynamic interactive capabilities of a digital mapping environment be utilized to design more effective map displays than were possible in a static environment? How well do these displays communicate spatial uncertainty information?*

This dissertation will address these issues with respect to the visualization of spatial data, in the form of online thematic maps, and metadata, in the form of attribute uncertainty, in the context of a previous forest management project on thematic accuracy I did in 1995.

1.2 Special Notes

While sample illustrations of the visualization tools are included in this dissertation, these tools are fundamentally dynamic in nature and static illustrations fail to convey their true functionality or utility. Please reference the indicated Web pages on the Internet for dynamic illustrations of key concepts. The use of a standard Web palette and a 256 colour setting will ensure consistency of viewing.

The use of printed colour representations of the online maps used in this study is problematic. The predominant artifacts of the colour printing process are the colour shifts towards magenta, and the irregular rosette patterns caused by the colour separations in the printing process, which can be seen in several images. Prior to printing, all images were corrected by desaturating their magenta components. However, the subsequent photocopying of these images resulted in the introduction of further colour shifts. The net effect is that in some images grey appears pinkish, and some of the distinction between the pink Pure Black Spruce category and the burgundy Spruce-Aspen category has been lost. While the rosette patterns have been largely eliminated, they are still present in some images. The issue of print quality is one that is common in any translation from screen display technology to printed matter. Recent advances in PostScript technologies have largely solved these problems, however, it will take some time until these algorithms are standard fare on low-end printers. While this issue remains one that plagues digital cartography, for the purposing of reviewing this document, I suggest you turn to the online representations that are referenced throughout the text.

1.3 Thesis Guide

In Chapter One, I outline how the rapid development of computer technology has impacted the fields of GIS and cartography. I discuss how the development of new techniques in scientific visualization and the incredible growth of the Internet have intensified existing issues and created new ones, giving birth (or rebirth) to the field of cartographic visualization. The design, development and experimental evaluation of computer-based tools that facilitate the visual exploration of spatial data and data quality information are key issues, and form the basis of the research questions I address in this thesis.

In Chapter Two, I summarize the background theories and previous research in the fields of cartographic visualization and cognitive cartography that provide the milieu for this study. I provide a theoretical basis for the design of my two new visualization tools. I also present a critique of existing studies, identifying areas and opportunities for improvement. I address this critique by explaining how my research approaches many of the issues. I conclude by presenting specific research questions.

In Chapter Three, I outline the experimental framework used to conduct my research. I begin by outlining the three primary hypotheses I formulated to address my research questions. In the remainder of the chapter, I describe the components that comprise the research environment, including the creation of the thematic and uncertainty maps, the development of the four visualization tools, and the specification of the task types. I address the issues of lack of context and user motivation, common to these types of studies, and then discuss how I measure map use effectiveness (response accuracy, user confidence, user preference).

In Chapter Four, I extend the framework given in Chapter Three by describing, in detail, the experimental methodology used to examine the relation between the visualization tools and the map use tasks. I provide information on the experimental setting, including a description of test subjects, hardware and software, test environment, lab setup, assumptions, the two pilot studies and subsequent main experiment. I conclude with a detailed discussion of the statistical analyses I used.

In Chapter Five, I provide statistical results and further details of the statistical analyses I used for each part of the main experiment.

In Chapter Six, I interpret and comment on these results in the context of my original research questions. I discuss the implications of these results with respect to the design, development and evaluation of interactive visualization tools for digital map exploration, and compare the results with related research in the field of cartographic visualization.

In the final chapter, Chapter Seven, I restate how this research contributes to the evolution of the field of Cartographic Visualization, provide insight into some of the limitations of my work, and present ideas for future research.

The Appendices provide details on the individual tasks that were used in the evaluation of the tools, the training script used to guide the lab sessions, a sample of the questionnaire used to gather subject feedback on preference, and three sample lab session handouts.

I hope you find reading this dissertation insightful, interesting and somewhat entertaining.

CHAPTER 2 – THEORETICAL BACKGROUND

2.1 Overview

This study is part of a growing body of research that has emerged in the late 1990's placed at the crossroads of the field of exploration-based cartographic visualization and the experimentally-based studies of cognitive cartography. In this chapter, I introduce the field of cartographic visualization and summarize research to date. I outline recent trends in cognitive cartography and discuss the successes and pitfalls of empirical studies in this field. Lastly, I describe the application of this theoretical and practical knowledge to the design of this research and pose specific research questions.

The study presented here involves the creation of two interactive visualization tools based on how users perceive and think about spatial information. These tools are compared in an experimental setting to two other tools in order to determine their effectiveness for the visual exploration of spatial data and data quality information in a resource management decision-making environment. The development of the experimental framework for this study attempts to take into consideration several key criticisms commonly found in the field of cognitive cartography: the lack of ecological validity, lack of a theoretical component, lack of tight experimental controls, failure to account for individual differences or user motivation. While not all these criticisms can be completely addressed in a meaningful study, they can be taken into account and results explained in light of their influences.

In the following chapter, Chapter Three, I outline the details of the experimental framework used to examine several key research questions in these fields.

2.2 What is Cartographic Visualization?

Cartographic research has undergone a shift away from the cartographic communication paradigm towards the scientific visualization paradigm. Cartographic communication, which developed out of theories of information transmission, views map interaction as a closed system in which a channel of information flows from the map producer through the map to the map user. The fundamental tenet of this paradigm is that there is an optimum way to communicate or present mapped information. The cartographic task is to produce the optimum map. However, there is no consensus among cartographers on what constitutes the optimum map or how optimality should be measured. The implicit assumption is that the map producer knows what information should be communicated to the user, and designs the map accordingly. Consequently, many studies in the cartographic communication era focused on the design of symbols (Keates 1993).

While the map-centered approach to design may be effective for specialized maps, it is becoming increasingly irrelevant. With the development and dissemination of spatial databases (e.g., via the Internet), data producers, map producers and users may have little knowledge of each other. Map producers and users have become removed from data collection and processing. Automated mapping environments are now so common that map producers are becoming increasingly unsophisticated. People who have little knowledge of the cartographic design process or map composition principles are frequently producing maps. These maps may be used for presentation purposes, but are often created for personal use (e.g., data exploration). Data producers are completely removed from the end-uses of their data. The shift away from the 'optimum' map toward understanding how people use maps reflects the shift from cartographic communication towards scientific visualization (MacEachren et al. 1992).

Visualization in scientific computing (ViSC), which developed out of the field of computer graphics, explores data graphically to gain understanding (Earnshaw and Wiseman 1992). Where the traditional use of graphics is for presentation, scientific visualization develops tools that facilitate the visual exploration of data. Scientific visualization, as visual data analysis, follows the traditions of Tukey's exploratory data analysis (1977), Huber's statistical graphics (1983), and Cleveland and McGill's dynamic graphics (1988).

When reviewing the literature one must be aware that visualization is a term that has very discipline-specific meanings. Thus, in computer science, visualization research is aimed at developing systems that imitate the way we perceive and interact with real objects. Visualization research in cartography and GIS is aimed at developing methods that allow us to perceive the unknown, to become aware of relations and patterns that otherwise we may miss, to know the data better. MacEachren et al. (1992) note that the panel on Graphics Processing and Workstations, in its report on visualization in scientific computing, characterize visualization as "a method of computing." From a cartographic viewpoint, "visualization is foremost an act of cognition, a human ability to develop mental representations..." (MacEachren et al. 1992, p.

101). Nonetheless, methods developed by computer scientists remain useful in cartographic visualization research.

Visualization in scientific computing is a component of knowledge discovery. The relatively new field of knowledge discovery in databases (KDD) (Brodley, Lane and Stough 1999) includes several processes such as data mining and visualization. Thus, the increased attention being paid to visualization in cartography should be seen as being part of a universal phenomenon. Scientists are becoming increasingly concerned with extracting the maximum amount of knowledge from databases of ever-increasing size. If maps are the best way to present 2D spatial data, and if they are the best means of representing spatial relations (two assumptions no cartographer would contest), then cartographers must monitor the advances being made in disciplines such as ViSC and KDD and apply them appropriately to cartographic problems.

In cartography, visualization has, of course, been a central tenet. The visual analysis of static maps has a long history beginning with Dr. Snow's map of the 1854 cholera outbreak in London (Tufte 1983). However, over the last decade, developments from ViSC, such as dynamic user interfaces, have been applied to create complex visual representations of two and three dimensional geographical data. The resulting subfield is called cartographic visualization or, more recently, geo-referenced visualization.

Unlike cartographic communication, which emphasizes map design, cartographic visualization emphasizes map use. The key in cartographic visualization lies in its emphasis on the use of graphics in the development of ideas, not, as in traditional (carto)graphics, in their presentation. While both are important and valid uses of map data, this research focuses on techniques for interacting with 2D spatial data. MacEachren (1994) states that cartographic visualization is not just about making maps, but is about the way users interact with them. Map use has been conceptualized by MacEachren as a three-dimensional cube in which user interaction (e.g., high, low), purpose (e.g., revealing unknowns, presenting knowns) and use (e.g., private, public) form the three axes (Figure 2.1). Cartographic visualization is concerned with map use that is highly interactive for the purpose of revealing unknowns in private use (Fauerbach et al. 1996). Cartographic visualization, being a more personal and spontaneous activity, may necessarily require a higher degree of user interaction than does cartographic communication.

Turk's (1990) model of cognitive ergonomics in GIS is also user-centered. The user's mental model--or in the case of spatial information, the user's cognitive map--is the focus of Turk's model. Visualization design is one component of this model which interacts directly with usability testing and user interface design. As Unwin et al. (1994) state, Bertin's (1983) "ideas of graphical information processing, and the monosemic nature of graphics, together with (his) confidence in the human eye/brain to find patterns in graphics which present as much of the data as possible, are almost a manifesto for ViSC".

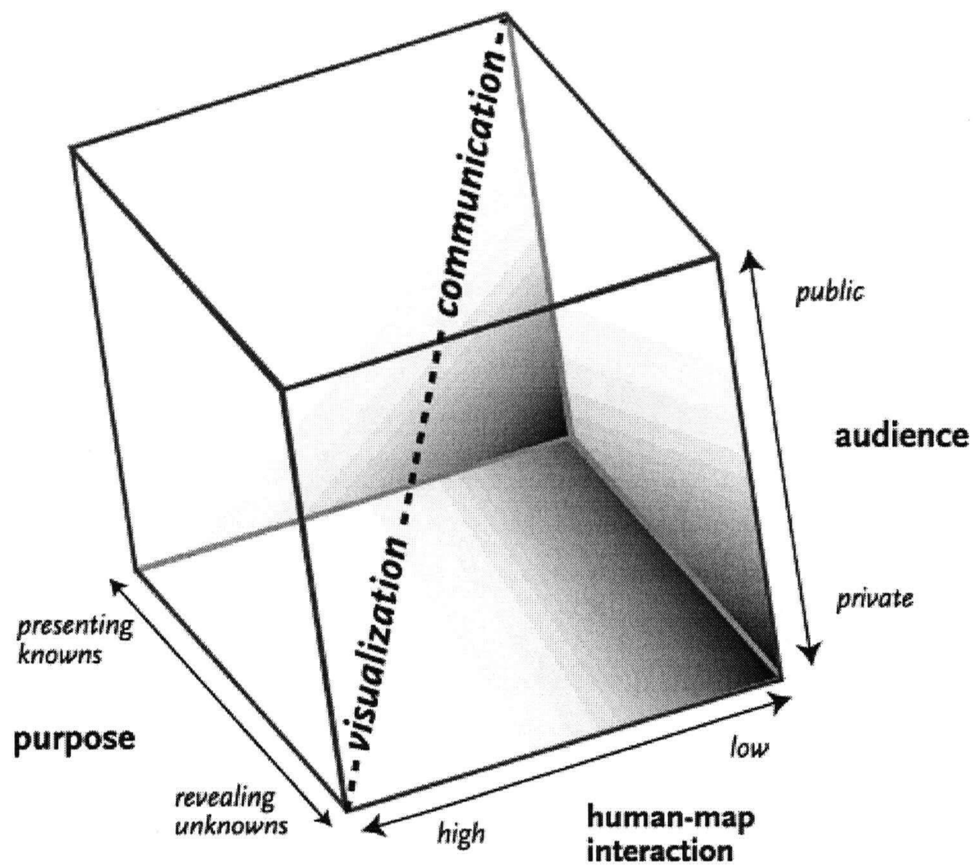

Figure 2.1 The realm of Cartographic Visualization.

2.3 Recent Studies in Cartographic Visualization

It is quite clear that scientific visualization and cartography have much in common. Given this, it should not be surprising that a number of studies that could be called cartographic or scientific visualization have emerged in the last few years. Cartographic visualization is a 1990's phenomenon, and like most technologically-driven areas, it is changing rapidly. Reviews by MacEachren et al. (1992), describing the work of a very active group based at the Pennsylvania State University and Syracuse, and by Fisher, Dykes and Wood (1993), outlining work at the UK Midlands Regional Research Laboratory at Leicester, provide good overviews of the field as it stood in 1992 and 1993, respectively.

The concept of multiple, linked windows is one tool commonly used in cartographic visualization. Display windows are linked to permit simultaneous viewing of the mapped data together with other graphics (e.g., histograms or scatter plots). For example, this allows the same data to be highlighted on a number of windows at the same time. Haslett et al. (1990) developed the system originally called SPIDER (now REGARD) to provide multiple linked windows in a Macintosh environment.

Some of the earliest cartographic experiments attempted to provide a means by which mapped data could be explored easily and interactively. For example, Egbert and Slocum (1992) developed a tool to provide instantaneous re-classification of choropleth mapping. Although originally developed in the Macintosh environment, the same notion of brushing (related to multiple linked windows) has recently been implemented on UNIX workstations. This has been done using a basic, public domain CASE tool and the GRASS geographical information system (Dykes 1994).

Techniques from scientific visualization have also been used to help designers improve the traditional choropleth map, as well as replacing them by theoretically and practically more useful alternatives (Bracken 1994). Although not usually thought of as such, the technique of density estimation yields a map that enables the scientist better to visualize density variations in a point pattern (Gatrell 1994). In terms of the spatial error distribution, computer technology has been used to visualize the residuals from kriging interpolation (Wood and Fisher 1993) or to produce spatial distributions of the uncertainty in classified remotely sensed images (rather than just producing a global accuracy estimate for the entire image) (Goodchild, Chih-Chang and Leung 1994).

Much of the above makes use of computer technology simply to improve on traditional graphics of the type that Bertin (1983) would recognize. More interesting uses of scientific visualization have been to create entirely new forms of map display. In these displays, the seven cartographic variables originally recognized by Bertin (eg. plan, size, shape, value, orientation, hue, texture) have been extended to include projection, animation and time, and sound.

Unconventional projection systems (eg., Hammer) have proven to be excellent graphical devices in themselves. Some of the more obscure standard projections that are seldom used have useful properties. Computerization enables maps to be drawn easily on almost any projection, and encourages more innovative uses. For example, maps based on the Hammer projection give an almost uninterrupted view of the seas (Spilhaus 1991). In addition, using unconventional views, data can be plotted onto an area-cartogram base where the size of an area reflects the value of some variable (say population) rather than geographical size (Dorling 1994).

Similarly, the effective use of animation to produce a series of maps has been known for a long time (Tobler 1970, Moellering 1976), but until recently has only been realized with difficulty. Most obviously, a sequence of maps can be ordered by time scale to display either faster (most common) or slower than the real world phenomenon they represent (Fisher 1993). However, MacEachren et al. (1992) suggest some alternatives. Animation can use a time ordering that is out-of-sequence, or sequence using the value of a selected variable. The dynamical nature of a variable can be emphasized (e.g., maps of streamlines of a flow). The Pennsylvania State group has developed the use of animation to include a whole series of new dynamic variables including: the duration, rate of change, ordering and phase (or rhythmic repetition of events) of the map sequence. Dorling and Openshaw (1992) use animation to visualize space-time patterns (e.g., the diurnal variation in crime incidence across a conurbation (Openshaw, Waugh and Cross 1994).

Although the concept of using animation for detecting patterns at the ideation stage of an investigation has been around for some time, several difficulties remain. The first is that while an investigator may gain insight from an animation, it is difficult to publish such insights in the usual way. An option common to the fields of engineering and computational fluid dynamics is for visualization sequences to be "published" as videos. The rise in Web-based publication has begun to facilitate the widespread publishing of multimedia. However, limited bandwidth outside government and corporate institutions and the lack of standardization among browser plug-ins (required to “play” most multimedia files) remain obstacles to dissemination. Second, there are very few rules for the use of graphic variables in animation. The concept of a graphic script developed by Monmonier (1992b, 1993) represents an important step towards remedying this deficiency. Finally, Fisher (1994) has experimented with the use of sound to increase the bandwidth of map to map user communication. Like time, sound is not a simple, single variable.

It has dimensions of tone, volume and rhythm that can be used. As with animation, we know almost nothing about good design in the use of these variables (Krygier 1994).

Visualization techniques including 3D modeling (e.g., of landscapes) and virtual reality have been experimented with, each allowing the user to experience a "reality-based" view of 3D data (Sheppard 2000). While 3D models are thought to improve the user's mental processing and understanding of spatial data and to place information in a visual context from which the broader implications of the data and courses of action can be seen, 3D models are computationally expensive and beyond the scope of this research.

2.4 Recent Trends: Cognitive Cartography

At the same time as this shift from cartographic communication towards scientific (and cartographic) visualization has occurred, there has been a resurgence of cognitive research and empirical studies in cartographic visualization. Cognitive-based empirical research, which aims to provide an understanding of the cognitive processes and human-computer interaction relative to online map use, is invaluable in guiding the design and evaluation of effective cartographic visualization tools. Effective map use involves two important linked processes: visual perception and spatial thinking. It has been argued that there is a strong similarity between perceived external objects and their corresponding internal representations (Shepard 1978). Much of this effect may be explained by the fact that the human visual system is very effective at processing images and understanding spatial relations (Legge and Luebker 1991). Subjects are able to make quicker and more accurate judgements of spatial relations when an appropriate mental image has been formed. Cartographic visualization research is therefore directed towards making spatial relations easier to visualize.

Early research that examined the relation between the map (paper) and the map-user adopted a psychophysical methodology in which the user's perception of one or more components of the map were examined. For example, Chang (1980) examined users' perceptions of the magnitudes of differently scaled map elements. However, psychophysical studies may be criticized on several fronts. First, since these experiments typically only examine a single map element, they ignore map-use context. Petchenik (1983) argues that the same symbol may vary in appearance against different backgrounds and in different map-use contexts. Thus, results from psychophysical studies are not easily generalized; they lack ecological validity. Second, users' expectations can affect psychophysical test results. Third, the failure to consider individual differences with respect to intelligence, previous map knowledge, and experience is also limiting (Blades and Spencer 1986). Finally, the emphasis of psychophysical studies on perception ignores the types of higher mental processes that are typically involved in map-use tasks (Gilmartin 1981).

In response to these criticisms, much map-use research in the 1970's and 1980's began to take a more cognitive approach. The role of cognitive processing of cartographic information was examined by several researchers (Petchenik 1977, Thorndyke 1981, Steinke and Lloyd 1985, Castner and Eastman 1984 & 1985, Peterson 1985, Blades and Spencer 1986, Lloyd 1989 and MacEachren 1991). These cognitive map-use studies were of two main types. One type examined the users' cognitive strategies in particular map-use tasks; for example, Thorndyke and Stasz's (1980) investigation of individual differences in spatial knowledge acquisition. The other type examined particular cognitive phenomena, such as Lloyd and Steinke's (1985) investigation of point symbols. Research in cognitive-based map-use has continued into the 1990's (e.g., Mersey 1990, Monmonier 1991, 1992, 1994, 1996, Peterson 1994, Parsons 1995, Wood and Gilhooly 1996, McGranaghan 1996, Nelson and Gilmartin 1996, Lloyd et al. 1996, Patton and Cammack 1996, Evans 1997, MacEachren et al. 1998), but not without criticism.

2.5 An Attack and Defense of Cognitive Cartography

Experimental studies in both cognitive cartography and human computer interaction (HCI) may be criticized on several fronts. First, and most importantly, many have lacked a theoretical justification that could be used to explain the users' performance (Petchenik 1983, Blades and Spencer 1986). Second, studies in which small alterations are made to the display environment have frequently failed to show significant results. Third, small distinctions in design elements that depend on fine perceptual adjustments are likely to be masked by individual differences and uncontrollable circumstances in the map reading environment (Monmonier 1980).

McGranaghan's study of choropleth maps on LCD panels (1996) exemplifies all three of these criticisms. The maps were designed with different combinations of background colour (dark or light), symbol order (whether dark or light symbols represented large data values) and legend arrangement (where the symbol for large values was at the bottom or top of the display).

First, a theory of choropleth map reading that explains how well or how rapidly a given individual would perform with a given map was not presented or inferred. Second, of the three main effects and their 2- and 3-way interactions, only one (background colour) was significant at the .05 alpha level. Third, McGranaghan found that performance depended largely on a combination of learned familiarity with symbolic conventions and available contrast within a map display.

Despite these criticisms, many authors maintain that understanding human cognition is essential if we wish to improve the design of effective visualization tools for spatial data (e.g., Lloyd and Steinke 1985, Blades and Spencer 1986, Lloyd 1989, MacEachren and Ganter 1990, DiBiase et al. 1992, Monmonier 1992, Turk 1990, Peterson 1994, Parsons 1995, McGranaghan 1996, Nelson and Gilmartin 1996, Patton and Cammack 1996). Rather than abandoning the cognitive approach, what are needed are well designed experimental studies that address the weaknesses of many of the previous studies. First, the studies must be ecologically valid. They must mimic real-world users in real-world situations so that results can be generalized. Second, they must have a theoretical component that attempts to explain the results. Third, influences that might undermine significant results (e.g., user experience, gender, display parameters) must be controlled in study designs. Since controlling so many variables at once is difficult, many studies will be needed to clarify results; anecdotal results may be valuable in designing these. Fourth, the studies should consider the contribution of individual differences to performance, and the results should be explained appropriately (for example, results given in terms of average users)[1]. Finally, user motivation in map use should be considered and accounted for in the analyses of the results.

2.6 Research Design

In keeping with this need for more focused research, a cartographic visualization study was conducted as part of the requirements for my doctoral degree. The simultaneous exploration and analysis of two related data sets is a common map-use problem. The development and empirical testing of online graphical methods that allow a user to explore two related data sets, where one is used to modify the other, represents an important advancement in cartography. Specifically,

[1] Personal correspondence with Janet E. Mersey, Ph.D., Professor, Department of Geography, University of Guelph, Ontario, November 1997.

the appropriate and simultaneous display of a thematic, spatially distributed data set and its associated uncertainty data is an important problem that remains largely unsolved (Beard and Buttenfield 1999). In response to this challenge, I imagined, designed, implemented and empirically tested a set of visualization tools that could be used for the simultaneous online exploration of a bivariate data set composed of thematic and uncertainty information.

Two original visualization tools were designed based on theories of spatial perception and cognition. An empirical study was then designed to examine the effectiveness of these tools (compared to a side-by-side map and a form of sequenced map) for the simultaneous visualization of thematic data and uncertainty data. The research proposal was explicitly designed in light of the experimental design principles presented above. In one component of this study, students were asked to answer a series of questions, the answers to which depend upon a careful examination of two maps. The two maps are a thematic map illustrating various forest species composition types and a certainty map illustrating one component of data quality for the data presented in the thematic map. Both maps are based on a quantitative study of thematic map uncertainty conducted by Michael Joy (Joy and Klinkenberg 1996, Klinkenberg and Joy 1994).

Theoretical Basis: Understanding the Processes

The theoretical basis of this study integrates the results of previous research in three specialized areas: spatial cognition, visualization of spatial information, and online map use, the findings of which are highly interwoven in the field. The first step in the design of effective visualization tools for spatial information is understanding the processes of spatial cognition–that is, understanding the map user's spatial cognitive abilities and limitations. Spatial cognition is the process by which humans understand or think about spatial information (Parsons 1995), and by which they represent and store information about the geographical space in which they live. An individual's understanding (i.e., cognition) of their environment is the result of both the direct and indirect interactions they have had with that environment. The development of cognitive maps is one component of spatial understanding. It is a highly complex process and results from different types of interactions (Kitchin 1994). Piaget has suggested that an individual's direct interaction with the physical environment is the most important component in this process (Hart and Moore 1973).

Development of spatial understanding through direct interaction can be extended to virtual environments. An individual who interactively explores online spatial information may be better able to form spatial cognition about that space than an individual presented with a static display.

Designing the Tools

Several visualization methods have been suggested for the dual presentation of spatial data and metadata. These include static tools such as bivariate maps and side-by-side maps (i.e., map pairs or Tufte's small multiples), and dynamic tools such as sequenced maps and animations (Fisher 1996). Interactive tools have recently been proposed and have gained much attention. However, the current state-of-the-art in cartographic visualization, as implemented in commercial GIS products, is not yet up to the standards available from ViSC-specific software. Until such time as sophisticated visualization tools become standard in GIS packages, prototype stand-alone tools–tools that can be easily developed and tested using software such as JAVA–will be a necessary part of any experimental study. While static side-by-side maps provide the baseline for exploration in a non-digital environment, they are cumbersome in an online environment due to limited display space and lack of interactivity.

To meet the exploratory goals of ViSC, tools must be interactive and the display dynamic. A sequenced map display can be created using a simple switch or toggle tool. The toggle tool controls a single map display and allows the user to alternately display two (or more) maps–in this study the two maps are a thematic map and an uncertainty map (as described above). Since the user controls the map display and can visually merge the maps by rapidly toggling between them, the toggle tool is classified (by me) as a more integrative tool than a simple side-by-side display. An interactive bivariate map display can be created using a data merger tool that allows the user to control the simultaneous presentation of two map variables (e.g., theme and uncertainty) in one display. If colour is used to represent one variable, and grey scale the other, then colour desaturation can be used to merge the two. A final level of interactivity is created by adding 'hypermap' functionality to the interactive bivariate map. A hypermap tool displays both attribute and uncertainty information about spatial locations in response to a mouse click query. In this study, these three tools (toggle tool, data merger and mouse click query) are used to create interactive sequenced, bivariate and bivariate hypermap map displays.

Testing the Tools

The added dimensions available through the use of dynamic visualization tools also places an increased burden on the cartographer. The message imparted through, for example, visualizing uncertainty information, is far less tangible (and familiar) than that imparted using traditional cartographic themes and methods and, therefore, far more difficult to evaluate. Thus, the means of examining the effectiveness of the experimental procedures must be carefully considered. Since the effectiveness of the visualization tools is to be determined relative to real-world map-use tasks, a relevant task-set must be developed to test the tools. Many taxonomies of map-use tasks have been used for both hardcopy and online map use studies. What is needed is an ecologically-valid taxonomy of tasks and a set of map-use tasks/questions that replicate the use of interactive visualization tools in a specific policy formation and decision-making environment.

Resource management can provide an appropriate task set for testing the tools. Interactive visualization tools play an integral role in the management of the earth's resources. Spatial data quality information is paramount in this setting. Natural resource managers and policy makers are often presented with visual representations developed with current GIS technologies, and they may have little or no understanding of how the information is developed, what its limitations are, and how different representations can be derived from the same basic data. With the increased use of GIS for spatial decision making and planning in forest management and conservation, understanding visualization tools for spatial data and metadata is crucial. For these reasons, the task set will be developed specific to a forest planning and conservation environment.

Thematic maps are one of the basic tools of forest management. Species composition maps provide spatial forest class and cover information that is used in a majority of forest management and planning tasks (e.g., land use allocation, site location and selection, access planning, change management and inventory monitoring). Information from thematic maps also serves as input to more complex modeling approaches (e.g., linear programming, multiple criteria decision making and spatial decision support systems) used to build ecological and forest models (e.g., succession models, determination of annual allowable cut, forest productivity models, harvest operations planning).

Each of these analyses involves the combination of many tasks. However, the exploration of spatial distributions and patterns is an important task present in all of the analyses outlined above and forms the basis of the task set used in this study. Specifically, I focus on tasks that incorporated spatial uncertainty information into the exploration of spatial patterns. In order to create a task set that mimicked a range of exploratory tasks commonly found in the analyses above, I included three types of tasks in my experimental design (see Section 3.5). Individual tasks were classified based on the type and relative complexity of the cognitive operations required to solve them (e.g., identify a specific location, search for a spatial pattern).

But how should we actually evaluate the effectiveness of visualization tools that can be used to perform these tasks? An effective map, or in this case, an effective visualization tool, is simply one that fulfills its required task (Wood 1993). Effectiveness is often equated with map-use performance, and accuracy (percentage of correct responses) and response time are the traditional measures used in such evaluations (Mersey 1990, McGranaghan 1996, Nelson and Gilmartin 1996). It is assumed that accuracy and response time are measures of the cognitive difficulty of performing tasks. Greater cognitive difficulty (i.e., less accuracy and longer response times) is associated with less appropriate display designs for a particular task. However, many experiments (e.g., McGranaghan 1988 &1989) fail to distinguish whether response time differences were attributable to map design, the individual's cognitive strategies for interpreting the maps, or familiarity with the task. It has also been suggested that accuracy and response time may be inversely related. Since accuracy tends to improve with longer response times and decline for shorter times, the two factors may be difficult to separate.

In this study, three criteria will be used to define map-use effectiveness, that is, to evaluate the tools. Accuracy is the first. When interacting with spatial data in a policy formation and decision-making environment, an accurate interpretation of information is critical. An accurate interpretation of information occurs when the user chooses the "best" solution from a set of alternatives based on a set of criteria. In this study, "best" is defined as the solution that most closely matches the information portrayed by the data.

Response time is less important, since in real life the map reader will not be subjected to restricted viewing. A user's confidence in her decisions is more relevant and can be used as a surrogate for accuracy on decision-making tasks (Koriat, Lichtenstein and Fischhoff 1980, Zakay 1997). Studies in Management Science have found that users put more weight on their own internal subjective feelings of confidence than almost anything else when decision-making (Oz, Fedorowitz and Stapleton 1993). Confidence is important since it is known to have a significant impact on not only a user's decision-making ability, it directly influences his behavior, actions and the credibility with which others perceive him, after the decision has been made (Sniezek and Buckley 1993, Zakay 1997). Thus, confidence is used in this study as the second measure of map-use effectiveness. Users will be asked to rate their confidence in their chosen "best" solution using a five-point confidence scale where the value one represents "most" confident and the value five represents "least" confident. Confidence scores were subsequently inverted and averaged across tasks to create an average confidence score out of a maximum of five (see Section 4.9).

Since users may have different cognitive strategies for solving map-use tasks, it is important to consider subjective preference when evaluating visualization tools. Preference is also key since it is important to consider users' reactions to and acceptance of visualization tools (Evans 1997, Beard and Buttenfield 1999). Thus, preference (recorded as a subjective ordinal rating from most preferred to least preferred) is the third criteria used to examine map-use effectiveness. Additional questions on preference (see Appendix C for a sample tool evaluation worksheet) and the collection of user profile information will be analyzed to address the issue of individual differences, which has been suggested by some authors to mask all of the other effects (Mersey 1990).

2.7 Research Questions

The key question I seek to answer in this research program is central to cartographic visualization research:

> *In an online geographic/cartographic environment, are displays that offer highly interactive visualization tools (e.g., hypermap, interactive bivariate map) more effective for communicating spatial information (e.g., thematic and uncertainty information) than less interactive tools (e.g., side-by-side, sequenced map)?*

The theory of spatial cognition and the formation of cognitive maps can be extended to virtual environments to suggest that increased interaction with geographic information will improve user accuracy and confidence on map-use tasks. This should particularly apply in situations where more complicated tasks are required of the user. Individual differences may lead to different user preferences relative to the visualization tools, which may or may not affect accuracy and confidence.

Specific questions that I address include:

1. *Do low level interactive visualization tools present spatial information as effectively (accuracy and confidence) as do highly interactive visualization tools?*
2. *Does increased interactivity result in increased accuracy?*
3. *Does the use of highly interactive tools result in the same accuracy but increased confidence, or decreased accuracy but the same or greater confidence?*
4. *How does the addition of textual attribute and uncertainty information (via hypermaps) affect accuracy and confidence in map-use tasks?*
5. *Within the subset of tools tested, is there an optimal "best" visualization tool for this type of spatial data and metadata, or are different visualization tools best suited to different task types? If so, which visualization tools are best suited to different task types?*
6. *Does increasing the level of interaction buy us anything? If so, for what task types?*

This research considers these questions. From my findings I will shed some light on the role that visualization tools can and should play in cartography. Furthermore, my results will allow others to design better interactive online visualization tools.

In Chapter Three, I outline the experimental framework used in this research, beginning with three key hypotheses that address the research questions posed above.

