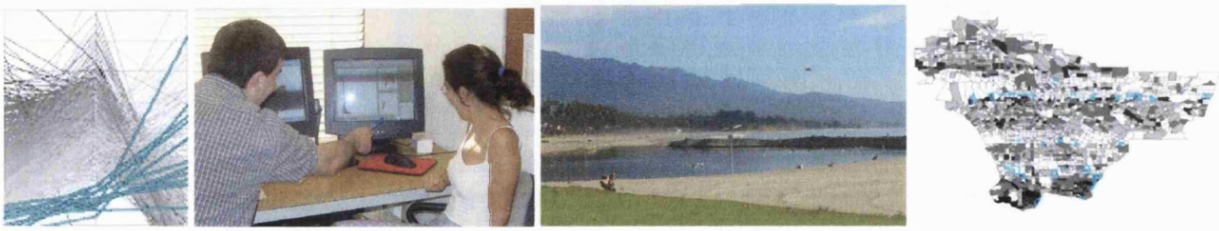

Evaluating Geovisualization and User Interaction: An Experimental Design Approach Based upon User Tasks

Carolina Tobón
Department of Geography
University College London

A dissertation submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in University College London, University of London, UK.

2003

ProQuest Number: U642483

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

ProQuest U642483

Published by ProQuest LLC(2015). Copyright of the Dissertation is held by the Author.

All rights reserved.
This work is protected against unauthorized copying under Title 17, United States Code.
Microform Edition © ProQuest LLC.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Abstract

Visualization is a cognitive activity the main purpose of which is to gain insight and understanding of a problem or dataset by representing the data in some visual form that can reveal otherwise hidden information. When applied to geospatial data, these techniques may facilitate the construction of hypotheses about relations between attributes in the data. Geographic visualization, or geovisualization, is the field concerned with exploring spatial data visually. However, graphical displays are of modest use if their users cannot interpret the information they represent or if the computer software tools where they may be embedded are not easy to use.

These issues are investigated in a series of three studies. The first study inspects usability limitations in geovisualization tools that may hinder the data analysis and cognitive processes they support. Due to the exploratory nature of geovisualization, designing the tasks with which to test the environment was a great challenge. Therefore, the second study suggests a methodology for defining a task typology which is used to evaluate a second geovisualization tool using formal experimental design. The third study further inspects the validity of the task taxonomy proposed and concentrates on learning about user strategies in approaching different types of tasks.

These evaluations are shown to contribute to the research on cognitive issues in geovisualization and geographic information science by facilitating the understanding of user tasks and requirements and how to support their work processes. Hence, the studies suggest an exciting possibility for cross-fertilisation between the evaluation methods reported and research on issues of spatial cognition, such as the representation in memory of geographical knowledge and its effects on geographical judgments. In addition, the combination of techniques used for the experiments are shown to be an effective means to evaluate whether the tools we develop are usable, useful and fit for purpose. These are crucial considerations if geovisualization environments are to be adopted by a larger set of users and support a wider range of work processes. The methods suggested for defining tasks can be extended to appraise a more ample range of data formats and users than those considered here in order to build taxonomies of tasks with which to formally evaluate geovisualizations. In this way, we would be able to accumulate evidence for understanding how these tools facilitate hypothesis formation and knowledge construction and thus contribute to the development of the field.

Table of Contents

ABSTRACT 2

TABLE OF CONTENTS 3

LIST OF FIGURES 7

LIST OF TABLES 11

DECLARATION 13

ACKNOWLEDGEMENTS 14

1 INTRODUCTION 17

2 GEOVISUALIZATION AS A COGNITIVE ACTIVITY 23

2.1 VISUALIZATION AS INFORMATION PROCESSING 24

2.2 ADVANTAGES OF VISUALIZATION 26

2.3 VISUALIZATION FIELDS AND METHODS 28

2.3.1 Scientific Visualization 28

2.3.2 Information Visualization 30

2.3.2.1 Focus+Context 31

2.3.2.2 Distortion 34

2.3.2.3 Dynamic Interaction and coordinated views 36

2.3.2.4 Iconographic Methods 38

2.3.3 EDA 40

2.3.3.1 Dotplots, Boxplots and Histograms 41

2.3.3.2 Geometric displays 43

2.3.3.3 Conditional Plots 47

2.4 CARTOGRAPHIC VISUALIZATION AND GEOVISUALIZATION 49

2.4.1 Issues of Cognition in Visual Representations 53

2.4.2 HCI in Geovisualization and GISc 57

2.4.2.1 Usability and Geovisualization 59

2.5 TOWARDS AN UNDERSTANDING OF USER TASKS 60

3 STUDY 1: USABILITY EVALUATION OF A GEOVISUALIZATION TOOL 64

3.1 LINKING PLOTS, LINKING SOFTWARE 64

3.2 USABILITY 68

3.3 USABILITY STUDY 70

3.3.1 Goals 70

3.3.2 Participants 72

3.3.3 Tailoring Usability Evaluation Techniques to Geovisualization 73

3.3.4 Invitation to Participate and Motivation 75

3.3.5 Test Design *76*
3.4 THE TASKS 77
3.5 RESULTS 79
3.5.1 User Observations about the software *79*
3.5.1.1 Observations of Satisfaction 79
3.5.1.2 Observations of Frustration 82
3.6 USER BEHAVIOUR 85
3.7 NEXT ITERATION: IMPROVING THE TOOL 89
3.7.1 The COM approach *90*
3.8 CHALLENGES 94

4 STUDY 2: GEOVISUALIZATION TASK TAXONOMY 96

4.1 LESSONS FROM THE FIRST EVALUATION: ABOUT TASK DEFINITION 96
4.2 USABILITY EVALUATION 97
4.2.1 Goals *97*
4.2.2 Participants *98*
4.2.3 Invitation to Participate and Motivation *99*
4.2.4 Evaluation Design *101*
4.2.4.1 Evaluation Phases 101
4.3 TASK DESIGN 103
4.4 EXPERIMENTAL DESIGN 107
4.4.1 Advantages of Factorial Experiments *108*
4.4.2 Standard Error and Standard Deviation *109*
4.5 ANALYSIS OF VARIANCE 112
4.5.1 A note on the assumptions of ANOVA and data transformations *114*
4.5.2 Main Effects and Interactions *117*
4.5.3 Main Effects *117*
4.5.4 Interactions *118*
4.5.5 Graphical Analysis of Main Effects and Interactions *120*
4.6 BINARY LOGISTIC REGRESSION 127
4.7 QUESTIONNAIRES 131
4.7.1 Ratings for perceived usefulness *134*
4.7.2 Ratings for perceived ease of use *134*
4.8 INTERVIEWS 135
4.9 PROBLEMS AND POSSIBILITIES FOR FURTHER DEVELOPMENT 137
4.10 CONCLUSION 140

5 STUDY 3: REFINING THE CHARACTERISATION OF GEOVISULIZATION TASKS 142

5.1 LESSONS FROM THE SECOND EVALUATION 142
5.2 EXPERIMENTAL DESIGN 144
5.2.1 Experiment Goals *144*

5.2.2 Profile and Number of Participants *145*
5.2.3 Evaluation Phases *147*
5.2.4 Tasks *149*
5.2.5 Latin Squares Design and Analysis *151*
5.2.5.1 Results of the second phase 155
5.2.5.2 Results of the fourth phase 160
5.3 USER STRATEGIES 164
5.3.1 User Behaviour *165*
5.3.1.1 Frequencies 166
5.3.2 Articulate and Evaluate search cycle *170*
5.3.2.1 IS task 173
5.3.2.2 DN task 174
5.3.2.3 DS task 175
5.3.2.4 CN task 176
5.3.2.5 CS task 176
5.3.2.6 AN task 178
5.3.2.7 AS task 178
5.3.3 Summary of Strategies *179*
5.4 THE RELEVANCE OF STRATEGIES FOR INFORMATION PROCESSING 181
5.5 CONCLUSION 185

6 DATA DESCRIPTION AND CONSIDERATIONS 187

6.1 UK 1991 CENSUS SMALL AREA STATISTICS AND LIFESTYLES DATA 188
6.1.1 UK Census data exploration with the environment *190*
6.2 U.S. CENSUS 2000 194
6.2.1 U.S. Census data exploration with the environment *195*
6.3 CONCLUSION 197

7 CONCLUSIONS AND FUTURE RESEARCH 199

7.1 REVIEW OF ASSERTIONS 200
7.2 ASSESSMENT OF THE MODULAR APPROACH TO BUILDING GEOVISUALIZATION ENVIRONMENTS 201
7.2.1 Limitations and Drawbacks *203*
7.3 ASSESSMENT OF THE USABILITY EVALUATION TECHNIQUES FOR EVALUATING GEOVISUALIZATIONS 204
7.3.1 Limitations and drawbacks *205*
7.4 ASSESSMENT OF THE EXPERIMENTAL DESIGN FOR EVALUATING GEOVISUALIZATION 206
7.4.1 Limitations and drawbacks *206*
7.5 EXTENDING THE RESEARCH 207
7.5.1 Collaborative Geovisualization *207*
7.5.2 Expertise *208*

7.5.3 Individual Differences, Strategies and Abilities *209*
7.6 CONCLUSION 210

REFERENCES **211**

APPENDIX A: EXPERIMENTS MATERIAL **229**

A.1 EXPERIMENT 1 229
A.1.1 Invitation to participate *229*
A.1.2 Instructions and Data *230*
A.1.3 The Tasks *232*
A.1.4 Interview Schedule *232*
A.2 EXPERIMENT 2 233
A.2.1 Invitation to participate *234*
A.2.2 Data Description: Lifestyles and 1991 UK Census data *235*
A.2.3 The Tasks *236*
A.2.3.1 Tasks administered: Day 1 238
A.2.3.2 Tasks administered: Day 2 239
A.2.4 Questionnaire *240*
A.2.5 Interview Schedule *241*
A.3 EXPERIMENT 3 242
A.3.1 Invitation to participate *242*
A.3.2 Data Description: US Census 2000 *243*
A.3.3 The Tasks *245*
A.3.4 Interview *245*

APPENDIX B: CODE EXAMPLES **247**

APPENDIX C: PUBLICATIONS **265**

C.1 REFEREED PAPERS 265
C.2 REFEREED BOOK CHAPTERS 265
C.3 REFEREED CONFERENCE PROCEEDINGS 265
C.4 CONFERENCE PAPERS 266
C.5 WORKING PAPERS 266

List of Figures

Figure 2.1 Surface of employment data for London 27

Figure 2.2 Magic Lenses: Left) For revealing details a road map. Right) For highlighting features in a map. Maps extracted from Stone et al. (1994) 32

Figure 2.3 Focus+Context example with georeferenced data (Komzak 2000) 33

Figure 2.4 Non-Contiguous, Contiguous and 'Simple Shape' Cartograms (maps taken form Bortins and Demers 2002) 35

Figure 2.5 Brushing and linking views with geospatial data 37

Figure 2.6 Parallel Coordinate Plot (Left) and Dynamic Query devices and Details-on-Demand (Right) in DecisionSite (Spotfire, Boston) 38

Figure 2.7 Details-on-Demand in ArcGIS (ESRI, Redlands) 38

Figure 2.8 Dorling (1994) cartogram showing the distribution of voting, housing, employment and industrial composition in the UK 1983 General Election. 39

Figure 2.9 Star Glyphs from XmdvTool (Ward 1993) 40

Figure 2.10 Dotplots and Boxplots 41

Figure 2.11 Histograms: Left) Narrow bins. Right) Broad bins classified by standard deviation. 42

Figure 2.12 Scatterplot of Inc_Score vs. Hhtcar. Qualml is encoded by the colour of the glyphs where yellow corresponds to a low value and black to a high value of the variable. Wkwife is encoded in the orientation of the glyph or the deviation from 90 degrees as the value of the variable increases. 43

Figure 2.13 (Half) Scatterplot Matrix and the use of brushing 44

Figure 2.14 Normal Q-Q Plots: Left) Log of Inc_Score. Right) Inc_Score. 45

Figure 2.15 PCP and use of colour for brushing ranges of data 46

Figure 2.16 Trellis Plot of Councl vs Buying conditioned upon values of 47

Figure 2.17 Comap with Eastings and Northings as the two conditional variables 48

Figure 2.18 Comap with Inc_Score and Qualml as the two conditional variables 49

Figure 3.1 DecisionSite Map IS (Spotfire, Boston). To the left of the image is ArcExplorer and DecisionSite appears to the right. The points on the map correspond to a set of records selected in DecisionSite which have been 'transferred' to the GIS as an image and laid on top of the map. 71

Figure 3.2 Four different implementations of Query Devices in DecisionSite (Boston, MA). From left to right: Far Left) Range Sliders: users move the end of the sliders or the

slider itself to select a data range they want to appear in the graphical displays. B) Checkboxes: allow users to select more than one value of an attribute to show in the graphical displays. C) Radio Buttons: allow users to select only one value of an attribute. D) Alpha Sliders: allow the querying of text data. 80
Figure 3.3 Visualization Process (Spence 2001) 86
Figure 3.4 Visualization Process II 87
Figure 3.5 Visualization Process III 89
Figure 3.6 Environment's data views for the visual exploration of spatial data 90