CHAPTER 3 – EXPERIMENTAL DESIGN — FRAMEWORK

3.1 Overview

Past cartographic visualization research has examined the affects of representational design factors, such as colour (e.g., Mersey 1990), map complexity (e.g., MacEachren 1982, Patton and Cammack 1996) and symbolization (e.g., McGranaghan 1996, Nelson and Gilmartin 1996) on map use. Other Human Computer Interface (HCI) studies investigate the interplay between physical display type and user performance (e.g., Marcus 1995). Many different visualization tools have been proposed for spatial data (see summary in Beard and Buttenfield 1999). However, few tools have been examined experimentally. This is one of the only known experiments to investigate the combined effects of tool type and task type on digital map use effectiveness for spatial data and metadata (others related studies include Slocum et al. 1990 and MacEachren et al. 1998).

One of the primary advantages of the online environment is that it allows the user to interact with a display and have the display respond dynamically. In this study, user-display interaction, as facilitated by online visualization tools, is examined. Specifically, subjects explored two online maps (section 3.3) with different visualization tools (section 3.4) and performed a series of map-use tasks (section 3.5). The four tools represent (and are classified according to) varying levels of interactivity. The tasks were classified as one of three types (specific, general or integrated) based on task complexity (a combination of the type of cognitive operation required and estimated difficulty). The maps were generated from a previous study by the author examining spatial data quality. One map is a thematic map of forest species composition (Figure 3.4) and the other is the corresponding certainty information (i.e., spatial data quality information) (Figure 3.5) for a subset of a ten km^2 area of boreal mixed wood forest in Northern Alberta. Subjects addressed a series of map-use tasks of varying complexity. (i.e., cognitively difficulty) in the context of a land use conflict for this forest area. This chapter describes the research hypotheses, maps, visualization tools, task set and framework for evaluating map use effectiveness. In Chapter Four, I describe the experimental methodology used to examine the relations among task, visualization tool, and map use effectiveness (user performance and preference).

3.2 Research Hypotheses

Four different visualization tools were examined to determine how effective they were for three different types of map-use tasks. Three hypotheses were examined by determining the interdependence among map-use task complexity, visualization tool, and three measures of effectiveness (accuracy, confidence and preference).

The first hypothesis examines how subjects' response accuracy (percent correct) on map-use tasks is affected by task complexity (i.e., relative cognitive difficulty) and visualization tool. The second hypothesis examines how subjects' confidence in their decisions is related to response accuracy, and to task complexity and to visualization tool. The third hypothesis examines how subjects' performance (i.e., accuracy and confidence) is related to user preference. Note that the use of the term *accuracy* refers to response accuracy on tasks, not map accuracy.

Hypothesis One: *Accuracy will decrease as the complexity of map-use task increases. The rate of decrease will vary depending on the degree of interactivity of visualization tool used.*

This hypothetical relationship is illustrated in Figure 3.1. The graph shows that for simple map-use tasks (specific), accuracy is high for all visualization tools used. As task complexity increases discretely along the X axis (general, integrated), accuracy scores will decrease most rapidly for non-interactive tools (side-by-side and toggle) and least rapidly for highly interactive tools (merger and hypermap-merger).

While low interaction tools are simple to create and use, they may not be as effective as tools that are more highly interactive. However, highly interactive tools are more difficult to create and may require training to use. If this hypothesis holds, interactive visualization tools would be preferable for complex tasks. Since all tools would yield accurate results for simple tasks, the choice of tool is less crucial if maps are only to be used for simple tasks.

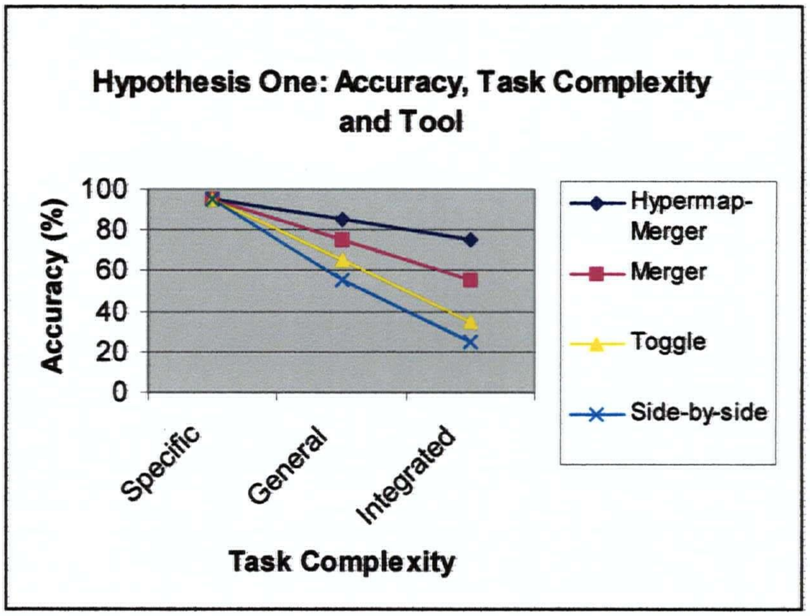

Figure 3.1 Hypothesis one: accuracy, task complexity and tool.

Hypothesis Two: *Subjects' confidence in their decisions will be proportional to accuracy and will decrease as the complexity of map-use tasks increases. The rate of decrease will vary depending on the degree of interactivity of visualization tool used.*

These hypothesized relations are shown in Figures 3.2 and 3.3. Subjects' confidence decreases as accuracy decreases (where low confidence is shown as low average values, see section 2.6), and subjects' confidence in their decisions also decreases as task complexity increases. Confidence has been suggested as a surrogate for accuracy (Koriat, Lichtenstein and Fischhoff 1980, Zakay 1997); however, this will only be true if other factors are not significantly affecting confidence (e.g., overall feelings of self-confidence, ease of use of tool, certainty values of data). It is also important to examine how accuracy and confidence are related for individual tools. If subjects' confidence is high but accuracy low for a particular visualization tool, then this tool may be producing unwarranted confidence. If interactivity provides an advantage in solving complex tasks, then as task complexity increases, confidence scores will decrease most rapidly for non-interactive tools and least rapidly for highly interactive tools.

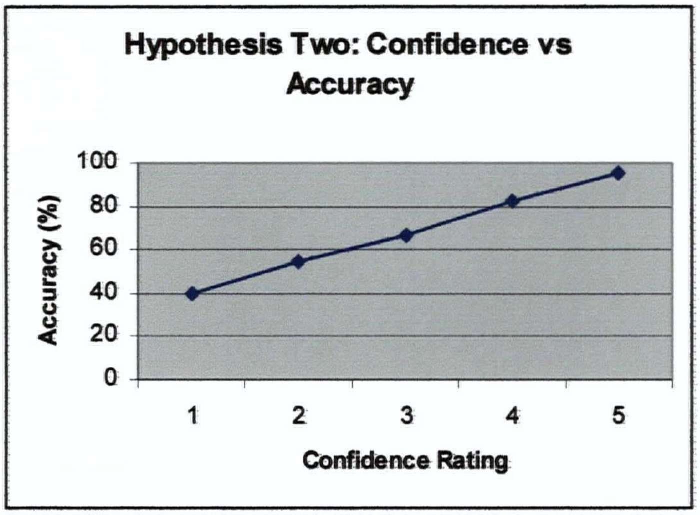

Figure 3.2 Hypothesis two: accuracy and confidence.

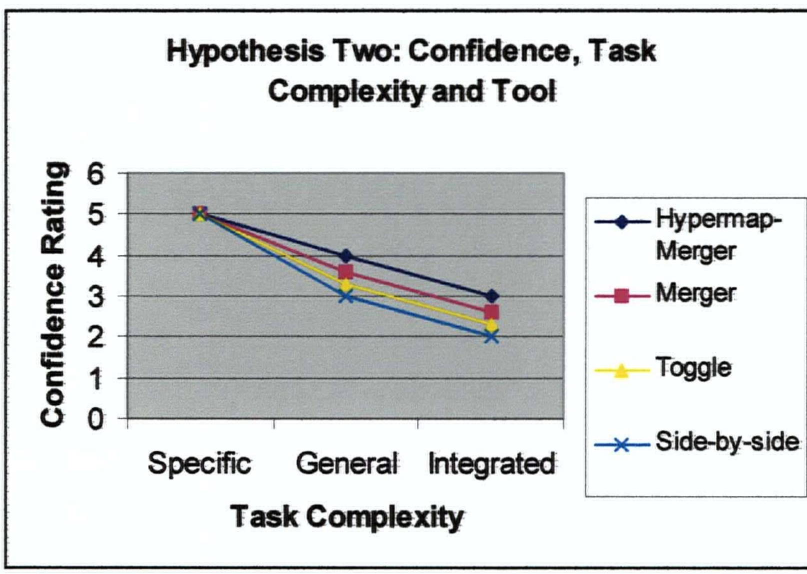

Figure 3.3 Hypothesis two: confidence, task and tool.

Hypothesis Three: *Accuracy and subjects' confidence in their decisions will be higher for visualization tools they prefer.*

There is a range of individual differences in visual and spatial abilities, preference, learning styles and cognitive strategies employed to solve spatial tasks (Blades and Spencer 1986, Mersey 1990). User preference, that is, which tool a user prefers, may have a strong effect on performance (accuracy and confidence). The other effects of individual differences were controlled as much as possible and, where they couldn't be controlled, were examined to determine if they affected accuracy and/or confidence significantly.

3.3 The Maps

The maps used in this study were taken from a small project on thematic accuracy done by the author in 1995 as part of a larger project that took place in a 73,000 km^2 region of the boreal mixedwood forest of northeastern Alberta, Canada (Cummings et al. 1995). A subset of the maps generated in the thematic accuracy project was selected for use in this experiment. This subset includes a thematic and uncertainty map for a subregion chosen from one township. The subregion (one quarter of a township i.e., 5 x 5 km) was chosen based on its size (suitable for viewing at 800 x 600 screen resolution) and the variability of thematic classes found in the region.

The base map (Figure 3.4) was created from the Alberta Vegetation Index (AVI) forest inventory. AVI data was collected at a scale of 1:15,000. The thematic class for the base map is species composition. The map contains seven key classes. Commercially, the Alpac Forest Management Agreement is used to produce aspen for pulp and black spruce for saw wood. The classification reflects the commercial interests (i.e., conifer content) in the area. These classes are: logged or roads, pure aspen, leading aspen-black spruce, pure black spruce, leading black spruce-aspen, other forested types, and water. Birch and poplar were aggregated with aspen. The classes were coded in seven hues: white, green, blue, magenta, red, yellow and cyan that were distinct at three saturation levels as explained in section 3.4.

The uncertainty map (Figure 3.5) was derived from vegetation survey plots that were created as part of a study on habitat fragmentation (Schmiegelow and Hannon 1993). A species composition map created from the information in the vegetation plots was taken to be "reality." For over 200 plots, species composition information provided a count of the number of stems, broken down by species and diameter class. This information was aggregated and used to calculate relative volume (basal area) of each species class in the canopy. Once the vegetation map was aggregated from plot data to vector data (corresponding to the AVI stand boundaries) the two maps were compared to generate an uncertainty map. The uncertainty map was recoded into three classes based on the seriousness of mismatch. For example, an aspen stand (according represents a severe mismatch and was therefore coded as high uncertainty since it is unclear what species type actually dominates that stand. A spruce-aspen stand coded as aspen-spruce is a minor mismatch and was coded as medium uncertainty. A direct match (i.e., agreement between vegetation survey and AVI map) was coded as low uncertainty. A map classification based on certainty (rather than uncertainty) was used to avoid the use of a double negative (e.g., less uncertainty) that may confuse users. The uncertainty values (high, medium and low) were inverted to create the certainty map used in this study.

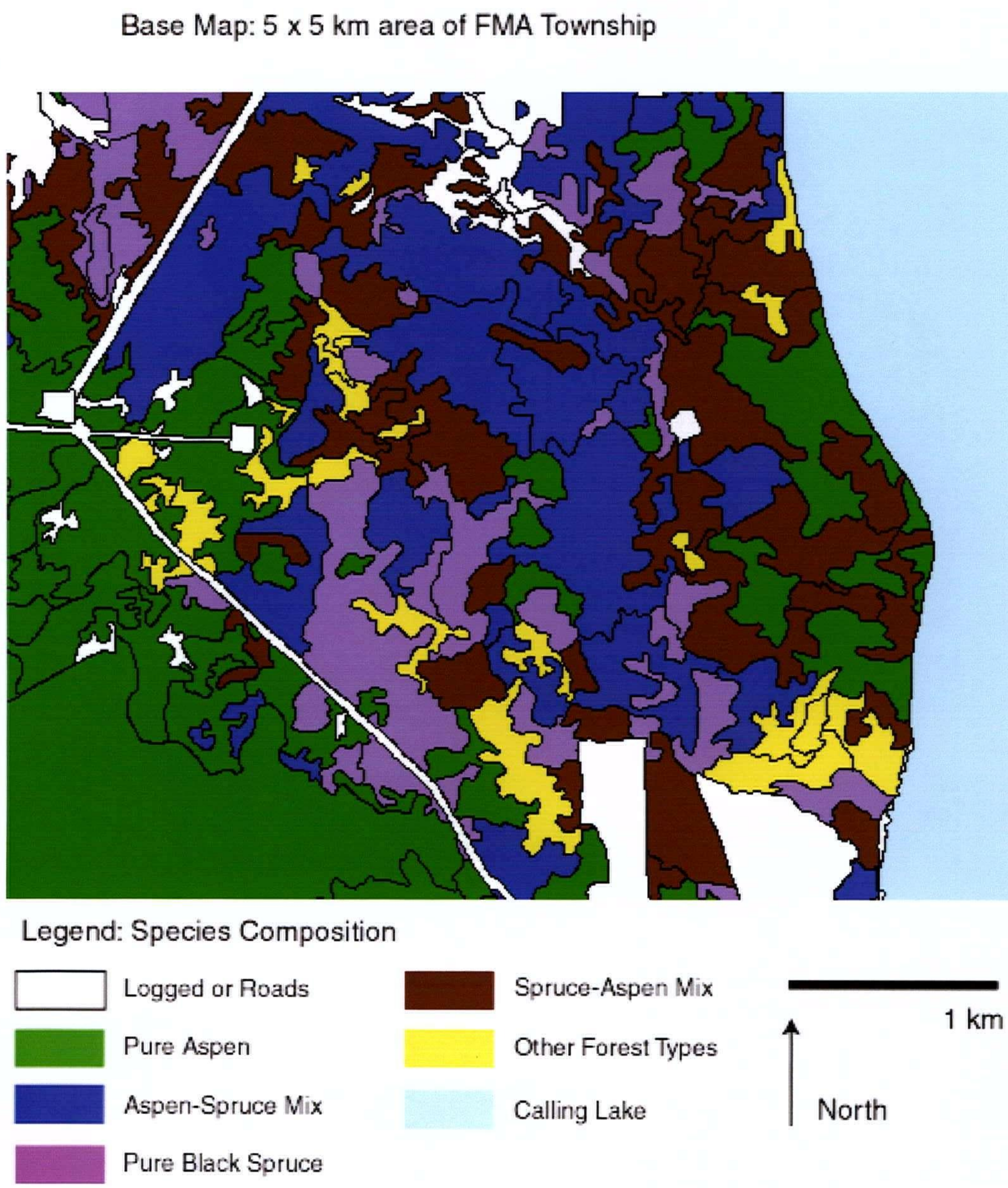

Figure 3.4 Base map: forest species composition.

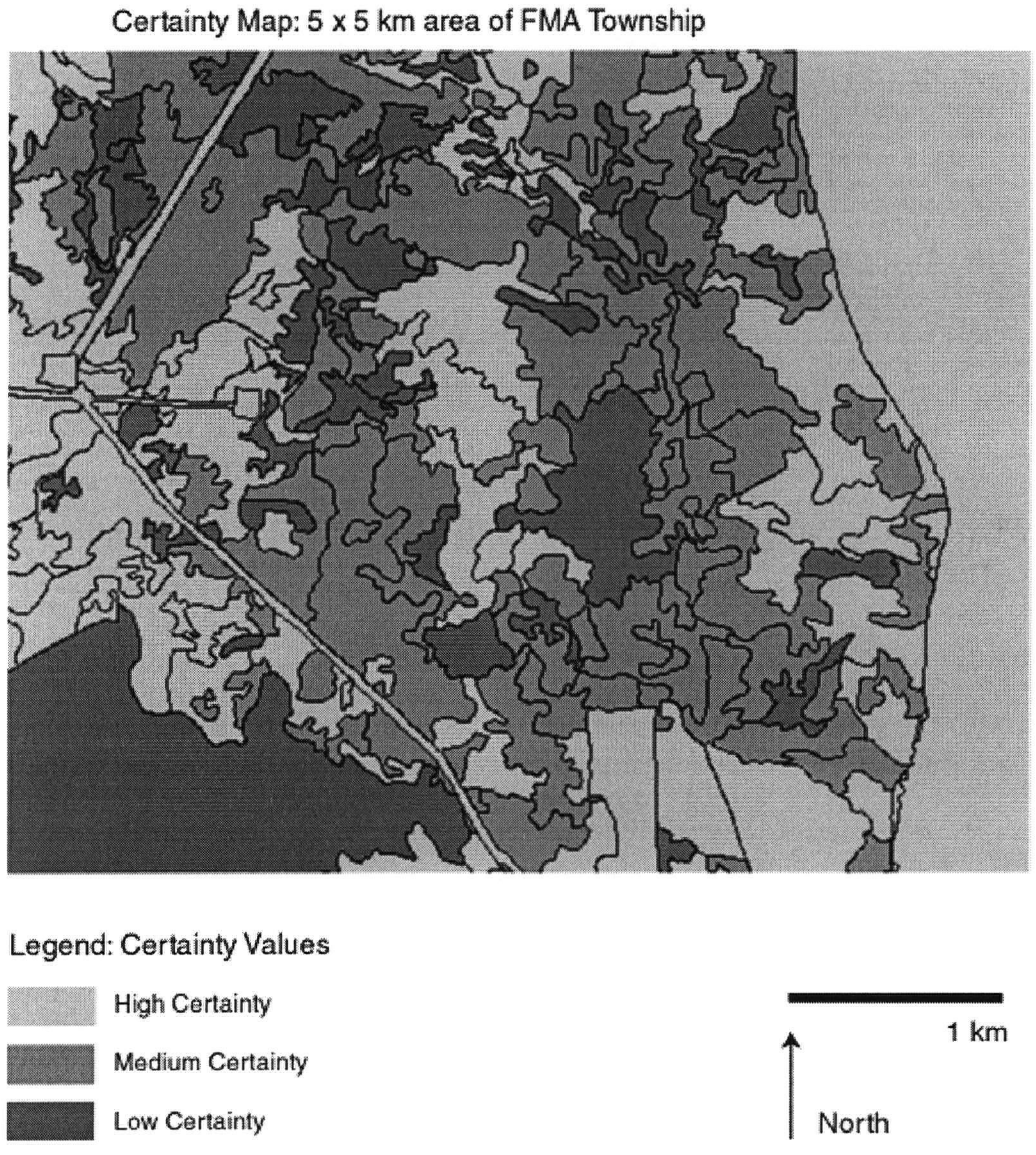

Figure 3.5 Certainty map.

3.4 The Visualization Tools

The visualization tools used in this study (in order of degree of interactivity) are: side-by-side scrolling (traditional/static), toggle (sequenced bivariate), merger (interactive overlaid bivariate) and hypermap-merger (interactive overlaid bivariate with attribute query). Map legends were included in all displays. The side-by-side and toggle tools, and hypermap functionality emerged from work by other authors (e.g., Slocum et al. 1990). The merger tool(s) was designed by the author based on the suggestion of MacEachren et al. (1993) to merge data and metadata into a bivariate display. However, instead of using symbolization to keep each component visually separate, hue and saturation were used. Data and metadata were represented dynamically in a display that can be interactively controlled to display each map layer separately or in a range of overlaid states.

Tool 1: Side-by-side

A traditional (print-based) display of corresponding maps is to place them side-by-side. This approach is cumbersome for computer displays, since they have limited screen real estate. It is included in this study as a baseline (the 'do nothing new' approach). In order to fit both maps, the legend and the task and response information into the display space, a vertical scroll bar was required (Figure 3.6). The scroll bar allows the user to move up and down the screen in order to view all the information presented. It does not alter the display of the static side-by-side maps, except to position them vertically on or off the screen. Both maps were visible at all times as displayed in Figure 3.6. A functional example of the side-by-side tool can be found at http://www.geog.ubc.ca/~antle/dissertation/test1/showa.htm.

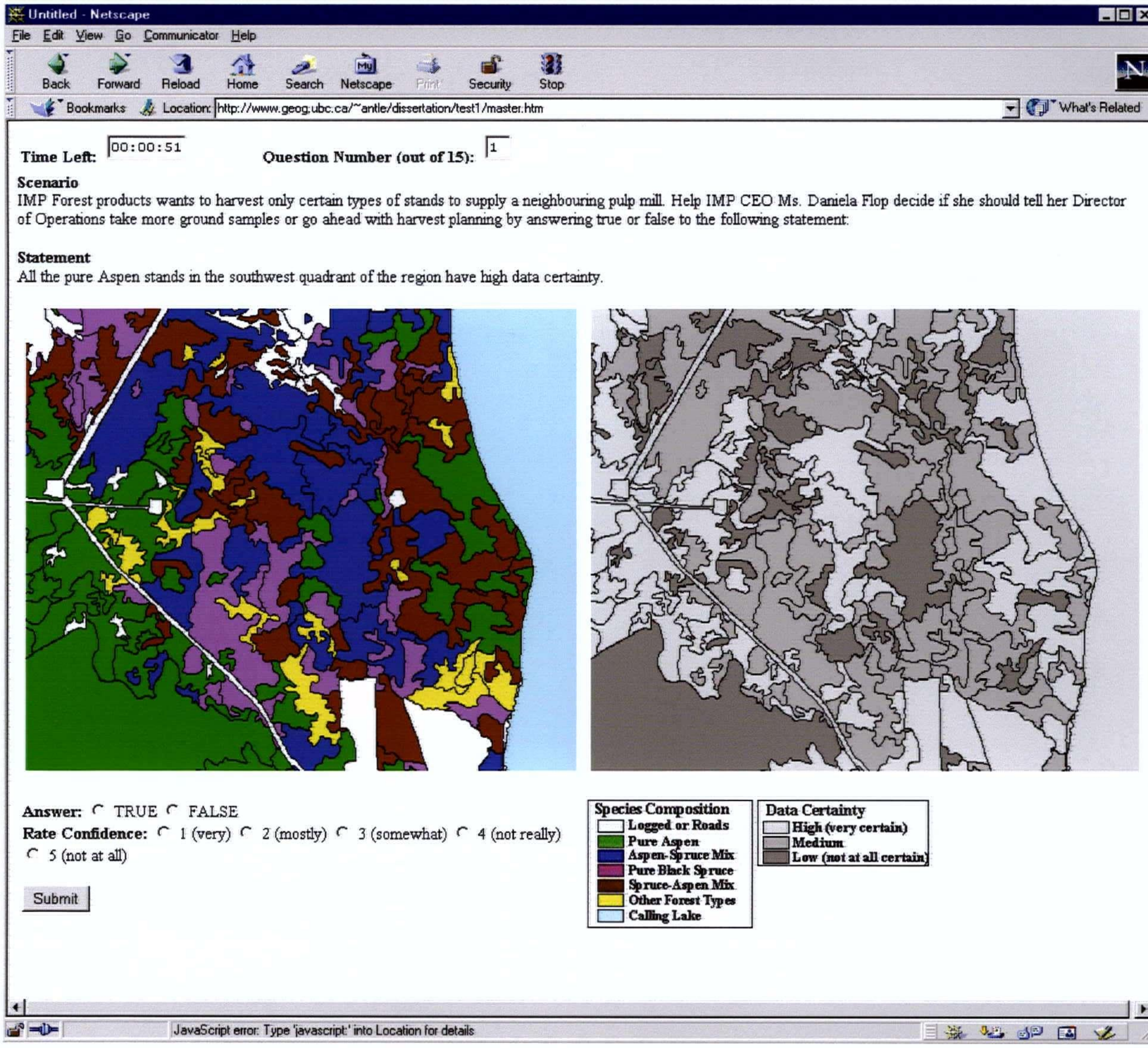

Figure 3.6. Side-by-side map display.

Tool 2: Toggle

A sequenced display (Slocum et al. 1990), in which the user controls the map view, can be created using a clickable button widget that controls the display (Figure 3.7). The user can alternately display each of the two maps, and can the toggle back and forth between the two maps. If the button is programmed to be responsive enough, the user can quickly toggle between maps creating an after-glow effect that may aid in the interpretation of the maps. The toggle display is minimally interactive, in that the user's only interaction is to click the toggle button. Other than that, the user has no control over map presentation, composition or information. A

functional example of the toggle tool can be found at
http://www.geog.ubc.ca/~antle/dissertation/test2/showa.htm.

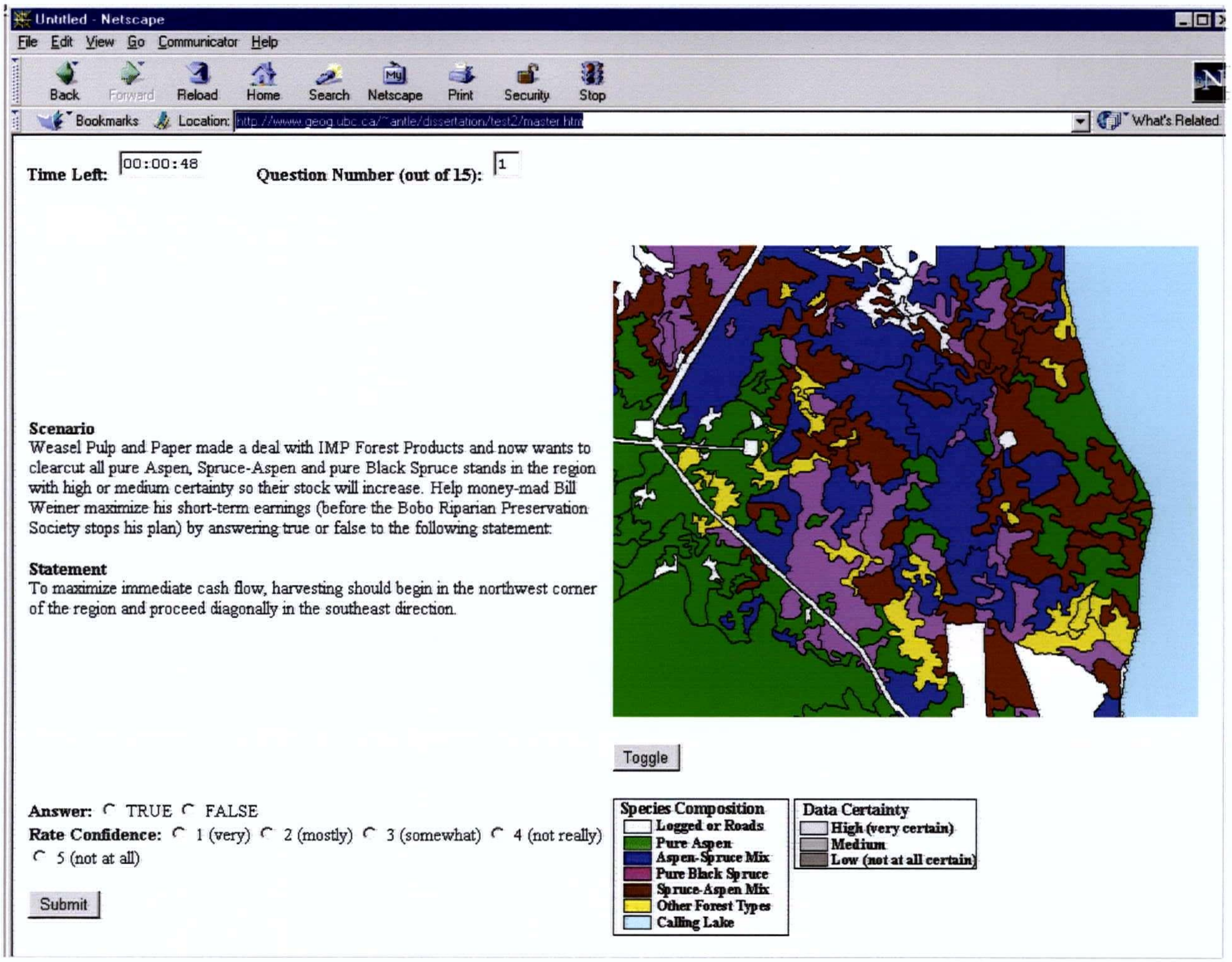

Figure 3.7 Toggle display tool.

Tool 3: Merger

For thematic data that has reliability information available, MacEachren et al. (1993) suggest merging or overlaying the data with the metadata using symbolization to keep each component visually separate. The concept of overlaying two maps can be implemented using colour theory instead of symbolization to keep the thematic data distinct from the certainty data.

Based on HSL colour theory, separation can be achieved by using two of the three perceptual dimensions of colour: hue and saturation (Brewer 1994). While the term colour is often used synonymously with hue, the term colour actually describes the combination of hue, saturation and lightness. Hue is the characteristic of colour described by the colour names we use (e.g., red, green, blue), and corresponds to the dominant wavelength. Saturation refers to the amount of hue in a colour and is also called purity, intensity or colourfulness. A bright red has a red hue and a high saturation. Conversely, a duller red still has a red hue, but has a lower saturation. Whites, greys and blacks are neutral colours that have no saturation or hue (Brewer 1994). Lightness may also be referred to as intensity, value, luminance, brightness and darkness, and the meaning is self-explanatory.

Hue and saturation can be used to represent both thematic and certainty information simultaneously. Hue can be used to represent thematic category and saturation to represent certainty. See Figure 3.8 for an example of the desaturation of a pure green hue using the HSL colour model where H:hue = 80, S:saturation varies and L:lightness = 60. Thematic information (and hue) is inherently discrete; therefore distinct colours can be assigned to each forest species composition class. Colours (hue) were chosen to be distinct at three different saturation levels and were adjusted in the first pilot study (see section 4.6). Unlike thematic attributes, (un)certainty (and saturation) is continuous. As discussed above (section 3.3), certainty was classified into three categories: low, medium and high. Three saturation values (high/pure, medium and low) were then used to (inversely) represent these discrete certainty categories (low, medium, high). This assumes that the desaturated hue represents more uncertainty, or low certainty (i.e., high desaturation = low certainty), and visa versa. Cartographically, there is a convention to represent larger data values with darker values (as colours are desaturated, they appear greyer and darker on a computer screen) (McGranaghan 1994). However, McGranaghan (1988, 1989) found that some users consistently associate darker with lower values and 25% of users associated symbols whose darkness was most different from the background with more. The association of darker with less certain makes intuitive sense (i.e., a map area is greyer/darker if it is difficult to know or "see" what the "real" value is there). This premise was informally tested in the first and second pilot tests.

In practical terms, saturation or desaturation of a pure hue can be achieved by "merging" or overlaying grey with a distinct hue. See Figure 3.9 for sample of a desaturated colour map.

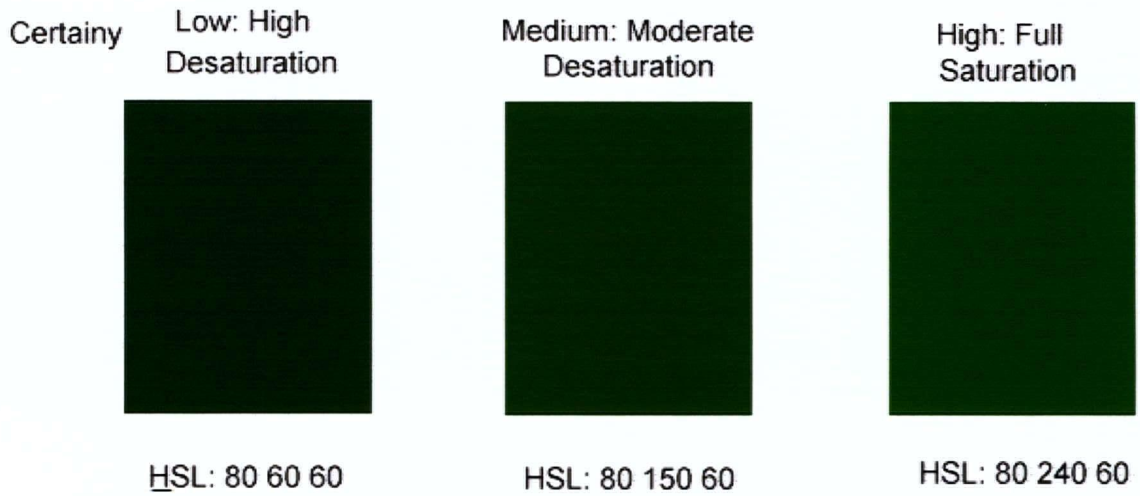

Figure 3.8 Desaturation of green.

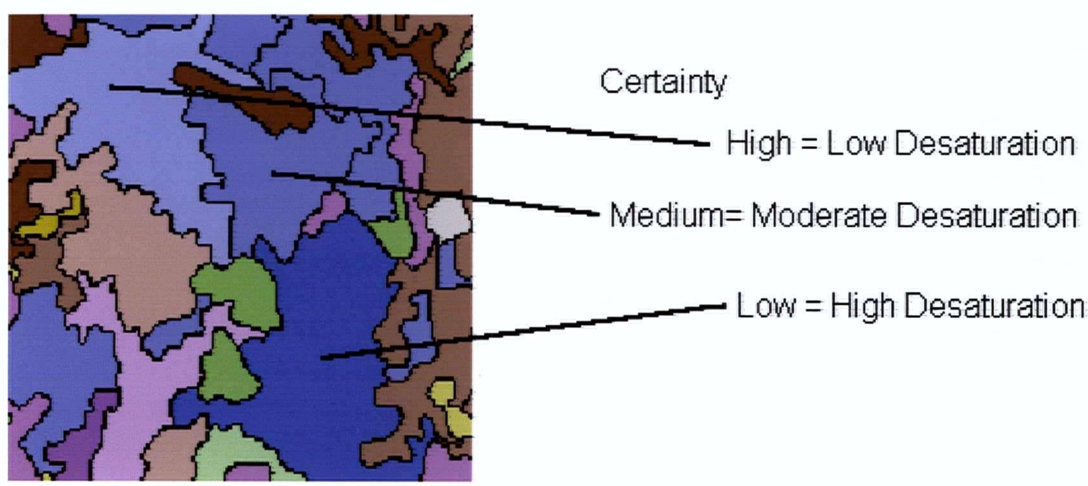

Figure 3.9 Sample of desaturated map.

The thematic and certainty maps can then be simultaneously displayed as a bivariate map in the same display space using an overlay function. One layer of pure hues represents the thematic data and the other layer of grey represents certainty. The display of these two maps can be interactively adjusted (or merged) in order to communicate only thematic information (only the coloured map is visible), only certainty information (only the grey map is visible) or both thematic and uncertainty information simultaneously. When the grey layer is overlaid on the hue layer, the overall effect is a variation of hue from pure to "greyed-out" (or de-saturated), depending on certainty values. Since blacks, whites and greys have zero saturation, this tool is not directly analogous to using hue and saturation together, but the overall visual effect is the same.

Using saturation to represent (un)certainty has been discussed in the literature (e.g., Schweizer and Goochild 1992) but not without some misgivings. Few studies have empirically examined the pairing of saturation and (un)certainty and the findings are not conclusive (e.g., Leitner and Buttenfield, 2000). The use of this pairing was based on the idea that perception of a “greyed out” or desaturated colour (hue) would be intuitively (or with some limited training) interpreted as “less sure” or uncertainty in the data. Brighter colours would be associated with more certainty. Anecdotal comments during the pilot studies confirmed this.

Howard and MacEachren (1996) suggested using slider widgets to allow the user to vary display characteristics. In their Reliability Visualization System (RVIS), a slider is used to successively fade out data values of a single hue as the slider is moved to higher and higher values. In this way, the user can focus in on a specific range of values interactively. An adaptation of this approach is to use the slider to control the simultaneous display of the species composition and certainty data using desaturation. At one extreme of the slider scale only thematic information (represented by different hues) is visible (Figure 3.10). At the other extreme only certainty values (represented by grey levels) are visible (Figure 3.11). As the user moves the slider from one extreme to the other, the hue and grey scale images are merged, (hue is desaturated) so that both thematic and uncertainty classes are visible to various degrees at once (Figure 3.12). In Figure 3.12, the user has positioned the slider (the rectangular object on the bar beneath the map) about two thirds of the way (from the left) along the bar in order to view the partly desaturated map. In this way, the user can view either data or uncertainty in their pure forms, or some combination of the two (Figure 3.13). A functional example of the merger tool can be found at http://www.geog.ubc.ca/~antle/dissertation/test3/showa.htm.

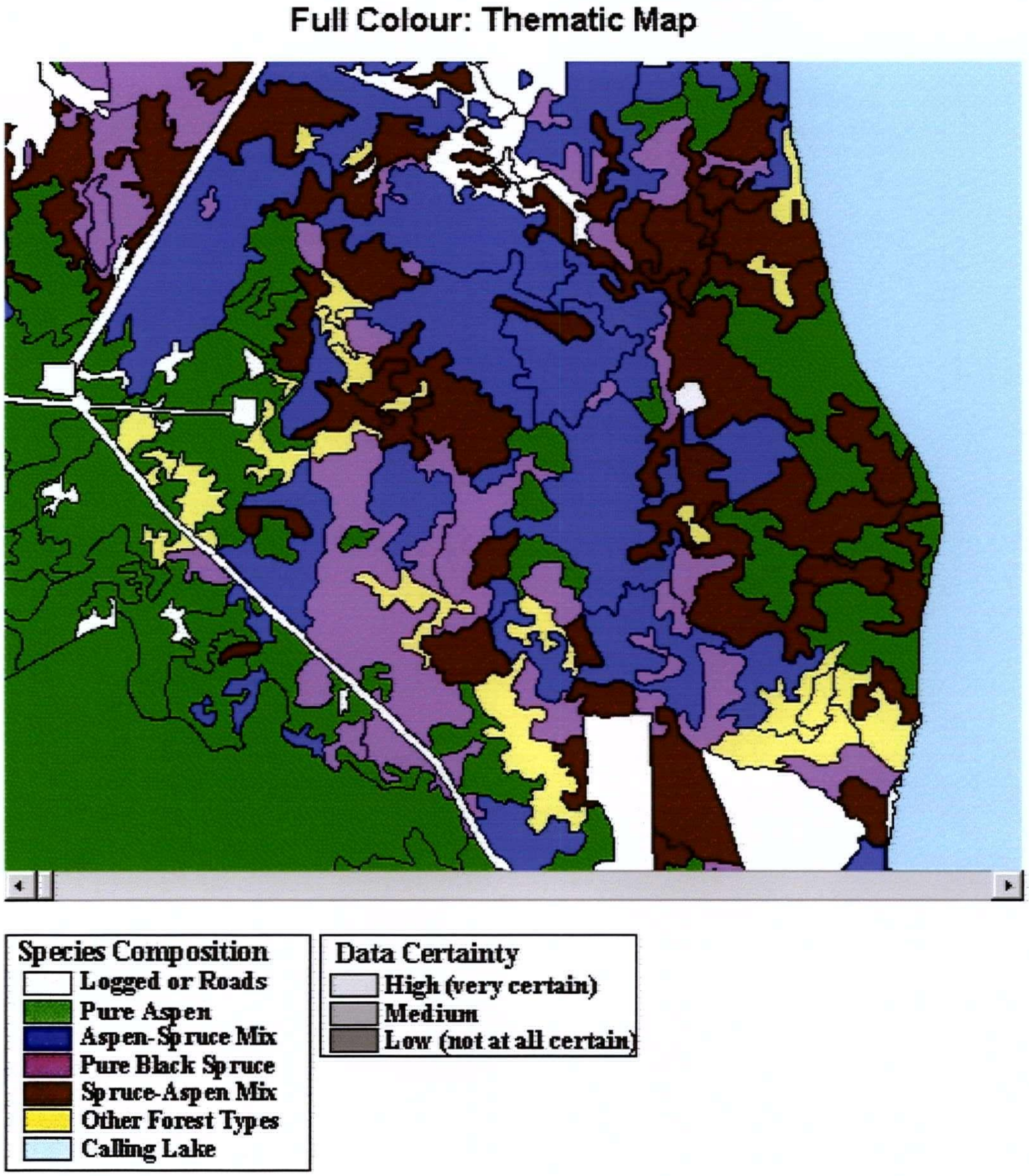

Figure 3.10 Full colour: thematic map.

Full Desaturation: Certainty Map

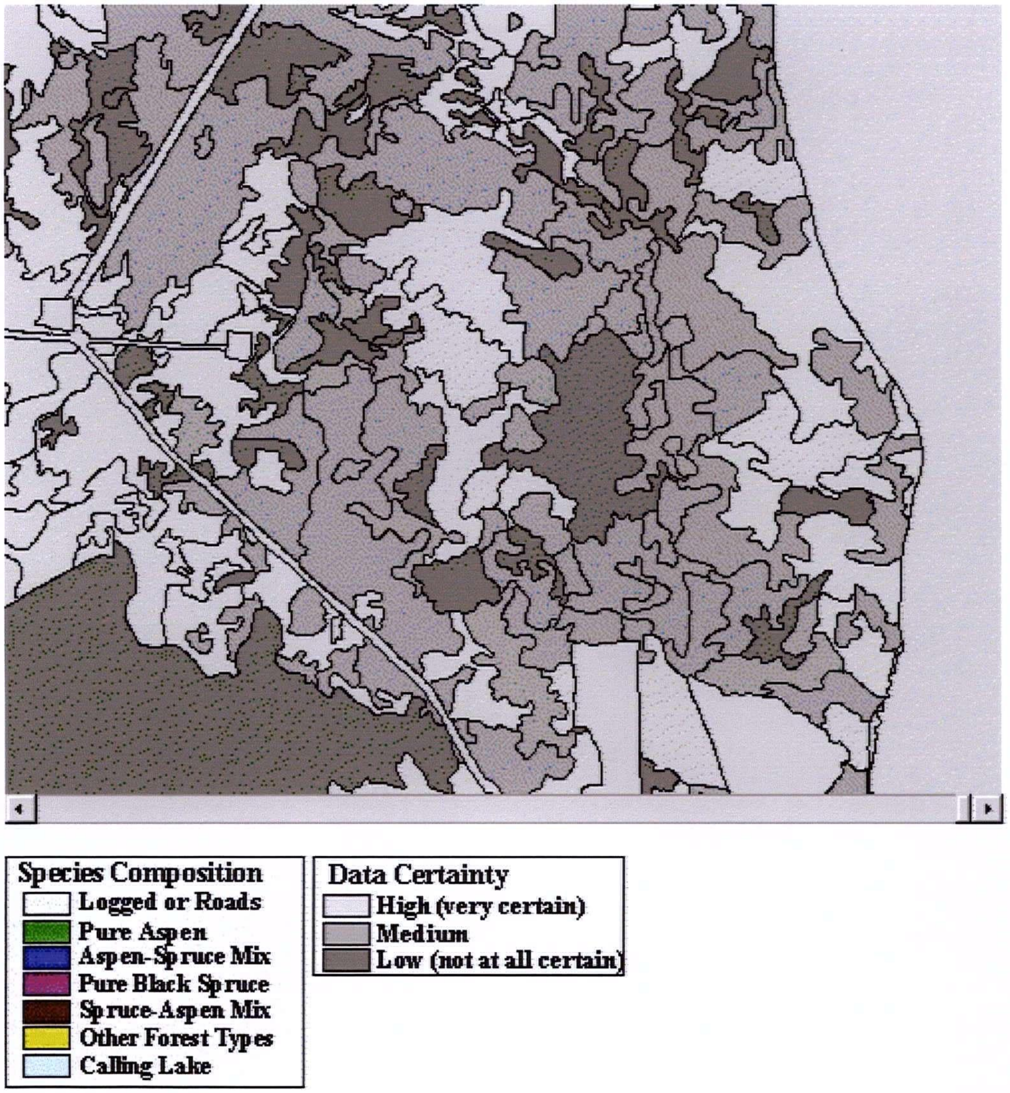

Figure 3.11 Full desaturation: certainty map.

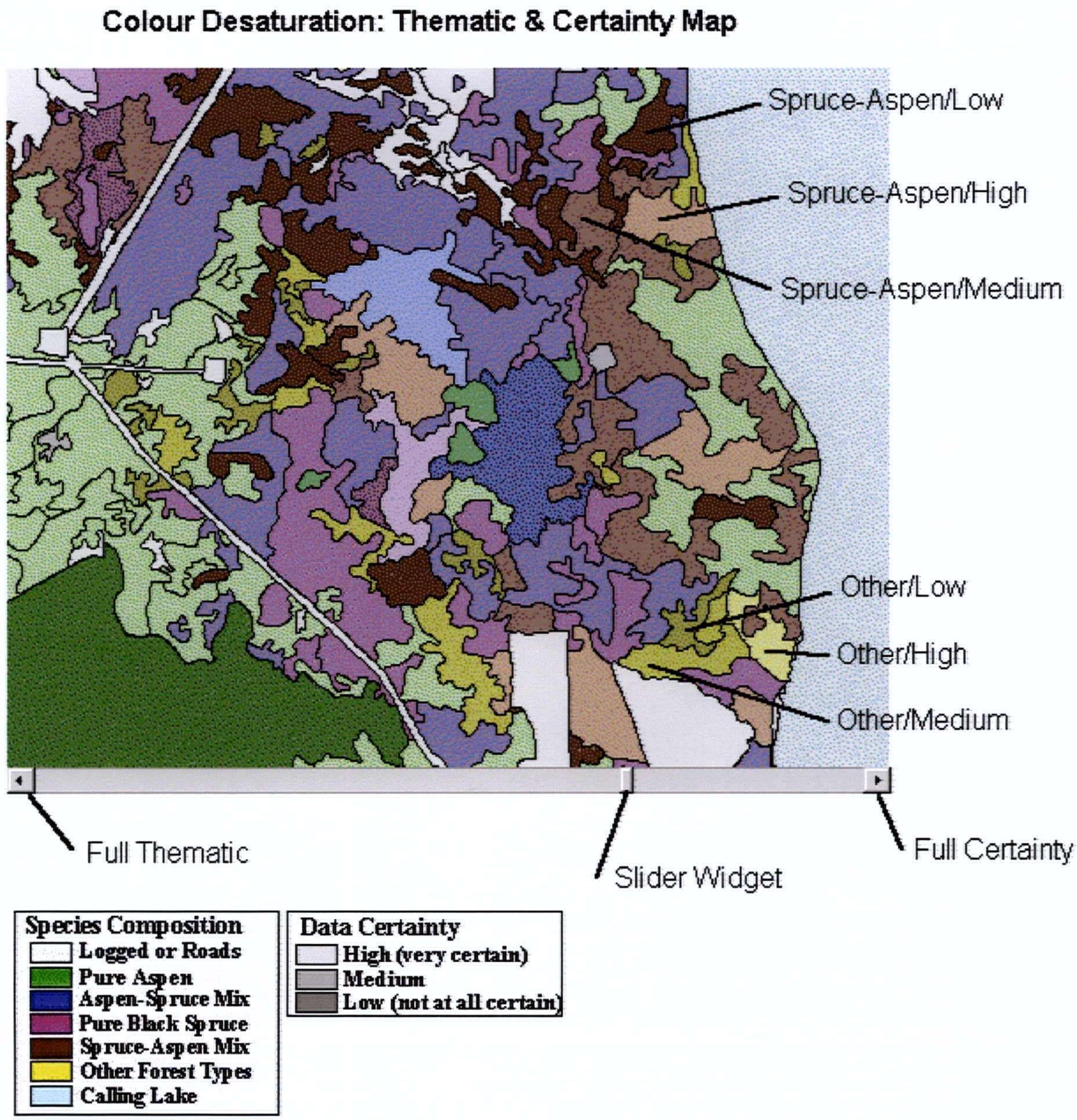

Figure 3.12 Colour desaturation: thematic and certainty map.

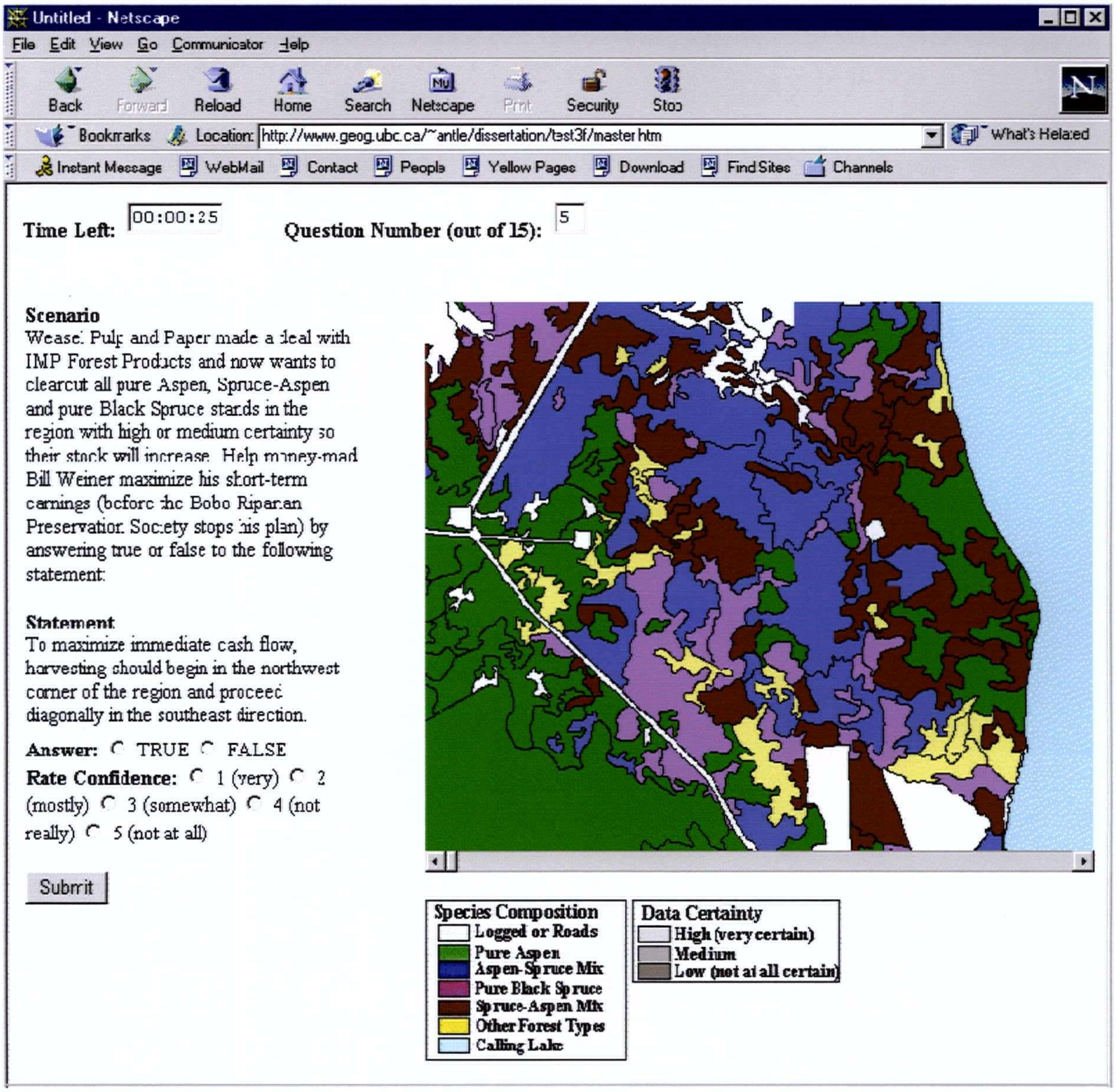

Figure 3.13 Merger tool.

Tool 4: Hypermap-Merger

Hypermaps are increasing in popularity as a tool of displaying attribute information in an online mapping environment. A hypermap is an interactive map in which the user controls the display of attribute information by clicking the mouse on a sub-region of the image. Attribute information (e.g., thematic class) is displayed either directly overlaid onto the existing image, or in another display window in response to the mouse position (see Figure 3.14). Tool 4, the hypermap-merger, is a variant of tool 3, the merger tool. Hypermap functionality was added to the interactive overlaid bivariate display tool (i.e., the merger tool) to create the hypermap-merger tool. In addition to controlling the display using the slider, the user can click on a portion of the map to display the thematic class and certainty information in a display window (Figure 3.15). This form of hypermap used in combination with the bivariate slider map is the most highly interactive tool used in this study. The attribute information that is displayed in response to the user's query is taken from an associated attribute database, making this tool the most complex from interactivity, usability and technical implementation perspectives.

A functional example of the hypermap-merger tool can be found at http://www.geog.ubc.ca/~antle/dissertation/test4/showa.htm.

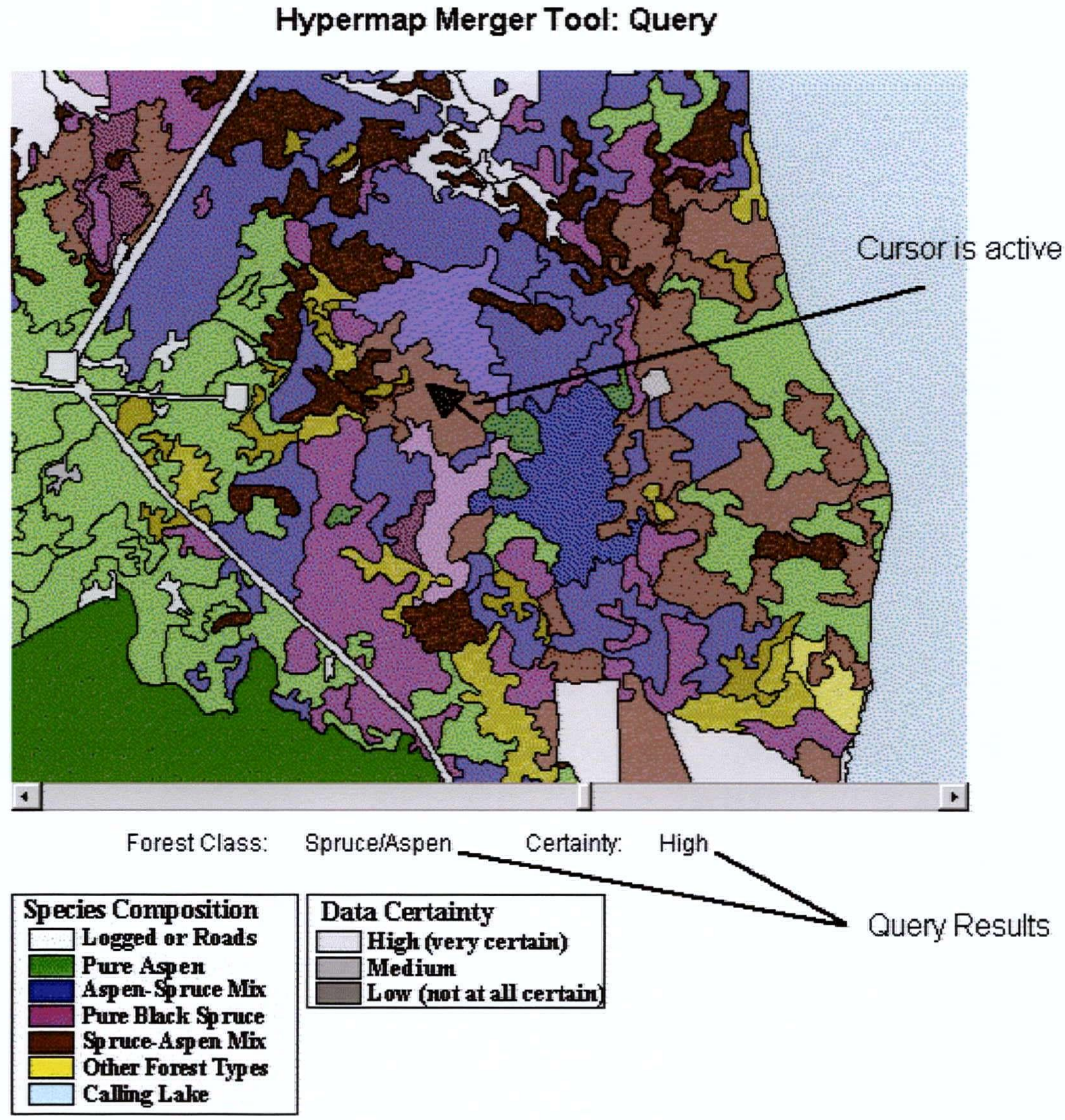

Figure 3.14 Hypermap-merger tool: query.

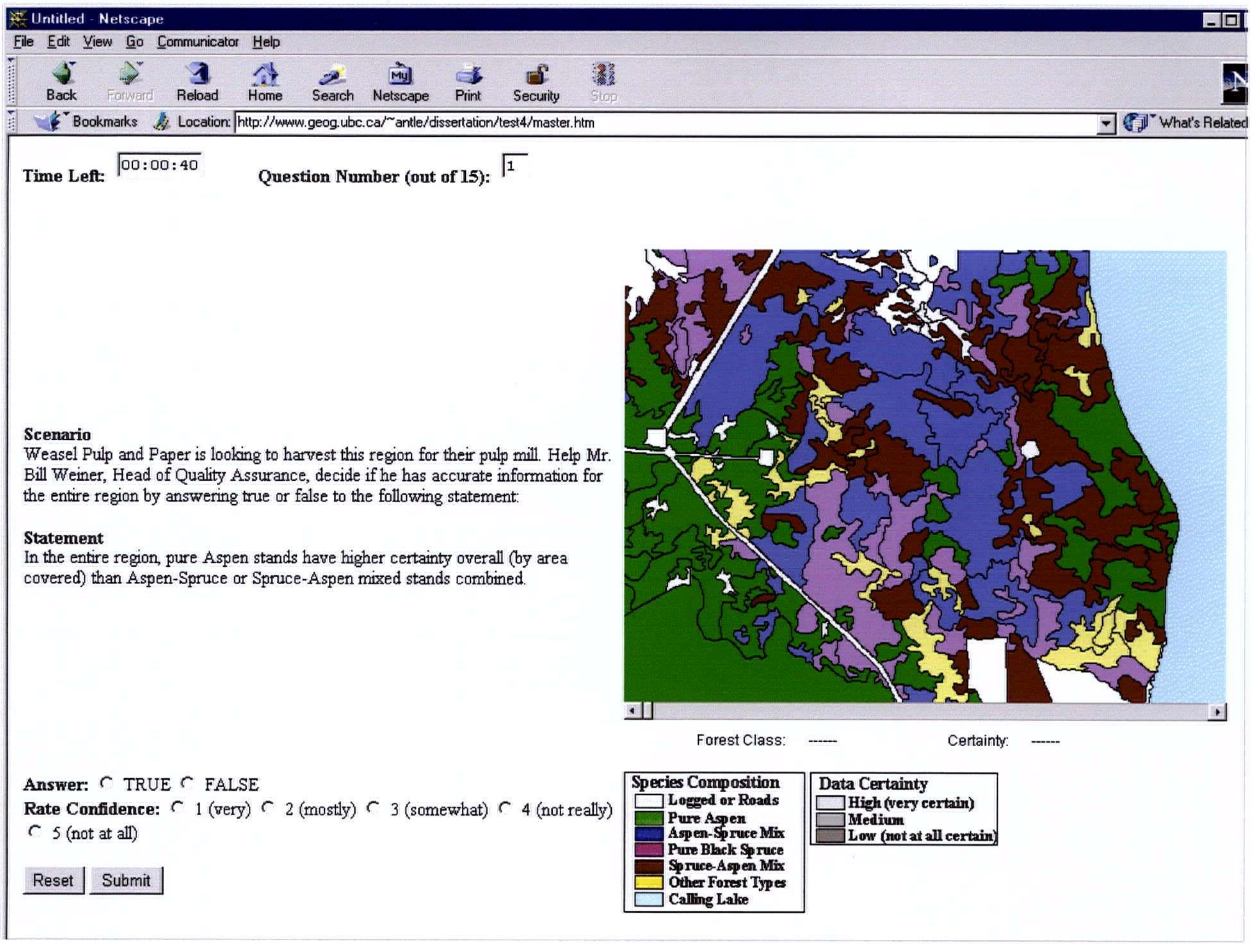

Figure 3.15 Hypermap-merger tool.

3.5 The Task Set

Thematic maps serve a variety of purposes, from static storehouses of spatial information to tools for sophisticated synthesis and analysis. The set of fifteen tasks used in this experiment were designed to replicate real-world tasks as much as possible from the field of forest resource management (see section 2.6). Since forest maps depicting thematic data and certainty information are used for many functions (e.g., inventory and monitoring of natural resources, planning, policy formation), a set tasks ranging from simple to complex was used. The fifteen tasks equally represent the task types (as described below). That is, the user was presented with five tasks from each of the three task types (see Appendix A for the final set of fifteen tasks used in the main experiment). The different task types were not identified to the user. The number of tasks was set at fifteen in order to ensure that an adequate number of tasks were given in each category, but not so many as to move beyond the attention span of users. As well, each lab session was one hour in length and imposed an arbitrary time limit on the total length of each session.

Based on the type of cognitive operation required and estimated difficulty, I categorized tasks as specific identification, general analysis, or integrative. Specific and general types can be classified (according Tan and Benbasat 1993) as data extraction tasks in that they involve retrieval and comparison of data (i.e., basic cognitive-perceptual tasks) and integrative tasks can be classified as decision activities that involve judgement, integration of information and inference (i.e., higher mental processes). It is important to note that the taxonomy I used includes both basic cognitive-perceptual and higher mental processes. Memory-based recall tasks (e.g., Mersey 1990) are not included in this taxonomy since exploratory visualization is more likely to include direct interaction tasks than recall-based tasks.

Several authors have suggested using true or false questions, rather than asking subjects to estimate numerical quantities or classifications, in order to avoid subjectivity and difficulty interpreting results (Nelson and Gilmartin 1996). Based on this, true or false questions were used. In order to avoid bias, the number of tasks whose answer was true was roughly equally to the number of tasks whose correct answer was false. Tasks involving data uncertainty were worded in terms of certainty to avoid double negatives, such as "low uncertainty," which may be confusing. Tasks were presented in a random order to reduce familiarity effects.

Task Type 1: Specific Identification

At the simplest level, subjects were asked to identify themes or classes for specific locations or regions. A user asked to solve a specific type task would likely scan a region looking for a sub-region (one or more adjacent polygons) that was uniformly classified. Once the area was found on both maps, the task was solved. Specific tasks concern what and where questions (worded as true or false questions). Sample true or false questions include:

- The central area of Aspen-Spruce mix has non-uniform data certainty.
- Just north of the centre of the region there is a single stand of Spruce-Aspen mix that has low certainty surrounded by stands of Aspen-Spruce mix.

Note that in order to ensure that wording such as "the central area" and "just north of" were not confusing or ambiguous, accuracy response data was analysed in the second pilot to look for an unusual number of wrong answers associated with these questions. Responses fell within the normal distribution expected for each task type.

Task Type 2: General Analysis

General analysis tasks are more complex tasks and include determining spatial patterns and correlations for the entire map. A user asked to solve a general analysis type task would likely repeatedly scan a region comparing several sub-regions (either adjacent or disparate) until the correct combination of sub-regions (polygons) was found. Sample true or false questions include:

- The largest contiguous area of pure Black Spruce stands with uniform certainty is located in the northwest corner of the region.
- All of the existing logged areas have high data certainty.

Task Type 3: Integrative

Integrative tasks are highly complex and involve picking out key information from the contextual information (presented through the text-base scenario information – see below) and integrating that with spatial thematic and certainty information to solve the task. A user asked to solve an integrative task type would likely repeat a pattern of scanning a region looking for a sub-region and comparing that to a set of characteristics taken from the contextual scenario and true of false question information (and possibly re-reading the textual information in each iteration) in order to make a decision or inference. A sample of an integrated task with requisite contextual information presented in the scenario and the associated true or false questions is:

> IMP Forest Products has now bought a Tree Farm License for the entire region. In order to maximize long-term income, the Director of Operations suggests a silvicultural treatment that includes thinning all pure Aspen stands now, in order to create income flow, and then to fully harvest these stands later for additional income. In order to minimize thinning costs, he wants to be sure to only thin marketable stands, and so will ground check stands with low certainty. Help IMP CEO Ms. Daniela Flop decide if she should follow this silvicultural treatment plan by answering true or false to the following statement:
>
> There is a sizable area that should be ground checked, after which most of the thinning should occur south and west of the

> logging road that runs through the southwest quadrant of the region and along the western edge of Calling Lake.

Note that terms such as "silvicultural treatment" and "thinning" were assumed to have been previously been encountered by students in the University of British Columbia's geography and forestry programs (i.e., the majority of students who participated in this study). Again, students did not seem to have undo difficulty with questions containing these types of forestry-specific terms.

It was initially assumed that the three task types outlined above increase in complexity or cognitively difficultly according to roughly equal amounts. While it is difficult to quantify task complexity, what is important is that the three types of tasks presented are distinct and involve different strategies that may or may not lend themselves to different tools. It is safe to assume that specific tasks are simpler than general or integrative tasks. The distinction between general and integrative is not as clear. It was also assumed that more difficult tasks would require more time to complete. In order to avoid the problem of students taking undue lengths of time to complete each task and in order to remain within the time limit of a lab session, time limits were set per task type. The allowable response time was determined based on the average amount of time students took to complete each task type in the pilot studies (as described in section 4.8).