Figure 4.1 Experiment setting: Left) User completing the test. Right) User being interviewed by the observer. 103
Figure 4.2 Framework for formulating user tasks as combinations of levels of three factors: The visual operation performed, the number of spatial units or the extent of the geographical area and the number of attributes investigated. 106
Figure 4.3 Normal QQ Plots: Left) Time. Right) LnTime. 115
Figure 4.4 Line graph showing the volatility of the variance for: Top) Time; Bottom) LnTime. 116
Figure 4.5 Main effects of the response means for each factor level. From left to right: Main effects of the number of spatial units explored; Main effect of the number of attributes investigated; Main effect of the visual operator; Main effect of the day in which the tasks were completed which reflects user practice with the environment. 118
Figure 4.6 (Left) NU*AE interaction for session 1 when VO is fixed to its first level; (Right) NU*AE interaction for session 2 when VO is fixed to identify 121
Figure 4.7 (Left) NU*NA interaction for session 1 when VO is fixed to its second level; (Right) NU*NA interaction for session 2 when VO is fixed to locate 122
Figure 4.8 (Left) NU*NA interaction for session 1 when VO is fixed to its third level; (Right) NU*NA interaction for session 2 when VO is fixed to compare 123
Figure 4.9 (Left) NU*NA interaction for session 1 when VO is fixed to its highest level; (Right) NU*NA interaction for session 2 when VO is fixed to associate 123
Figure 4.10 NU*NA*VO: Interactions of Spatial extent, Attribute extent and Visual Operators with Spatial extent set to the low level (results are averages of both sessions for all users). 124
Figure 4.11 NU*NA*VO: Interactions of Spatial extent, Attribute extent and Visual Operators with Spatial extent set to many units (results are averages of both sessions for all users). 125

Figure 4.12 NU*NA*VO: Interactions of Spatial extent, Attribute extent and Visual Operators with Attribute extent set to its lower level (results are averages of both sessions for all users). 126

Figure 4.13 NU*NA*VO: Interactions of Spatial extent, Attribute extent and Visual Operators with Attribute extent set to many units (results are averages of both sessions for all users). 127

Figure 4.14 Colour coding multiple selections in six linked data view DataDesk (DataDescription Inc., Ithaca): The top three graphic displays show Normal QQ plots for three variables. The lower three displays show two scatterplots and boxplots for the same variables. The colour coding (green, red and blue) show were the selected records (outliers as shown by the boxplots) are in all data representations. 139

Figure 5.1 The environment to be evaluated closely coupled ArcGIS (ESRI, Redlands) and DecisionSite (Spotfire, Boston). 146

Figure 5.2 Usability Lab: Left) Room where participants completed the tasks using the computer to the left. They used the computer to the right to transcribe the answers to the tasks. In the middle is the camera that was used to record participants face in order to trace the direction of their gaze. Right) This room was separate from the one in which users completed the evaluation and was used to monitor their progress. The monitors on the top show the video streamed by the camera, as well as the computer screen used by participants. These two video sources were mixed using the equipment above and recorded onto VHS tape for subsequent analysis. 148

Figure 5.3 Time performance for all tasks and users: Top) Raw times in seconds for all tasks. Bottom) Log of times taken to complete all tasks. 157

Figure 5.4 Main effects of factors on the performance variable (log of time to complete 'identify or 'distinguish' tasks): Left) Main effects of the visual operator. Centre) Main effects of the position in which tasks were administered. Right) Main effect of subjects. 159

Figure 5.5 Main effects of factors on the performance variable (log of time to complete 'identify or 'distinguish' tasks): Left) Main effects of the visual operator. Centre) Main effects of the position in which tasks were administered. Right) Main effect of subjects. 163

Figure 5.6 Articulate and evaluate search cycle 173

Figure 5.7 User 5 sequence of graphic display use for task IS 174

Figure 5.8 User 1 sequence of graphic display use for task DN 175

Figure 5.9 User 7 sequence of graphic display use for task DS 176

Figure 5.10 User 4 sequence of graphic display use for task CN 176

Figure 5.11 User 6 sequence of graphic display use for task CS 177
Figure 5.12 User 5 sequence of graphic display use for task CS 178
Figure 5.13 User 4 sequence of graphic display use for task AN 178

Figure 6.1 Data Exploration of UK 1991 Census data for Bristol: Left) High proportion of Wkwife compared to Inc_Score, Unemp and Child. Right) Low proportions of Wkwife compared to the same four variables. 192
Figure 6.2 Data Exploration of UK 1991 Census data for Bristol: Left) High proportion of Child compared to Inc_Score, Unemp and Wkwife. Right) Blunders in the dataset. 193
Figure 6.4 Data Exploration of U.S. 2000 Census data for Los Angeles: Left) High proportion of White compared to Black, Hisp and AggSalaInc. Right) Blunders in the dataset. 198

List of Tables

Table 2.1 Example Software for Geovisualization 62

Table 3.1 GIS and ESDA implementations with ESRI's products 65
Table 3.2 Description of participants 73
Table 3.3 Summary of goals, benefits and foreseen costs (Dumas and Redish 1999). 75
Table 3.4 Summary of Test's Resources 77
Table 3.5 Tasks 78
Table 3.6 Tasks aims and description. 78
Table 3.7 Frequently of feature acceptance 81
Table 3.8 Users' frequent problems 85

Table 4.1 Goals, benefits and foreseen costs (Dumas and Redish 1999). 97
Table 4.2 Users' Description 100
Table 4.3 Summary of Test Resources 102
Table 4.4 Task Typology 107
Table 4.5 User log of times (in seconds) during first evaluation 110
Table 4.6 Squared Residuals 111
Table 4.7 Normality of the response variable 115
Table 4.8 ANOVA results summary 119
Table 4.9 General Linear Model: LnTime versus NU, NA, VO, D 120
Table 4.10 Logit Regression Results 129
Table 4.11 Questionnaire's Items and Corresponding Constructs being Measured 132
Table 4.12 Questionnaire Results 135
Table 4.13 User's Perceptions 137
Table 4.14 Users' frequent problems 138

Table 5.1 Users' description 147
Table 5.2 Summary of Experiment's Resources 149
Table 5.3 Latin Squares (Sequence by position) 152
Table 5.4 Latin Squares (Sequence by visual operator) 153
Table 5.5 Within Subjects ANOVA: (V*S) design 154
Table 5.6 Latin Squares 154
Table 5.7 Results for the visual operators equal to Identify and Distinguish 157
Table 5.8 Results organised by task position (Phase 2) 156

Table 5.9 Latin Squares Analysis for operators Identify and Distinguish 158
Table 5.10 Estimates of effects for 'identify' and 'distinguish' tasks 158
Table 5.11 Results for visual operators equal to Compare and Associate 161
Table 5.12 Results organised by task position (Phase 4) 161
Table 5.13 Latin Squares Analysis for operators Compare and Associate 162
Table 5.14 Estimates of effects for 'compare' and 'associate' tasks 164
Table 5.15 User Behaviours 166
Table 5.16 Frequency of user operations from computer logs and videos 167
Table 5.17 Frequency of user operations from computer logs and videos (Compare and Associate tasks) 168
Table 5.18 User Behaviours frequencies for Compare and Associate tasks (Identify and Distinguish tasks) 169

Table 6.1 Income Ranges 189
Table 6.2 1991 UK Census SAS variables provided in the experiments 191
Table 6.3 U.S. Census 2000 variables provided in the experiments 196

Declaration

The work presented here is my own, unaided work except where acknowledgement is given and has not been submitted for a higher degree in this or any other university.

Carolina Tobón

Acknowledgements

This is the one piece of this thesis where I get to sound as if I was a movie director receiving an award for a life time achievement. The last three years haven't really felt that long. Still, thanking all the people who have helped me get to where I am now must be done in true and grandiose style because each and every one has been tremendously generous, helpful, patient and supportive.

First, big 'thanks' must go to my family. Although they were never too happy about me being so far away, they gave me their support and help to first arrive to London. Unfortunately, some of them may not be able to see what I have accomplished, but I know they feel proud nevertheless, especially since some of them think I've been working at the British equivalent of NASA. Well, it does rhyme with CASA. I never said that was the case but then again, explaining what GIScience is to a Musician, a Radio Broadcaster, an Electronic Engineer, a Dentist, a Designer and an Economist is not an easy task. Thanks to all of them and especially to Carlos Andres, for putting up with my dreams and giving up what he longs for the most—a home in the tropics—so we could live on the same Latitude and Longitude.

'Many thanks' to my supervisor and colleagues at CASA, or Centre for Advanced Spatial Analysis at UCL, where I've been working for the past three years. Professor Paul Longley has always been a guiding light and still managed to let me do whatever I wanted. Then again, maybe he thought it was not such a good idea to disagree with a Colombian... I really hope he enjoyed reading earlier versions of this document because he did so more than once. So Paul, thank you for your support, trust and very generous help. 'Many thanks' to Professor Michael Batty who introduced me to CASA in the first place. The boys at the office have to be mentioned now for keeping me like a Popsicle while working at the office and thus awake during the past three years. Paul Torrens, Muki Haklay and Daryl Lloyd have given me emotional support on request, Soleros in the summer, very useful conversation sessions on demand, and most of all, their friendship every single day. I must thank Mark Thurstain-Goodwin for being a fantastic person to work with and for his efforts to learn Spanish. He did manage to say "Scochio!" Whatever that means. It is crucial to acknowledge Ian Greatbatch's demonstrations that explained the meaning of a googly and a China man; as well as Torsten Schietzelt's lessons on how to deal with frustrating customer services. Their teachings did not go unnoticed. And last but not least, I must thank Ita McDonnell for her good energy and coffee, which brought me back to life every morning. To all the people at

CASA, East and West End, it was a real pleasure working with so many bright individuals from all over the world.

Being at UCL has been an exciting experience. The sheer amount of bright people all around makes you crave for knowledge and want to pick people's brains at the slightest chance they give you. Special thanks go to Professor Tom Fearn, a wonderful statistician and fluent Spanish speaker, who had the patience to guide me through unfamiliar grounds. (I must ask him to liaise with Mark). The staff at Geography, especially Peter Wood and Suse Keay, were always very welcoming and helpful. The students and staff at Computer Science and the Human-Computer Interaction Centre (UCLIC) were of great help when I was planning and conducting experiments. They gave me some great opportunities to present and discuss ideas, get to grips with all the HCI literature and a hold of a real usability lab.

I met Professor Ben Shneiderman during the first summer of my PhD. I was a bit desperate then, ready to quit the degree, possibly like any other first year PhD student. He was tremendously helpful. Not only did my research benefit greatly from meeting him one afternoon, he actually put me in contact with Spotfire who supplied me with the visualization software used for this research. On that note, many thanks are due to Martin Nilsson and Johanna Altenstedt from Spotfire with whom I worked back in 2001. Discussions with other very famous people were illuminating too. Professors Robert Spence, David Unwin and Jason Dykes were always fun to talk to. They always left me with heaps of homework to do though.

I had the great pleasure of spending a few months at the University of California, Santa Barbara during the last year of my PhD. Going there was incredibly useful as I found out that I was not so alone in thinking about the things discussed in the pages that follow (plus, I was really missing some sunshine!). Thanks to Professor Michael Goodchild and Sara Fabrikant for inviting me there. Sara, thank you so much for your hospitality and for so many useful discussions. Thanks too to Daniel Montello for his help and for a challenging graduate seminar. Many thanks indeed to all the people at NCGIA who were all very welcoming. But most of all, thanks to Matt B. Canham for helping me with my research, for getting me through the last bit of this ordeal, for his understanding and his generosity. Sharing thoughts and ideas with all these people was exciting and it inspired me to find the last bit of the puzzle I was missing to wrap up my dissertation. They also reminded me of how stimulating this period of my life has been and how fortunate I am for having had the chance to meet so many brilliant people in the past few years.

And finally, I shall never forget all the good souls that subjected themselves to my experiments for a cup of good coffee or a pint of beer at best. I most definitely could not have managed it without you!

Many thanks indeed to you all. I very much enjoyed it. I think I might even miss it.

Carolina Tobón

London, 30th August 2003

1 Introduction

"The purpose of graphical interaction is to facilitate the use of a computer, to enhance the user's powers and to accomplish or to create" (Baecker 1987: 471).

The last fifteen years have seen a rapid advance in the technology for collection, transmission and display of georeferenced data at all geographical scales of measurement. This in turn has massively increased the availability of data for investigating a wide range of problems. In the initial stages of analysis, it is commonplace that we need to sift through large amounts of data in order to extract and make sense of the relevant information that they may contain, by selecting, preparing or organising the data for a particular application or task. In this context, analysis may be best accomplished by using multiple representations of variables in a dataset in a variety of media or graphical displays. In this way, complementary views of attribute values can be provided with ease and speed that can facilitate data inspection and the potential understanding of relations between them.

It has been suggested that the computer environment designed to support the exploration of data in this manner must sustain an iterative process of user interaction in order to facilitate uncovering of meaningful information through visual displays (Wachowics 2001; MacEachren *et al.* 1999; Dykes 1997; Unwin 1994; Wood 1994; MacEachren 1994). It is by facilitating the integration of user knowledge and experience into the computer-assisted exploration process that we may extract meaning from data. User domain expertise, coupled with the power of human vision, can be used in harness with computational methods in order to obtain information and insight from masses of data (MacEachren 1999). Thus, it is in the interpretation of this information, based on a particular context or purpose, where the user adds value to the exploratory process and transforms the information into insights or knowledge (Longley *et al.* 2001).