Context and Motivation

To mimic real world tasks as much as possible, each task was presented in a context. For example, a forestry company plans to build a pulp mill and wants to locate it near large aspen stands. To provide user motivation, these tasks were embedded in a story – a land use conflict – complete with likeable and not so likeable characters. It was assumed that students from the University of British Columbia's geography and forestry programs were generally aware of the issues presented through the story (e.g., environmental conservation, site planning). The element of humour was injected to provide an engagement factor and to keep students interested in the tasks as they proceeded. Ongoing work (e.g., www.brainium.com) with younger students in online environments has shown that the use of a narrative theme increases user motivation and performance over more traditional presentation methods. (Cash for right answers would have worked too, but was untenable to the Ethics Review committee). For example,

> The Bobo Riparian Preservation Society is planning to lobby for a protected riparian area for the rare Bobo salamander. Help

Dr. Jessie Jones decide if the lake in this region meets the habitat requirements for the Bobo salamander by answering true or false to the following statement:

The largest distribution of high certainty pure Black Spruce in the region is along the western shore of Calling Lake.

and

Weasel Pulp and Paper is questioning the accuracy of the data they received from the Bobo Riparian Preservation Society. Help Mr. Bill Weiner, Head of Quality Assurance for Weasel Pulp and Paper, decide if he can trust the environmentalists' data by answering true or false to the following statement:

Just north of the center of the region, there is a single stand of Spruce-Aspen mix that has low certainty surrounded by stands of Aspen-Spruce mix.

Each task was presented with a preliminary scenario, followed by a true or false question. In specific and general tasks the scenario provides context but no crucial information needed to solve the task. In integrative tasks, the scenario contains information that must be integrated with the information in the true or false question and the information contained in the map to be answered accurately.

In answering tasks, subjects were instructed to assume that north is toward the top of the screen, that the topography of the study area was flat, that the stands were mature (and therefore harvestable), that the species composition information given was for the canopy, and that the whole area had the same fire history. All tasks require the user to get information from both the species composition map and the certainty map.

3.6 Measuring Map Use Effectiveness

As I discussed earlier in Chapter Two, the response time and accuracy with which a task is performed are traditionally used to measure map use effectiveness (Mersey 1990, McGranaghan 1996, Nelson and Gilmartin 1996). The use of response time is problematic, since many experiments (e.g., McGranaghan 1988 &1989) that utilize response time as a measure of performance fail to distinguish whether response time differences were attributable to map design or to individuals' cognitive strategies for interpreting them. In addition, accuracy and response time may be inversely related. Finally, response time (in terms of seconds to complete a task) is not a significant factor in most real-world decision-making environments. More important are accuracy and the confidence with which subjects make their decisions (Oz, Fedorowitz and Stapleton 1993). Unlike previous studies, this study uses accuracy and confidence to measure performance or online map use effectiveness. In order to avoid students taking undue time on tasks and to ensure that all questions would be completed within a lab session, an allowable response time was set based on average times as determined in the pilot studies.

Accuracy can be used as a measure of the cognitive difficulty of performing each task with each map display type. It is assumed that more cognitive difficulty (i.e., less accuracy) is associated with less appropriate designs for a particular task. In order to eliminate subjective evaluation of answers, and following the lead of Mersey (1990) and Slocum et al. (1990), tasks are presented as true or false questions. Accuracy is then recorded simply as right or wrong for each task and totaled as a percent correct for each task type.

In addition to answering true or false questions, subjects were also asked to rate their confidence in their decisions using a five-point scale:

- 1 (very confident)
- 2 (mostly confident)
- 3 (somewhat confident)
- 4 (not really confident)
- 5 (not at all confident/guess)

The design of scale used for rating opinions or attitudes is a complex issue that has received much attention in the social sciences literature, particularly in the choice of whether or not the scale should have a midpoint. The standard British social attitude scale is a five-point scale (Moser and Kalton 1979). The difficulty with not including a midpoint is that you may be forcing an artificial response – the student is mixed but they are forced to lean toward more or less confidence than they may be truly feeling.

Although time was controlled, it was still recorded since subjects could submit their answers when they were finished each task (as well as a time-out submission function). The time remaining was visible at all times, and they were encouraged to answer before time ran out, even if they were unsure. In this case, they were encouraged to rate their confidence as low (i.e., not really confident or not at all confident).

Since user preference also plays an important role in user performance, it was also examined in this study. In a post-test, subjects were shown all four tools, given some time to explore a set of maps with them, and then asked a series of open and closed format survey questions. Questions included rating the tools from most to least preferred (closed) and giving subjective feedback on tool design and usability (open) (see Appendix C for a sample feedback worksheet).

In Chapter Four, I outline the experimental methodology used to examine the relation among tasks, tools and map-user effectiveness (accuracy, confidence and preference).

CHAPTER 4 – EXPERIMENTAL DESIGN — METHODOLOGY

4.1 Overview

To test the three hypotheses outlined in the previous chapter, a map-use experiment was devised. The experiment examined user interaction with online spatial data and metadata. Spatial data took the form of a digital thematic map and an associated certainty map displayed in a Web environment. Subjects used one of four visualization tools to explore the maps and perform fifteen tasks. A two factor between-within subject design was used. The two factors are task type and visualization tool. There are three task types and four visualization tools (as outlined in the previous chapter).

Two pilot studies and one subsequent main experiment were conducted. Information from the pilot studies was used to fine-tune the design of the subsequent main experiment. Elements that were examined in the pilots included tool functionality and ease of use, technical performance and stability of the test environment, wording of the tasks and scenarios, interface design, the data submission and collection process, average task completion time and overall experimental flow (see more on this below). The second pilot served as an acid test to verify and validate the test environment and test process. The results from the second pilot were analyzed to look for trends and identify anomalies as discussed below.

The main experiment began with a brief orientation and training session (see Appendix B for a training script) in order to familiarize subjects with the online environment. Subjects then participated in two sequential activities. In the first activity, subjects were assigned one of four visualization tools, which they used to perform a series of fifteen time controlled map-use tasks. Subject scores on accuracy (true or false) and confidence (self-rated) were automatically recorded. In the second activity, subjects used all four tools to explore the maps, then rated the tools by preference and answered several questions comparing and contrasting the tools (see Appendix C for a sample worksheet). These two activities were followed by a short debriefing session explaining the purpose of the experiment and soliciting further feedback.

4.2 Subjects

Overall, 161 subjects participated in this experiment. There were 12 subject sessions in the first pilot study, 45 in the second and 104 in the main experiment. In the first pilot study, subjects were volunteers from fourth year geography or geography graduate students. Subjects were chosen due to their advanced level of understanding of digital spatial data and computer software applications since the goal of the first pilot was to refine design.

Since the goal of the second pilot was to be an acid test, subjects for the second pilot and main experiment were chosen based on the same criteria. These subjects were volunteers from Geography 315: Environmental Inventory and Classification, Geography 370: Introduction to GIS, Geography 372: Cartography and Geography 470: Advanced GIS, which are offered to Forestry and Geography undergraduate students. To promote ecological validity, these subjects were chosen to match the profile of the target audience for geographic visualization tools. Members of the target audience include: environmental interest group members, third party consultants and resource-based company analysts, managers and policy-setters. Subjects were also chosen to ensure the formation of a relatively homogeneous group, thus minimizing some of the effects of individual differences. For example, it was assumed that based on program and coursework, subjects would have a similar understanding of spatial data, uncertainty data, digital maps, related computer application software and online environments, and have similar experience concerning geographical information and environmental issues (statistical analysis of subjects did not refute this).

Subjects were randomly partitioned into four groups, with roughly equal numbers of males and females in each group in order to minimize any possible gender effects (which have been shown to be significant in visualization recall tasks in experimental cartography) (Gilmartin 1986). Subject groups contained an even distribution of students from Geography, Forestry and other disciplines. Prior to the session, all subjects were asked if they were colour blind. Three subjects that indicated that they were colour blind were removed from the study prior to the first activity. Three subjects whose native language was not English and who scored significantly lower (by more than two standard deviations) on accuracy scores than the mean for their group were removed.

4.3 Software and Equipment

Subject sessions took place in the Geography Undergraduate Computer Lab (Room 115) at the University of British Columbia. Most subjects were familiar with this lab and its equipment from their previous lab work in Geography courses. The lab consists of 22 Pentium

IBM-compatible personal computers running Windows 95 and Netscape Gold 3.01. User interaction was through the keyboard (subject profile information) and mouse (task responses). The PC monitors have VGA 1064 x 768 pixel resolution and were set to 1024 x 768 resolution. Screen colours were set to 256 in order to ensure consistent colours across all machines and sessions. An LCD overhead projector, connected to a PC workstation, was used for the group training sessions.

Since no commercial GIS available at the time of this study had the capabilities required, a stand-alone online prototype environment was created using HTML, JavaScript, Java and CGI/PERL scripts. The digital maps were created with ARCVIEW (ESRI) and exported as bitmaps. HTML was used to create the display environment, including titles, tasks (text), legend graphics, radio buttons for true or false, and the five confidence ratings, and to position the maps. The side-by-side tool was created using straight HTML. The toggle tool was created with JavaScript. The merger and hypermap-merger tools were created using the Java object-oriented programming language. JavasScript was also used to control time and to randomly sequence the task set. CGI files created with PERL recorded and processed user profile information, tasks responses and to record actual time on task (versus maximum time limit allotted to that task).

4.4 The Test Environment

The online test environment was distributed and displayed through Internet technology. An interactive form was used to collect subject profile information (see Figure 4.1). The form was followed by a series of fifteen Web pages, one for each of the fifteen tasks. Task pages were presented in random order. Each task Web page was created with the same layout. The layout consisted of a top frame containing a timer (time remaining per task) and question counter (e.g., question 5/15). The main window contained the integrated map display and visualization tool, legend, title, scenario, task, true or false radio buttons, one through five confidence rating and a submit button. Answers were submitted automatically when the timer ran out, or by clicking the *submit* button at any time. The only differentiating factor between pages was the content of the scenario and task. The basic page layout used for the study is shown below in Figure 4.2.

In preparation for the first pilot, the environment was tested repeatedly (by myself and several acquaintances) until major quirks were ironed out. Most problems involved backend programming and interface implementation problems.

4.5 Assumptions

This study assumes that the affect of using different visualization tools for map-use tasks can be isolated. Since data representation (i.e., maps), user interface, and visualization tool all influence map use, representation and interface were kept the same across all tool groups. The thematic map was coded with seven visually distinct colours (which were also distinct at three grey scale or saturation levels) that were easily distinguishable (verified in the first pilot study).

The Netscape browser interface was familiar to all subjects and is a relatively simple interface when evaluated based on user interface literature (e.g., Nielson 1994, Marcus 1995). For example, with the location and tool bars minimized, the entire content window is taken up by the experiment display. Informational features (e.g., timer and question counter) were displayed prominently at the top of the screen. Subject responses were recorded using clickable radio buttons located directly under the task question. The *submit* button was located to the right of the radio buttons, encouraging a flow sequence: answer true or false, rate confidence and submit. In addition, if the subject did not click the submit button, responses were submitted automatically once the timer ran out.

Subject performance in this experiment reflects a blend many factors, including the subject's familiarity with online environments, digital maps, uncertainty data, and their past experience, knowledge, gender, native language, age, problem-solving abilities and motivation level. By choosing subjects with a similar academic history and past experience, the group was relatively homogenous in terms of familiarity with the subject matter and the environment of the study. In addition, information, including: department; program; program year; computer experience (number of years); Internet experience (years); and professional experience, were recorded as part of subject profiles (Figure 4.1) in order to determine if these factors had an effect on performance. Gender, age and native language were also recorded and analysed to determine if they were correlated to performance variables. Data analysis showed no significant correlations between these demographic variables and subject responses (see Chapter Five –

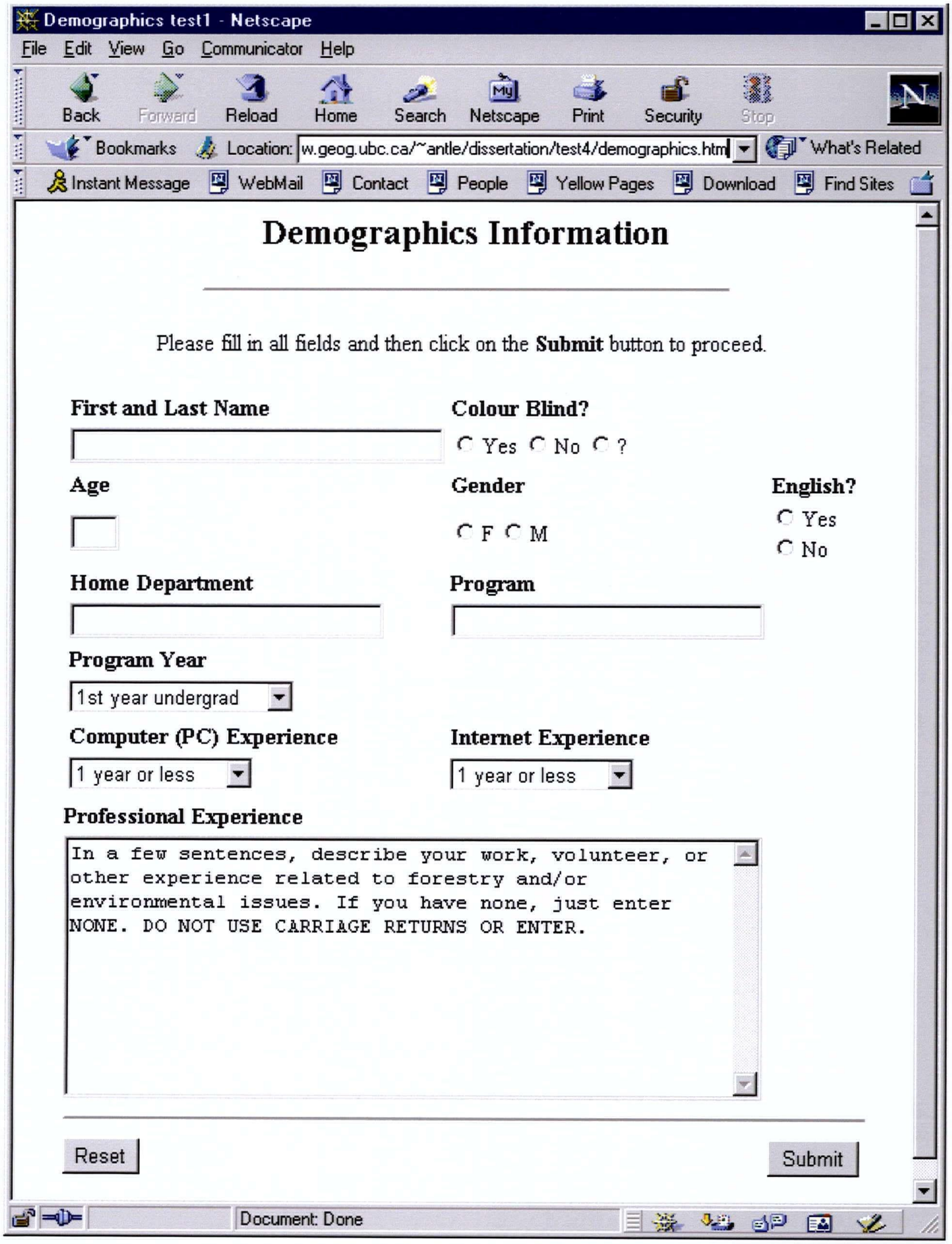

Figure 4.1 Subject profile information screen.

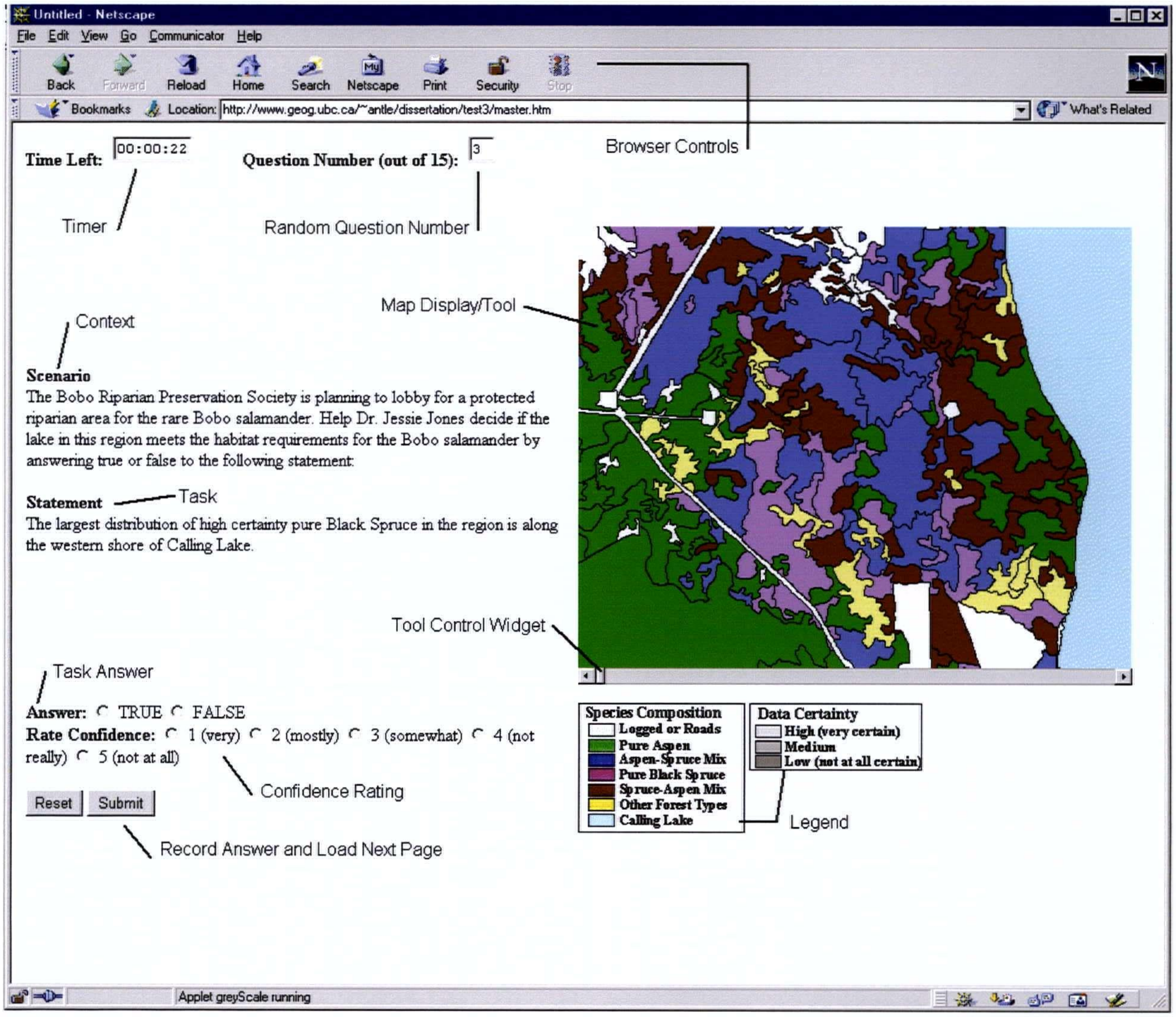

Figure 4.2 Basic screen layout.

Results).

It is also possible that subjects' differing cognitive strategies for solving map use tasks could have masked all other effects. While examining cognitive strategies is out of the scope of this experiment, it was hoped that looking at the correlation between tool preference and performance might identify a relation between these three factors. Subjects were asked to rate the tools from most to least preferred in the second activity. The results were analysed in relation to the tool group the subject was in, to determine if there was a relation between performance, assigned tool and preferred tool. A positive relation could indicate that the assigned tool reflected a preferred cognitive strategy.

Motivation is another factor that is difficult to control. In the main experiment, subject sessions were conducted as part of student lab sessions. Although students were not assigned a mark based on performance, they were encouraged to do their best and did receive either a pass or bonus mark for completing the session. In the second pilot students were given cookies once they had completed their session. In addition, by creating a context for the tasks (i.e., a playful story pitting do-gooder environmentalists against money-grubbing loggers), it was hoped that subject motivation would be enhanced. This approach is used successfully in children's online educational environments (e.g., www.brainium.com). Although times were chosen to ensure that most subjects had adequate time to answer, a prominently displayed countdown timer created a sense of urgency (not the same as motivation but related).

In terms of the order of task type, it was assumed that integrated tasks were more difficult than general tasks and thus would require more time. While this was found in the first pilot study, the results indicate that these two types of tasks may not be as different in terms of the cognitive difficulty (reflected in task time and accuracy) as originally thought. However, this assumption is not critical to the experiment. In terms of generalizing this work to interactive visualization tool design, what is key is that different kinds of tasks were examined and that performance on these tasks with different tools was analysed, regardless of relation between task types or ordering of the task classification hierarchy.

4.6 Pilot Study One

The first pilot study was used to tune the experiment. Twelve subjects were chosen based on advanced experience with computer tools, digital map data and online map use. Subject sessions took place in a one-on-one environment. Subjects were oriented to the goals of the first pilot and experiment in general. In addition to performing the tasks, subjects were asked a series of questions, either verbally or on paper, to gather information on all aspects of experimental design and implementation. Feedback on question wording, display and tool usability, average time required to complete tasks, as well as preliminary accuracy and confidence scores, was collected from the pilot studies and used to fine-tune the subsequent experiment. Technical, functional and performance issues were also ironed out during and after the pilot studies.

Specifically, the goals of the first pilot study were as follows:

- Determine if the test environment performed as expected;

- Determine if the test environment was stable;
- Determine if there were any usability or technical problems with the implementation of the visualization tools;
- Fine-tune resolution, colours and saturation levels;
- Determine key points where the subjects would need training;
- Determine the appropriate time limits for each task;
- Determine if any of the tasks caused undue difficulty for subjects because of unclear or ambiguous wording of the tasks or instructions;
- Determine if any of the tasks were unduly easy or difficult.

Of special interest was the time required to complete each task type. Subjects were asked to perform each task quickly but not to race through them. It was expected that each task type would require a different amount of time, and that simple tasks would require less time than complex tasks. In addition, integrative tasks were expected to inherently require more time to complete since they require more reading than either specific or general tasks. Median times based on seven subjects were determined for each task type (Table 4.1) and tested with the remaining five subjects in pilot one. These times were slightly adjusted (increased by roughly 20%) and used in the second pilot with the aim of giving most subjects enough time to perform each task without giving them overly long times (i.e., given enough time, most subjects would perform most tasks correctly). These adjusted times were further tested in the second pilot study to determine if they gave at least 80% of subjects enough time to finish each task. The median times shown in Table 4.1 show the expected increase from specific to integrative task types.

Task Type	**Median Time (seconds)**	**Adjusted Time (seconds)**
Specific	25	30
General	40	45
Integrative	62	75

Table 4.1 Pilot one: median task times.

Accuracy scores were also examined in the first pilot to determine the distribution of correct and incorrect answers. Since a significant number of subjects got tasks 7 and 10 wrong and tasks 6, 12 and 14 correct, the tasks were modified (for final task set see Appendix A: Tasks). Tasks 7 and 10 were subsequently made easier and tasks 6, 12 and 14 made harder. Confidence scores were not examined in the first pilot study. Modifications based on the results of the first pilot included:

- Modified difficulty of tasks 6, 7, 10, 12 and 14;
- Modified wording for tasks 3, 4, 5, 7, 10, 13, 14, 15;
- Modified implementation of merger tools to improve real time performance (used LUT (look up tables) to decrease processing speed time);
- Modified timer (sporadically failed to count down);
- Modified page layout by putting timer and question counter in separate top frame;
- Modified PERL scripts to perform enhanced preprocessing (standardize coding of missing values, create comma delimited data and record timeouts);
- Adjusted saturation level;
- Modified positioning of content blocks;
- Created separate environment for the second activity (performance evaluation);
- Adjusted category hues to maximize distinction at all saturation levels;
- Created training scripts and sample interfaces for four tools;
- Created print subject feedback forms.

4.7 Pilot Study Two

The second pilot study was used as an acid test. Since the test environment and process ran smoothly, and preliminary statistical analysis showed expected trends and did not show any anomalies, the second pilot ended the pilot phase. Minor modifications were made to the environment and process as outlined below.

Forty-five subjects were chosen according to the same criteria as for the main experiment, as outlined above in section 4.2. Subjects were split into four groups and taken through the full test procedure: orientation and training; activity one: fifteen tasks; activity two: tool comparison; and debriefing as outlined above in section 4.1.

Specifically the goals of the second pilot were as follows:

- Determine if test environment performed as expected;
- Determine if the administration of the test ran smoothly;
- Determine if there were any further technical problems with the visualization tools;
- Determine if the training session needed any further modification;
- Perform preliminary analysis (frequency) to determine how many subjects didn't complete tasks in the allotted time;
- Perform preliminary analysis (means and distribution) to see if any individual tasks within a task type were unduly easy or difficult;
- Perform preliminary analysis (means) to see if accuracy differed across tools;
- Perform preliminary analysis (means) to see if accuracy differed across tasks;
- Perform preliminary analysis (means) to see if confidence differed across tools;
- Perform preliminary analysis (means) to see if confidence differed across tasks;
- Examine comments from activity two: tool comparison.

Minor changes based on the results of the second pilot were as follows:

- Converted map display area to a frame to force a one-time preload of the Java merger tools;
- Preloaded the whole test environment for all four tools (cached);
- Ensured that tool groups 1 and 2 had adequate experience using the merger tools since they were not trained on them initially;
- Modified page contents for training on second activity (removed scenario as it was distracting);
- Modified PERL scripts to look for and delete multiple submissions of same tasks (subjects were click happy!);
- Modified activity two: tool evaluation worksheet (See Appendix C for final version);
- Adjusted times (from adjusted times of 30, 45, 75 to final times of 35, 50, 80 seconds)

4.8 The Main Experiment

One hundred and four volunteer students registered in Geography 315, 370, 372 or 470 at the University of British Columbia took part in the main study. The test sessions were incorporated into their tutorials (optional participation). Subjects were split by gender and then randomly assigned to one of the four visualization tool groups, ensuring that equal numbers of males and females were in each group. The visualization tool was the only known factor varying between the four groups.

The experiment was integrated into a lab exercise for each course. Although the title and introduction to the labs for Geography 315, 370, 372 and 470 differed slightly, all information relevant to the experiment itself was identical (see Appendix D for sample lab handouts). The lab was titled *Online Forest Inventories* for Geography 315, *Internet Maps* for Geography 372 and 472, and *Data Visualization* for Geography 370. There were no significant differences in performance between the course groups.

I conducted all lab sessions. The lab was introduced by a brief orientation. Subjects were instructed to have pre-read the lab (they were given time after the training to do so, if they hadn't). Subjects were assigned to tool groups to ensure equal numbers of males and females and equal numbers of subjects from different disciplines were in each group. Other than these two factors, the assignment was random. The orientation was followed by a training session (see Appendix B for training script). The goal of the training was to familiarize subjects with the exact test environment (using sample content and maps) and with the process. The training included familiarization with forest species composition thematic maps and related (un)certainty information, basic features of the Netscape Web browser, instructions for using the display interface (e.g., how to select and submit answers) and instructions for using the assigned map visualization tool. A demonstration of the test environment and process was given using the LCD overhead projector. To ensure subjects understood how to use the assigned tool and to reduce the novelty effect of the toggle, merger and hypermap-merger tools, subjects were given a brief set of questions to answer using the assigned tool. Answers were checked and questions encouraged. Subjects were asked to take time to read the lab (if they hadn't) and ask further questions. Subjects could begin whenever they felt ready.

After the orientation and training session, subjects were asked to complete an online user profile questionnaire (Figure 4.1) in order to record demographic and background information. The categories on the questionnaire were explained in the training session. The question "English?" was explained as "Is your native language English?"

In the first activity, subjects were asked to perform a series of time-controlled map-use tasks. Subjects were sequentially presented fifteen map-use tasks (presented in a random order), each for a specified period of time (based on task type), and asked to answer true or false and rate their confidence in their answer. A complete listing of the tasks is included in Appendix A. A running clock was a visible part of the display screen to indicate the length of time allocated for each task and show the subject how much time had passed for a particular task (see Figure 4.2). The tasks were presented in random order to balance sequential learning and reduce affects across stimuli. Subject information, true or false answers, confidence ratings, and time were automatically recorded. Subjects were encouraged to answer each task rather than leave them blank, even if they were unsure of the answer. They were also encouraged to rate their confidence as "not at all confident" if they were unsure of their answer. Subjects were also instructed to assume that that north was towards to top of the display, that the topography of the study area was flat, that the stands were mature (and therefore harvestable), that the species composition information given was for the canopy, and that the whole area had the same fire history.

After the fifteen tasks were completed, subjects were shown all four visualization tools, given the chance to experiment with the tools to answer similar tasks, and then asked to rate their preferred tool and give comments on the design and usability of the tools. The preference (tool evaluation) worksheet was submitted in print form (see Appendix C). A five-minute debriefing and brainstorming session concluded the sessions.

The main experiment was designed to test the effectiveness of the four visualization tools and the effects of the three levels of task complexity, in terms of accuracy and confidence; the effects of subject preference, in terms of accuracy and confidence; and the effects of individual differences (e.g., age, gender, native language). In addition, interaction effects were tested visualization tool type and task complexity as outlined below.

4.9 Statistical Analysis

The total number of subjects before preprocessing was 104. Missing values can be dealt with by exclusion or by substitution with the mean value for that question, a value determined using multiple regression, or by the group mean. Of the total subject sessions, 7 sessions had no demographic data (used as the basic identifier) and were excluded, leaving 97 subjects. Three outliers (subjects who scored more than two standard deviations lower than the mean for their group and whose native language was not English) were also removed from the accuracy data, leaving 94 subjects with complete data. In addition, for each analysis, subjects also were also removed for that analysis if response data showed that there were missing values.

As stated above, the experimental design (hypothesis one and two) was a two-factor between-within subject design where both independent variables (task and tool type) are discrete, and the dependent variable is either accuracy (percentage of correct responses) or confidence (average value out of five), which are both continuous values. Note that to facilitate the comparison of accuracy and confidence, confidence scores were inverted so that larger values represented more confident. This recoding did not affect results other than making analyses simpler. The between-subjects factor was the visualization tool (side-by-side, toggle, merger, hypermap-merger), because a subset of the subjects was assigned to each of the four tools. The within-subjects factor was the task complexity type (specific, general, integrated), since all subjects were given the complete set of fifteen questions (five of each task type).

At first glance, traditional univariate analysis of variance (ANOVA) seems the appropriate tool to use to examine the relation between tool type, task and accuracy or confidence (hypotheses one and part of two) (Phillips 1996). However, ANOVA assumes that the data were randomly assigned to a group and that each cell in the design is independent. The first assumption is valid, since subjects were randomly assigned to groups. If each task is considered a factor, then this second assumption is unrealistic since each subject performs all of the tasks. If a subject tends to answer accurately one on task, they might be more likely to respond accurately on all tasks. The same could be said for confidence. A solution to this possible correlation between tasks is to perform a multivariate analysis of variance (MANOVA) (Slocum et al. 1990). In MANOVA, the accuracy responses can be considered a series of dependent variables, with the tool type the factor in the design. So, the first step in analysis was to determine whether the responses were independent (Bray and Maxwell 1985). The Bartlett-Box test of equality of covariance matrices can be used to determine if the responses are correlated (Slocum et al. 1990). For accuracy responses, the null hypothesis failed to be rejected and we can assume independence. For confidence responses, the null hypothesis was rejected, however, the equations used were robust since sample sizes for the four tool groups are roughly equal. Therefore, I relaxed this assumption (SPSS).

Although ANOVA is robust to departures from normality (Arney 1990), the data should be symmetric (SPSS). Histograms of the accuracy and confidence scores showed normal distributions. Other assumptions that were tested for include homogeneity (i.e., that the variances across all four tool groups are roughly equal) and sphericity (i.e., that the covariances are constant). Levene's test of equality of error variances was used to test for homogeneity (SPSS). Note that, in this study, sample sizes for the four tool groups are roughly equal so the homogeneity requirement can be relaxed if needed. Sphericity can be tested using Mauchly's sphericity test (SPSS). If the condition of sphericity is not met, a correction factor must be added and the results adjusted.

Since both hypotheses one and two involve a between-within effect, a two-way (repeated measures) balanced design ANOVA was first used to examine the main and interaction effects. Where no interaction effects exist, a one-way ANOVA can be used to examine the effect of a single independent variable on a dependent variable. A priori contrasts can be used for hypothesis testing on factors. After ANOVA has shown significance, post hoc tests (e.g., Tukey's honestly significantly difference test) can be used to evaluate differences between specific means (Arney 1990).

The relation between accuracy and confidence as outlined in hypothesis two were examined using the Pearson product moment correlation coefficient since accuracy and confidence responses were both tabulated as averages (i.e., continuous) values (Phillips 1996). Pearson's product moment correlation coefficient, r, is used to determine the measure of association for two continuous variables. If confidence had been left in its ordinal form, the Spearman rank-difference coefficient could have been used. Since the Spearman statistic is more conservative than Pearson, any result found using Pearson's statistic would have been verified with Spearman's. Before using Pearson coefficient, data should be scanned for outliers (recall that three were removed).

Preference (hypothesis three) was examined using frequency analysis and an independent-sample t-test. Frequency distributions were created from subjects' ratings of their most and least preferred tools. In order to perform a t-test, subjects using their preferred tool were coded one and subjects not using their preferred tool were coded zero. The independent-sample t-test compares the means of a continuous variable for two groups of cases (Arney 1990, Phillips 1996). The test was used to compare the accuracy and confidence means for the preferred versus non-preferred tool groups. This t-test requires that the assumptions of normality and equality of variances be met. Levene's test can be used for equality of variance (SPSS). The two groups were then compared in terms of mean accuracy and mean confidence scores.

In order to examine possible relations between demographic variables with two discrete values or categories, and accuracy, confidence or preference, the independent-sample t-test was again used. Demographic variables with two values included: gender (male and female) and native language English (yes, no). In order to compare accuracy and confidence means for each demographic variable with more than two discrete values, one-way ANOVA, as described above, was used. Variables included: age, home department, computer/internet experience (combined) and professional experience. Age was grouped into the following five-year intervals.

- Group 1: Under 20
- Group 2: 20-25
- Group 3:26-30
- Group 4: 30+

Home departments were grouped as follows:

- Group 1:Geography
- Group 2: Forestry
- Group 3: Arts and Economics
- Group 4: Biology, Earth Science, Science and Human Kinetics

Computer/Internet experience was grouped as follows:

- Group 1: Novice User: Under 2 years combined experience
- Group 2: Typical User: Between 2 and 5 years
- Group 3: Expert User: Over 5 years

Professional experience was grouped as follows:

- Group 1: No related experience
- Group 2: Limited related experience (e.g., tree planter)
- Group 3: Significant related experience (e.g., field work)

In Chapter Five – Results, I provide the complete set of statistical results, as well as the specific details of the statistical analyses used in the main part of this study. In Chapter Six – Discussion, I interpret these statistical results in order to examine the implications of these findings, discuss the generalization of them to the design of interactive visualization tools and compare them with the results of related studies.

CHAPTER 5 – RESULTS

5.1 Overview

To test the three hypotheses outlined in Chapter Three, the data collected in the main experiment was analyzed. This chapter details this analysis and presents the results obtained. Specifically, the effects of task complexity and visualization tool on subject performance—accuracy and confidence as outlined in hypotheses one and two—are examined. Examination of hypothesis two continues with analysis of the relation between accuracy and confidence. Then the relation between accuracy and tool preference, and confidence and tool preference, as outlined in hypothesis three, are presented. Finally, the effects of demographic variables on accuracy and confidence are determined.

Examination of the data for hypothesis one—accuracy—showed that task complexity had a significant effect on accuracy and that this differed significantly across the four tool groups as predicted. The exact nature of these effects is detailed below in section 5.2. Examination of the data for hypothesis two—confidence—showed that confidence was significantly correlated to accuracy across all tools and task types. It was also found that confidence decreased as task complexity increased (as expected) but the rate of decrease did not vary significantly across tool groups. Further details are described below in section 5.3. Examination of the data for hypothesis three—preference—showed no significant relation between user preference and accuracy or confidence. An overwhelming number of subjects preferred the hypermap-merger tool. Further details are described below in section 5.4. Analysis of demographic variables showed that the test subjects did form a relatively homogeneous group. None of the demographic variables examined (gender, native language, program, computer experience, professional experience) had significant effects.

5.2 Hypothesis One: Accuracy

As discussed previously, the subjects' ability to answer tasks accurately is one measure of map-use effectiveness. This section will examine how task complexity and the visualization tool used affected subjects' ability to perform fifteen map-use tasks accurately.

Predictions

The first hypothesis examined how subjects' response accuracy on map-use tasks was affected by task complexity and the visualization tool used. It was predicted that accuracy would decrease as the complexity of map-use task increased, and that the rate of decrease would vary depending on the visualization tool used.

Analysis

Preliminary analysis showed that the conditions of independence and normality were met, therefore, ANOVA was used rather than a multivariate method (see section 4.9). A 4 x 3 between-within (two-way) ANOVA based on the percentage of correct answers was conducted in order to examine how accuracy was affected by task complexity and visualization tool. The within-subjects factor was the task complexity (specific, general, integrated), since all subjects were given the complete set of fifteen questions (five of each task type). The between-subjects factor was the visualization tool (side-by-side, toggle, merger, hypermap-merger), because a subset of the subjects was assigned to each of the four tools. Levene's and Mauchly's test results showed that the conditions of homogeneity and sphericity were met. Therefore, no corrections were required to the ANOVA results (as discussed in section 4.9).

Preliminary Results: Main Effects

Task Complexity

Two-way ANOVA results indicated that task complexity had a significant effect (at an $\alpha = .0001$ level) on accuracy ($PR > F(6,180) = 9.65$, $p < 0.0001$) (as predicted). The group means for accuracy (see Table 5.1) showed that accuracy was highest for specific type tasks (72.2 percent correct) followed by integrated type tasks (63.5 percent correct) and general type tasks (61.1 percent correct). This finding was expected since specific identification tasks are cognitively simpler than either general pattern analysis tasks or integrated tasks. The similar means for general and integrated tasks could be a result of the longer time given for the integrated tasks that may or may not have been cognitively more difficult than general tasks (see section 3.5). The prediction that accuracy would decrease as the complexity of map-use task increased held.

Task complexity	Mean accuracy	Std. Error
1 – Specific	72.2%	.024
2 – General	61.1%	.023
3 – Integrated	63.5%	.020

Table 5.1 Mean accuracy (%) by task complexity.

Visualization Tool

ANOVA results also showed that visualization tool had a significant main effect on (at an $\alpha = .01$ level) accuracy ($PR > F(3,90) = 4.53$, $p = 0.005$). Since results also showed a significant interaction effect (as predicted), the interaction effect of visualization tool and task complexity is examined in tandem below.

Interaction Effects

It was predicted that task complexity and visualization tool would have a combined effect on accuracy (that the rate of decreased accuracy due to increased task complexity would vary depending on which visualization tool was used). ANOVA results showed exactly this. The effect of task complexity on accuracy differed significantly (at the $\alpha = .05$ level) across the four tool groups ($PR > F(6,228) = 2.21$, $p = .044$) .

The means for each tool group (Table 5.2) showed a similar overall pattern as the group means for task complexity (Table 5.1), with some departures (see Figure 5.1).

Tool / Task	Side-by-Side	Toggle	Merger	Hyper-map/Merger
Specific	59.1	60.8	83.9	81.5
General	55.0	60.8	60.0	67.2
Integrated	60.0	58.3	67.7	67.2

Table 5.2 Mean accuracy (%) by task complexity and tool type.

The means showed that the side-by-side tool subjects answered specific and integrated tasks more accurately (59.1% and 60.0%) than general tasks (55.0%). Overall, side-by-side tool subjects had lower accuracy on all tasks than the other tool groups (as expected).

The toggle tool subjects answered all tasks with similar accuracy (60.8%, 60.8% and 58.3%), with slightly higher results for specific and general tasks. The toggle tool subjects' scores for specific and integrated tasks were roughly equal to the side-by-side tool group's scores. However, the toggle group's scores for general tasks were much higher than the side-by-side tool group's. Why was this? For specific and integrated tasks, toggle group subjects probably used each map separately in much the same way as the side-by-side group. But for general spatial pattern identification tasks, they were able to use the flickering effect of toggling between the two maps to answer questions more accurately. This effect, a type of map sequencing, has been suggested by Slocum et al. (1990) to be effective for examining general spatial patterns (see section 6.3 for a thorough discussion). These results confirmed than the toggle tool (sequencing) was more effective than a side-by-side map display for identifying general spatial patterns.

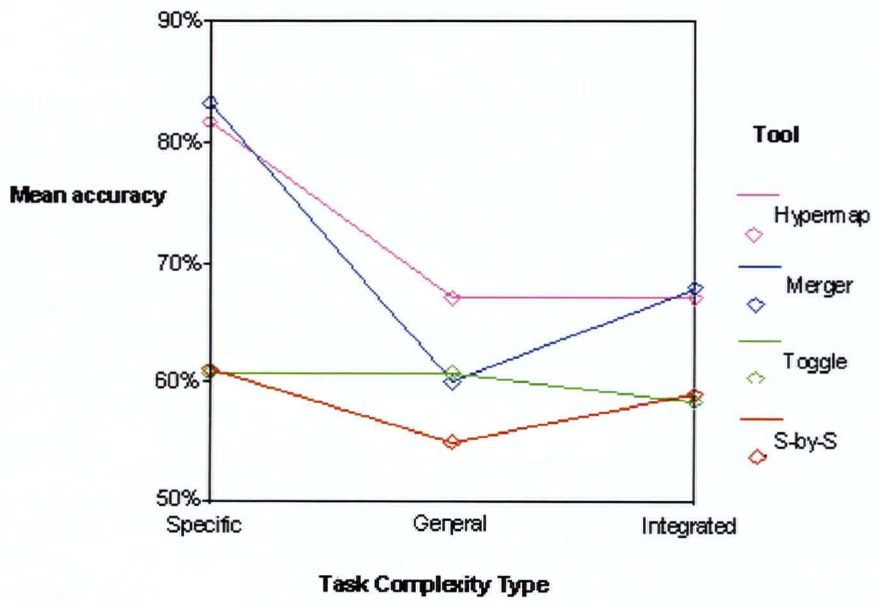

Figure 5.1 Mean accuracy by task complexity and tool.

The merger and hypermap-merger tool groups showed much higher accuracy for all tasks types, with one exception. For general tasks, the merger tool group's responses were about the same as the toggle tool groups'. Overall, the means for the merger and hypermap-merger tool groups showed that the effect of increased interactivity was increased accuracy.

How did the merger and hypermap-merger tools differ? For specific and integrated tasks, both tool groups showed similar accuracy. However, for general tasks, the hypermap-merger tool group far out-performed the merger group. This was interesting, since the hypermap component of this tool was geared at identification of specific attributes (through point and click querying), not general patterns. One explanation is that subjects repeatedly used the point and click aspect of the tool to help clarify patterns. This 'mad-mousing' behavior was exhibited by several subjects during the study. It is also possible that the merger tool was difficult to use without the added hypermap functionality as some of the post-test questionnaires indicated.

More detailed interaction effects are described below.

Simple Effects

The significant interaction effect indicated that task complexity and tool used had a combined effect on accuracy. Simple effect analysis was conducted in order to provide further information on the nature of these effects.

A simple effect analysis consisted of conducting repeated measures ANOVA's separately for each tool in order to examine the simple effects of task complexity on accuracy levels. Results from this analysis (see Table 5.3) shows that for the side-by-side and toggle tool groups, accuracy was not significantly different across the three levels of task complexity (($PR > F(2,38) = .707$, $p = .499$ and $PR > F(2,46) =.114$, $p = .892$) for side-by-side and toggle groups respectively). Subjects performed at approximately the same levels on all tasks with these two tools. The results for the merger and hypermap-merger tool groups showed that accuracy differed significantly across task complexity (($PR > F(2,48) = 14.21$, $p < .0001$ and $PR > F(2,48) = 5.45$, $p = .007$) for the merger and hypermap-merger groups respectively).

Tool	F	Significance	Differ?
Side-by-side	.707	.499	No
Toggle	.114	.892	No
Merger	14.21	.000	Yes
Hypermap	5.45	.007	Yes

Table 5.3 ANOVA results: accuracy differences across task complexity types for each tool.

Post hoc analysis using Tukey's honestly significant difference test (HSD) was conducted in order to examine how accuracy scores differed across complexity types for the merger and hypemap-merger tools (see section 4.9 for details on HSD). The HSD statistic was calculated to be .12 or 12 % at an α =.05 significance level. This indicates that means must differ by more than 12% for the difference to be significant. By comparing means for the merger tool, the HSD statistic showed that accuracy significantly differed between specific and general tasks (1 & 2) and between specific and integrated tasks (1 & 3), but not between general and integrated tasks (2 & 3). (See Figure 5.4).

Tasks	Means	Difference	HSD? (diff > 0.12)
1 & 2	.839 .600	.239	Yes
1 & 3	.839 .677	.162	Yes
2 & 3	.600 .677	.077	No

Table 5.4 Tukey's HSD for merger tool.

The HSD test for the hypermap-merger tool showed similar results. Note that the calculation of the HSD statistic is the same for the merger and hypermap-merger analyses since there were 25 subjects in both groups. Accuracy differed between specific and general tasks (1 &

2) and between specific and integrated tasks (1 & 3), but not between general and integrated tasks (2 & 3). (See Table 5.5).

Tasks	**Means**	**Difference**	**HSD?** (diff > 0.12)
1 & 2	.815 .672	.138	Yes
1 & 3	.815 .672	.138	Yes
2 & 3	.672 .672	.000	No

Table 5.5 Tukey's HSD for hypermap-merger tool.

The HSD findings in combination with the merger and hypermap-merger means showed that subjects are able to perform significantly better on specific tasks using the merger and hypermap-merger tools than on either general or integrated tasks. There was, however, a trend for subject to perform better on integrated tasks than general tasks using the merger tool (as seen in Figure 5.1).

Overall, the prediction that the rate of decrease in accuracy due to increased task complexity would vary depending on the tool type was seen most significantly for the merger and hypermap tools (as seen in Table 5.2). These two highly interactive tools showed better accuracy for all task complexity types. This effect was significant for specific type tasks. It is also interesting to examine how the effect of complexity on accuracy was influenced by tool type (the reverse scenario). One-way ANOVA analysis conducted separately for each task type showed that, for specific tasks, accuracy differed significantly across tool groups ($PR > F(3,96) = 9.01$, $p < .0001$) (see Table 5.6). Accuracy did not differ significantly across tools for general and integrated tasks. The tool means for general and integrated tasks (Table 5.2) showed a weak trend toward increased accuracy for the more interactive tools.

Tasks	F	Significance	Differ?
Specific	9.01	.000	Yes
General	1.10	.354	No
Integrated	1.41	.244	No

Table 5.6 One-way ANOVA results: accuracy differences across tools for each task complexity type.

Post hoc analysis using Tukey's honestly significant difference test (HSD) was again conducted to examine how accuracy scores differed across tools for the specific tasks (see section 4.9). The HSD statistic was calculated at an $\alpha = .05$ significance level for each pairing (since the number of subjects between groups varied). The merger (3) and hypermap (4) tools both had significantly higher accuracy than the side-by-side (1) or toggle (2) tools (see Table 5.7). There was no significant difference between the side-by-side or toggle tools (1 & 2), or between the merger and hypermap tools (3 & 4).

Tools	Means	Difference	HSD?
1 & 2	.591 .608	.017	No
1 & 3	.591 .839	.248	Yes
1 & 4	.591 .815	.224	Yes
2 & 3	.608 .839	.231	Yes
2 & 4	.608 .815	.207	Yes
3 & 4	.839 .815	.024	No

Table 5.7 Tukey's HSD for specific tasks.

Summary

The results show that the map-use task complexity had a significant effect on subjects' accuracy and that the nature of these effects varied across different tool groups as predicted. Subjects performed most accurately on specific tasks, compared to general or integrated tasks. Accuracy results were similar for integrated and general tasks. The ability to perform these three types of tasks in an accurate manner was significantly affected by the tool used. Accuracy was about the same for the side-by-side and toggle tools, although the toggle tool subjects showed optimum performance on general pattern analysis tasks. The merger and hypermap-merger tool groups showed a weak trend toward higher accuracy on general and integrated tasks. Accuracy was significantly higher on specific tasks for merger and hypermap tool subjects.

5.3 Hypothesis Two: Confidence

As discussed previously, the factor confidence refers to subjects' confidence in their decisions and has been shown to be an important factor in determining the quality of decisions in decision-making. In this study, the subject's confidence in their decisions was used as another measure of map-use effectiveness. This section will examine the relation between accuracy and confidence and the effects of map-use task complexity and visualization tool on the subjects' confidence when performing tasks.

Predictions

The second hypothesis examines the correlation between accuracy and confidence, and suggests how subjects' confidence in their decisions on map-use tasks was affected by task complexity and visualization tool used. It was predicted that confidence is directly related to accuracy. It was also predicted that confidence would decrease as the complexity of map-use task increased, and that the rate of decrease would vary depending on the visualization tool used.

Analysis

The bivariate relation between confidence and accuracy can be examined using Pearson's correlation coefficient (see section 4.9). A 4x3 between-within ANOVA based on the percentage of correct answers was conducted to examine how confidence was affected by task complexity and visualization tool. The within-subjects factor was the task complexity (specific, general, integrated), since all subjects were given the complete set of fifteen questions (five of each task type). The between-subjects factor was visualization tool (side-by-side, toggle, merger, hypermap-merger), because a subset of the subjects was assigned to each of the four tools. Preliminary analysis using Bartlett's Box test (for independence) showed that the variance-covariance matrices were not equal across all groups. However, this violation can be ignored since the test is robust under conditions of equal numbers of subjects in each group (which was met) (see section 4.9 for details). Mauchly's test results showed that the condition of sphericity was not met either. The Greenhouse-Geisser and Huynh-Feldt coefficients (.937 and .998 respectively) suggest minor corrections are required when calculating the F value (SPSS). However, since the results of the ANOVA were highly significant, these corrections had little effect.

Note that confidence values were recoded so that the highest confidence corresponded to a value of five and the lowest to a value of one in order to make comparison with accuracy data easier to follow. This recoding did not in any way affect the results.

Correlation Results

Pearson's correlation analysis for all tool groups over all tasks showed that confidence is positively correlated to accuracy at an $\alpha = 0.01$ level (two-tailed) and has a value of $r = .53$. While this correlation value is not high (i.e., close to +1), it does explain some of the variance seen in the dependent variable (as expected). However, given the value of this correlation (i.e., $r = .53$), it is likely that other factors influenced subjects' confidence rating.

How does the relation between subject confidence and accuracy vary across tool or tasks types? Across tool groups, correlation analysis shows that confidence and accuracy are correlated for the side-by-side (at the 0.05 level), toggle (at the 0.01 level) and hypermap-merger (at the 0.05 level) tools but not for the merger tool (see Table 5.8). Again, the correlations are not high (i.e., close to a value of +1), but some relation is evident. Likely, there are other factors at work. The relation between confidence and accuracy is highest for the toggle tool.

Correlation analysis across task complexity levels shows that the relation between confidence and accuracy is strongest relation for specific tasks (Table 5.9). For specific tasks,

Tool Type	Pearson Coefficient	Significant?
1 – Side-by-side	+.51*	Yes
2 – Toggle	+.57**	Yes
3 – Merger	+.20	No
4 – Hmap-merger	+.45*	Yes

* correlation is significant at the 0.05 level (two-tailed)
** correlation is significant at the 0.01 level (two-tailed)

Table 5.8 Correlation between accuracy and confidence across tool types.

subjects are generally confident on tasks they got right. The correlation for general and integrated tasks is not strong (i.e., the correlation coefficient is not close to +1). The relation for harder tasks is weaker, indicating, again, that other factors may be at work.

Task Type	Pearson Coefficient	Significant?
1 – Specific	+.72	Yes**
2 – General	+.30	Yes**
3 – Integrated	+.26	Yes*

* correlation is significant at the 0.05 level (two-tailed)
** correlation is significant at the 0.01 level (two-tailed)

Table 5.9 Correlation between accuracy and confidence across task types.

Preliminary Results: Main Effects

Task Complexity

Two-way ANOVA results indicated that task complexity had a highly significant effect on confidence ($PR > F(6,228) = 13.4$, $p < 0.0001$) (as predicted). However, there was no significant interaction effect. Thus, the pattern of confidence across the three task types was similar for all tools. That is, tool type does not show evidence of affecting confidence differently for each level of task type.

The group means for confidence (Table 5.10) shows that subjects' confidence in their decisions decreases for more complex tasks (as predicted).

Task complexity	Mean confidence
1 – Specific	2.88
2 – General	2.68
3 – Integrated	2.39

Table 5.10 Mean confidence (5 – highest) by task complexity.

Post Hoc analysis (Tukey's HSD) of the main effect of task complexity on confidence was conducted in order to examine how confidence differed across task types across all tools. HSD was calculated to be 0.22. Thus, means must differ by more than 0.22 (on a five-point scale) for the difference to be significant (at the $\alpha = .05$ level). By comparing means for all tools, HSD showed that confidence significantly differed between specific and integrated tasks (1&3), and between general and integrated tasks (2&3), but not between specific and general tasks (1&2), although there was a strong trend for confidence to be less on general than specific tasks (see Table 5.11).

Tasks	Means	Difference	HSD? (diff > 0.22)
1 & 2	2.88 2.68	.20	No
1 & 3	2.88 2.39	.49	Yes
2 & 3	2.68 2.39	.29	Yes

Table 5.11 Tukey's HSD for all tools.

Visualization Tool

Since there was no interactive effect, one-way ANOVA was conducted to examine the effect of tool type on confidence across all tasks types. Results indicated that tool type had a highly significant effect on confidence ($PR > F(3,93) = 9.27$, $p < 0.0001$). Mean confidence is highest for the merger tool, followed by the mean confidence for the hypermap-merger, as shown in Table 5.12. The means are lowest for the side-by-side and toggle tools, which have similar means. A trend reveals that the subjects seem more confident when using the merger and hypermap-merger tools, regardless of task type. This result is similar to the previous results that showed that overall accuracy was higher for the merger and hypermap-merger tools (Table 5.2), although there was a significant interaction effect. However, it is interesting to notice that while accuracy and confidence are both high for the merger group, the correlation between accuracy and confidence was not high (Figure 5.8).

Post hoc analysis using Tukey's HSD test ($\alpha = .05$ level) shows that confidence results are significantly different between the side-by-side tool and both merger tools, and between the toggle tool and both merger tools. The results for the side-by-side and toggle and the two merger tools showed no significant differences.

Tool	**Mean confidence**	**Std. Error**
1 – Side-by-side	2.25	.165
2 – Toggle	2.36	.136
3 – Merger	3.04	.087
4 – Hmap Merger	2.86	.114

Table 5.12 Mean confidence (5 – highest) by tool.

Summary

Confidence levels were low overall. They were correlated to accuracy (as predicted) but not as highly as was expected. In general, subjects who were more confident performed better, or subjects who performed better were more confident (no causal relation implied). The strongest correlation between confidence and accuracy was for the toggle tool and for specific tasks. The weakest correlation was for the merger tool. This is interesting since both mean accuracy and confidence scores were high for this group.

Confidence decreased as task complexity increased (as expected) but the rate of decrease did not vary significantly between tool groups. For all tool groups, there was a significant difference in confidence between specific and integrated tasks, and between general and integrated tasks, but not between specific and general tasks (although there was a trend). Across task types, confidence levels were similar for side-by-side and toggle tools. Overall, subjects using the merger tools had higher confidence for all tasks and this confidence was higher on specific tasks than general or integrated ones.

5.4 Hypothesis Three: Preference

As discussed previously, it was proposed that the subjects' differing cognitive strategies for solving map use tasks might be reflected in tool preference. Or, more simply, subjects might perform more accurately and feel more confident using tools they prefer. This section will examine the relation between preference and accuracy, and preference and confidence.

Predictions

For the third hypothesis, I examined how subjects' accuracy and confidence levels were related to tool preference. It was predicted that subjects assigned to a tool group that corresponds to their preferred tool would perform more accurately and feel more confident.

Analysis

Recall that in the second activity (after subjects had answered all fifteen task questions), they were given additional sample tasks that they could solve using each of the four visualization tools. They were they asked to rank the tools from most to least preferred and to provide comments on why and what they liked or disliked about each tool. For this set of analyses, subject data was divided into two groups: those that were assigned their preferred tool (based on their post test response) and those that weren't. A t-test was then used to compare the means of these two independent samples (see section 4.9). A t-test requires equality of variances, which can be tested using Levene's Test. The assumption of equal variance could not be rejected for both the accuracy and the confidence groups.

Results

Of the 86 subjects that completed the second activity, 27 subjects were assigned their preferred tool and 59 were not. Examination of accuracy and confidence means for the two groups showed that the mean scores were very similar (see Table 5.13).

Not surprisingly, t-test results showed no significant difference between the preferred and not-preferred groups for either accuracy or confidence. However, frequency analysis of subject ratings of their most and least preferred tool showed that an overwhelming 68 out of 86 subjects rated the hyper-map merger tool their most preferred tool. Analysis of least preferred showed that roughly equal numbers of subjects rated the side-by-side, toggle and merger tools as least preferred. No subjects rated the hypermap-merger tool as least preferred.

Subject scores on accuracy and confidence were also examined for subjects assigned their least preferred tool. No significant results were found.

Preference	**N**	**Mean Accuracy**	**Std. Error**
Preferred	27	66.2 %	2.44
Not-preferred	59	65.5 %	2.32

Preference	**N**	**Mean Confidence**	**Std. Error**
Preferred	27	2.64 (out of 5)	.125
Not-preferred	59	2.68 (out of 5)	.089

Table 5.13 Mean accuracy and confidence for preferred and not-preferred tools.

Summary

The third hypothesis did not hold. Subjects did not perform more accurately or feel more confident using tools they preferred. This was primarily a result of the fact that over 75% of subjects preferred the hypermap-merger tool, whether they had been assigned it or not.

5.5 Demographic Variables

Demographic data was collected and analyzed to ensure that the test subjects formed a relatively homogeneous group and to determine if any of the demographic variables were related to performance, particularly accuracy.

Results

Gender

T-test analysis showed that the means did not vary significantly between the female and male groups for either accuracy or confidence. Males did perform slightly more accurately and with greater confidence (see Table 5.14). This may be more a result of males higher comfort levels in computer environments than to any factor specific to spatial ability or cognitive strategy.

Gender	**N**	**Mean Accuracy**
Female	48	63.7 %
Male	46	68.0 %

Gender	**N**	**Mean Confidence**
Female	48	2.54 (out of 5)
Male	46	2.78 (out of 5)

Table 5.14 Mean accuracy and confidence per gender.

Native Language

After removing three outliers (see Section 4.9), T-test analysis showed that the means did not vary significantly between the native English speakers and non-native English speaking groups for accuracy (see Table 5.15).

Native Language	N	Mean Accuracy
English	77	65.7 %
Not	17	68.2 %

Table 5.15 Mean accuracy per native language.

Other

Similarly, there were no significant differences found between accuracy means for groupings based on age, department, computer experience or professional experience.

Summary

The sample population formed a relatively homogeneous group in terms of demographic variables including gender, native language, age, department, computer experience and professional experience. None of these factors significantly affected either accuracy or confidence.

5.6 Timeouts

Timeouts are cases where the student either ran out of time (in which case the automatic submission function recorded responses) or students purposely waited for the timer to run out rather than clicking the submit button. Since there is no way to separate these two cases, any analyses of the distribution of timeouts are conjectural only. Keeping this in mind, the frequency of timeouts was examined in order to determine if any one task type or tool group had a significantly higher distribution of timeouts.

A frequency distribution of average time spent per task was calculated for each task type (Table 5.16). The timeout value is the number of occurrences of the maximum time allowed for specific, general and integrated tasks (35, 50 and 80 seconds respectively as outlined in section 4.7).

	N	Range	Timeouts	Percent
Specific	94	24 – 35 seconds	20	21.1 %
General	94	31 – 50 seconds	13	13.8 %
Integrated	94	42 – 80 seconds	12	12.8 %

Table 5.16 Frequency of timeouts per task type.

The results show that across all tools, there are more timeouts for specific tasks (21.1%) than general (13.8%) or integrated (12.8%) tasks. This is likely due to the shorter allowable time given for specific tasks (30 seconds). However, it is not possible to say whether more subjects ran out of time on specific tasks or if more subjects let the timer run out. Recall that allowable response times were determined based on the expectation that they would provide enough time for at least 80% of subjects to complete the tasks (described in sections 4.6 and 4.7). All the timeout percentages fall close to this range.

For specific tasks, a frequency distribution of average time task was calculated for each tool group (Table 5.17).

The results show that for specific tasks, there were more timeouts for the side-by-side (30.0%) and toggle (37.5%) tools groups than timeouts for the merger groups.

For general type tasks, a frequency distribution of average time on task was calculated for each tool group (Table 5.18).

	N	Timeouts	Percent
Side-by-side	20	6	30.0%
Toggle	24	9	37.5%
Merger	25	2	8.0%
Hypermap-merger	25	3	12.0%

Table 5.17 Frequency of timeouts on specific tasks per tool group.

	N	Timeouts	Percent
Side-by-side	20	4	20.0%
Toggle	24	6	25.0%
Merger	25	2	8.0%
Hypermap-merger	25	1	4.0%

Table 5.18 Frequency of timeouts on general tasks per tool group.

The results show that for general tasks, there were again more timeouts for the side-by-side (20.0%) and toggle (25.0%) tools groups than timeouts for the merger groups.

For integrated specific tasks, a frequency distribution of average time on task was calculated for each tool group (Table 5.19).

	N	Timeouts	Percent
Side-by-side	20	3	15.0%
Toggle	24	7	29.2%
Merger	25	1	4.0%
Hypermap-merger	25	1	4.0%

Table 5.19 Frequency of timeouts on integrated tasks per tool group.

The results show that for integrated tasks, there were more timeouts for the side-by-side (15.0%) and toggle (29.2%) tools groups than timeouts for the merger groups. Overall, the toggle tool group had more timeouts, and the merger tool groups had less timeouts.

In Chapter Six – Discussion, I provide the context for these findings by addressing the research questions posed in the preliminary chapters, discussing the implications of the results of this experiment with respect to the design of interactive tools for digital map exploration and comparing the results with related research in the field of cartographic visualization.

CHAPTER 6 – DISCUSSION

6.1 Overview

This is one of the only known experiments to investigate the combined effects of tool type and task type on digital map use effectiveness for spatial data and metadata (other studies include Evans 1997, Slocum et al.1990, MacEachren et al. 1998). The two merger tools developed to simultaneously display class data and associated certainty data are unique and subject to patent. This chapter addresses the research questions posed in the first two chapters, gives an overview of how the results outlined in Chapter Five relate to the hypotheses posed in Chapter Three, discusses the implications of the detailed results of this experiment with respect to the design of interactive tools for digital map exploration, and in relation to previous research in the field of cartographic visualization, and brings forward some of the limitations this research.

The challenge in any field of experimental study is to first pose good questions based on both theory and on previous research. Next, these questions must be translated into the design of an experiment that can be used to test one or more hypotheses. The results from the experiment must be translated from the language of statistics in order to address the original questions and to pose new ones.

In the second chapter I posed one general and six specific research questions. In general, in an online geographic/cartographic environment, are displays that offer highly interactive visualization tools (e.g., hypermap, interactive bivariate map) more effective for communicating spatial information (e.g., thematic and uncertainty information) than less interactive tools (e.g., sequenced map)? Performance results (both accuracy and confidence) show that overall the highly interactive merger tools were more effective for communicating thematic and certainty information. However, a sequenced map (toggle) was also found to be as useful as the interactive bivariate map (merger tool) for the analysis of general patterns. The combination of a hypermap with the interactive bivariate map to create a hypermap-merger tool proved to be highly effective and was preferred by the majority of users for all tasks.

If theories of cognitive map formation can be extended to virtual environments we would expect the following specific research question to be false. Do low level interactive visualization tools present spatial information as effectively (accuracy and confidence) as highly interactive visualization tools? Performance results indicated that low level interactive visualization tools do not present information as effectively as more interactive tools. By adding interactivity, users have more control over their environment and can explore it in order to gain a better understanding of the information the data represents and thus make more accurate and more confident decisions about map use tasks.

Does increased interactivity result in increased accuracy? While this was found to be true in this study, I cannot discount the effect of good theory-based design. Software developers and others that simply add interactivity for interactivity's sake, or to add flashy novel features to an existing tool set, should be cautioned that users' performance will likely not be improved, nor will users necessarily prefer such tools.

Do highly interactive tools result in the same accuracy but increased confidence, or decreased accuracy but the same or greater confidence? In this study highly interactive tools resulted in both increased accuracy and confidence. However, the correlation between accuracy and confidence was not high. In particular the correlation between accuracy and confidence for the merger tool was not significant. However, examination of mean confidence scores shows that confidence was slightly higher than for the other groups. This is interesting, and although no causality can be proven, the combined effect of a novel and "flashy" new tool might have resulted in overconfidence. The hypermap-merger tool groups' accuracy scores were the highest, suggesting that the addition of the hypermap functionality facilitated exploration and allowed users an alternative method to check answers (resulting in a better match of accuracy and confidence). With the hypermap-merger tool, what you think you get is what you get. I would strongly argue that inclusion of the hypermap functionality with the merger tool should be standard in a visualization tool set.