Much of the appeal and power of visual interfaces for exploring data at the initial stages of analysis comes from their ability to support the ways in which we normally think about a problem. Direct manipulation of graphical widgets and displays "allow us to think in the familiar terms of the application domain rather than those of the medium of computation" (Hutchins *et al.* 1987: 486). This semantic directness increases the ease for performing some activities that makes graphical interfaces a vehicle for data understanding (Unwin and Hofmann 1999).

In the context of geographical information systems (GIS), the use of multiple graphical displays is typically restricted to the presentation of results. GIS are not currently strongly supportive of multiple representations whether they are a result of "the accumulation of multiple representations over time, due to collection from multiple sources, or as accumulation of multiple representations from the transformation of a single source" (Beard 2000: 42). GIS lack the flexibility needed to re-express spatial data in multiple but complementary representations or the capacity to support processes of interactive visual data analysis.

Nevertheless, the strengths of GIS rely in their ability to collect, store, manage, query and present geographic information (Longley *et al.* 2001), among others. Therefore, linking analytical tools to GIS has been a major area of development and research which aims to provide the analyst with spatial modelling tools and more interactive and flexible display capabilities while taking advantage of the GIS technology available. However, if the coupling is not done carefully, "the result is a system that simply automates the process of data retrieval, analysis and display" (Densham and Amstrong 1995: 180). In a loosely coupled system, for instance, users would interact with the GIS to select data and then export them to the coupled software in order to perform some modelling or analysis. The results would need to be imported back into the GIS for display. Loosely coupled systems may be easy to build but they do not support high levels of interaction with the software or large amounts of information to be processed.

The pages that follow explore the close coupling of GIS to software tools that allow the visual information processing of data. Two computer-based environments were used to investigate how to better support tasks that include the identification, location, analysis and comparison of spatial patterns and features and their spatial relations (MacEachren and Ganter 1990; Dorling 1992; Gahegan 2001). These systems enable individuals to interactively engage in low level tasks such as sorting, highlighting, filtering, and otherwise transforming data in a search for patterns and relations between attributes (Unwin 1996).

Some geographic visualization tools have been proposed as a flexible and interactive method for the exploration, analysis, synthesis and presentation of geospatial data (Dykes 1997; Andrienko and Andrienko 1999; Takatsuka and Gahegan 2002; among others). They have been built on principles of geovisualization, which can be defined as a process that involves humans searching for insight from georeferenced datasets through their manipulation via visual displays that provide representations of the data (MacEachren *et al.* 1999). The work

presented here builds upon this previous research and concentrates on understanding user tasks and behaviour to better support their information processing.

By synthesising perspectives on visual representation and analysis from a number of fields— including visualization in scientific computing (ViSC), cartography, GIS, exploratory data analysis (EDA) and information visualization— geovisualization attempts to provide tools, theory and techniques for the investigation of spatial data (MacEachren and Kraak 2001; MacEachren and Kraak 1997). However, in order for geovisualization tools to be an effective method for knowledge construction, we require a more complete understanding of how people conceptualise problems and interact with computers. This in turn suggests that one of the biggest challenges in geovisualization is that of defining the range of tasks that may need support (Fairbairn *et al.* 2001; Slocum *et al.* 2001). It is in this context that this thesis constructs three experiments using geovisualization systems in order not only to investigate an appropriate task characterisation method but also to develop an understanding of user approaches and strategies when exploring spatial data visually. These experiments investigated the tasks in which users need support when using area-based socio-economic data (Chapter 6), as well as the cognitive demands of geovisualization (Chapter 2). The first of the three studies (Chapter 3) employed usability testing techniques in order to elicit information from users about difficulties in handling interfaces and to assess how, in practice, such difficulties can hinder the visualization process. It also provided better insights as to how tasks could be defined. The second experiment (Chapter 4) proposes a methodology for task characterisation that allows the adoption of a formal experimental design for testing the effectiveness of tools and understanding user behaviour. The last experiment (Chapter 5) concentrates on methods for refining this task taxonomy and for analysing user strategies when interacting with geovisualization tools.

It is appropriate to begin stating what this research did not seek to do. The purpose of these evaluations was not to establish whether graphics facilitate information processing tasks. This has already been demonstrated by studies in Psychology, such as Larkin and Simon's (1987), which suggests that graphics can reduce the cognitive work that users must perform. Graphics have also been shown to aid user search for information "by grouping related information in a single locality, and by employing encoding techniques such as color, shading, and spatial arrangement that support preattentive [...] visual search" (Casner 1991: 112). Thus, graphical displays allow logical inferences such as numerical comparisons or mental arithmetic to be replaced by faster perceptual inferences such as distance and colour comparisons and spatial coincidence judgements. A simple example is that of a bar chart

where the lengths of the bars on the display represent real numbers. It is the lengths of the bars that are compared and perceptually processed and which can reduce the amount of computation and search needed to understand the data, especially when compared to tasks such as the comparison of numbers that may be more cognitively demanding. In addition, research on the utility of graphics has provided evidence that they facilitate problem-solving and discovery by providing an efficient structure for expressing data (Lohse *et al.* 1994). Furthermore, they

> *"appear to help improve recall for presented data, reduce short-term memory loads during problem solving, provide information about the state of a problem solution, help users organize their knowledge about a problem, increase motivation and user satisfaction" (Casner 1991: 118).*

Neither was the purpose of the evaluations reported here to investigate effective graphic design principles. Cleveland (1994), Tufte (1983, 1990, 1997) and Bertin (1983) provide general design practices. The empirical work conducted by these authors provides guidelines to avoid the design mistakes that can make graphics ambiguous or confusing. Tufte (1983) focuses on the intended purpose and use of graphics and hence his classification of graphs and images does not reflect their physical structure (Lohse *et al.* 1994). Bertin (1983) on the other hand focuses upon the form of the image rather than on its content. Very little is known about why the graphs that Bertin, Tufte or Cleveland investigate are successful in facilitating user cognitive work. This is partly because their work is more concerned with the information that is actually presented than with the user tasks that the graphs can help to address. Hence, this research provides limited guidance when new and innovative graphics are to be created, when different data are to be used, or when a specific task must be supported. Bertin's (1983) guidelines, for instance, are followed by most interface designers but their assertions have not been tested empirically. To the best of our knowledge, at the time of writing this dissertation only one research group at the Department of Geography in the University of California, Santa Barbara (UCSB) was addressing some of these issues (Fabrikant *et al.* 2002). For example, they are examining the effect of some of Bertin's (1983) visual variables in the information processing of spatial displays such as colour. They are for instance investigating how colour coding can alter the perception of membership of a spatial feature to a group based on proximity.

The work presented here also did not seek to evaluate the usefulness of individual graphics for supporting information processing tasks. This has been undertaken to some extent *inter alia* by Casner (1991). He studied the task-specific advantages that graphs offer and showed that different presentations of the same information better support different tasks. In a geovisualization context, MacEachren and colleagues have explored the use of individual

graphical displays for the exploration of georeferenced data and their relative advantages for supporting different tasks (MacEachren *et al.* 1998; MacEachren and Ganter 1990; Edsall *et al.* 2001).

Research has also been carried out in automating the design of graphs for keeping the intervention of the designer to a minimum by providing design rules that can suggest logical choices between alternative presentations. For instance, Mackinlay (1986) developed an application that automatically designed two-dimensional (2D) static representations of relational information. The purpose of his work was to address the graphical presentation problem or the choice of the graphical display to use for particular data and the actual representation of data objects on the screen. Mackinlay codified graphic design criteria in a 'graphical language' that described the syntactic and semantic properties of graphical presentations with propositional formalism. This research, however, is also based on the analysis of the information presented to the user rather than on identification of the tasks that need support.

Casner (1991) extended this work to propose a graphic design and presentation tool that builds graphics based on the analysis of tasks that the graph is to support. He describes "abstract operations [which are] logically adequate to perform the overall task, and selects representations that replace these abstract operations with easy perceptual operations" (Wehrend and Lewis 1990: 140). Roth and Matis (1990) also extended Mackinlay's (1986) work to include more quantitative data types. In addition, they were concerned with analysing data properties based on user information seeking goals. Thus, they defined visual operations to distinguish problems where the goal for using a particular representation differed. Wehrend and Lewis (1990), in turn, built on the work by Roth and Matis (1990) in order to develop a classification of visual techniques. These describe problems in terms of the operations to be supported by the representation as well as by the objects to be represented. In addition, Zhou and Feiner (1996, 1998) have tried to characterise data in order to understand the properties of the information to be visualised and how they are related to the presentation design and to design constrains. Moreover, they were concerned with how these presentation-related data properties relate to user goals and visual display preferences. In addition, they aimed at extending the aforementioned research in terms of developing systems that can automatically design visual presentations for more data types and for dynamic displays.

It is on research concerned with the definition of task typologies, based on user requirements, that we develop the work presented here. The motivation throughout this dissertation is to investigate a range of tasks and to characterise how users address them in a geovisualization domain. In both geovisualization and geographic information science (GISc), understanding how users conceptualise problem domains, process visual displays and link mental schemata to actions through interface tools has seen a revived interest. Nevertheless, only very limited research has been undertaken in characterising either the users or the tasks that they may need to accomplish with computer-based systems. Therefore, there is a lack of theory to guide a user centred design or evaluation of geovisualizations. The following pages illustrate the benefits of drawing from methods and techniques developed within other disciplines including cognitive science, human computer interaction and experimental design for understanding the relation between users, visual data representations and geovisualization tasks. It is argued that, in this way, we may contribute to the development of more solid theoretical foundations of geovisualization. This in turn would allow a more systematic evaluation of geovisualization techniques, tools and implementations that can assist the human capacity for visual inference.

2 Geovisualization as a Cognitive Activity

"The world's total yearly production of print, film, optical, and magnetic content would require roughly 1.5 billion gigabytes of storage. This is the equivalent of 250 megabytes per person for each man, woman, and child on earth" (Lyman and Varian 2000)

Although we rarely use but a small fraction of all the data that are collected, we still face the problem of identifying and sifting through data to search for information. "Lacking the ability to adequately explore the large amounts being collected, and despite its potential usefulness, the data [become] useless and the databases datadumps" (Keim 2001: 39). Hence, we are confronted with the challenge of learning how to explore and make sense of the increasing amount of available data.

The term visualization was mainly used to describe the construction of a visual image in the mind until recently. However, "it has come to mean something more like a graphical representation of data or concepts" (Ware 2000: 1). Visualization has also been defined as the use of visual data representations that engage the human information processing abilities associated with vision for 'making sense' of such data and supporting decision-making (MacEachren *et al.* 1992). Visualization techniques are typically implemented in software systems or computer-based tools that allow user interaction with data and methods of envisioning information, in order to facilitate investigation and thence understanding. However, if these tools do not support the tasks users require to address, if users do not find them useful, or if users are unable to interpret the information encoded in the graphic displays, then the adoption of these tools and the techniques they serve is likely to be hindered, or even fatally undermined.

This chapter defines the scope of visualization as it is relevant to this research, namely as a technique for data exploration and an information processing activity. It discusses a number of visualization methods that are applicable to the geospatial domain but which have largely been developed within other disciplines. This in turn allows us to understand their development, as well as the advantages and limitations of using visualization techniques for investigating datasets and understanding the information they contain. By recognising the purpose and application of visualization for data investigation, we set the scene for the following chapters in which we investigate some of the tasks visualization supports and the way in which users interact with spatial data representations.

2.1 Visualization as Information Processing

Visualization is a cognitive activity (Hearnshaw 1994; MacEachren 1995) the main purpose of which is to gain insight into a dataset by representing it in some visual form that can reveal otherwise hidden information. Visualization can facilitate learning about relations that may be present in data by re-expressing them through "interpretable graphic summaries of spatial information" (MacEachren 1995: 12). Thus, visual data exploration may be particularly useful when solving ill-defined problems where little is known about a dataset or the goals of its exploration (Keim 2001). These are common circumstances in the initial stages of data analysis when users may attempt to interrogate an unfamiliar dataset only "guided by knowledge of the subject under study" (Cleveland 1993: 12). The purpose of visualization is to aid the analyst by facilitating the creation of hypotheses and deriving knowledge from the available data.

Data visualization may serve a multiplicity of purposes, including facilitating learning of known information by an individual or communicating a message to an audience (see §2.3). The emphasis in this dissertation is upon the use of visualization by individuals to discover and understand the information content of a dataset. The motivation is to identify and characterise some of the tasks for which the visualization of spatial data in a digital environment, as well as the interaction of users with such environment and data, is useful. By this we mean that it facilitates the discovery of unknowns and the interpretation of data represented on a computer screen.

It is therefore important to make clear that the position adopted here is that of visualization as an information processing task. In other words, this approach to visualization focuses upon the ways in which information is processed and how stimuli that are perceived or remembered are interpreted. This can be regarded as a top-down approach to processing that emphasizes the importance of concepts, expectations and memory (Matlin 2002) in the understanding of visual displays. This is to say that information perceived visually is interpreted through or matched to some knowledge.