How does the addition of textual attribute and uncertainty information (via hypermaps) affect accuracy and confidence in map-use tasks? The accuracy and confidence results show that users assigned to the hypermap-merger tool group either outperformed or were not significantly different from all other tools on all tasks. Given the warning of overconfidence that should be tied to the merger tool, the addition of a text-based attribute query tool would improve the performance of the merger tool as well as a variety of other visualization tools (e.g., side-by-side).

Is there an optimal "best" visualization tool for this type of spatial data and metadata, or are different visualization tools best suited to different task types? If so, which visualization tools are best suited to different task types? While the merger tools enhanced performance and were significantly preferred, the toggle tool was useful for general tasks. Other studies have suggested that side-by-side displays are useful for the display of spatial data and reliability data (Evans 1997, MacEachren et al. 1998). Other factors, including data type, scale and resolution and classification scheme, will affect which tools are most appropriate for particular tasks.

Does increasing the level of interaction buy us anything? If so, for what task types? The effect of increased interactivity were most pronounced for specific identification tasks (i.e., simple). The advantage of increased interactivity was less salient for more difficult tasks (more on this in section 6.3).

6.2 Research Hypothesis: Overview

Overall, the results of this experiment were close to those outlined in the Research Hypotheses (see section 3.2), with a few interesting findings. As predicted, the accuracy of subjects' responses decreased as task complexity increased, and the rate of decrease differed across tool groups. By allowing subjects to interact more with their spatial data, subjects performed more effectively on their map use tasks. This is key. However, simply adding interactivity to tools is not enough. Tools must be designed that consider how users process and interact with data. The ability to present both spatial data and metadata simultaneously seems to provide an advantage for tasks involving both sets of data.

Another thing that is interesting about the accuracy results is that performance was most significantly improved on the simpler specific type tasks. Perhaps the edge given by interactivity is less salient for more difficult tasks. It may also be the case that the merger tools were simply more effective for specific identification type tasks. It would be interesting to uncover users' cognitive strategies for solving these tasks and to see if there was a relation between them and the functionality provided by the merger tools. In any case, this last finding was not necessarily what was expected. Since it may be difficult to identify and isolate exact tool-task relations, and many real-world tasks are a combination of task types, providing the user with several kinds of tools is probably the best option for creating the optimal visual exploratory tool set. See section 6.3 for a detailed discussion of the implications of these of accuracy findings.

Subjects' confidence ratings were directly related to the accuracy of their responses except for the merger tool (as noted above). This is important since confidence could have been influenced by external factors (e.g., data uncertainty values involved in the task) or individual factors (e.g., general level of computer experience). Confidence decreased as task complexity increased across all tool groups, as expected, although there was no interaction effect. Thus, task complexity plays a stronger role than tool type in determining confidence. It is interesting to note that subjects did not feel less confident with the more difficult-to-use merger tools, even though there was probably some kind of novelty effect. With some training, the merger tools increase both user confidence and accuracy. See section 6.4 for a detailed discussion of the implications of these confidence findings.

Preference was examined to determine if users performed better with the tool they choose as most preferred in the post-test. This was not found to be the case. Users' accuracy and confidence ratings did not vary significantly (at an $\alpha = .05$ level) between the two groups. Perhaps preference was not a good indicator of individual differences in cognitive strategies, or the overwhelming acceptance of the hypermap-merger tool clouded the affect of other factors. The three-way split in least favourite tool (between side-by-side, toggle and merger) does show some individual variation. Again, a well-designed visual exploration tool set should contain these and others. See section 6.5 for a further discussion of these findings.

Other user-specific variables (e.g., gender, age, computer-related experience) were also examined to determine if they played a significant role in performance. However, none of these demographic variables proved to have a significant impact on accuracy, confidence or performance. While individual differences may outweigh design factors in determining map-use effectiveness (as suggested by Mersey, personal correspondence), this study did not concur with those findings. Good design seems to be good design for everybody.

6.3 Discussion: Accuracy

Hypothesis one stated that users' accuracy scores would decrease for more complex tasks, and that the rate of decrease would vary depending on visualization tool used. This is based on the assumption that more complex tasks are more difficult and, therefore, fewer users should be able to solve these tasks accurately in the allowable response time. It is important to test tools across different types of tasks, whether the tasks types are more difficult or not, since most real world map use involves a combination of task types. Preliminary results (Main Effects) did show that users' accuracy scores were higher for specific than general tasks, and higher for general than integrated tasks as expected (see section 5.2). Even with the increased time for general and integrated tasks, accuracy scores were still lower. The difference between general and integrated tasks was smaller than that between specific and other tasks. This is perhaps due to the additional time given for integrated tasks, which may have allowed users to compensate for additional difficulty. It also may be that while integrated and general are different types of tasks, one is not significantly more difficult than the other. What is important is that the three types of tasks are different in terms of cognitive difficulty and are representational of the kinds of tasks found in a real world resource management exploratory environment. For example, an environmentalist in the exploratory phase of data analysis might use thematic and certainty data to determine the likelihood that identification of a stand (specific task) or group of stands (general pattern analysis) is correct. A policy setter might use the same information and perform the same tasks to determine if the information is certain enough to make a particular decision or if more data need be collected.

The second part of the hypothesis suggests that users' performance on the different types of tasks would vary across tool types. A well-designed exploratory analysis tool set should provide users with tools appropriate to different tasks types as well as combinations of tasks. This information did emerge from the analysis of tools across tasks. Preliminary analysis showed that accuracy varied significantly across tools types (see section 5.2). Analysis of interaction effects also showed that the task complexity and the visualization tool had a combined effect on accuracy (as predicted). Overall, the means for the side-by-side, toggle, merger and hypermap-merger tool groups showed that the effect of increased interactivity was increased accuracy. Subjects using the more interactive tools performed better on most tasks. It is unlikely that interactivity alone is the sole factor responsible for increased performance, but it is an important one. The design of the merger tools was based not only on increasing the amount of interaction and control the user could have over the data, but also incorporated a new way to represent bivariate data (particularly data and the associated uncertainty data) based on colour theory. This arose from thinking about how users might form cognitive maps from digital spatial data, and previous work by MacEachren (1994c). It seems likely that some combination of interaction and theory-based design was responsible for the overall better performance with the merger tools. By allowing users to interact with their data in a meaningful way, interactive tools can help them better explore, understand and solve tasks involving multivariate spatial data.

It has also been noted that giving users control over the display of data is important. In a study on animated sequencing, Monmonier and Gluck (1994) found that users found flickering between maps "offensive" unless they were given control over this sequencing. In a recent study by Evans (1997), she found that some users assigned the static display found a way to toggle between map views. Those that improvised in this way, performed better. In children's educational interactive multimedia, interactivity has been seen to increase engagement, which in turn increases motivation and the achievement of required learning objectives (e.g., www.brainium.com/science). Visualization tool developers should take notice and focus their efforts on the development of tools based on these kinds of design principles.

I also found that the merger tools gave users a significant advantage over the side-by-side or toggle tool on specific tasks. The advantage on simple tasks was not expected, based on the idea that giving users increased interaction with the data would improve performance most on complex tasks. Perhaps the advantage afforded by increased interactivity and functionality was mitigated by the increased cognitive difficulty of general and integrated tasks. This is definitely worthy of further investigation, as several authors agree that task characteristics play a key role in determining the appropriateness of a given representation (e.g., Tan and Benbasat 1993).

It is also interesting that accuracy scores on general tasks using the merger and toggle tools were roughly the same. While previous studies have shown that the toggle tool (i.e., sequencing) was not usually as effective as other means (e.g., Slocum et al. 1990), my results show that the toggle tool was as effective as the merger tool for general spatial pattern analysis tasks. However, sequencing as defined by Slocum et al. involved the sequential display of map parts rather than map layers. This may explain the difference, as Slocum et al. suggests that a major factor in "learning" or gaining understanding from maps is that all of the map "schemes" or layers need to be interrelated mentally, instead of learning one scheme at a time. The toggle tool developed in this study supports the interrelation of the thematic and the certainty layers by allowing users to rapidly switch between the two layers, creating an afterglow effect. In addition, polygon boundaries were identical in the two layers, helping users form a mental map of the data. In a study done by Evans (1997) examining static and flickering display types, she found that many users who had higher scores surreptitiously created a toggle effect between a land use map and a reliability map (although this display type was not explicitly created as part of the methodology). The tasks employed in this study can be classified as general pattern analysis tasks, thus confirming the notion that the toggle tool is effective for general type tasks.

Overall the hypermap-merger tool was excellent for all tasks, which points to successful design since most real world map use tasks are a combination of specific, general, integrated and other tasks types. The high accuracy scores for the merger and hypermap-merger tools are in direct contrast to the results from a recent experimental study by MacEachren et al. (1998) examining how users perform on tasks related to the exploration and interpretation of health statistics (mortality rates) using three visualization display types. The three tool types used in MacEachren's study are termed adjacent, coincident (visually separable) and coincident (visually integral) where the first is a side-by-side display, the second uses symbolization (texture overlay of diagonal line work) to depict reliability information overlaid on class data and the third uses a shift in saturation (or hue) to depict reliability (i.e., merging of mortality data and reliability metadata through colour representation). The third method is a similar approach to that used in the merger tools in my study, although there are only two classes of reliability (reliable or not) in the mortality study. In MacEachren's study it was found that users performed worse with the merger display for regional and national level tasks. It was also found that users least preferred the merger display. While tasks were similar specific/local tasks, general pattern analysis or regional/national tasks, the tasks predominantly involved looking for numbers of clusters of a certain characteristic (e.g., clusters of high mortality rates, clusters of low reliability data). Tasks in my study all involve both classified thematic data and certainty data. The differences in task type alone do not likely account for the differences in findings since there is a significant difference in display design between the two merger displays.

One key difference in the MacEachren et al. (1998) study is that it was noted users had difficulty distinguishing between various saturation levels of specific hues and that "hue shifting" was used instead of desaturation for some colour schemes. The nominal colour scheme used in my experiment was, in part, chosen to ensure that for each hue, three distinct levels of desaturation (corresponding to three certainty levels) were visually separable (as tested in the first pilot study). In my study, consistent representation of colours was ensured by the use of a standard Web palette and the use of 256 colours for all monitors. In addition, the merger and hypermap-merger tools developed in my study also allowed the user to control the data presentation in order to create both visually separable (examine only thematic data or only certainty data) and integral displays (merge two data layers) as desired. However, in the separable display mode, users cannot view both maps simultaneously. This might be a useful feature that could be added. One last point is that the addition of hypermap functionality to the second merger tool can also account for the improved performance, and may explain the contrast with MacEachren's findings. Allowing users to visually examine values and patterns as well as display the values per stand/polygon provides a significant advantage over simple merger tools.

Results from the MacEachren et al. (1998) mortality study also showed that users were less likely to incorporate reliability information into their answers when it was depicted on an adjacent map (side-by-side). Since accuracy was lowest for the side-by-side tool in my study, it is possible that lack of attention to reliability was, in part, responsible for the decreased scores. It is also possible that by embedding the tasks in a playful story, users' attention was captured and led to higher motivation than for users engaged on context-less task testing. (See the reference above on the effect of interaction and engagement on children's learning and motivation).

A recent study examined the impact of the use of the visual variables: value, texture and saturation to represent attribute certainty on decision-making (Leitner and Buttenfield 2000). The authors recorded response accuracy (on an easy and a difficult site location task), response time and confidence. Analysis of accuracy scores showed that value was preferable to texture for representing certainty information about nominal class data, which, in turn, was preferable to saturation. Self-rated confidence levels were uniformly high for all visual variables. Analysis of response time showed that the use of saturation to represent certainty yielding the shortest response times. In the future, comparison of the use of saturation to other visual variables including value and texture would be worthwhile.

While accuracy was higher for the merger tools, as predicted, it seems likely that there were some novelty effects (i.e., tools might have been more difficult to use since they were not familiar). Although subjects were given a training session, it is likely that these tools took longer to use since they were more complicated than the side-by-side or toggle tools. Did novelty significantly reduce the interactive advantage of using these tools? This is an interesting question that cannot be answered without further investigation. Humans are remarkably adaptable and can often learn how to use new tools very quickly, especially those designed with human cognition and perception in mind.

In practical terms, what does all this mean? Nothing new. Good design of map tools should be based on a real understanding of how we perceive, interact with,and think about spatial information. The merger tools are simply one step in the right direction. To follow good engineering design practices, the tools should probably go through a second iteration of design and user feedback in order to fine tune them further, before they are incorporated into a map analysis package. I will leave this up to cartographic software developers.

It is also apparent that the different tools are useful for different tasks, and good software design should incorporate several well-designed tools into any package designed to aid in the exploration and understanding of spatial data. And, since individual differences play a significant role in decision-making in general (although this was not found in this experiment), customization may increase user performance by allowing users to choose tools that are inherently more suitable to their own style. One step further is to allow users to adapt tools to their own needs, depending on their own unique characteristics. Adaptive interfaces have been suggested as a response to this problem (Norcio and Stanley 1989).

In future studies it might be interesting to apply other methods of task evaluation to determine which cognitive strategies were used per task. For example, by using protocol analysis in which subjects "think out loud" as they perform tasks, it is possible to determine if they used random search, systematic search, focused search/selective search or combinations for each task (Howard and MacEachren 1996). By looking at accuracy scores relative to actual cognitive strategy (versus task type), it is possible to learn more about which tools might be most appropriate and gain clues into useful designs by creating tools that closely mimic actual cognitive strategies. It would also be interesting to see if these results apply to other types of bivariate data. The same techniques could be used to represent mean and standard deviation values.

6.4 Discussion: Confidence

Hypothesis two stated that user's confidence scores would be directly related to their accuracy scores. It also suggested that confidence would decrease as task complexity increased, and that the rate of decrease would vary across tool groups. Why is this important?

Users' confidence levels play a key role in their exploration and interpretation of data. It is often a determinant of subsequent behavior and interpretation of results (Zakay 1997). Confidence (i.e., the degree to which a decision maker believes that the decision will bring out the desired outcome) is also related to the perceived credibility of that decision maker (Oz, Fedorowitz and Stapleton 1993). In both cases, it is important that confidence is neither under- nor over-rated. The degree to which a map-use task can be carried out accurately and with accurate self-ratings of confidence is key. As such, users' confidence ratings are an important measure of the effectiveness of a map use tool (Antle and Klinkenberg 1999).

However, there are many factors that may affect confidence besides users' feelings of confidence about their solution to a particular task. It is often difficult to discriminate between these factors (Zakay 1997). Confidence can be a result of both external and internal factors including: the users' comfort level with the environment or display tool they are using, their comfort level with their understanding of the data and the decisions they must subsequently make, the amount of information presented, the amount of uncertainty represented in the metadata, individual differences in intelligence, the users' general internal feelings of confidence based on personality traits, the motivational context and users' feelings of dissonance as a result of a particular decision. How much of a users' confidence is independent of the process of solving tasks is difficult to determine. For example, do more confident users perform better or did users that knew they were right subsequently feel more confident?

In light of the many factors influencing confidence ratings, it was important to determine that the dominant factor affecting confidence scores in this study was, in fact, the users' feelings of confidence about the decision they were making (i.e., their ability to accurately solve a particular task). By looking at the relation between confidence and accuracy we saw that the two were significantly correlated, verifying this relationship for three of the four tool groups. These findings verify the use of confidence as an important measure of map effectiveness. However, the correlation values were not exceptionally high (see section 5.3). Given the moderate value of these correlations, it is likely that other factors played some role in determining confidence. Subjects could be guessing the correct answer without confidence or, more likely, subjects chose the incorrect answer but were very confident about it. It would be useful to perform further analysis to determine the exact nature of the correlation between accuracy and confidence.

Once the relation between accuracy and confidence was verified, it was important to examine if any of the tools were causing an under- or over-estimation of confidence relative to accuracy. The correlation between confidence and accuracy was strongest for the toggle tool group, followed by the side-by-side group and the hypermap-merger group. The correlation was not significant for the merger group as discussed previously. While the toggle group had lower accuracy overall than either of the merger tool groups, the relation may have been stronger since the toggle tool was easier to use than the merger tools. This became evident in the training period and on the preference worksheets, and was not unexpected. The high correlation may reflect the fact that when subjects were correct, they felt more confident because the tool was easy to use. For the other tool groups, even when they were correct, subjects may have been slightly less confident since the tools were harder to use. This fits with the post-test questionnaire results that the merger tool was the most difficult tool to use. It is likely that the correlation between accuracy and confidence would increase with increased exposure to and familiarity with the merger tools.

The second part of hypothesis two predicted that confidence would decrease for more difficult tasks. The results indicate that confidence did significantly decrease for increasing task complexity. Confidence was highest (where five was the maximum value) for specific tasks (mean = 2.9), followed by general (mean = 2.7) and integrated (mean = 2.4). There was a significant difference in confidence (i.e., decrease) between specific and integrated tasks, and between general and integrated tasks, but not between specific and general tasks (although there was a trend). This suggests that the confidence values did reflect the cognitive difficulty of performing tasks. This is important since one of the goals of this study was to test visualization tools across a range of tasks.

However, there was no significant effect of tool type on the differences in confidence found between the task types. Nor was there an interaction effect. That is, the rate of decrease of confidence relative to increased complexity did not alter significantly by tool type. This suggests that task complexity is the dominant of the two factors influencing confidence. In terms of designing tools that do not cause users to over- or under-estimate their feelings of confidence, the fact that task complexity is the dominant factor is good. If tool type overpowered task complexity for a particular tool design, the tool might unduly influence users' confidence in their decisions.

I also examined if tool type, by itself, had a significant impact on confidence. It did. Confidence scores were significantly higher for the merger and hyper-map merger groups than the side-by-side or toggle groups (as expected since accuracy scores were also higher). This suggests that confidence is an appropriate surrogate for accuracy, as suggested in the information management literature (e.g., Oz, Fedorowitz and Stapleton 1993).

Overall, confidence has been shown to be an important measure of map-use effectiveness. Tools that are easy to use (i.e., are intuitive or facilitate normal perceptual and/or cognitive processes) may increase feelings of confidence when compared to more difficult tools, stressing the important of user training and help features for novel tools. The merger tool design supports users accurately rating their confidence about their decision and, of all the tools, the merger tools best help users understand the data.

6.5 Discussion: Preference

While it was expected that individual differences might contribute to results in numerous ways, this was not the case. Most notably, I had thought that subjects' different cognitive strategies for solving tasks might lead to a preference for the tool that best reflected their personal strategy. This was not found. Instead, a highly significant number of subjects across all groups and all tasks preferred the two merger tools. These findings support several studies in management and information science that suggest that task type is more important that individual differences (e.g., Benbasat, Dexter and Todd 1986).

Although individual differences (as expressed through tool preference) did not significantly affect accuracy or confidence, I would suggest that the development of software aimed at cartographic visualization for the exploration of spatial data and metadata should include both traditional and more interactive tools, such as the merger tools. Traditional tools provide familiarity and will always be preferred to more complex tools by some individuals. However, the development of new tools should continue to draw from a wide range of theories and experimental results in order to create tools that enable users to accurately and easily explore data.

Comments from the user preference worksheets (Appendix C) provided invaluable insights into reasons for preference, likes and dislikes, and comments on tool design. Included below are comments that were repeatedly made.

Even though users thought that the side-by-side maps would be good for general (comparison) tasks, accuracy results did not support this. Users were mixed in their ratings of the usefulness of this tool. Many thought it was great, while other thought it was too difficult to visually scan back and forth between maps. Users agreed that this tool would be difficult to use for specific tasks, as the results showed.

Users were mixed in their acceptance of the toggle tool, although many agreed it would be useful for identifying general patterns (which it was). The ease and quickness of use were repeatedly cited as advantages. However, only being able to view one map at a time was problematic for most users. I wonder if toggle tool users would have paid attention to the certainty map at all if they had not been forced to address data quality through the tasks.

Many users suggested that the merger tools would be excellent for specific tasks (which they were), since you could select one polygon and then simultaneous display both sets of information about it (versus visually comparing in the side-by-side display or having to toggle). Reactions to and acceptance of the merger tool without hypermap functionality were definitely mixed. Users either "got it" and loved the tool or found it confusing. Many found the slider functionality slow, however, they did like being able to control the display using the slider tool. Suggestions were made to change the slider function to discrete points that could be clicked on.

It would still be possible to accommodate a range of visual preferences with a selection of clickable points. This comment was also made in the first pilot study. It was also suggested that training was critical in terms of being able to distinguish the three levels of each colour.

The hypermap functionality proved an invaluable tool to initially learn these distinctions. Once learned, the hypermap tool was no longer needed in this respect. Several users commented that using grey to code the certainty map made it hard to distinguish between the levels, a finding collaborated by the results of an experimental study by Schweizer and Goodchild (1992). There is no doubt that the hypermap functionality was a welcome addition to the merger map, allowing users two forms of access to information. It was often used to verify information once a decision had been made using the merger (colour) function alone and was cited as a factor in increased confidence.

It is important to include anecdotal components to any map-use study, since this type of feedback gives us insights that statistical results cannot possibly match. User comments on the usefulness of the merger tools for specific tasks give an understanding of the perceptual and cognitive strategies used to solve these tasks. Good design requires a thorough understanding of these processes and of the factors that are important to users. In addition, if users' reactions to and acceptance of display methods of certainty information are important factors determining whether users' will pay attention to data quality (Evans 1997), it makes sense to determine their opinions. The comments above will prove invaluable in the redesign and fine-tuning of these tools, and again point to the suggestion to include multiple tools in any toolbox.

6.6 Discussion: Demographics

It has been suggested that individual differences may outweigh design factors in determining map-use effectiveness. However, the subjects chosen for this study seemed to form a relatively homogeneous group. In terms of verifying results, it was important to check for any significant differences in demographic variables that might have affected accuracy, confidence or preference results. The variables examined included: age, gender, home department, program, program year, computer (PC) experience, Internet experience and professional experiences. No significant differences were found between these factors and accuracy or confidence. I think this may have been, in part, because the subjects chosen had many similar characteristics. In addition, I tried to minimize the impact of some these factors through good experimental design. For example, all users were provided with pre-training in map representations and quizzed on this. They were also trained in the use of the visualization tools with practice questions. The map-use environment, that has been shown to impose a significant effect on performance (both online and surroundings) (Carter 1988), was chosen to be simple and similar for all subjects.

While studies have shown that there are clearly gender differences in spatial abilities (e.g., Schofield and Kirby 1994), these differences do not seem to carry over into the ability to perform on online map use tasks. Studies by Gilmartin and Patton (1984), Gilmartin (1986) and Evans (1997) found no significant differences between genders. One possibility is that while females may have lower spatial abilities in general, those that enter the field of geography do not. Males did have slightly higher confidence ratings than females, although their accuracy scores were slightly lower. This may reflect their level of confidence in a digital environment as it has been found that males are typically exposed to computers at an earlier age and subsequently are more confident in that environment than females.

Some experimental cartography studies split users into expert and novice users. The subjects who participated in this study fall in between the two groups. They were certainly not novice users, but were they experts? Possibly some of them were (for example, five subjects had over three years computer experience and over five years professional experience). In any case, there were no significant differences in accuracy or confidence between less experienced and more experiences subjects in this experiment. This was found to be the case by Evans (1997), who also found no differences for relatively straightforward map-use tasks in a familiar environment with some pre-training.

6.7 Summary

One of the goals of this experiment was to create ecological valid results that might be applied in the design of interactive tools for visualization. Both statistical and anecdotal data provided invaluable information towards achieving this goal. Well-designed and theory-based interaction helps users, in the exploratory stages of map use, form better mental models of spatial data and uncertainty data, leading to increased performance. There is solid evidence that some tools (e.g., hypermap-merger) are better matched to particular task types (e.g., specific) than others. A visualization toolbox should provide users with both traditional and innovative tool options. Users' overall feelings of confidence in their decisions were more related to task complexity than any particular tool used, although simpler tools seem to instill more confidence when answers were correct. The effects of user preference were minimal. A huge majority of users preferred the hypermap-merger tool with its combination of user controlled visual overlay of bivariate data and point and click functionality. Individual differences in age, gender or experience were not found to be significant factors in determining map-use performance. This information provides valuable insight into the design of interactive tools for the visualization of spatial data and metadata.

In Chapter Seven – Conclusion, I summarize the importance of these findings, mention some of the limitations of this study, and offer suggestions for future research.

CHAPTER 7 -- CONCLUSION

7.1 Overview

Currently, there are three main categories of research issues in Cartographic Visualization. The first is the introduction of new concepts and theoretical perspectives. The second is the development of new tools. The third is the application and evaluation of visualization tools. The research presented here addresses all three areas.

In Chapter One – Introduction, I posed several general questions that need to be addressed in the field of cartographic visualization. First, *How do we effectively represent both data and metadata in a digital display?* Many visualization methods have been suggested that can be used to represent both spatial data and data quality information. These include: side-by-side, sequenced, animated, and other bivariate approaches. However, few have been experimentally tested in order to determine their effectiveness. Visual exploration of data is one dimension of interaction with maps that is key in decision-making and policy formation. In this study, four exploratory visualization tools were tested in order to determine how well they represent both spatial data (thematic forest cover information) and data quality (associated uncertainty information). Two tools were traditional (side-by-side, sequenced/toggle) and two were newly developed (merger, hypermap-merger). Effectiveness, as measured by subject accuracy (frequency of correct responses) and confidence (self-rated assessment of feelings of confidence in the decision), was shown to be a relative measure in this experiment. To some degree, all of the tools proved to be effective ways of representing theme and certainty information. However, the merger and hypermap-merger tools proved better than the others for all tasks types, and were highly effective for specific identification tasks. The toggle and merger tools were equally effective for general comparison tasks. None of the tools caused subjects to significantly under-or over-estimate their confidence levels. However, the lack of a strong correlation between accuracy and confidence for the merger tool warrants further investigation. While a majority of subjects preferred the hypermap-merger tool for this type of bivariate data representation, subjects were split on their least favourite tool. In summary, we can effectively represent spatial data and data quality information by presenting both sets of data, either in individually or in tandem, through exploratory visualization tools that allow users to identify, compare, contrast, and problem solve using the data.

How can the dynamic interactive capabilities of a digital mapping environment be utilized to design more effective map displays than were possible in a static environment? An online environment affords tremendously opportunities to influence the interpretation and subsequent understanding of spatial data through the dynamic and interactive capabilities of the medium. However, simply adding interactivity does not guarantee any improvement over static displays, and, in many cases, may even detract from display effectiveness. New tools must be designed based on theories about how humans perceive and think about spatial information, as well as on practical guidelines taken from solid experimental studies. A traditional static or less-interactive tool (i.e., side-by-side tool) was included in the experiment to serve as a baseline. A version of a previously developed sequencing tool (i.e., toggle tool) with user control over map viewing frequency was included in order to act as a contrast and to compare against earlier studies. Two unique tools were designed and implemented for this experimental study. In contrast to several cartographic design experiments that have been critiqued for their lack of a theoretical component, the design of these two new tools was based on ideas about how users might visually perceive and think about visual information in the exploration process. The merger tools were created based on the notion that user controlled colour desaturation and enhanced interactivity could provide the means for users to visually explore spatial data and data quality in order to gain better understanding of the data and make better decisions. The idea that "greyed" out colour might be used to communicate uncertainty was verified. The idea that increased user control and interaction might allow users to form better mental models and thus make more accurate decisions was also verified. It has been shown that users engaged in spatial tasks often supplement spatial with verbal strategies for problem solving. The success of the hypermap-merger tool may, in part, be attributed to the addition of text-based attribute information that could be used to explore and verify information presented through shape, colour and location. While it is difficult to separate out specific factors that contributed to the increased performance in the more interactive merger and hypermap-merger groups, some general design and research guidelines can be gleaned from this and other experiments.

Guidelines for software developers:

1. Increasing the users' ability to interact in a meaningful way with data will help increase understanding of spatial information.
2. Facilitating user control, choice and adaptation of tools are important considerations in the design process.
3. The design of all aspects of visualization tools should be based on theories of how users perceive and process spatial information.
4. All new designs should be tested for both usability (recall pilots one and two) and effectiveness (main study) and the design iteratively improved.
5. Training (and online help), which explains both concepts and functionality, is an important feature that must accompany new and unfamiliar tools.
6. Users choice and preference should be accommodated by creating adaptive tool boxes that can be used in a variety of modes and are matched to common map-use tasks.
7. Controlled experiments should test a range of tasks that resemble, or form a component of, those found in the "real world."
8. Subjects should be chosen to resemble "real world" map users as much as possible, so that results can be generalized to the appropriate population.
9. The test environment should be simple and standardized.
10. Task context and user motivation are other key factors required to ensure that the results are ecologically valid.
11. Confidence is a valuable measure of performance, and should be included as a measure of effectiveness when users' confidence plays an important factor in the outcome of the map-use task.
12. Setting an allowable time limit can introduce difficulties in comparing different task types, but this approach may be more valuable that measuring the response time.
13. User feedback and observational data are invaluable aspects of user testing.

And lastly, *How well do these displays communicate spatial uncertainty information?* Like McGranaghan (1996) and others, the approach used here assumes that one of the best ways to improve the design of interactive maps is through controlled user experiments. However, few studies have found significant results that can be applied to the design of visualization tools. This study is one of a handful that examines the effectiveness of representations that display both spatial data and spatial uncertainty information. In order for a display to be effective, both sets of data must be accessible. All of the tools used in this experiment present this bivariate data, either through linked individual views or in tandem. Map-use effectiveness was examined through time controlled subject responses (accuracy and confidence) to a range of tasks and a post-test preference questionnaire. The experimental results show that both merger tools increased subjects' ability to accuracy examine and interpret spatial data and associated uncertainty information when performing all kinds of tasks. The advantage afforded by the merger tools was greatest for specific identification tasks. While some users found the use of desatuation to simultaneously represent thematic and certainty data confusing, results show optimal performance with the merger tools. Having multiple avenues of accessing data is possible in digital world and best accommodates individual styles and preferences, and different tasks. In terms of user performance (accuracy and confidence) and preference, this study showed that the merger tools are very effective at representing spatial uncertainty information.

7.2 Limitations

In this study, I used three measures (accuracy, confidence and preference) to evaluate map-use effectiveness on a range of tasks involving a specific dataset with four visualization tools. There are many other effectiveness measures, tools, map and task types that could be examined in an experimental setting. The permutations are endless. Accuracy is a standard measure of map-use effectiveness. Confidence and preference are important measures. The use of self-rated feelings of confidence has been shown to be problematic (Zakay 1997). By controlling and testing for as many factors as possible that might affect confidence ratings (e.g., simple test environment and display, explicit training and explanation), I was at least sure that none of the tools created obviously unwarranted feelings of under- or over-confidence in the users' ability to solve the tasks. Other measurement options that could have been used include response time, user satisfaction, and user acceptance.

The tasks used in this study were all relatively simple compared to the complexity of real world land use conflicts, which are often a combination of many sub-tasks, much more data and more conflicting interests. Resource management involves a huge range of map-based activities, many of which occur in a digital environment (see section 2.6). By including two data extraction task types and one decision activity (Tan and Benbasat 1993), it was hoped that the results would be applicable to a range of simple and complex map-use tasks that form a part of managing natural resources. For example, the exploration of spatial locations, distributions and patterns of thematic information forms an integral part of many management activities (e.g., inventory planning, change monitoring). The inclusion of uncertainty information into these explorations cannot be undervalued and forms an important component of a variety of tasks (e.g., identifying the need for updated data, identifying missing data, examining results for biases, isolating areas which may require further investigation). There is a place for all of the exploratory visualization tools examined in this study in many software packages including general mapping packages, GISs, spatial multiple criteria decision-making models and spatial decision support systems.

By embedding the tasks in an imaginary land-use conflict, context and motivation were at least to some degree addressed. In addition, subjects reported enjoying the story-telling aspect. The tasks used also made explicit reference to the certainty map, a situation not likely to occur in the real world. Tools that enhance accessibility and draw attention to data quality information are important, as it has been shown that many users ignore data quality (e.g., Evans 1997). The assumptions associated with the tasks imply a simplification that is also unrealistic (e.g., flat topography, similar fire history). It is more likely that a series of maps would exist and be used in this type of decision making. The tools presented in this study would be equally useful for a series of bivariate maps as a single pair. User controlled combinations of themes could aid decision-making.

The merger tools were designed specifically for choropleth maps, with thematic and certainty data in mind. The thematic data was nominal and had to be represented using seven hues that were distinct under three levels of desaturation. This classification requirement could be very limiting in cases where the data type was ordinal or quantitative, the colour scheme was chosen for pre-existing reasons or the number of classes was significantly larger. In addition, it is not always possible to code data certainty information in three levels. It is likely that using more than three levels of desaturation would have confused users, making the tools less useful. The data volume and scale (i.e., two maps covering a 5 x 5 km geographic region) are not realistic in a digitial map-use world. Despite these criticisms, the results of this study are an important step forward in terms of determining solid methodologies for interactive visualization tool design and cartographic experimentation.