This is not the only approach to visual perception and is in direct contrast to Gibson's (1979) *ecological* approach. Gibson regarded perception as a visually active but cognitively passive process (MacEachren 1995) in which visual scenes *afford* meaning to the viewer. He argued that, under natural conditions, there is sufficient evidence or constancy in the environment to enable a person to *directly* and immediately 'pick-up' something without any mediation by internal representations. This is because the relations between things we perceive are not

only highly organised but also objective and rather constant, even when there is movement or change in the viewer's perspective. Therefore, what we perceive are affordances—or that for which things are good for—to which our senses are attuned or adapted and do not require conceptual processes to interpret them. In other words, we perceive possibilities for action such as surfaces for walking or handles for pulling, which are "physical properties of the environment we directly perceive" (Ware 2000: 23).

Arguably, humans not only perceive stimuli in natural environments. For instance, the visualization tools discussed in subsequent sections and chapters use visual graphic displays in digital environments and utilise "abstract symbolic two-dimensional representations" (MacEachren 1995: 23). Thus, even if perception is direct, graphical displays of data on-screen are hardly natural or organised in an objective and constant manner. On the contrary, processing is usually required to decode the layers of data transformations that may be embedded in a display. Therefore, these representations can hardly be said to afford meaning and rather require training from the part of the user to be effectively interpreted (Andrienko *et al.* 2002).

Nevertheless, the idea of affordance has been applied to the domain of interface design. Norman (1988) suggests that we can build interfaces where graphical widgets clearly and evidently afford some physical operation—such as a button which affords to be pushed. However, their affordances are only indirect in the sense that they derive "their power from our ability to represent pictorially, or otherwise, the affordances of everyday world" (Ware 2000: 25).

Despite its criticisms, Gibson's work suggested that, in order to understand visual perception, it is necessary to consider what vision is for. Marr's (1982) influential work on information-processing systems develops this insight. He discerned three levels of understanding in an information processing task which must be considered when evaluating how various processes function:

'The level of computational theory (at which we describe what a process must do and why, along with a logical strategy by which the process might be carried out), the level of representation and algorithms (dealing with how the theory might be implemented), and the level of processing device or hardware implementation (that considers how a particular representation might be implemented in the available device)" (MacEachren 1995: 27).

Starting from the last level upwards, for this dissertation we considered two computer-based systems as implementations of geovisualization tools for exploring spatial data and evaluated whether they supported geovisualization tasks. These computer systems in turn implement

representations and techniques discussed later in this chapter which have been identified as successful in allowing geovisualization. Finally, we considered how people utilised these tools and for what purpose to derive an understanding of tasks and user behaviour in a geovisualization context.

In addition, Marr's work suggested a model of how vision functions (Marr and Nishihara 1978; Marr 1982; Marr 1985). The argument is that we first extract primitive features of the image and their orientation in a *primal sketch* which contain image elements with attribute values "derived by looking for contiguous areas of relatively uniform color and light intensity characteristics and the discontinuities between them" (Peuquet 2002: 98). These elements are then grouped to derive a 2.5-dimensional (2.5D) sketch which contains a viewer-centred representation of the properties of the visible surfaces such as their orientation, illumination conditions and distance from the viewer (Marr 1985). Thus, like Gibson's ecological approach discussed above, Marr (1982) argues that there is a pre-cognitive processing up to this point, where there is no "global information about shape, only depths and orientations of local pieces of surfaces" (MacEachren 1995: 30). In the final stage, an object-centred 3D model is derived, which contains information about shapes, as well as about the structure and arrangement of each part of the scene with respect to each other.

Marr's modular approach suggests that different processes are at work during early vision. They are also at work during higher levels of processes where the output of early vision is used in combination with existing knowledge (Pinker 1984). The concern here is with higher processes and, more specifically, with how visualization tools can serve as a "brainstorming environment" (Wood 1994: 10) during the initial stages of scientific enquiry.

2.2 Advantages of Visualization

One of the advantages of visualization is the amount of data that can be represented and readily interpreted. An example follows. Figure 2.1 shows a representation of employment data for an area of London, UK. The original dataset, collected by the Office for National Statistics (ONS) for the Annual Business Inquiry (ABI), contains approximately three million records of information georeferenced to the unit postcode level. These data were used to define a methodology for identifying town centres in London (Office of the Deputy Prime Minister 2002; Thurstain-Goodwin and Unwin 2000). When the data are converted into a surface with its height representing the value of a variable such as employment, a number of issues become evident. The first is the very small or nearly negative values which show as blue spots in the surface. They are mainly caused by *non*-town centre employment, like

warehouses and factories. In addition, the visualization gives a quick impression of data features such as how major concentrations of employment tend to be located along main roads.

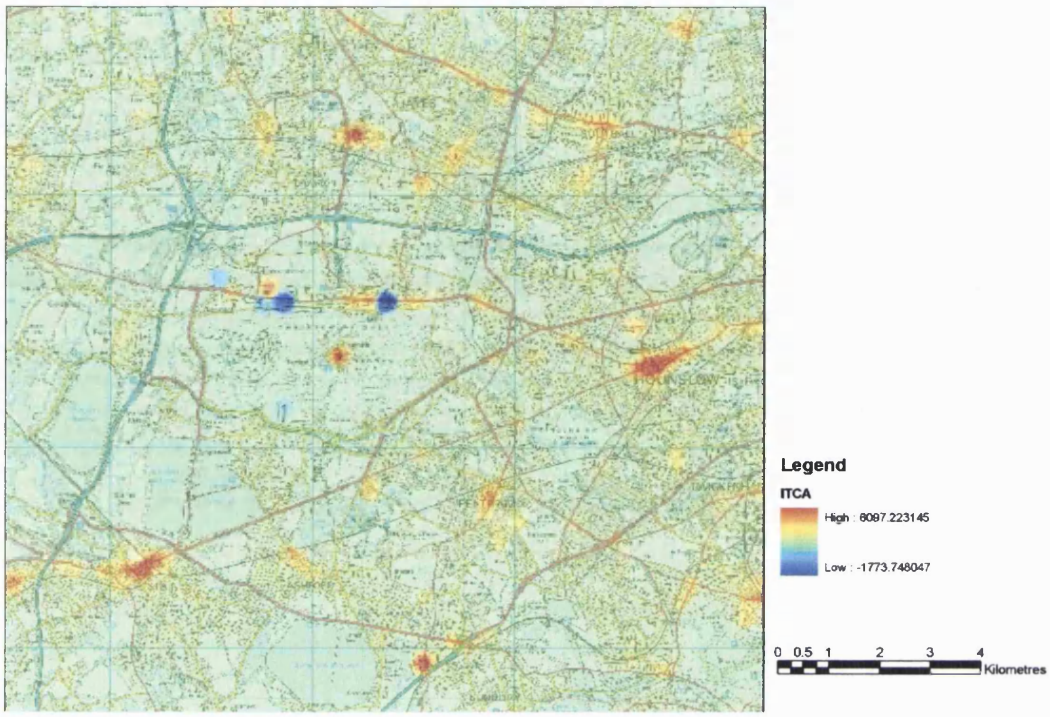

Figure 2.1 Surface of employment data for London

Figure 2.1 illustrates some of the advantages of visualization. First, it can allow the representation and understanding of large amounts of data, as aforementioned. Second, it commonly reveals problems with the data such as errors or artefacts that may not be readily obvious. For instance, apart form aiding he identification of non-town centre activity, representations like that in Figure 2.1 prompted the discovery of issues in the data collection process. These have been investigated as part of a research project studying the effect of error propagation in the town centres definition (Office of the Deputy Prime Minister 2002; Lloyd 2004). Third, visualization can support a process of hypothesis formation as it typically reveals features or properties that where not anticipated. For instance, the visualization of an Index of Town Centre Activity (ITCA), which combines a number of surfaces including that in Figure 2.1, showed how the spatial location and extent of town centres are temporally unstable. This inspired their investigation and definition as fuzzy objects (Lloyd *et al.* 2003). Finally, visualization can also be effective when communicating these outputs. In the town centres project, a major concern was transmitting the methodology used for the town centres definition to stakeholders. Lloyd *et al.* (2003) discuss how some representations such as

graduated points or contours where less effective in communicating the results of the project. On the other hand, the authors discuss how the surface metaphor was appropriate for explaining and discussing the index, which was a crucial step for facilitating stakeholders' views on the methodology. This in turn allowed the consideration of their concerns and gave stakeholders confidence about the validity of the town centres index (DETR 1998; Office of the Deputy Prime Minister 2002).

2.3 Visualization Fields and Methods

A number of fields have built on the idea of using visualization for data exploration to propose a number of methods, techniques and principles to represent data in graphical displays. This section introduces developments in three fields—Scientific Visualization, Information Visualization and Exploratory Data Analysis—where a variety of visualization methods have been developed. The techniques reviewed share the common objective of uncovering information by representing data visually that would have otherwise remained hidden. These methods have been studied at great length in the literature (Tukey 1977; Cleveland 1994; Jacoby 1998; Orford *et al.* 1998; Card *et al.* 1999; Ware 2000; Spence 2001). Hence, the main purpose of this section is to provide a description of some well-known visualization techniques and graphic displays and provide examples of how they have been used with georeferenced data. It particularly considers concepts, "types of graphs and ways of encoding quantitative information" (Cleveland 1994: 119) that are discussed or drawn upon in subsequent sections in this dissertation.[1]

2.3.1 Scientific Visualization

In 1987 a formal initiative, developed by the U.S. National Science Foundation Panel on Graphics, Image Processing and Workstations, pointed to the need for new methods of analysing for the vast amounts of complex, often multivariate data, streaming from computer models, remote sensors and other automatic observational instruments:

> *"The deluge of data generated by supercomputers and other high-volume data sources [...] makes it impossible for users to quantitatively examine more than a tiny fraction of a given solution. That is, it is impossible to investigate the qualitative global nature of numerical solutions" (DeFanti et al. 1999: 39).*

The initiative, known as Scientific Visualization or Visualization in Scientific Computing (ViSC), was a response to the demands from a broad range of fields for techniques that enabled the exploration of large datasets. ViSC was defined as "a tool utilizing the human

sensory modalities to understand complex, abstract information and gain insight that would otherwise be unavailable" (McCormick *et al.* 1987). This included the use of computer graphics to provide graphical representations of datasets in more flexible computer environments aimed at fostering the generation of hypotheses (Fisher *et al.* 1993). For these reasons, ViSC was "recognised as a method and product that integrated the power of digital computers and human vision and directed the result towards facilitating insight across the sciences" (McCormick *et al.* 1987).

ViSC is mostly concerned with representing physical objects and the results from complex models and functions. ViSC's approach is that of displaying "sufficient detail to highlight critical features" (Unwin 1998: 2) while still handling large amounts of scientific data to make it an effective analysis technique. Surface and volume visualizations are commonly used representations for exploring different features and properties of the data represented. Surfaces, for instance, have been used for modelling physical terrain which can typically be navigated by rotating a display or by moving across the surface to examine its changes across space.

In geography, surfaces have been extensively used as a metaphor for representing data attributes such as population density in order to examine their pattern across an area (Wood *et al.* 1999). They have also served to assess the output of simulations that aim at understanding physical processes. Such is the case of landscape simulation problems where surfaces are used as an exploratory tool to investigate the implications of algorithms that form the basis of the simulation models (Mitas *et al.* 1997). As a last example, surfaces have also been used to understand the role of uncertainty in the output of spatial analysis. By dynamically viewing transformations or realizations of a surface, users may investigate and visually perceive the probabilistic nature of uncertainty (Ehlschlaeger *et al.* 1997; Beard *et al.* 1991).

Volume visualization, on the other hand, can aid in the understanding of inner structures or of amorphous phenomena. Examples of the former use of volume visualization include the use of imaging techniques such as magnetic resonance to produce images of the human body. Examples of the latter include flow simulation of phenomena in areas such as aerodynamics, meteorology or oceanography. In those fields, animation is frequently used or

---

[1] The datasets used in most of the visualizations are explained in detail in Chapter 6. Briefly they are two vector datasets with socio-economic census data for Bristol, UK (from the 1991 UK Census of

made available to the user for exploring results that can be better "expressed and taught visually over time" (DeFanti *et al.* 1999: 40). Visualization may therefore permit the analyst to rapidly derive some information about the patterns or processes that might be in operation although buried in a large dataset.

Fisher and Unwin (2002: 1) suggest that scientific visualization in geography is closely tied, if not equal, to virtual reality (VR) defined as "the ability of the user of a constructed view of a limited digitally-encoded information domain to change their view in three dimensions causing update of the view presented to any viewer". Coupled with GIS, VR systems have been used to address a myriad of problems where the exploration or analysis of large georeferenced datasets has been the underlying motivation. In urban and regional planning for instance, VR has been applied to the production of large-scale urban models to assess the visual impact of planning schemes or communicate ideas throughout the planning process (Haklay 2002; Miller *et al.* 2002; Batty *et al.* 1998). Similarly in environmental modelling, VR has been used to assess the impact of phenomena such as land contamination or air pollution (Bishop and Karadaglis 1996), and more generally, to assess changes in the landscape and the social acceptability of such changes (Brown *et al.* 2002). In GISc, VR has been utilised to construct and represent digital terrain models, as well as the visualization of the structure and scale dependency of landscape surfaces (Wood 2002). VR has also been used in education to present information graphically, "in an immersive and engaging format, whereby the user's experience of operating in the real world can be used successfully to interact with the information" (Dykes 2002: 68). For instance, the Virtual Field Course initiative built an application that combined spatial information and photographic imagery to produce an environment that supports student fieldwork (Dykes *et al.* 1999). Thus, VR tools and applications have been of interest to academic geography because of their potential to aid the description, explanation and communication of concepts of varying complexity to a large range of users (Brown *et al.* 2002).

2.3.2 Information Visualization

Card *et al.* (1999: 6) define visualization as "the use of computer-supported, interactive, visual representations of data to amplify cognition" and information visualization as "the use of computer-supported, interactive, visual representations of *abstract* data to amplify cognition" (emphasis added). Hence, information visualization deals with abstract data that are not necessarily mappable to the geographical space and therefore may not have a "natural and

---

population) and for Los Angeles, USA (from the US 2000 Census).These same datasets are used in the evaluations discussed in Chapters 3-5.

obvious physical representation" (Gershon *et al.* 1998: 10). Moreover, while scientific visualization is concerned primarily with the representation of physical objects, the interest of information visualization is in abstract concepts. However, both application fields comprise methods for rearranging—interactively and in real-time—how the data are displayed visually to prompt additional insight into the information they provide.

Nevertheless, the key research issues in information visualization deal with classes of data that may lack "inherent 2D or 3D semantics" (Keim 2001: 40) and therefore require innovative visual metaphors in order to represent the information they may provide. Some examples include making sense of information contained in text documents, flat file databases with business information or properties of software programs. Basic techniques and their development into applications for the handling of abstract data are described below. Their relevance or application to the spatial domain is also discussed where appropriate since some of these concepts will appear recurrently in subsequent chapters.

2.3.2.1 Focus+Context

Focus+Context techniques are intended to display a large amount of information at the user's interest point. Thus they build on the idea of analysing a portion of the data while still retaining an overview. They take into account the fact that computer displays have a limited extent of screen space and hence cannot display the entirety of the data especially if it is large in size (Card *et al.* 1999). Furthermore, they also consider that users have a limited resource of attention over time in terms of the data elements they can hold in memory at any one time (Miller 1956).

Providing context aims at reducing search, allowing the detection of overall patterns and aiding the user in choosing what to attempt next in their exploration process (Card *et al.* 1999). A technique known as *Overview+Detail* serves this purpose and has been used in a large number of application areas for both abstract and physical data. Magic Lenses are an example implementation for discerning detail from a small portion or area of a dataset while keeping the surrounding information intact for context. Magic Lenses combine an arbitrarily-shaped viewing frame with an operator that changes the look of objects viewed through that shape or 'lens'. Their function is to access "the underlying data structures of an application and reformatting the data to generate a modified view" (Spence 2001: 125). Magic Lenses can be used to reveal more detail, for instance to provide a more detailed road map as in Figure 2.2 (Left), or highlight features of interest such as the main roads in Figure 2.2 (Right). This technique has been implemented in at least one commercial GIS package,

Cadcorp SIS (Stevenage, UK). It can be used for instance to edit data on-screen without losing the wider geographic context or to view spatial layers of different scales simultaneously.

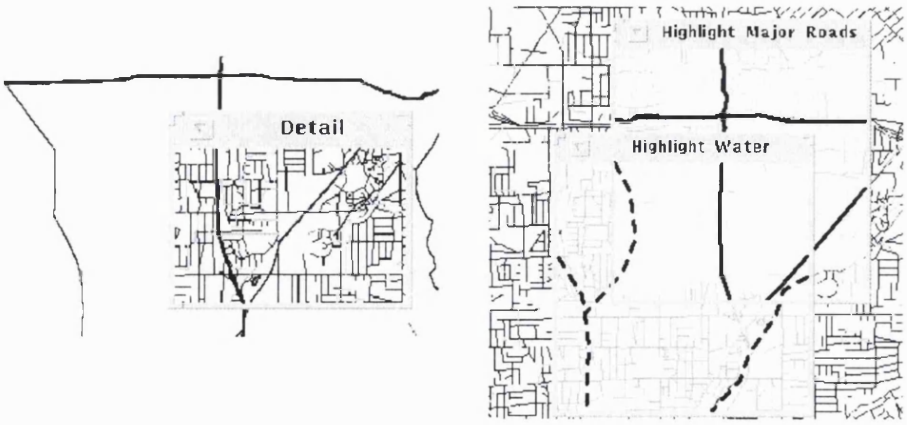

Figure 2.2 Magic Lenses: Left) For revealing details a road map. Right) For highlighting features in a map. Maps extracted from Stone *et al.* (1994)

A fisheye lens (Furnas 1986) is another Focus+Context example which implements a slightly different operation: it magnifies an area of interest and de-magnifies the surrounding area while leaving it visible. It distorts the image in a way similar to a wide-angle camera lens, which makes the magnification higher in the centre of the lens than in the borders. Therefore, Focus+Context methods perform operations on the data to achieve new representations that may prompt information of interest.

Zooming, or the "increasing magnification of a decreasing fraction (or vice versa) of a two-dimensional image under the constraint of a viewing frame" (Spence 2001: 130), is standard functionality in any GIS. Similarly panning, or the "movement of a frame over a two-dimensional image of greater size" (Spence 2001: 130), can aid in the local examination of areas of interest. Most GIS allow for these functions to be customised. For example, the zoom factor can be pre-defined in a map display to take the user to a layer showing a particular level of detail. Moreover, the visibility of spatial data layers can be defined, very much like a Magic Lens, according to how zoomed in or out the user is. These techniques reveal details as users zoom in or conceal it as they change their focus of interest to an overview of the area. This functionality is enhanced when context maps are provided, as shown in Figure 2.3. The upper left corner of the display in Figure 2.3 provides an overview of the location where the user is performing some analysis relative to the smaller scale area of

interest. Thus, Overview+Detail may be usefully applied when dealing with large geographic areas by providing a visual cue of the spatial background.

Figure 2.3 Focus+Context example with georeferenced data (Komzak 2000)

These examples illustrate three premises of Focus+Context techniques (Card *et al.* 1999). First, users may need both an overview of the data and detailed information simultaneously. For instance, when exploring detail in a digital map covering a large area, this allows combining piecemeal observations into a single overhead perspective. Second, the kind of information needed in the overview might be different to that required from the detail. For example, changing the visibility of features in a GIS data layer as a function of the zooming operation may only be relevant in the focus (small scale) map of Figure 2.3 where the change can be observed. On the contrary, the red square in the upper left corner which provides the spatial context of the analysis would not be of use in the main (smaller scale) view as it would occlude the map and not add any useful information. Third, both overview and detail information can be combined within a single display, as in the examples illustrating the use of Magic Lenses on a geographic data layer (Figure 2.2).

Risden *et al.* (2000) show that for improving search performance, integrating the ability to view the overall structure of an information space with easy access to global and local relations is a key feature. However, there might be a competition for resources between providing detailed views and relevant context as they both require screen space. This competition must be resolved carefully for the display to be effective, that is, serve its

purpose of amplifying cognition. One approach suggested by Card *et al.* (1999) is to selectively reduce the information for the contextual information by means of techniques such as distortion.

2.3.2.2 Distortion

Distortion can make more effective use of display space while representing information in a way that is clearer or more obvious for the user to interpret (Spence 2001). It can be achieved by changing the size, position or perspective of objects or, more generally, by changing the number of pixels devoted to represent objects in display space (Card *et al.* 1999: 308). A famous example is the map of the London Underground transport network that was originally created by draughtsman Harry Beck in 1931(Garland 1994). Although the map does not represent the exact spatial arrangement of the lines and stations, it conserves their sequence and connectivity, thus allowing the map to gain a degree of independence of scale while still providing sufficient information for the navigational purposes of its users.

A further example using abstract data is the Bifocal Display (Spence and Apperley 1982) later developed into the Perspective Wall (Mackinlay *et al.* 1991). The Bifocal Display is a 3D perspective technique for linear information which follows the analogy of a "strip of paper to be pulled back across two posts" (Spence 2001: 116). The view comprises a central panel, which provides detailed information, and perspective panels on either side, which display contextual information (for this reason it is also considered an application of Focus+Context). By distorting information in its x dimension, the side panels make users aware of other data surrounding their point of focus. Furthermore, the use of colour as well as (x, y) position on the panel can be employed to encode various attributes of the data. Finally, the 'strip of paper' can be scrolled to allow a change of detailed view and browse through the information.[2]

Distortion techniques are also used as a means of mapping n-dimensional datasets into 2D space or a 3D volume. Projection Pursuit and the RADVIZ methods (see Table 2.1) provide examples of how linear and non-linear projection algorithms have been used for this purpose (Brunsdon *et al.* 2000). Projection methods are typically useful for identifying clusters in the data and outliers from those clusters. They have been used extensively in geography for exploring compositional data in which variables represent the proportion or percentage of the attribute for an area (Dorling 1990; Coombes *et al.* 1993).

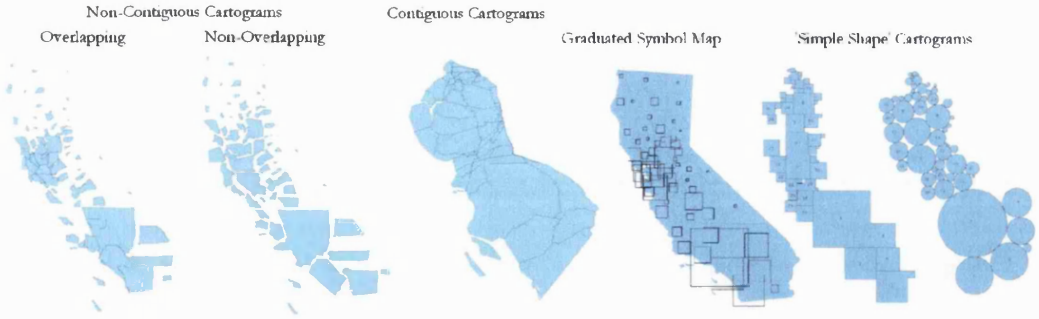

Figure 2.4 Non-Contiguous, Contiguous and 'Simple Shape' Cartograms (maps taken form Bortins and Demers 2002)

A good example of the use of distortion in the context of geographical data are cartograms, a type of graphic that depicts attributes of geographic objects by distorting the object's area, shape, distance or even topology. A number of cartograms have been proposed including contiguous and non contiguous cartograms as shown in Figure 2.4. In non-contiguous cartograms, geographic objects do not have to maintain connectivity with adjacent objects in order to allow them to grow or shrink in size according to some attribute value. Their shape, however, may be preserved (Bortins and Demers 2002). Contiguous cartograms on the other hand do preserve their topology but distort the object's shape. Dorling (1996) discusses a number of methods for creating cartograms including those to the far right of Figure 2.4 where the geographical object is replaced by some geometrical shape such as a circle or a square. They of course do not maintain their shape and may not maintain their topology. The purpose of these distortions, however, is for the cartogram to serve as a base map upon which other statistics are shown.

Cartograms have been extensively used for displaying census or socio-economic information for a good reason. People are unevenly spread across the space and areas with high population and high population density tend to be very small in map based (areal) representations. Hence, traditional maps have a propensity to allow large and sometimes unimportant areas (e.g. where few people live) to visually dominate the map. In other words, maps tend to "highlight the uneven distribution of people over the land rather than the differences between groups of people, which are of more interest" (Dorling 1996: 7). Therefore, by making area proportional to some attribute, cartograms may overcome major

[2] See <http://www.parc.xerox.com/istl/projects/uir/images/perspectivewall.gif> and <http://kogs-www.informatik.uni-hamburg.de/~moeller/pwall/examples.html> for examples.

draw backs of using traditional cartographic representations for portraying socio-economic information. Cartograms have been criticised because slight alterations to the algorithms used to produce them can result in very different outputs even for the same places and data (Dorling 1996). However, it is also the different perspectives they show about a process being represented which can be seen as their main strength for exploring data.

2.3.2.3 Dynamic Interaction and coordinated views

Interactive and dynamic properties of visual representations may give users the experience of being in direct contact with the data (Ware 2000). "Interaction techniques allow dynamic changes of the visualization according to the exploration objectives, but they also make it possible to relate and combine multiple independent visualizations" (Hinnenburg *et al.* 1999: 24). Allowing users to relate visualizations has been identified as an important feature of effective information visualization (Haslett *et al.* 1991). For instance, North and Schneiderman (2000) have pointed at how multiple coordinated visualizations enable the rapid exploration of complex information and improve user performance. Therefore, data objects or different attributes of one row or database record often appear in more than one display view. Nevertheless, these various graphical displays must be related in some way so as to allow the user to interpret the complementary information that the representations provide.

Brushing is a technique which utilises the dynamic highlighting of all the objects in visual displays that correspond to the same record in a dataset (Cook *et al.* 1997; Monmonier 1989; Becker and Cleveland 1987; MacDonald 1982). The first reference to brushing is attributed to Newton (1978) who defined brushing as interactively changing the representation of a group of points to a unique colour or symbol. In other words, when a group of data objects is highlighted in one graphical display, brushing allows highlighting of the same group in all other views in which they appear. Figure 2.5 shows an example with three data views: a map, a scatterplot and a parallel coordinate plot (PCP). The three graphic displays show attributes from a geographical dataset. Note that a set of objects was selected on the scatterplot and highlighted in light blue. Using brushing, all corresponding objects are highlighted in each of the linked views using the same colour, which facilitates relating the information provided by each graphical display (MacDonald 1982). Thus, the technique allows the investigation of multivariate data as different attributes can be explored simultaneously.