7.3 Future Research

There are many permutations of effectiveness measures, tasks, tools, data sets and user groups that could be experimentally tested. There are also as many designs for visualization tools as there are imaginations to create them. Which are the most valuable? The experimental setting of this study had some limitations that could be immediately addressed in order to fine-tune the design of the toggle and merger tools. The tools should be tested on an expanded user group (experts and novice users) over a wider range of tasks. In addition to response data, observational techniques and focus groups could be employed in order to gain further insight into possible enhancements. It would be invaluable to directly address the visual, verbal and cognitive strategies employed by users to solve particular tasks. This could be done by using protocol analysis or by simply having users verbalize their thoughts, the reasons for their actions and difficulties as they proceed. This information could then be used to further tailor the design of tools to specific spatial and/or verbal strategies (in addition to matching tool design to task types). The task set could also be expanded to include memory and recall tasks, and a wider range of data extraction tasks and decision activities. Results from previous map-use studies involving memory and recall tasks, have shown differences between appropriate colour schemes, map complexity (as represented by the number of classes) and performance (e.g., Mersey 1990). Although exploratory tasks tend to involve direct tasks, I cannot anticipate all the uses of these tools and it is possible they are effective for other kinds of tasks. The tools should also be tested on an expanded data set that could include other choropleth maps, different classification and colour schemes, and other types of map presentations (e.g., perspective maps such as DEM with certainty data).

7.4 Last Words

The design of tools that allow effective visual exploration of spatial data quality information in a resource management setting is critical if decision-makers and policy setters are to make accurate and confident decisions that will have a positive long-term impact on the environment. This study is one step in that direction and contributed significantly to the existing body of research. In this study, I designed and implemented two visualization tools based on ideas about colour theory and spatial cognition. In order to test their effectiveness, I developed a methodology that used a range of tasks that resembled simple tasks found in a real world resource management setting, verified the use of confidence as a measure of map-use effectiveness, provided a context and motivation for subjects, showed limited affects of individual differences through careful experimental design, and produced both statistically-based and anecdotal information that could be generalized and used by GIS and cartographic software developers to design effective visualization tools.

I hope that this study will guide software developers to create practical and innovative cartographic visualization tools for spatial data and metadata, and act as a catalyst for further experimental cartographic studies into their appropriate use.

BIBLIOGRAPHY

Antle, A.N. and Klinkenberg, 1999. Shifting paradigms: from cartographic communication to scientific visualization. *Geomatica*, 53(2): 149 -155.

Arney, W.R. 1990. Understanding Statistics in the Social Sciences. W.H, Freeman and Company, NY.

Aspinall, R.J. and Pearson, D.M. 1995. Describing and managing uncertainty of categorical maps in GIS. In: Fisher, P. *Innovations in GIS 2*, Taylor and Francis Ltd, London.

Beard, M.K. and Buttenfield, B.P. 1991. *NCGIA Research initiative 7: Visualization of spatial data quality*. NCGIA Technical Paper: 91-26.

Beard, M.K., Buttenfield, B.P. and Clapham, S.B. 1991. Visualization of spatial data quality: scientific report for the specialist meeting, Santa Barbara, CA: NCGIA Technical Paper: 26 – 91.

Beard, M.K. and MacKaness, W. 1993. Visual access to data quality in geographic information systems. *Cartographica,*30(2-3): 37 - 45.

Beard, M.K. and Buttenfield, B. 1999. Detecting and evaluating errors by graphic methods: In: P. Longley, M. Goodchild, D.Maguire and D.Rhind (eds.) *Geographical Information Systems: Volume 1: Principles, Techniques and Issues.* Wiley and Sons, NY: 219 - 233.

Benbasat, I., Dexter, A.S., and Todd, P. 1986. An experimental program investigating color-enhanced and graphical information presentation: an integration of findings. *Communications of the ACM*, November, 29: 1094 – 1105.

Bertin, J. 1981. Graphics and graphic information processing. Berlin and New York: deGruyter.

Bertin, J. 1983. Semiology of graphics. Madison: University of Wisconsin Press. (Translation by W.J.Berg of Semiologie Graphique, 1967)

Blades, M. and Spencer, C. 1986. The implications of psychological theory and methodology for cognitive cartography. *Cartographica*, 23: 1 - 13.

Bracken, I. 1994. Towards improved visualization of socio-economic data. In: Hearnshaw, H.H. and Unwin, D.J. (eds.) *Visualization in Geographical Information Systems*. Chichester, John Wiley & Sons.

Bray, J.H. and Maxwell, S.E. 1985. Multivariate Analysis of Variance. Sage Publications, Beverly Hills, CA.

Brewer, C.A. 1994. Color Use Guidelines for Mapping and Visualization, In: MacEachren, A.M. and Taylor, D.R.F. (eds.) *Visualization in Modern Cartography,* Pergamon/Elsevier Science, New York.

Brodley, C. E., T. Lane and T. M. Stough. 1999. Knowledge discovery and data mining. *American Scientist*, 87(1): 54 - 61.

Buttenfield, B.P. and Beard, M.K. 1992. Graphical and geographical components of data quality. In: Hearnshaw, H.H. and Unwin, D.J. (eds.) *Visualization in Geographical Information Systems.* Chichester: John Wiley and Sons.

Buttenfield, B. 1993. Representing data quality, *Cartographica,* 30(2 -3): 1 - 7.

Carter, J.R. 1988. The map viewing environment: A significant factor in cartographic design. *The American Cartographer*, 15(4): 379 – 385.

Castner, H.W. and Eastman, J.R. 1984. Eye-movement parameters and perceived map complexity – I. *The American Cartographer*, 11(2): 107 - 117.

Castner, H.W. and Eastman, J.R. 1985. Eye-movement parameters and perceived map complexity – II. *The American Cartographer,* 12(1): 20 - 40.

Chang, K. 1980. Circle size judgement and map design. *The American Cartographer,* 7: 155 - 162.

Cleveland, W.S. and McGill, M.E., 1988. *Dynamic graphics for statistics.* Wadsworth and Brooks/Cole Advanced Books and Software, Belmont, CA.

Cumming, S.G., Burton, P.J., Joy, M., Klinkenberg, B., Schmiegelow, F.K.A. and Smith, J.N.M. 1995. *Experimental Habitat Fragmentation and Simulation of Landscape Dynamics in the Boreal Mixedwood: A Pilot Study.* Joint publication of Canadian Forest Service and Land and Forest Service pursuant to the Canada-Alberta Partnership Agreement in Forestry.

DiBiase, D., MacEachren, A.M., Krygier, J.B. and Reeves, C. 1992. Animation and the role of map design in scientific visualization. *Cartography and Geographic Information Systems,* 19(4): 201 – 214, 265 – 266.

Dorling, D. 1994. Cartograms for visualizing human geography. In: Hearnshaw, H.H. and Unwin, D.J. (eds.) *Visualization in Geographical Information Systems.* Chichester, John Wiley & Sons.

Dorling, D. 1995. Visualizing changing social structure from a census. *Environment and Planning A*, 27: 353 - 378.

Dorling, D. and Openshaw, S. 1992. Using computer animation to visualize space-time patterns. *Environment and Planning B: Planning and Design*, 19: 639 - 650.

Dykes, J. 1994: Area-value data: new visual emphases and representations. In: Hearnshaw, H.H. and Unwin, D.J. (eds.) *Visualization in Geographical Information Systems.* Chichester, John Wiley & Sons.

Earnshaw, R.A. and Wiseman N. (1992) An Introductory Guide to Scientific Visualization, Springer-Verlag, New York.

Egbert, S.L. and Slocum, T.A. 1992. EXPLOREMAP: an exploration system for choropleth maps. *Annals, Association of American Geographers*, 82: 275 - 288.

Evans, B.J. 1997. Dynamic display of spatial data reliability: does it benefit the map user? *Computers & Geosciences,* 23(4): 409 - 422.

Fauerbach, E. D., Edsall, R.M., Barnes, D. and MacEachren, A. 1996. Visualization of uncertainty in meteorological forecast models. *Advances in GIS Research II, Proceedings 7th International Symposium on Spatial Data Handling,* August 12 - 16, Delft, the Netherlands, 1: 17 - 28.

Fisher, P.F. 1992. Animation and sound for the visualization of uncertain spatial information. In: Hearnshaw, H.H. and Unwin, D.J. (eds.) *Visualization in Geographical Information Systems*, Chichester: John Wiley and Sons: 181-185.

Fisher, P.F. 1993. Visualizing uncertainty in soil maps by animation. *Cartographica*, 30: 20 - 27.

Fisher, P.F. 1994. Animation and sound for the visualization of uncertain spatial information. In: Hearnshaw, H.H. and Unwin, D.J. (eds.) *Visualization in Geographical Information Systems*, Chichester, John Wiley & Sons.

Fisher, P.F. 1996. Animation of reliability in computer-generated dot maps and elevation models. *Cartography and Geographic Information Systems*, 23(4): 196 –205.

Fisher, P., Dykers, J. and Wood, J. 1993. Map design and visualization. *The Cartographic Journal*, 30: 136 – 142

Gatrell, A. 1994. Density estimation and the visualization of point patterns. In: Hearnshaw, H.H. and Unwin, D.J. (eds.) *Visualization in Geographical Information Systems.* Chichester, John Wiley & Sons.

Gilmartin, P. 1981. The interface of cognitive and psychological research in cartography. *Cartographica*, 18(3): 9 - 20.

Gilmartin, P. 1986. Maps, mental imagery, and gender in the recall of geographic information. *The American Cartographer*, 13: 335 - 344.

Gilmartin, P. and Patton, J.C. 1984. Comparing the sexes on spatial abilities: map-use skills. *Annuls of the Association of American Geographers*, 74(4): 605 – 619.

Goodchild, M. 1988. Stepping over the line. *The American Cartographer.*

Goodchild, M.F. 1991. Issues of quality and uncertainty. In: Muller, J.C. (ed.) *Advances in Cartography*, Elsevier Applied Science Series, New York.

Goodchild, M. and Gopal, S. 1989. *Accuracy of spatial databases.* Taylor and Francis, Ltd., London.

Goodchild, M., Chih-Chang, L. and Leung, Y. 1994. Visualizing fuzzy maps. In: Hearnshaw, H.H. and Unwin, D.J. (eds.) *Visualization in Geographical Information Systems.* Chichester, John Wiley & Sons.

Hall, S.S. 1992: Mapping the Next Millennium: The Discovery of New Geographies. New York: Random House.

Hart R.A. & Moore G. 1973. The development of spatial cognition: a review. R. M. Downs and D. Steas (eds.) *Image and Environment*, Chicago, IL. Aldine: 246 - 288.

Haslett, J., Willis, G., and Unwin A. 1990. SPIDER: an interactive statistical tool for the analysis of spatially distributed data. *International Journal of Geographical Information Systems*, 4: 285 - 296.

Hocking, D. and Keller, C.P. 1992. A user perspective on atlas content and design. *The Cartographic Journal*, 29: 109 - 117.

Howard, D. and MacEachren, A.M. 1996. Interface design for geographic visualization: tools for representing reliability. *Cartography and Geographic Information Systems*, 23(2): 59 – 77.

Huber, P.J. 1983. Statistical graphics: history and overview. *Proceedings of the Fourth Annual Conference and Exposition of the National Computer Graphics Association*, Fairfax, VA: 667 - 676.

Jiang, B, Brown, A and Ormeling, F.J. 1996. Some perceptual aspects of colouring uncertainty. *Advances in GIS Research II, Proceedings 7th International Symposium on Spatial Data Handling,* August 12 - 16, Delft, Netherlands, 1(8A): 29 - 42.

Joy, M. and B. Klinkenberg. 1996. Managing uncertainty in forest inventory data: analysis and visualization. M. Heit, H. D. Parker, and A. Shortreid (eds.) *GIS Applications in Natural Resources 2.* Fort Collins, Colorado, GIS World Books: 145 - 153.

Keates, J.S. 1993. Some reflections on cartographic design. *Cartographic Journal*, 30(2): 199 - 210.

Kitchin, R.M. 1994. Cognitve maps: what they are and why study them, *The Journal of Environmental Psychology,* 14: 1 - 19.

Klinkenberg, B. and Joy, M. 1994. Visualizing uncertainty: succession or misclassification? *GIS/LIS Proceedings*, 25 - 27 October, Phoenix, Arizona.

Koriat, A., Lichtenstein, S. and Fischhoff, B. 1980. Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory*, 6(2): 107 – 117.

Krygier, J.B. 1994. Sound variables, sound maps and cartographic visualization. MacEachren, A.M. and Taylor, D.R.F. (eds.) *Visualization in Modern Cartography.* London, Pergamon.

Legge, G., Y. Gu, and A. Luebker. 1991. Efficiency of graphical perception. S. R. Ellis, M. K. Kaiser, and A. C. Grunwald (eds.) *Pictorial communication in virtual and real environments.* London, Taylor and Francis: 111 - 130.

Lloyd, R. 1989. The estimation of distance and direction from cognitive maps. *Cartography and Geographic Information Systems,* 16(2): 109 - 122.

Lloyd, R. and Steinke, T.R. 1985. Comparison of qualitative point sysmbols: the cognitive process, *The American Cartographer,* 12: 152 - 162.

Lloyd, R., Rostkowska-Covington, E. and Steinke, T. 1996. Feature matching and the similarity of maps. In: C. Wood and C.P. Keller (eds.*) Cartographic Design: Theoretical and Practical Perspectives*, John Wiley and Sons, London.

MacEachren, A.M. 1982. The role of complexity and symbolization method in thematic map effectiveness. *Annals of the Asociation of American Geographers,* 72(4): 495 - 513.

MacEachren, A.M. 1991. The role of maps in spatial knowledge acquisition. *The Cartographic Journal,* 28(2): 152 - 162.

MacEachren, A.M. 1991b. Visualization of data uncertainty: representational issues. *NCGIA Specialist Meeting on Visualizing the Quality of Spatial Data*, June 1991.

MacEachren A.M. 1992b. Visualizing uncertain information. *Cartographic Perspectives*, 13: 10 – 19.

MacEachren, A.M. 1993. Visualizing uncertain information. *Cartographic Perspectives,* 13: 11 - 17.

MacEachren, A.M. and Taylor, D.R.F. 1994 (eds.) Visualization in Modern Cartography, Pergamon/Elsevier Science, New York.

MacEachren, A.M. 1994b. Some Truth with Maps. Association of American Geographers, Washington, D.C.

MacEachren, A. 1994c. How Maps Work: Issues in Representation, Visualization and Design, Guilford Press, New York.

MacEachren, A.M. and J. H. Ganter. 1990. A pattern identification approach to cartographic visualization. *Cartographica*, 27(2): 64 - 81.

MacEachren A.M., Bishop, I., Dykes, J., Dorling D., and Gatrell, A. 1992. Introduction to advances in visualizing spatial data. In: Hearnshaw, H.H. and Unwin, D.J. (eds.) *Visualization in Geographical Information Systems*. Chichester: John Wiley and Sons.

MacEachren, A., Howard, D., von Wyss, D. and Taormino, T. 1993. Visualizing the health of Chesapeake Bay: an uncertain endeavor. GIS/LIS Proceedings, 2 - 4 November, Minneapolis, Minnesota.

MacEachren A.M., Brewer, C.A. & Pickle L.W. 1998. Visualising georeferenced data: representing reliability of health statistics. *Environment and Planning A,* 30: 1547 - 1561.

McGranaghan, M. 1988. Symbolizing quantitative differences on colour CRTs. Report Number 1115, Naval Submarine Medical Research Laboratory, Submarine Base Groton, CT.

McGranaghan, M. 1989. Ordering choropleth map symbols: the effect of background. *The American Cartographer*, 16(4): 279 - 285.

McGranaghan, M. 1996. An experiment with choropleth maps on a monochrome LCD panel. In: C. Wood and C.P. Keller (eds.) *Cartographic Design: Theoretical and Practical Perspectives*, John Wiley and Sons, London.

Marcus, A. 1995. Principles of effective visual communication for graphical user interface design. In: Baecker, R.M., Grudin, J., Buxton, W.A.S. and Greenberg, S. (eds.)*Readings in Human-Computer Interaction: Toward the Year 2000*, Morgan Kaufmann Publishers, Inc., San Francisco, CA.

Mersey, J. 1990. Colour and thematic design: the role of colour scheme and map complexity in choropleth map communication. *Cartographica*, 27(3): Monograph 41.

Moellering, H. 1976. The potential uses of a computer animated film in the analysis of geographical patterns of traffic crashes. *Accident Analysis and Prevention*, 8: 215 - 227.

Monmonier, M. 1980. The hopeless pursuit of purification in cartographic communication. *Cartographica*, 17(1): 24-37.

Monmonier, M. 1991. How to Lie with Maps, Chicago, The University of Chicago Press.

Monmonier, M. 1992. Summary graphics for integrated visualization in dynamic cartography. *Cartography and Geographic Information Systems,* 19(4): 247 - 260.

Monmonier, M. 1992b. Authoring graphic scripts: experiences and principles. *Cartography and Geographic Information Systems*, 19: 247 - 260.

Monmonier, M. 1993. Exploring the quality of enumeration-area data with graphic scripts. *Cartographica*, 30: 28 - 36.

Monmonier, M. 1994. Minimum-change categories for dynamic temporal choropleth maps. *Journal of the Pennsylvania Academy of Science*, 69: 42 - 47.

Monmonier, M. 1996. Cartographic complementarity: objectives, strategies and examples. In: C. Wood and C.P. Keller (eds.) *Cartographic Design: Theoretical and Practical Perspectives*, John Wiley and Sons, London.

Monmonier, M. and Gluck, M. 1994. Focus groups for design improvement in dynamic cartography. *Cartography and Geographic information Systems*, 21(1): 37 – 47.

Morrison, J.L. 1995. Spatial data quality. In: Guptill, S. C. and Morrison, J.L. (eds.) *Elements of Spatial Data Quality*, Elsevier Science, Oxford, UK.

Moser, C. and Kalton, G. 1979. *Survey methods in social investigation* (2nd ed.). Aldershot, Gower.

Nelson, E.S. and Gilmartin, P. 1996. An evaluation of multivariate quantitative point symbols for maps. In: C. Wood and C.P. Keller (eds.) *Cartographic Design: Theoretical and Practical Perspectives*, John Wiley and Sons, London.

Nielson, J. 1994. Usability Engineering, Morgan Kaufmann, SF.

Norcio, A.F and Stanley, J. 1989. Adaptive human-computer interfaces: a literature survey and perspective. IEEE Transactions on Systems, Man and Cybernetic, 19(2).

Olson, J. 1981. Spatially encoded two-variable maps. *Annals of the Association of American Geographers*, 71(2): 259 - 276.

Openshaw, S., Waugh, D. and Cross, A. 1994. Some ideas about the use of map animation as a spatial analysis tool. In: Hearnshaw, H.H. and Unwin, D.J. (eds.) *Visualization in Geographical Information Systems,* Chichester, John Wiley & Sons.

Oz E., Fedorowitz, J. and Stapleton, T. 1993. Improving quality, speed and confidence in decision making. Information and Management, 24: 71 - 82.

Parsons, E. 1995. GIS visualization tools for qualitative spatial information. In: Fisher, P. *Innovations in GIS 2*, Taylor and Francis Ltd, London.

Patton, D.K. and Cammack, R.G. 1996. An examination of the effects of task type and map complexity on sequenced and static choropleth maps. In: C. Wood and C.P. Keller (eds.) *Cartographic Design: Theoretical and Practical Perspectives*, John Wiley and Sons, London.

Peterson, M.P. 1985. Evaluating a map's image. *The American Cartographer*, 12(1): 41 - 55.

Peterson, M.P. 1994. Cognitive issues in cartographic visualization. In: MacEachren, A.M. and Taylor, D.R.F. (eds.) *Visualization in Modern Cartography,* Pergamon/Elsevier Science, New York.

Petchenik B. 1977. Cognition in cartography. In: Guelke (ed.) Monograph No. 19. *Cartographica*, 117 - 128.

Petchenik B. 1983. A mapmaker's perspective on map design research 1950-1980. In: Graphic Communication and Design in Contemporary Cartography, D.R.F. Taylor (ed.) London: Wiley and Sons, Ltd., 37 - 67.

Phillips, J.L. 1996. How to Think About Statistics, Fifth Edition. Freeman and Company, USA.

Slocum, T.A., Robeson, S.H. and Egbert, S.L. 1990. Traditional versus sequenced chloropleth maps: an experimental investigation. *Cartographica* , 27(1): 67 - 88.

Slocum, T.A. 1999. Thematic Cartography and Visualization. Prentice Hall, Upper Saddle, New Jersey, USA.

Steinke, T.R. and Lloyd, R.E. 1981. Cognitve integration of objective choropleth map attribute information. *Cartographica*, 18(1): 13 - 23.

Schmiegelow, F.A.K. and Hannon S.J. 1993. Adaptive management, adaptive science, and the effects of forest fragmentation on boreal birds in Northern Alberta. *Trans. of the 58th North American Wildlife and Natural Resources Conference*, 584 - 598.

Schofield, N.J. and Kirby, J.R. 1994. Position location on topographical maps: effects of task factors, training and strategies. *Cognition and Instruction*, 12(1): 35 – 60.

Schweizer, D.M. and Goodchild F.M. 1992. Data quality and choropleth maps: an experiment with the use of color. *GIS/LIS Proceedings* (ASPRS/ACSM): Bethesda, MD., 686 – 699.

Shepard, R. N. 1978. The Mental Image. *American Psychologists*, 33(2): 125 - 137.

Sniezek, J.A. and Buckley, T. 1993. Becoming more or less certain. In: Castellan, N.J. (ed.) *Individual and Group Decision Making*, Lawrence Erlbaum Associates, Publishers, Hillsdale, New Jersey.

Spilhaus, A. 1991. Atlas of the World with Geophysical Boundaries, Philadelphia, American Philosophical Society.

SPSS for Windows – Help Files, 1990. Release 9.0.0.

Tan, J.K.H. and Benbasat, I. 1993. The effectiveness of graphical presentation for information extraction: a cumulative experimental approach. *Decision Sciences*, 24: 167 – 191.

Taylor, D.R.F. 1987. Cartographic communication on computer screens: the effect of sequential presentation of map information. *Proceedings of the 13th International Cartographic Conference*, Morelia, Mexico, 1: 591 - 611.

Thorndyke, P. 1981. Distance estimations from cognitive maps. *Cognitive Psychology*, 13: 526 - 550.

Thorndyke, P. and Stasz, C. 1980. Individual differences in procedures for knowledge acquisition from maps, Co*gnitive Psychology*, 12: 137 - 175.

Tobler, W.R. 1970. A computer movie simulating urban growth in the Detroit region. Economic Geography, 46: 234 - 240.

Tufte, E.R. 1983. The Visual Display of Quantitative Information, Cheshire, Connecticut.

Tufte, E. R. 1990. Envisioning Information, Graphics Press, Cheshire, Connecticut.

Tukey, J.W. 1977. Exploratory Data Analysis, Reading, Mass: Addison-Wesley.

Turk, A.G. 1990 Towards an understanding of human-computer interaction aspects of geographic information system, Cartography, 19(1): 31 - 60.

Unwin, D. J., J. A. Dykes, P. F. Fisher, K. Stynes, and J. D. Wood. 1994. WYSIWYG? Visualization in the spatial sciences. Paper presented at the AGI conference, Birmingham, 1994.

van der Wel, F.J.M., Hootsman, R.M. and Ormeling, F. 1994. Visualization of data quality. In: MacEachren, A.M. and Taylor, D.R.F. (eds.) *Visualization in Modern Cartography,* Pergamon/Elsevier Science, New York.

Weibel, R. and Buttenfield, B. 1992. Improvement of GIS graphics for analysis and decision-making. *International Journal of Geographical Information Systems*, 6(3): 223 - 245.

Wood, J. and Fisher, P.F. 1993. Assessing interpolation accuracy in elevation model. *IEEE Computer Graphics and Applications*, 13: 48 - 56.

Wood, M. 1993. Maps and mapmaking. *Cartographica*, 30(1): 1 - 9.

Wood, M. and Gilhooly K. 1996. The practitioner's view? A pilot study into empirical knowledge about cartographic design. In: C. Wood and C.P. Keller (eds.) *Cartographic Design: Theoretical and Practical Perspectives*, John Wiley and Sons, London.

Wright, J.K. 1942. Map makers are human. American Cartography Society of New York, reprinted *Cartographica*, Monograph 19, 1977: 8 - 25.

Zakay, D. 1997. Post-decisional confidence – can it be trusted? In: Ranyard, R., Crozier, W.R. and Svenson, O. (eds.) *Decision Making: Cognitive Models and Explanations*, Routledge, New York.

APPENDIX A -- TASKS

1.0 TASK TYPE ONE: SPECIFIC IDENTIFICATION TASKS

Question 1: Harvetsting aspen stands
Scenario
IMP Forest products wants to harvest only certain types of stands to supply a neighbouring pulp mill. Help IMP CEO Ms. Daniela Flop decide if she should tell her Director of Operations take more ground samples or go ahead with harvest planning by answering true or false to the following statement:

Statement
All the pure Aspen stands in the southwest quadrant of the region have high data certainty.
Answer: False

Question 2: Pico spruce gruff habitat
Scenario
Pico Poe Bird Watchers Association is planning an overnight outing to photograph the rare Pico Spruce Gruff. Help Mary Fidget, long-time member of the Pico Poe Bird Watchers, find a campsite located in the Pico Spruce Gruff's habitat by answering true or false to the following statement:

Statement
The southwest quadrant of the region is dominated by low certainty Aspen-Spruce mix.
Answer: False

Question 3: Bobo salamander riparian habitat
Scenario
The Bobo Riparian Preservation Society is planning to lobby for a protected riparian area for the rare Bobo salamander. Help Dr. Jessie Jones decide if the lake in this region meets the habitat requirements for the Bobo salamander by answering true or false to the following statement:

Statement
There is a stand of high certainty pure Black Spruce bordering the western shore of Calling Lake.
Answer: False

Question 4: Road construction
Scenario
Monster Construction Inc. has been hired by IMP Forest Products to build a road through the region. Help Mr. Bobby Smiley, Head of Operations, decide if he has enough information to start road construction right away by answering true or false to the following statement:

Statement
The central area of Aspen-Spruce mix has non-uniform data certainty.
Answer: True

Question 5: Conflict of interests and data verification
Scenario
Weasel Pulp and Paper is questioning the accuracy of the data they received from the Bobo Riparian Preservation Society. Help Mr. Bill Weiner, Head of Quality Assurance for Weasel Pulp and Paper, decide if he can trust the environmentalists' data by answering true or false to the following statement:

Statement
Just north of the center of the region there is a single stand of Spruce-Aspen mix that has low certainty surrounded by stands of Aspen-Spruce mix.
Answer: True

2.0 TASK TYPE TWO: GENERAL ANALYSIS TASKS

Question 6: Locating a pulp mill
Scenario
IMP Forest Products is considering building a pulp mill in the region. Help IMP CEO Ms. Daniela Flop decide which part of the region would be a good location for the new mill by answering true or false to the following statement:

Statement
Spruce-Aspen mix with either medium or high data certainty is the dominant species in the northeast quadrant of the region.
Answer: False

Question 7: Bird watching
Scenario
The Pico Poe Bird Watchers Association are planning a hike through the region to track the secretive male Poe Gruff as it chases the female Gruff in mating season. Help Mary Fidget, a Poe Gruff expert, plan the group's hike by answering true or false to the following statement:

Statement
Moving northward from the middle of the centermost southern logged area, the hikers will first pass through Spruce-Aspen mixed stands, Aspen-Spruce mixed stands and then pure Black Spruce with all non-logged stand types having either medium or low data certainty.
Answer: True

Question 8: Bobo salamander habitat
Scenario:
The Bobo Riparian Preservation Society is mapping the region to find stands that meet the habitat requirements for the rare Bobo salamander. Help Dr. Jessie Jones with his task by answering true or false to the following statement:

Statement
The largest contiguous area of pure Black Spruce stands with uniform certainty is located in the northwest corner of the region.
Answer: False

Question 9: Harvesting and stand certainty
Scenario
Weasel Pulp and Paper is looking to harvest this region for their pulp mill. Help Mr. Bill Weiner, Head of Quality Assurance, decide if he has accurate information for the entire region by answering true or false to the following statement:

Statement
In the entire region, pure Aspen stands have higher certainty overall (by area covered) than Aspen-Spruce or Spruce-Aspen mixed stands combined
Answer: True

Question 10: Road construction
Scenario
Monster Construction Inc. is beginning to build additional roads in the region to increase accessibility. In order to determine where the new roads should be located, they must be sure about the locations of existing logged areas and roads. Help Mr. Bobby Smiley, Head of Operations, locate the old logged areas and roads by answering true or false to the following statement:

Statement
All of the existing logged areas in the region have high data certainty.
Answer: False

3.0 TASK TYPE THREE: INTEGRATED TASKS

Question 11: Silvicultural treatments
Scenario
IMP Forest Products has now bought a Tree Farm License for the entire region. In order to maximize long-term income, the Director of Operations suggests a silvicultural treatment that includes thinning most pure Aspen stands now, to create income flow, and then fully harvesting these stands later for additional income. In order to minimize thinning costs, he wants to be sure to only thin marketable stands and so will ground check stands with low certainty. Help IMP CEO Ms. Daniela Flop decide if she should follow this silvicultural treatment plan by answering true or false to the following statement:

Statement
There is a sizable area that should be ground checked, after which most of the thinning should occur south and west of the logging road that runs through the southwest quadrant of the region, above the east-west running road, and along the western edge of Calling Lake.

Answer: True

Question 12: Protesting and forest tours
Scenario
The Pico Poe Bird Watchers Association is planning to picket the construction of the proposed pulp mill near the western edge of the central region, since this is the mating ground for the secretive Poe Gruff. In order to bring in as many supporters as possible, they are offering a forest tour on their way to the picket site. The tour should pass through as many stand types as possible. Since no one knows the roads well, they will only use roads that have high certainty. Help Mary Fidget, long-time activist, decide which road to bring the supporters in on by answering true or false to the following statement:

Statement
The supporters should be brought in on the road that runs from the south central region to the west central region.
Answer: True

Question 13: Preservation of the Bob Salamander
Scenario
Thanks to the Bobo Riparian Preservation Society, the rare Bobo salamander is now on the endangered species list. However, IMP Forest Products is planning to log the entire region, including Aspen mixed stands for pulp and Spruce mixed stands for saw wood. To minimize risk, IMP will only log areas that have high or medium data certainty. The Bobo salamander's habitat is Spruce-Aspen mixed Riparian stands along the shore of Calling Lake. Help Dr. Jessie Jones preserve all the stands scheduled for logging that contain Bobo habitat by answering true or false to the following statement:

Statement
There are eight stands that IMP plans to log that meet the habitat requirements for the Bobo salamander.
Answer: False

Question 14: Clearcutting
Scenario
Weasel Pulp and Paper made a deal with IMP Forest Products and now wants to clearcut all pure Aspen, Spruce-Aspen and pure Black Spruce stands in the region with high or medium certainty so their stock will increase. Help money-mad Bill Weiner maximize his short-term earnings (before the Bobo Riparian Preservation Society stops his plan) by answering true or false to the following statement:

Statement
To maximize immediate cash flow, harvesting should begin in the northwest corner of the region and proceed diagonally in the southeast direction.
Answer: False

Question 15: Harvesting from existing roads and lake

Scenario

The head quarters for Monster Construction Inc. was bombed by some unknown environmental group. Any road construction in the region will be delayed by six months. IMP Forest Products wants to immediately harvest all pure Aspen stands in the region that have high data certainty. They have an idea and ask Bobby Smiley, Head of Monster Operations, for his advice. Help Bobby keep his job by answering true or false to the following statement:

Statement

By using existing logging roads running through stands and barges on Calling Lake to access shoreline stands, over half the targeted stands can be removed before new roads are built.

Answer: True

APPENDIX B – TRAINING SCRIPT

1.0 Introduction

This study examines four different spatial interactive visualization tools that can be used for environmental decision-making with mapped data and metadata, where the map and metadata are displayed simultaneously in a Web-based environment.

For the main part of this session, you'll be using one of these four tools to answer a set of fifteen true or false questions. At the end, you'll have a chance to look at all four tools and comment on them on the worksheet.

2.0 Maps

There are two maps. The first is a species composition or forest class map of an area of mixed wood forest in Northern Alberta. There are seven classes coded as seven distinct colours. The forest classes are: pure aspen, aspen-spruce, pure spruce, spruce-aspen, other forested, logged and roads and water/Calling Lake. There will be a legend under the map.