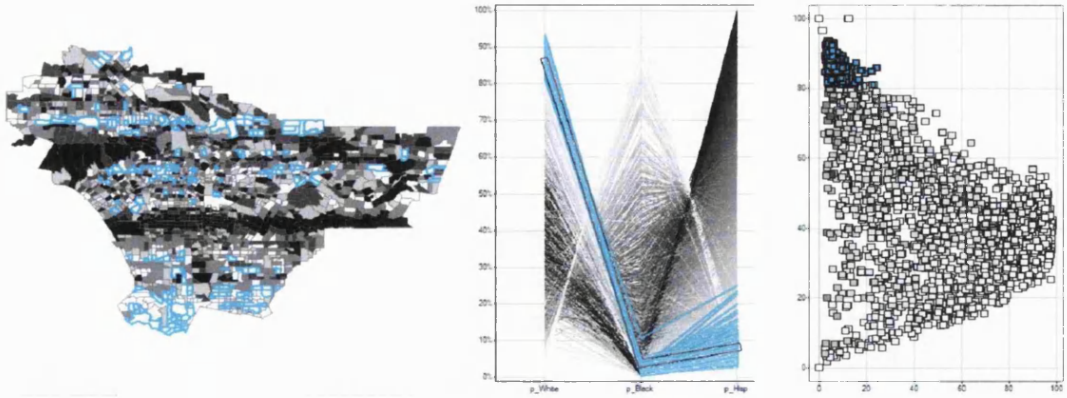

Figure 2.5 Brushing and linking views with geospatial data

Dynamic querying (Williamson and Shneiderman 1992) is a technique that allows users to perform structured queries on-the-fly using graphic devices such as scrollbars. Each scrollbar in Figure 2.6 (Right) represents one attribute in a dataset. By dragging the ends of any of the scrollbars, users can interactively manipulate the limits of the attributes of data records and, in that way, the number of data objects that are visible in a graphic display. Note that in the example of Figure 2.6, three scrollbars have been manipulated so that only records within certain value ranges for the three attributes are visible in the PCP. If the response to the dynamic query occurs in near real time, the technique allows users to pose and solve questions concurrently, as they can manipulate the graphic device and see the result of the operation at the same time. Tobón (2003) describes an environment that implements these devices in a geospatial domain, which is also discussed in detail in Chapter 4.

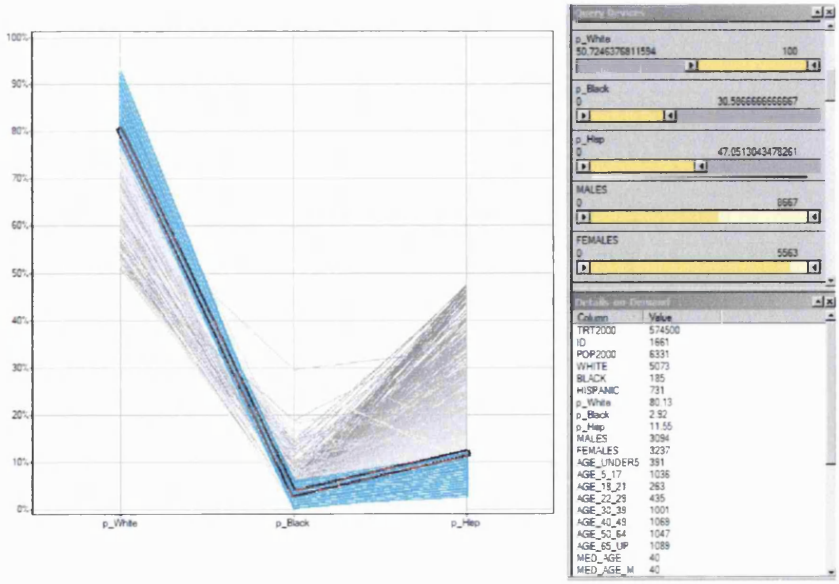

Figure 2.6 Parallel Coordinate Plot (Left) and Dynamic Query devices and Details-on-Demand (Right) in DecisionSite (Spotfire, Boston)

Details-on-Demand (Ahlberg and Shneiderman 1994) is a technique by which details of all attributes associated with a particular data object may be required by the user. Figure 2.6 (Right) shows an example where the specific values of attributes in a record of interest are queried. By clicking on the object in the graphic display, the Details-on Demand window is populated with the relevant information. This concept has also been implemented as off-the-shelf functionality in many GIS packages. Figure 2.7 shows an example from ArcGIS (ESRI, Redlands). By clicking on a spatial object, all attribute values associated with its corresponding record in the database may be presented to the user. This technique is useful in data exploration for a number of reasons including assessing the relevance of selected subsets of data objects when the user is trying to address some task or reach an answer to a particular problem scenario.

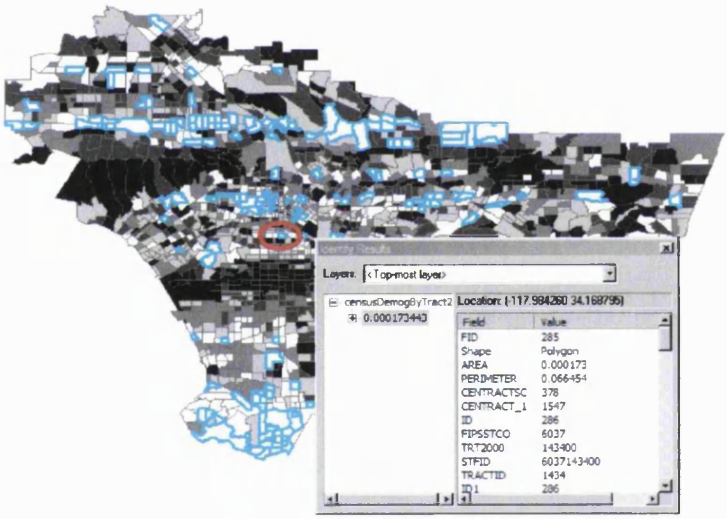

Figure 2.7 Details-on-Demand in ArcGIS (ESRI, Redlands)

2.3.2.4 Iconographic Methods

Chernoff faces (Chernoff 1973) are representations in which data attributes are mapped onto the features of a simplified image of a human face, such as the size of the eyes or the curvature of the lips. There are a number of reasons why faces have been used. First, faces are an important class of object in the human environment which can be rapidly distinguished (Ware 2000). Second, this type of display is said to reduce visual clutter by integrating many variables into a single object which, in addition, facilitates the integration of multiple sources of information (Wickens 1992). Dorling (1994) provides an example in the

context of georeferenced data for illustrating British voting behaviour relative to other socio-economic variables (Figure 2.8). Nevertheless, a number of problems can arise when using faces. Ware (2000), for instance, argues that the combination of variables may result in stereotypical faces such as a sad or a happy face which are identified more readily. This could make some of the faces visually more dominant. In addition, we may be more sensitive to some features than to others such as the height of the eyebrows, which implies that the perception of Chernoff faces is non-linear.

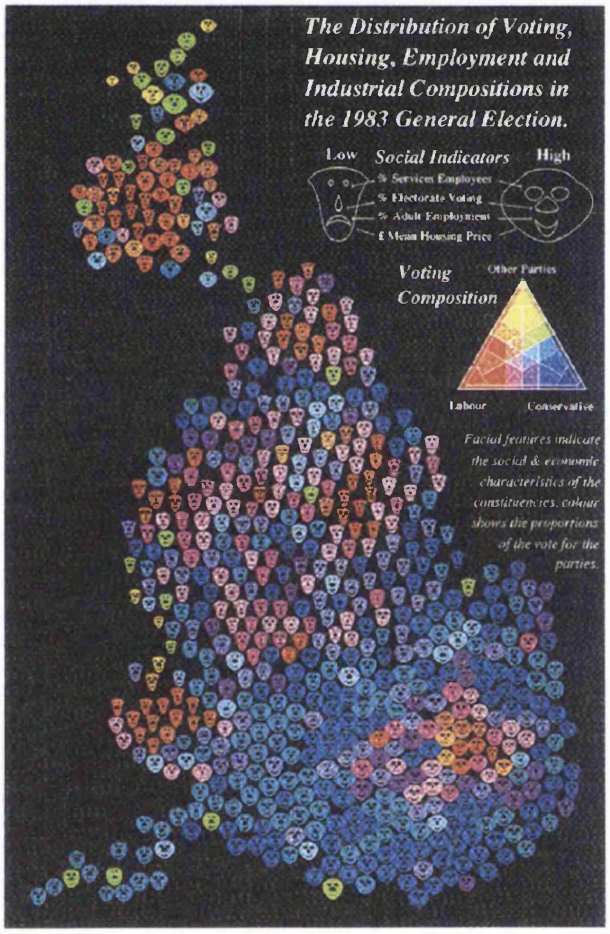

Figure 2.8 Dorling (1994) cartogram showing the distribution of voting, housing, employment and industrial composition in the UK 1983 General Election.

Star glyphs provide another example of iconographic methods for representing multivariate data. They create n rays emanating at equal angles from a point on the screen, where n is the dimensionality or number of attributes in the dataset (Figure 2.9). Hence, a number of objects each represented by a star plot can be compared according to their shape. Therefore, mapping multiple attributes onto a single icon by merging them into a single display may

provide more information than is observable from viewing the individual attributes displayed side by side (Erbacher and Grinstein 1995).

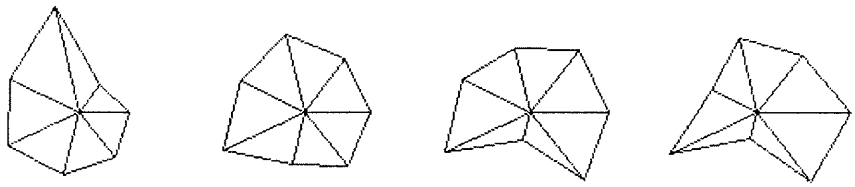

Figure 2.9 Star Glyphs from XmdvTool (Ward 1993)

2.3.3 EDA

Exploratory Data Analysis (EDA) "consists of a collection of descriptive and graphical statistical tools intended to discover patterns in data and suggest hypotheses by imposing as little prior structure as possible" (Anselin 1998: 78). Like visualization, these techniques are an aid for acquiring insights about data rather than a confirmatory analysis technique. As Tukey (1977: vi) suggests:

> *"Pictures that emphasize what we already know—"security blankets" to reassure us—are frequently not worth the space they take. Pictures that have to be gone over with a reading glass to see the main point are wasteful of time and inadequate of effect.* ***The greatest value of a picture*** *is when it forces us to notice* ***what we never expected to see****" (original emphasis).*

Arguably, both exploratory and confirmatory techniques—such as statistical analysis—should go side by side (Tukey 1977), but only the former methods are discussed here. Emphasis is placed on EDA approaches that take advantage of the interactivity provided by computer systems through direct manipulation interfaces, as well as the vast human capacity for visual information processing (Ahlberg and Schneiderman 1994) to allow the exploration of larger datasets in shorter periods of time (Schneiderman 1998).

Closely related to EDA are the exploratory spatial data analysis (ESDA) techniques defined by Anselin (1998: 79) as "a subset of [EDA] that focuses on the distinguishing characteristics of geographical data", which explicitly accommodate issues of spatial autocorrelation and spatial heterogeneity. Their emphasis is on graphical displays that can aid in the detection of spatial patterns, clusters or outliers through visual exploration. They however are not discussed further as they transcend the scope of this dissertation. Nevertheless, some graphical displays commonly utilised for both EDA and ESDA are discussed next.

2.3.3.1 Dotplots, Boxplots and Histograms

One variable displays such as dotplots, boxplots and histograms are frequently used to show the empirical distribution of a variable or the variation in its measurements (Cleveland 1994). Dotplots are graphs that represent the statistical distribution of a single variable using symbols located along one axis. Figure 2.10 shows an example of their use in 'cdv' (Table 2.1) or the Cartographic Data Visualizer (Dykes 1999), where dotplots can are utilised to enhance the visualization of attribute distribution. Three variables are plotted *(Buying, Unemp, Council)*[3] together with a map of Bristol. The boxed items in the middle bar, for instance, suggest that some values of *Unemp* are unusually high and may be outliers in the data.

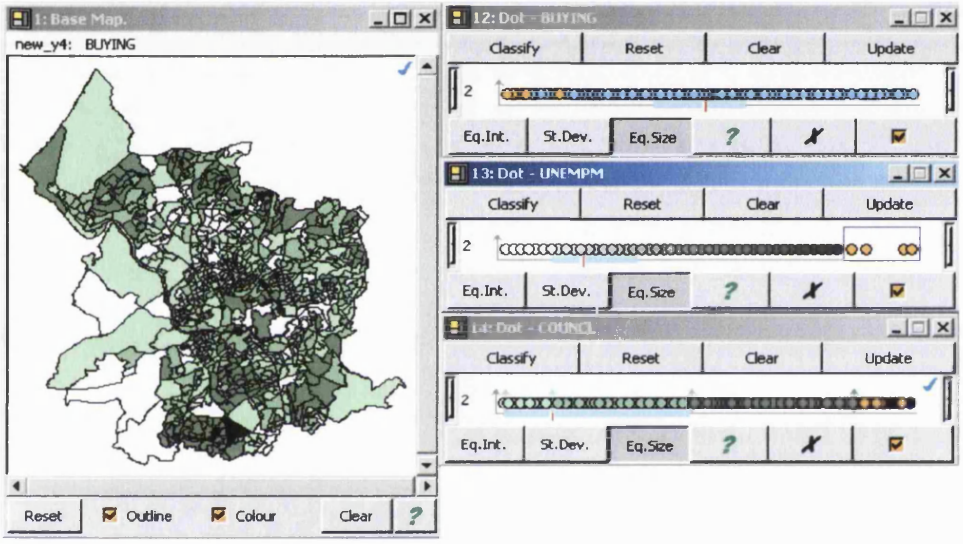

Figure 2.10 Dotplots and Boxplots

The information provided by these dotplots may be enhanced by adding a boxplot to their axes as indicated by the small blue boxes underneath the dotplots in Figure 2.10. Boxplots are a visual representation of the so called 'five number summary' of a variable which describes its distribution according to its minimum, maximum, median, first quartile and third quartile values.[4] These values are represented in Figure 2.10 by the extreme points of the horizontal grey line (minimum and maximum), the red bar (mean), and the ends of the light blue box (interquartile range). Interpreting the dotplots together with the boxplots

[3] *Buying* refers to the percentage of people in the area who are buying the property in which they live; *Unemp* is the percentage unemployed and *Council* the percentage living in council housing. See Chapter 6 for a detailed description of the dataset.

[4] Each quartile value is approximately calculated as the proportion of data that have smaller values than a given observation. Hence, 25% of the data are smaller then the lower quartile, 50% are smaller than the median, and 75% are smaller than the upper quartile.

indicates that while *Buying* is evenly distributed along its median but with long tails, *Unemp* and *Council* are highly skewed.

Combined with brushing, these graphic displays can be used to inform the visual analysis of the distribution of variables in space. In Figure 2.10, the map is colour coded according to the five number summary of the *Council* variable. Note that the four extreme values in the *Unemp* dotplot have been selected and coded in a different hue to explore their relation to other attributes. The corresponding objects in all other views are highlighted with the same colour. This shows unemployment and residence in council housing to be positively related. In addition, it shows that the enumeration districts (ED) with the highest proportion of these variables are in central Bristol.

Histograms are a common way of showing the distribution and spread of data. However, their shape is very sensitive to the (sometimes arbitrary) way in which data are grouped into bins used to categorise continuous data values. The histograms of *Buying* in Figure 2.11, generated in DecisionSite (Spotfire, Boston; Table 2.1), show a manifestation of this problem. If the bins are too narrow as in the histogram to the left, they can exaggerate local variation but, if they are too broad, they can hide it as in Figure 2.11 (Right).

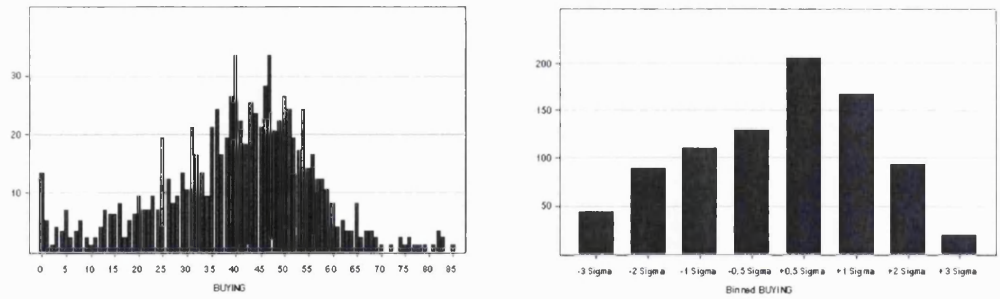

Figure 2.11 Histograms: Left) Narrow bins. Right) Broad bins classified by standard deviation.

The apparent weakness of histograms—their dependency on the binning—can actually be a useful investigation tool. However, a further criticism argues against the very nature of histogram displays because they can hide the richness of continuous data. Dividing the attribute values into discrete bins can occlude its internal variability and produce discontinuities that are often not intrinsic to the original data. Therefore, other representations such as probability, quantile or percentile plots are regularly used instead as discussed in the following section.

2.3.3.2 Geometric displays

Very frequently, a number of cases may be recorded for more than three variables. Hence, a point representing each record in a dataset is said to be multidimensional. Scatterplots use two perpendicular axes each associated to a variable in the dataset. Each item in a view corresponds to combinations of values of the two variables plotted for the same record in the dataset. This graph is a useful tool for identifying relations, clusters, outliers and trends between at least two variables. More than two dimensions or variables can be encoded in a scatterplot by altering the representation of its objects or glyphs. See for instance the scatterplot in Figure 2.12. While the x and y axes are associated to two attributes in the dataset *(Hhtcar* and *Inc_Score)*[5] the glyph colours are used to encode the value of a third variable *(Quaml)*[6] and their orientation the values of a fourth *(Wkwife)*.[7] If necessary, their shape could also correspond to the values of a fifth attribute.

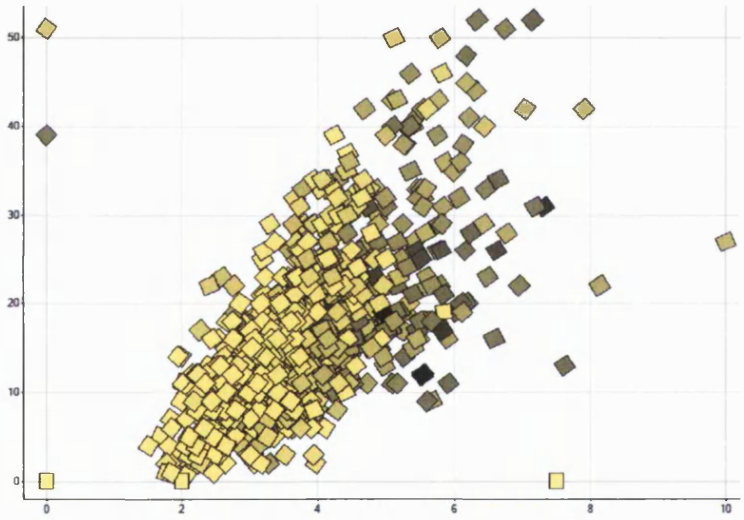

Figure 2.12 Scatterplot of *Inc_Score* vs. *Hhtcar*. *Qualml* is encoded by the colour of the glyphs where yellow corresponds to a low value and black to a high value of the variable. *Wkwife* is encoded in the orientation of the glyph or the deviation from 90 degrees as the value of the variable increases.

When organised as a matrix, a collection of scatterplots can be used to explore the relations between larger numbers of variables. In a dataset with n variables, a scatterplot matrix arranges all $n(n-1)$ possible bivariate scatterplots into a square symmetric table. In such an

---

[5] *Hhtcar* indicates the percentage of households with more than two cars in Bristol. *Inc_Score* is a measure of household income derived from a lifestyles dataset, which ranges from 1 to 10 (see Chapter 6 for details).

[6] *Qualml* is the percentage of qualified males in an ED in Bristol (see Chapter 6).

[7] *Wkwife* is the percentage of married women in an ED in Bristol who work (see Chapter 6).

array, "each scatterplot in a row has the y variable in common, and each scatterplot in a column has the x variable in common" (Fotheringham *et al.* 2000: 75). Cleveland (1993) argues that facilitating comparison is the main reason for the redundancy of scatterplots in some matrices where, for every two variables, say l and m, two scatterplots are includad in the array: l vs. m and m vs. l. In Figure 2.13, only half a scatterplot matrix for three variables was plotted using DataDesk (DataDescription Inc., Ithaca; Table 2.1) to illustrate how the linking of all the views enables the comparison of the various relationships between attributes. See for instance how the brushing (red objects) allows seeing that high *Inc_Score* values are related to high *Hhtcar* and *Qualml* values.

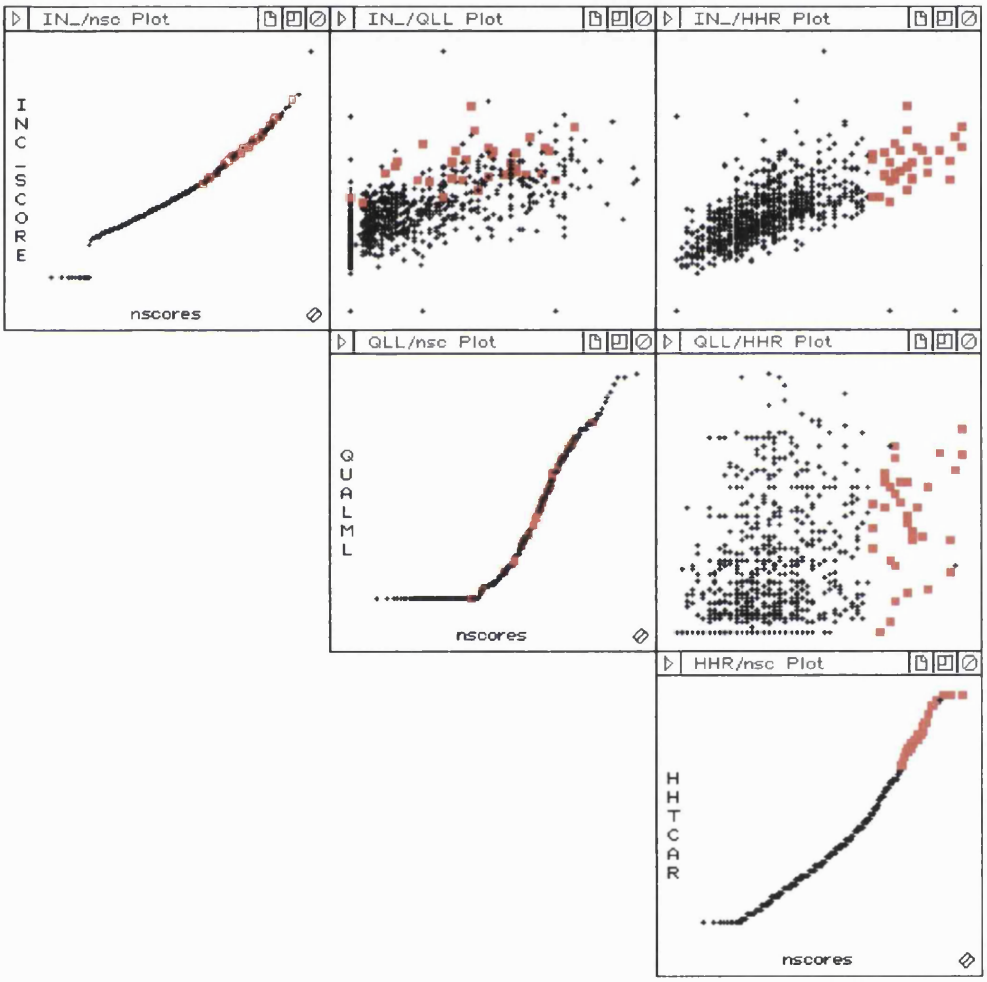

Figure 2.13 (Half) Scatterplot Matrix and the use of brushing

The plots in the diagonal of Figure 2.13 are Normal Probability Plots—also known as Normal Quantile-Quantile (Q-Q) Plots. Normal QQ plots provide a simple way to identify

whether a variable is approximately Normally distributed.[8] They belong to a larger family of probability plots called percentile plots. In such graphical displays, the x values correspond to the values of the attribute of interest while the y values are their percentiles.[9] The percentile plot is therefore created by drawing the points in the x, y plane and connecting them with straight line segments. A quantile-quantile plot graphically compares two batches of values by matching corresponding percentiles in each batch. A normal QQ plot therefore plots the ordered values of an attribute and compares its percentiles against the same percentiles for the standard Normal distribution. This facilitates the visual assessment of the normality of the values by imposing some "structure on the data in an attempt to help us decide what variation [from the Normal distribution] appears meaningful and what variation it is better to ignore" (Cleveland 1993: 28). If the plot shows a straight or nearly straight line, the distribution of the variable can be said to be normal or nearly normal and the slope of the line estimates the standard deviation of the variable. However, judging the 'straightness' of the probability plot with any degree of confidence might be an issue. Ryan and Joiner (1976) suggest an objective measure such as calculating its correlation coefficient. Cleveland (1993) suggests visually assessing the deviation of the curve from a 45° or solid line, as shown in the plots in Figure 2.14 calculated using S-Plus (Insightful, Seattle).

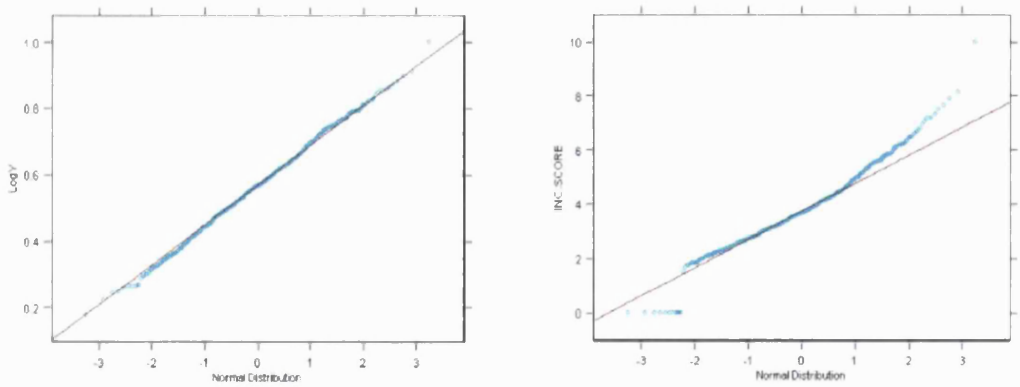

Figure 2.14 Normal Q-Q Plots: Left) Log of *Inc_Score*. Right) *Inc_Score*.