The second map represents metadata in the form of certainty or accuracy about the first map. For each stand in the first map, the second map shows how accurate the data is. This is done using a three point certainty scale. If the data is very accurate (from ground checking) then it is classed as high certainty and coded a light grey colour — for an aspen stand with high certainty -- highly certainty means we are absolutely sure it is really pure aspen and not any other class, not aspen-spruce mix, not spruce, not logged, not any other forest type. Medium certainty means we are mostly certain it is pure aspen. Low certainty -- coded dark grey -- means we're not sure at all — it could likely be something else. Think of dark as it's hard to see what's going on -- we don't know what's there -- the data has low certainty. So the second map shows these three levels of certainty about the first forest class map. Light grey for high certainty — dark grey for low.

3.0 Assumptions (on board)

When answering the questions, assume

1. Flatish topography
2. Mature stands -- all stands are ready to harvest
3. Species composition information is for canopy.
4. Assume north is to the top of the screen.

Any questions?

4.0 Interfaces and Tools (on LCD overhead projector)

Each question page has several parts. The map display for the two maps. Some text that describes the scene. A question you'll be answering. Place to answer using radio buttons (T or F). A place to give your confidence rating 1 through 5. At the top will be a countdown timer and a question out of display. I'll explain all these parts as we go.

Sometimes the map display loads slowly. Do NO CLICK RELOAD. Just wait. If something looks weird — say the colours are messed up — ask me to come over.

Side-by-Side
Having the two maps side by side is the traditional solution. It's not too elegant so you'll have to use the scroll bar. This is intended.

Toggle
Use the button labeled Toggle to switch back and forth between the two maps.

Data Merger
How the data merger tool works is this. Both map are displayed, one overlaid on top of the other. With the slider at the far left endpoint, all you see is the coloured forest class map. Now, how to show the other grey certainty map? One way to show certainty information, is to desaturate (desaturate means to make the colour duller) colours by an amount that depends on how certain the data is. If the data is highly certain (i.e., accuracy is high), you only desaturate or grey out the colour a little bit. If the data has low certainty (i.e., accuracy is low), you grey out the colour or darken it a lot. So if the forest class is represented by yellow — a light yellow would represent this class and high certainty. A dark yellow means this class and low certainty. By moving the slider to the right, the coloured map is greyed-out using the certainty map to show both maps at the same time. At the far left, only the grey uncertainty map is shown.

There's a place about 2/3 of the way to the right where it is easy to distinguish all seven colour classes and the three greys. (Show them this.) So you can see forest class and certainty maps at the same time. Some questions will be easier if you can do this, others will require you to look at one map or the other. That is, moving the slide to the endpoints.

Hypermap-Merger
In addition to the data merger, your map display is a hypermap. This means, if you click on a stand (polygon) with your mouse, the forest class and certainty will be displayed textually in the bar at the bottom of your map display. So you can use the data merger to look at overall patterns and the mouse click to determine exact stand attributes.

5.0 Tasks
There are fifteen questions, some easier than others. The questions are presented in two parts. First you'll read a scenario — this explains who the parties involved in the landuse conflict are and what their interests are. In each question, you will represent the interests of one group member. So you always want to answer in their interest. That is don't purposely answer wrong to mess up a party you don't like! After you read the scenario, you'll see a T or F statement. You'll answer T or F to this using the radio buttons at the bottom of the page. Make sure the button is clicked in. You won't have much time in some cases. The idea is to answer as quickly and accurately as you can.

Confidence

After you answer T or F to the statement, you must also rate how confident you are in your answer. The rating is a scale of 1 to 5. So 1 means you are super confident — you are absolutely certain you got the right answer. You would put money on it, a lot of money. 2 — means you are mostly certainty but wouldn't bet your life savings on it. 3 — you are kind of certain. 4 — you have serious doubts about your answer. 5 — a complete guess, you have no idea or you are about to run out of time. If you guess, either because you don't know, or you are about to run out of time, please click on 5. (Put this on board)

NOTE: the confidence rating is about how confident you are in your answer.

6.0 Timing

You'll have a set amount of time for each question, so answer as quickly as you can. You'll see the time in the countdown timer at the top of your screen. The timer shouldn't start counting down until the images/applets are fully loaded. If it does put up your hand and I can restart you or make a note of it. When the timer reaches 0, the next question will be loaded and your answer will be automatically submitted. So, to answer in time, you'll need to select your answer and confidence rating before the timer reaches 0:00:00. So you'll have to keep an eye on the timer and when there's about 10 seconds left, click your answer and your confidence rating.

This study is designed to evaluate the tools, not your map use skills. Please answer as accurately as possible in the time allotted. It is more important that you answer something and rate your confidence low than you don't answer at all.

7.0 Submitting

After you answer T or F and select your confidence rating using the radio button, click the SUBMIT button. Don't click reset. And make sure the buttons are clicked before you click submit. ONLY click submit once ... then be patient ... something will happen. Do NOT click it again if nothing happens, raise your hand and I'll come over. No room for click happy mousers here.

If you are running out of time, just click you answer and confidence and don't worry about submit ... when the timer runs out, it will automatically submit your answers.

8.0 Demographics

Before you start the questions, you'll get a demographics page. Please fill in your name, age etc ... fill in all the boxes. English? means Is English your native language? Colour Blind? means Are you colour blind? If you don't know, click ? Fill in your department and program, computer experience and Internet experience. In the text box fill in any experience you have in forestry or land-use. Do not use <CR> just type in and erase the note there first! Click SUBMIT (once) when you are done to start the questions. The first question may take a moment to load in.

9.0 Follow-up

At the end, you'll get a sheet with hot links to the four tools. Please go to each tool, check it out for a few minutes and then, keeping in mind all the questions you just answered, pick the one you prefer and the one you least prefer. Please explain why in a few sentences and then answer the other questions for each tool. Ask me if you have any questions. When you are done, hand in your worksheet. THANKS.

10.0 Reminders (on board)

Check that you clicked radio buttons.
DO NOT RELOAD.
Keep an eye on the timer

Plus assumptions.
Confidence rating.
Outline. (Training/introduction, demographics, 15 questions, follow-up on worksheet).

APPENDIX C – SAMPLE TOOL EVALUTION (PREFERENCE) WORKSHEET

Worksheet for Alissa's Interactive Visualization Study

Name ________________________________
E-mail: ________________________________

Course: ________________________________

Session:	4a	Tuesday Nov 18 11:30 - 12:30
(circle one)	4b	Tuesday Nov 18 12:30 - 1:30
	5a	Tuesday Nov 18 1:30 - 2:30
	5b	Tuesday Nov 18 2:30 - 3:30
	6a	Tuesday Nov 18 3:30 -4 :30
	6b	Tuesday Nov 18 4:30 -5 :30

Display Preference:
Which tool did you like the MOST? Tool 1 Tool 2 Tool 3 Tool 4
(circle only one) Why?________________________________

Which tool did you like the LEAST? Tool 1 Tool 2 Tool 3 Tool 4
(circle only one) Why?________________________________

What did you like/dislike about each tool? What would you change about them? Did each make it easier to answer questions? Harder? What types of questions do you think are suited to each tool? What would you change about each tool? What other ways could you display the maps simultaneously?

Tool 1: Side-by-side
Comments:________________________________

Tool 2: Toggle
Comments:________________________________

Tool 3: Data Merger
Comments:________________________________

Tool 4: Data Merger + Hypermap

Comments:____

Please return to Alissa Antle.

APPENDIX D – SAMPLE LABS HANDOUTS

G315 - LAB

ONLINE FOREST INVENTORIES

Lab Objective

The objective of this lab is to introduce you to issues concerning the use of online forest inventory maps and to familiarize you with interactive tools that can be used to aid forestry-related decision-making in an online environment.

Technical Notes

First, use the QuickRes icon to set your screen resolution to 1024 x 768. If you double click the QuickRes icon on your desktop it will appear in the bottom right corner of your screen. Single click this mini-icon and then set the resolution. You will be using the Netscape web browser. Open Netscape by double-clicking on the Netscape icon on your desktop. Once Netscape has opened, (this may take a second) turn off the tool bar and directory buttons by de-selecting them under the Options menu in the top menu bar (de-select = there should NOT be a check mark next to them). Click in the Netscape location bar to type in the URLs listed below as outlined in the lab. The first URL will be http://www.geog.ubc.ca/~antle/dissertation/test1/showb.htm

Introduction

This exercise consists of using forest inventory information for forest decision making (specifically, a land-use conflict) in an online (Internet) environment. As more map data becomes available through the Internet and on the Web, decision makers are increasingly using this online information. Since the Internet map producer and map user are rarely the same person, it is crucial to have metadata. Metadata is data about data. It should document the full history of the map including a reference to the original data source, the steps taken and assumptions made during map production, spatial and attribute accuracy information, currency, scale and resolution. If this information is available for an online map, the question remains: How can this information be integrated into the decision-making process?

Internet tools which facilitate interaction with online maps are rapidly being developed. In this lab you will explore four interactive tools which simultaneously display a forest inventory map and attribute accuracy information about that map in an online environment.

Attribute accuracy refers to the accuracy of thematic class assignment (versus spatial accuracy which refers to the accuracy of class boundaries or location). For example, for a stand of aspen trees, attribute accuracy concerns the accuracy of assigning the forest inventory class aspen while spatial accuracy concerns the location and boundary of the stand itself. In reality, both the attribute class aspen and the spatial boundary of the stand marked on a map are only man-made generalizations of a non-homogenous continuously changing mixedwood forest taken in a snap-shot of time.

SECTION ONE *The Maps: Forest Inventories and Metadata*

Uncertainty mapsUncertainty mapsUncertainty maps

You will be using two maps of the same area in this lab. The first is a forest inventory or forest class map of an area of mixed wood forest in Northern Alberta. The original map data has been generalized into seven classes useful land-use allocation conflict between conservation, and pulp and saw wood harvesting. The seven classes are coded as seven distinct colours (for reasons that will be clear later). The forest classes are: pure aspen, aspen-spruce, pure spruce, spruce-aspen, other forested, logged and roads and water/Calling Lake. The legend on the online map will show which colours correspond to which forest classes.

When creating the forest map some assumptions were made. Therefore, when making decisions assume that the map represents the following:

1. Flat topography -- all stands are equally accessible with respect to hill slope.
2. Mature stands -- all stands are ready to harvest.
3. Species composition information is for the forest canopy.
4. Assume north is to the top of the map.

The second map represents **metadata** in the form of accuracy information about the first map. For each stand in the first map, the second map shows how accurate the forest inventory classification data is. This information is determined by manual ground-truthing the aerial maps that were used to create the digital forest inventory map. This information is crucial when digital maps are being used for decision-making, especially when the decision-makers are not familiar with the actual area and uncertainty must be incorporated into the decision-making process.

The accuracy (in this case data certainty) map in this lab is classified based on a **three point certainty scale**. If the data is very accurate then it is classed as "high certainty" and coded a light grey colour. For example, an aspen stand coded with high certainty means that we are absolutely sure it is really pure aspen and not any other forest class. "Medium certainty" is coded a medium grey colour. For example, an aspen stand coded with medium certainty means that we are mostly certain it is pure aspen. "Low certainty" means we are not at all certain that the data is accurate and the stand could likely be something other than what it is coded. Low certainty is coded dark grey. The accuracy map shows these three levels of certainty with respect to the first forest inventory map.

Familiarize yourselves with the two maps at:

http://www.geog.ubc.ca/~antle/dissertation/test1/showb.htm

(2) What are the key differences and similarities between the forest inventory map and the certainty map?

(2) How important is attribute accuracy as part of metadata? Do you think are other types of metadata that are more important in forest land-use decision-making?

SECTION TWO *Interactive Online Tools for Forest Inventory Maps*

Interactive Visualization Tools

You will be familiarizing yourself with four different interactive visualization tools which can be used to display a forest inventory map and the corresponding uncertainty map simultaneously. In this section, for each of the four tools, look at each one and familiarize yourself with how it works. You will be coming back to this after you have completed Section Three. (You will be asked to evaluate all four tools in Section Four.)

1. Side-by-side Map **http://www.geog.ubc.ca/~antle/dissertation/test1/showb.htm**
The traditional, static (i.e., non-interactive) approach is simply to display the two maps side by side as shown. The disadvantage to this technique in an online environment is that typically, both maps and all the auxiliary information will not all fit on one display screen. If you are assigned this tool for SECTION 3, you will have to use the right hand side scroll bar to see all of the information and both maps.

2. Toggle Map **http://www.geog.ubc.ca/~antle/dissertation/test2/showb.htm**
With the toggle tool, the two maps are displayed in the same display space. However, only one is visible at a time. By using the button labeled **Toggle** you can switch back and forth between the two maps. If you toggle back and forth quickly, the afterglow of one map can still be seen while the next is displayed.

3. Data Merger **http://www.geog.ubc.ca/~antle/dissertation/test3/showb.htm**
With the data merger tool, both maps are displayed in the same display space, one overlaid on top of the other. By using the slider bar (at the bottom of the display) you can control which map is visible. When the slider is at the far left endpoint, you'll see only the coloured forest inventory map. To show the grey certainty map, move the slider to the extreme right.

This tool also allows you to display both maps at once. The certainty map can be used to desaturate (desaturate means to make a colour duller or greyed-out) each coloured forest stand by an amount proportional to data certainty for that stand. If the stand has high certainty, its forest class colour will appear only slightly desaturated (since high certainty is coded a light grey). If the stand has low certainty, its forest class colour will be significantly desaturated or darkened. For example, if aspen forest class was coded yellow, a stand on the combined map would show a *light* yellow to indicate aspen and high certainty, a *medium* yellow to indicate aspen and medium certainty, and a *dark* yellow to indicate aspen and low certainty.

Using the data merger tool, move the slider from left to right and see how all the colours of the forest inventory map are desaturated (greyed-out) based on values from the certainty map. There may be a small time lag as you move the slider, so the best approach is to drag the slider box, release it and wait for the map to change.

(1) At what point on the slider can you distinguish all the forest inventory classes and their corresponding certainty values? (example, at the far right end, far left end, half way, one third from the left end etc.)

4. Hypermap **http://www.geog.ubc.ca/~antle/dissertation/test4/showb.htm**
The fourth method involves the addition of a hypermap function to the data merger (above). A hypermap is an interactive map that allows you to query attribute values using the mouse. When you click on a stand (polygon) with your mouse, the forest inventory class and certainty value for that stand (i.e., the attributes) will be displayed textually in the bar at the bottom of your map

display. By combining the hypermap with the data merger, you can use the data merger to look at overall patterns and the cursor to determine specific stand attributes.

(1) What is the certainty value of Calling Lake? Does this make sense to you? Why?

SECTION THREE *Decision-Making with Interactive Online Forest Inventory Maps*
Decision-Making Using Interactive Visualization Tools
In this section, you'll be using one of the four tools above to answer a set of timed fifteen true or false questions involved in a forestry land-use decision-making scenario. The TA will assign one of the four tools for you to use and give you a specific URL (web page address) for this part of the lab. *Please see Alissa for the correct URL before continuing.*

Read this entire section before starting.

After you type in the URL, you'll see a demographics information page. This information required is so that we can automatically record your answers (and I can use this information for my study). Please fill in all the boxes. "English?" Means Are you fluent in English? "Colour Blind?" means Are you colour blind? If you don't know, click "?" Fill in your department and program, computer experience and Internet experience. In the text box fill in any experience you have in forestry or land-use management. Do not use ENTER at the end of each line, it will automatically wrap. When you are done, click **SUBMIT** (once). The timed questions will start immediately so be ready to go.

There are fifteen questions, some easier than others. The questions are presented in two parts: a scenario and a statement. The scenario sets the context and explains who the parties are in the land-use conflict. The statement refers to the two maps and the scenario. For each question, you will represent the interests of one group. So, you always want to answer in their interest. That is, don't purposely answer wrong to mess up a party you don't like!

You will answer True or False to each statement. To answer, use the radio buttons at the bottom of the page. Make sure the button is clicked in. You won't have much time in some cases. The idea is to answer as quickly and accurately as you can. Always try and answer, even if it means guessing at the last moment.

You will also be asked to rate your confidence in your answer. This is important because in decision-making your confidence in your decision is sometimes more important than your actual decision. The confidence rating is a scale of 1 to 5:

1 Very Confident
2 Mostly Confident
3 Somewhat Confident
4 Not Very Confident
5 Not At All Confident (i.e., guess)

NOTE: The confidence rating is not directly or necessarily related to the data certainty information. It is about how confident you are in your answer.

The questions are timed, so you'll only have a set amount of time for each question. Answer as quickly as you can. You'll see the time in the countdown timer at the top of your screen. When the timer reaches 0, the next question will be loaded and your answer will be *automatically* submitted. To answer in time, you'll need to select your answer and confidence rating *before* the timer reaches 0:00:00. *To make sure you answer in time, you'll have to keep an eye on the timerand when there's about 10 seconds left, check your answer and your confidence rating.*

To submit your answers click the **SUBMIT** button. Make sure the T and F, and Confidence buttons are clicked in before you click submit. *ONLY click submit once* ... then be patient ... something will happen. Do NOT click it again if nothing happens. Wait. No room for click happy mousers here. If you run out of time, the page will automatically be submitted.

(2) 2 Marks given simply for completing this section.

SECTION FOUR *Evaluation*

Evaluating the Tools

After you've answered all fifteen questions, you'll see a page that has hot links to all the four tools.

(2) Evaluate the four online map tools by filling in the attached worksheet.

(5) Write a 500 word report outlining the importance of using various forms of forest inventory metadata in land-use decision-making.

INTERNET MAPS

Lab Objective

The objective of this lab is to introduce you to issues concerning Internet maps and to familiarize you with four interactive visualization tools that can be used to examine interactive maps in an online environment.

Technical Notes

First, log in as usual. Then, use the QuickRes icon to set your screen resolution to 1024 x 768. `If you double click the QuickRes icon on your desktop it will appear in the bottom right corner of your screen. Single click this mini-icon and then set the resolution. You will be using the Netscape web browser. Open Netscape by double-clicking on the Netscape icon on your desktop. Once Netscape has opened, (this may take a second) turn off the tool bar and directory buttons by de-selecting them under the Options menu in the top menu bar (de-select = there should NOT be a check mark next to them). Click in the Netscape location bar to type in the URLs listed below as outlined in the lab
(e.g., URL = http://www.geog/ubc.ca/~antle/dissertation/test1/showb.htm).

Introduction

This exercise consists of using uncertainty information for forest decision-making (specifically, a land-use conflict) in an online (Internet) environment. As more map data becomes available through the Internet and on the WWW, decision makers are increasing using this online information. Since the Internet map producer and map user are rarely the same person, it is crucial to have metadata available for Internet maps. Once digital maps have been downloaded, they can be imported into other applications, such as a GIS, or used directly in an Internet environment. Internet tools that facilitate interaction with online maps are rapidly being developed. In this lab you will explore four interactive visualization methods for simultaneously displaying map data and metadata in an online environment.

Metadata = data about data. For digital maps this includes uncertainty (accuracy) information, currency, scale, resolution, author and other aspects of the digital map's history.

SECTION ONE *Uncertainty maps*

Uncertainty maps

You will be using two maps of the same area in this lab. The first is a species composition or forest class map of an area of mixedwood forest in Northern Alberta. There are seven classes coded as seven distinct colours. The forest classes are: pure aspen, aspen-spruce, pure spruce, spruce-aspen, other forested, logged and roads and water/Calling Lake. The legend on the online map will show which colours correspond to which forest classes.
When creating the forest map some assumptions were made. Therefore, when making decisions assume the following:

1. Flat topography -- all stands are equally accessible with respect to hill slope.
2. Mature stands -- all stands are ready to harvest.
3. Species composition information is for the forest canopy.
4. Assume north is to the top of the map.

The second map represents **metadata** in the form of certainty or accuracy information about the first map. For each stand in the first map, the second map shows how accurate the forest classification data is. This information is determined by manually ground-truthing the aerial maps that were used to create the digital forest class map. This information is crucial when digital maps are being used for decision-making, especially when the decision-makers are not familiar with the actual area and must be incorporated into the decision-making process.

The accuracy (in this case data certainty) map in this lab is classified based on a **three point certainty scale**. If the data is very accurate then it is classed as "high certainty" and coded a light grey colour. For example, an aspen stand coded with high certainty means that we are absolutely sure it is really pure aspen and not any other forest class. "Medium certainty" is coded a medium grey colour. For example, an aspen stand coded with medium certainty means that we are mostly certain it is pure aspen. "Low certainty" means we are not at all certain that the data is accurate and the stand could likely be something other than what it is coded. Low certainty is coded dark grey. The accuracy map shows these three levels of certainty with respect to the first forest class map.

Familiarize yourselves with the two maps at:

http://www.geog.ubc.ca/~antle/dissertation/test1/showb.htm

(2) What are the key differences and similarities between the forest class map and the certainty map?
(2) Why is it important to have a certainty (or uncertainty) map?

SECTION TWO *Interactive Visualization Tools*

Interactive Visualization Tools

You will be familiarizing yourself with four different interactive visualization tools that can be used to display the forest map and the corresponding uncertainty map simultaneously. In this section, for each of the four tools or methods, look at each one and familiarize yourself with how it works. After you have completed section three, you will be asked to evaluate all four tools (in section four.)

1. Side-by-side Map **http://www.geog.ubc.ca/~antle/dissertation/test1/showb.htm**
The traditional, static (i.e., non-interactive) approach is simply to display the two maps side by side as shown. If you are assigned this tool for SECTION 3, you will have to use the right hand side scroll bar to see all of the information and both maps.

2. Toggle Map **http://www.geog.ubc.ca/~antle/dissertation/test2/showb.htm**
With the toggle tool, the two maps are displayed in the same display space. However, only one is visible at a time. By using the button labeled **Toggle** you can switch back and forth between the two maps. If you toggle back and forth quickly, the afterglow of one map can still be seen while the next is displayed.

3. Data Merger **http://www.geog.ubc.ca/~antle/dissertation/test3/showb.htm**
With the data merger tool, both maps are displayed in the same display space, one overlaid on

top of the other. By using the slider bar (at the bottom of the display) you can control which map is visible. When the slider is at the far left endpoint, you'll see only the coloured forest class map. To show the grey certainty map, move the slider to the extreme right.

This tool also allows you to display both maps at once. The certainty map can be used to desaturate (desaturate means to make a colour duller or greyed-out) the coloured forest map by an amount proportional to data certainty. If the forest class data has high certainty, its colour will appear only slightly desaturated (since high certainty is coded a light grey). If the forest class data has low certainty, its colour will be largely desaturated or darkened. For example, if aspen were coded yellow, the com bined map would show a *light* yellow to indicate aspen and high certainty, a *dark* yellow to indicate aspen and low certainty, and a *medium* yellow to indicate aspen and medium certainty.

Using the data merger tool, move the slider from left to right and see how all the colours of the forest class map are desaturated or greyed-out based on values in the certainty map. There may be a small time lag as you move the slider, so the best approach is to drag the slider box, release it and wait for the map to change.

(1) At what point on the slider can you distinguish all the forest classes and their corresponding certainty values? (example, at the far right end, far left end, half way, one third from the left end etc.)

4. Hypermap **http://www.geog.ubc.ca/~antle/dissertation/test4/showb.htm**
The fourth method involves the addition of a hypermap function to the data merger (above). A hypermap is an interactive map that allows you to query attribute values using the mouse. When you click on a stand (polygon) with your mouse, the forest class an d certainty value for that stand (i.e., the attributes) will be displayed textually in the bar at the bottom of your map display. By combining the hypermap with the data merger, you can use the data merger to look at overall patterns and the mouse-click to determine exact stand attributes.

(1) What is the certainty value of Calling Lake? Does this make sense to you? Why?

SECTION THREE *Decision-Making Using Interactive Visualization Tools*
Decision-Making Using Interactive Visualization Tools
In this section, you'll be using one of the tools above (the TA will assign a tool for you to use) to answer a set of fifteen true or false questions involved in a forestry decision-making scenario.

Before you start the questions, you'll get a demographics page. This information required is so that we can automatically record your answers (and I can use this information for my study). Please fill in all the boxes. English? means Is English your native language? Colour Blind? means Are you colour blind? If you don't know, click ? Fill in your department and program, computer experience and Internet experience. In the text box fill in any experience you have in forestry or land-use management. Do not use ENTER at the end of each line, it will automatically wrap. When you are done, click **SUBMIT** (once) when you are done to start the questions. The first question may take a moment to load in.

There are fifteen questions, some easier than others. The questions are presented in either one or two parts. Each question will have a scenario which explains who the parties involved in the land-use conflict are. In each question, you will represent the interests of one group member. So you always want to answer in their interest. That is, don't purposely answer wrong to mess up a party you don't like!

For all the questions, you'll get a true or false statement. You'll answer T or F to this using the radio buttons at the bottom of the page. Make sure the button is clicked in. You won't have much time in some cases. The idea is to answer as quickly and accurately as you can. Always try and answer, even if it means guessing at the last moment.

After you answer T or F to the statement, you must also rate how confident you are in your answer. In decision-making your confidence in your decision is sometimes more important than your actual decision. The rating is a scale of 1 to 5:

1 Very Confident
2 Mostly Confident
3 Somewhat Confident
4 Not Very Confident
5 Not At All Confident (i.e., guess)

NOTE: The confidence rating is not directly or necessarily related to the data certainty information. It is about how confident you are in your answer.

You'll have a set amount of time for each question, so answer as quickly as you can. You'll see the time in the countdown timer at the top of your screen. When the timer reaches 0, the next question will be loaded and your answer will be automatically submitted. To answer in time, you'll need to select your answer and confidence rating before the timer reaches 0:00:00. *To make sure you answer in time, you'll have to keep an eye on the timer and when there's about 10 seconds left, check your answer and your confidence rating.*

To submit your answers click the **SUBMIT** button. Make sure the buttons are checked before you click submit. *ONLY click submit once* ... then be patient ... something will happen. Do NOT click it again if nothing happens. Wait. No room for click happy mousers here. If you run out of time, the page will automatically be submitted.

SECTION FOUR *Evaluating the Tools*

Evaluating the Tools

After you've answered all fifteen questions, you'll see a page that has hot links to all the four tools.

(4) Evaluate the four Internet map tools by filling in the worksheet.

DATA VISUALIZATION

INTRODUCTION

This exercise consists of using uncertainty information for forest decision making (specifically, a land-use conflict) in an online (Internet) environment. As more map data becomes available through the Internet and on the WWW, decision makers are increasing using this online information. Tools which facilitate interaction with online maps are rapidly being developed. In this lab you will explore four interactive visualization methods for simultaneously displaying map data and uncertainty information about that map data in an online environment.

SECTION ONE: *Uncertainty maps*

You will be using two maps of the same area in this lab. The first is a species composition or forest class map of an area of mixed wood forest in Northern Alberta. There are seven classes coded as seven distinct colours. The forest classes are: pure aspen, aspen-spruce, pure spruce, spruce-aspen, other forested, logged and roads and water/Calling Lake. The legend on the online map will show which colours correspond to which forest classes.

When creating the forest map some assumptions were made. When making decisions assume the following:

1. Flat topography -- all stands are equally accessible with respect to hill slope.
2. Mature stands -- all stands are ready to harvest.
3. Species composition information is for the forest canopy.
4. Assume north is to the top of the map.

The second map represents **metadata** in the form of certainty or accuracy about the first map. For each stand in the first map, the second map shows how accurate the forest classification data is. This information is determined my manual ground-truthing the aerial maps which were used to create the digital forest class map. This information is crucial when digital maps are being used for decision-making, especially when the decision-makers are not familiar with the actual area, and must be incorporated into the decision-making process.

The accuracy (in this case data certainty) map in this lab is classified based on a three point certainty scale. If the data is very accurate then it is classed as "high certainty" and coded a light grey colour. For example, an aspen stand coded with high certainty means that we are absolutely sure it is really pure aspen and not any other forest class. "Medium certainty" is coded a medium grey colour. For example, an aspen stand coded with medium certainty means that we are mostly certain it is pure aspen. "Low certainty" means we are not at all certain that the data is accurate and the stand could likely be something other than what it is coded. Low certainty is coded dark grey. The accuracy map shows these three levels of certainty about the first forest class map.

SECTION TWO: *Interactive Visualization Tools*

You will be familiarizing yourself with four different interactive visualization tools that can be used to display the forest map and the corresponding uncertainty map simulataneously. In this section, for each of the four tools or methods, look at each one and familiarize yourself with how it works. After you have completed section three (the following section) you will be asked to evaluate all four tools (in section 4.)

1. Side-by-side Map
Having the two maps side by side is the traditional solution. Since the two maps are too large to display side-by-side on one screen, you'll have to use the scroll bar to see all of both maps.

2. Toggle Map
With the toggle tool, the two maps are displayed in the same display space. Only one is visible at a time. By using the button labeled Toggle you can switch back and forth between the two maps. If you toggle back and forth quickly, the afterglow of one map can still be seen while the next is displayed.

3. Data Merger
With the data merger tool, both maps are displayed in the same display space, one overlaid on top of the other. By using the slider bar (at the bottom of the display) you can control which map is visible. When the slider is at the far left endpoint, you'll see only the coloured forest class map. To show the grey certainty map, move the slider to the extreme right.

This tool also allows you to display both maps at once. The certainty map can be used to desaturate (desaturate means to make a colour duller or greyed-out) the coloured forest map by an amount proportional to data certainty. If the forest class data has high certainty, its colour will appear only slightly desaturated (since high certainty is coded a light grey). If the forest class data has low certainty, its colour will be largely desaturated or darkened. For example, if aspen were coded yellow, the combined map would show a light yellow to indicate aspen and high certainty, a dark yellow to indicate aspen and low certainty, and a medium yellow to indicate aspen and medium certainty.

Using the data merger tool, move the slider from left to right and see how all the colours of the forest class map are desaturated or greyed-out based on values in the certainty map. There may be a small time lag as you move the slider, so the best approach is to drag the slider box, release it and wait for the map to change.

There's a place about two thirds of the way to the right where it is easy to distinguish all seven colour classes and the three greys. Try and find the point where you can distinguish all the forest classes and their corresponding certainty values at the same time.

4. Hypermap
The fourth method involves the addition of a hypermap function to the data merger (above). A shypermap is an interactive map that allows you to query attribute values using the mouse. When you click on a stand (polygon) with your mouse, the forest class and certainty value for that stand

(i.e., the attributes) will be displayed textually in the bar at the bottom of your map display. By combining the hypermap with the data merger, you can use the data merger to look at overall patterns and the mouse-click to determine exact stand attributes.

SECTION THREE: *Decision-Making Using Interactive Visualization Tools*

In this section, you'll be using one of the tools above (the TA will assign a tool for you to use) to answer a set of fifteen true or false questions involved in a forestry decision-making scenario.

Before you start the questions, you'll get a demographics page. This information is so that we can automatically record your answers (and I can use this information for my study.)Please fill in your name, age etc ... fill in all the boxes. English? means Is English your native language? Colour Blind? means Are you colour blind? If you don't know, click ? Fill in your department and program, computer experience and Internet experience. In the text box fill in any experience you have in forestry or land-use management. Do not use ENTER at the end of each line, it will automatically wrap. When you are done, click SUBMIT (once) when you are done to start the questions. The first question may take a moment to load in.

There are fifteen questions, some easier than others. The questions are presented in either one or two parts. Some questions will have a scenario which explains who the parties involved in the land-use conflict are. In each question, you will represent the interests of one group member. So you always want to answer in their interest. That is don't purposely answer wrong to mess up a party you don't like!

For all the questions, you'll get a true or false statement. You'll answer T or F to this using the radio buttons at the bottom of the page. Make sure the button is clicked in. You won't have much time in some cases. The idea is to answer as quickly and accurately as you can. Imagine deadlines pressing in on you as you work!

After you answer T or F to the statement, you must also rate how confident you are in your answer. In decision-making your confidence in your decision is sometimes more important than your actual decision. The rating is a scale of 1 to 5. So 1 means you are super confident, you are absolutely certain you got the right answer. You would put money on it, a lot of money. 2 means you are mostly certainty but wouldn't bet your life savings on it. 3 means you are "kind of" certain. 4 means you have serious doubts about your answer. 5 means you are making a complete guess, you have no idea or you are about to run out of time.

NOTE: The confidence rating has nothing to do with the certainty of the data — it is about how confident you are in your answer.

You'll have a set amount of time for each question, so answer as quickly as you can. You'll see the time in the countdown timer at the top of your screen. When the timer reaches 0, the next question will be loaded and your answer will be automatically submitted. To answer in time, you'll need to select your answer and confidence rating before the timer reaches 0:00:00. To make sure you answer, you'll have to keep an eye on the timer and when there's about 10 seconds left, click your answer and your confidence rating.

To submit your answers click the SUBMIT button. Make sure the buttons are clicked before you click submit. ONLY click submit once ... then be patient ... something will happen. Do NOT click it again if nothing happens. Wait. No room for click happy mousers here. If you run out of time, the page will automatically be submitted.

SECTION FOUR: *Evaluating the Tools*

1. After you've answered all fifteen questions, you'll see a page that has hot links to all the four tools. Answer the questions on the attached worksheet

2. Write a one page report evaluating the usefulness of incorporating uncertainty information into the decision-making process.

Hand in your evaluation of the tools and a one page report.