---

[8] Normal distributions are sometimes desirable due to the characteristics of this class of statistical distributions. First, they have a single peak at the centre of the distribution which occurs at the mean. Second, the distribution is symmetrical about the mean. Third, the spread of the distribution is given by the standard deviation. In addition, *all* normal density curves satisfy the property that 65% of the observations fall within one standard deviation of the mean; 95% lie within two standard deviations of the man and 99.7% of all observations fall within three standard deviations of the mean. This empirical rule applies only to normal distributions.

[9] A percentile is any of the 99 numbered points that divide an ordered set of scores into 100 parts each of which contains one-hundredth of the total. For instance, if the percentile value can have values [0,1], then the 0.5 percentile of a range is its mean, the 0 percentile of a range is its minimum value and the 1 percentile is the maximum value.

Parallel coordinate plots (PCP) are another way of "visualizing n-dimensional geometry and multivariate problems" (Inselberg 1999: 53). They map "n-dimensional space onto the two display dimensions by using n equidistant axes parallel to one of the display axes" (Hinnenburg *et al.* 1999: 23). Each axis corresponds to a variable in the dataset. Values for each variable are sorted from minimum to maximum, scaled in some way (linearly or by any other transformation) and finally plotted onto an axis, usually from bottom to top of the display. Each one of the values plotted on an axis is joined by a line to values represented on the other axes, which correspond to attributes in the same dataset record. In this way, "a point in Euclidian N-space, R^N, is represented by a polygonal line" (Inselberg 1999: 53). Figure 2.15 shows an example from GeoVista Studio (Geovista Center, PennState University; Table 2.1) where colour coding is used to group ranges of values interactively selected by the user in one axis in order to rapidly see their relation to other variables. Hence, Figure 2.15 illustrates how groups of objects selected in the PCP form profiles or depict relations between several variables in the dataset. Variations of the PCP have been proposed for instance using radial axes as opposed to parallel axes (Hoffman *et al.* 1997). These provide graphic displays similar to the star glyphs mentioned above. PCPs have also been adapted for hierarchical data by Fua *et al.* (1999).

Figure 2.15 PCP and use of colour for brushing ranges of data

2.3.3.3 Conditional Plots

A conditional plot (Figure 2.16) is a rectangular array of displays that show the bivariate relationship between two variables conditioned upon the values of one or more other variables (called the conditioning variable). The conditioning variable may be either a variable that takes discrete values or a continuous variable that is divided into a limited number of subsets. Following Cleveland (1993), the x axis of each display corresponds to the panel variable and the y axis to the dependent variable. These plots follow a principle known as 'small multiples' (Tufte 1983, 1990) where all diagrams are the same in terms of scale and visual cues except for the data itself, so that the user can scan the panels and watch "how the data 'move', relative to the fixed references" (Jacoby 1998: 54) of the conditional variable. The strength of the conditional plots is their ability to show interaction effects between variables at certain values of another variable. Few displays can give an indication of this type of relation between variables as "some methods are too heavily based on subjective decisions on the part of the analyst (e.g., iconic displays), other methods are problematic for graphical perception (three-dimensional plots), and still other methods break the multivariate data down into a series of strictly bivariate displays (scatterplot matrices)" (Jacoby 1998: 53).

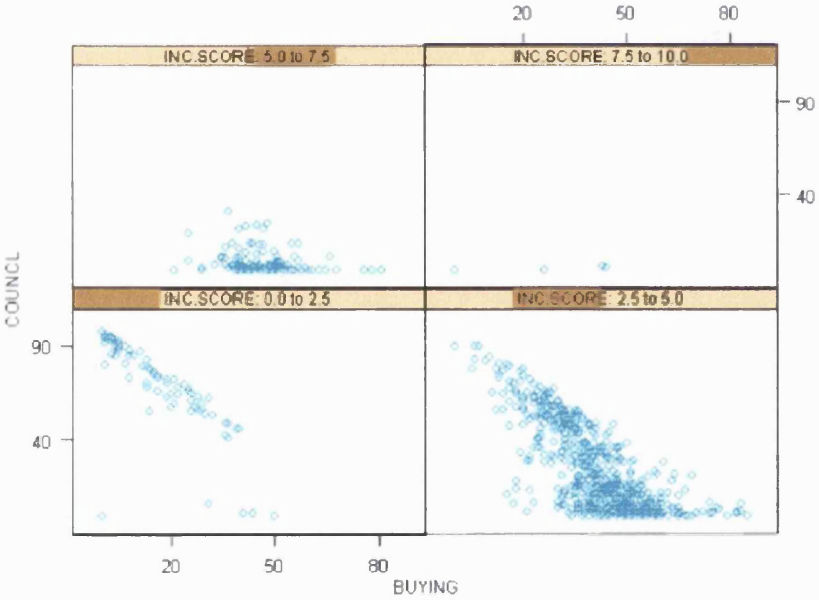

Figure 2.16 Trellis Plot of *Council* vs *Buying* conditioned upon values of *Inc_Score* divided into bins of equal range

Figure 2.16 shows a Trellis Plot, a form of conditional plot that has the information of the conditioning variable in a bar on top of each display in the array. Note that the displays are organised according to the ranges of the conditioning variable so that the lowest occupies the lowest left hand side corner, followed on the same row by contiguous ranges. Rows towards

the top correspond to higher ranges organised in the same fashion. The Trellis Plot in Figure 2.16 shows how the *Council* and *Buying* variables are negatively related for lower to middle ranges of *Inc_Score*. For middle to higher ranges there is no distinctive relation that can be discerned.

Conditional plots can be extended to incorporate two conditioning variables, one on each axis of the plot. Thus, they have been used by Brunsdon (2000) to investigate geographical data by making the two conditioning variables to be Latitude and Longitude. Thus, *Eastings* and *Northings* are used to divide the observations between the plots. Figure 2.17 shows an example where the bars on the top of the plots show the ranges for the first conditioning variable (*Eastings*) and the bars on the side indicate the ranges for the second conditioning variable (*Northings*). The graphical displays show the relation between *Inc_Score* and *Qualml* for different combinations of *Eastings* and *Northings* ranges. Similarly to the case with one conditioning variable, the first row of plots corresponds to observations in the lowest *Eastings* and *Northings*. This modification of the conditional plot is known as a conditional map (or comap) and it can include any "contextual information, such as place names, or area boundaries" (Brunsdon 2000: 99) in all the panels to respect the 'small multiples' principle. A feature that becomes apparent is the more prominent positive relation between *Inc_Score* and *Qualml* in the north-western and north-central areas of the study region.

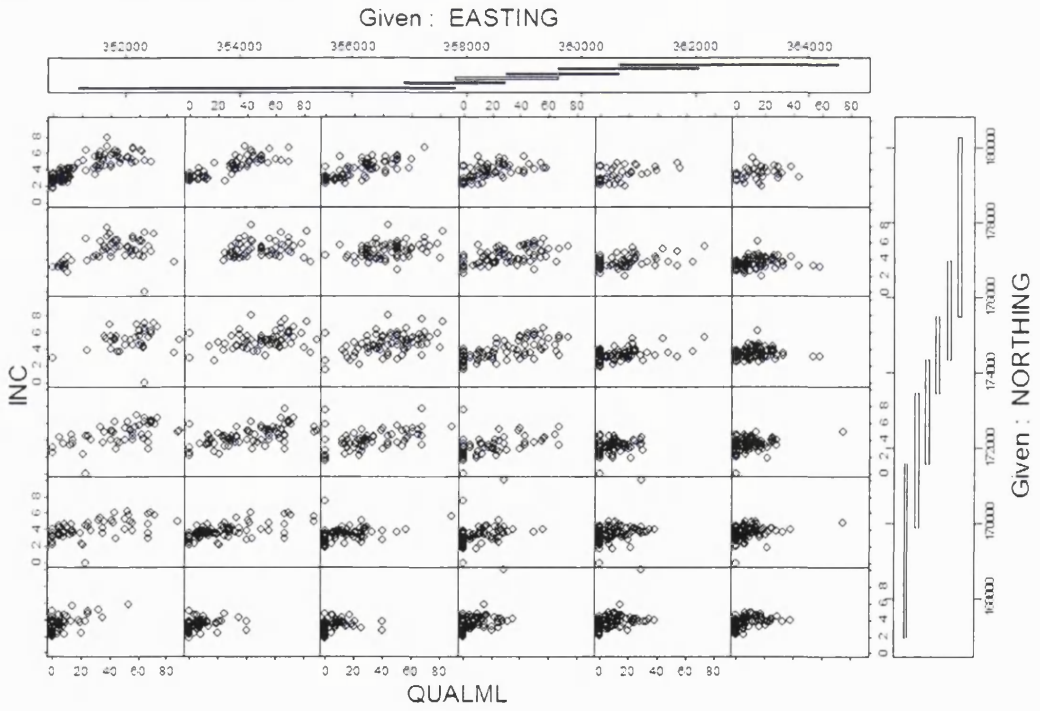

Figure 2.17 Comap with *Eastings* and *Northings* as the two conditional variables

Brunsdon (2001) also discusses comaps were two attributes of the dataset are used as the conditioning variables and the *Eastings* and *Northings* are ploted on the displays (Figure 2.18). This in effect creates point maps of the contidioning attributes. Brunsdon (2001) argues that they may be useful for showing spatial locations in more detail and identifying small scale patterns.

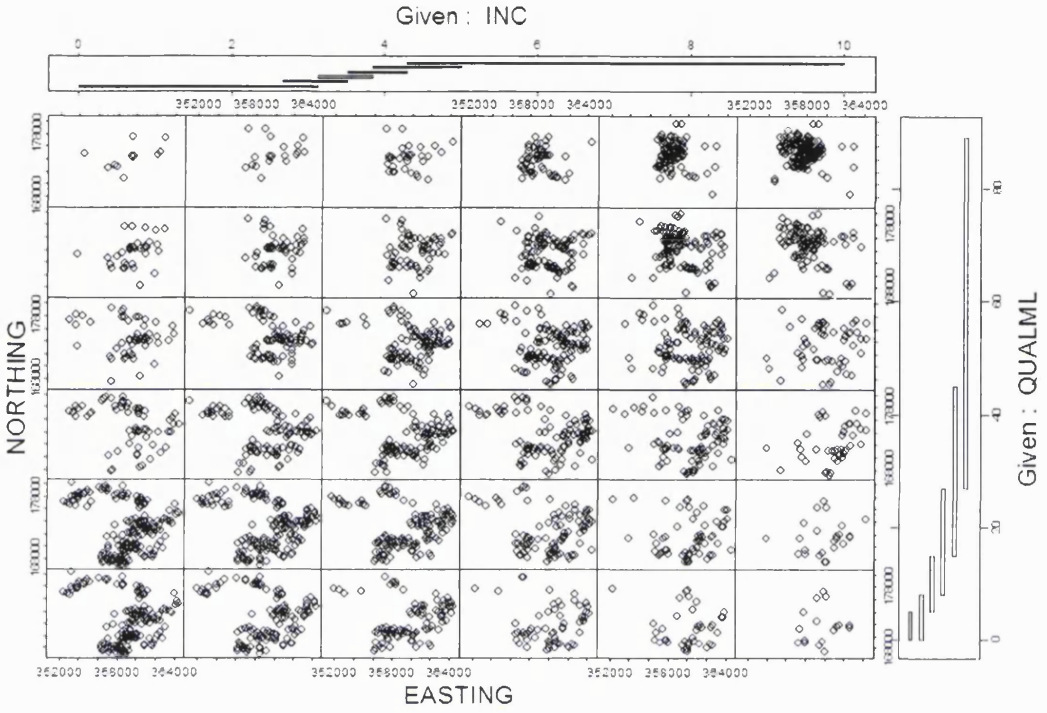

Figure 2.18 Comap with *Inc_Score* and *Qualml* as the two conditional variables

2.4 Cartographic Visualization and Geovisualization

Representing physical objects has long been an important issue in the discipline of cartography with its concern to "make aspects of the world visible" (MacEachren and Kraak 97: 335). However, after World War II, the emphasis of cartography "shifted from production efficiency and graphic design" (MacEachren 1995: 2) to creating maps with a clear function. That function was to convey a message, which was often defined or articulated by the cartographer who selected the content that met the map's purpose (Robinson 1952). The cartographer would therefore aim to encode the world on the map using abstraction methods such as projection, generalization, classification or symbolisation. Using his expertise and ability he would then output a map which was suitable for some pre-defined use, while still delivering a product that was visually pleasing and informative to average users (Unwin 1998).

This paradigm, dominant during the 70s and 80s, viewed cartography as a process of communicating spatial information. As such it was conceivable that the map, as the vehicle of the communication, could be improved so as to represent a geographical reality more objectively and facilitate a more accurate interpretation of the map (Kolácny 1969). Although this broadened the view of cartography as mapmaking, it has been heavily criticised by those who do not believe in the contention that there is such thing as an objective representation of reality, as well as by those who argue that the communication paradigm ignores many other ways in which people use maps (MacEachren 1995).

This last point has been an active topic of recent research as cartography has become concerned with how maps are used and with studying the human perceptual and cognitive processes that affect how maps are interpreted. This has broadened the function of maps from transmitting tailored meanings to a specific audience, to maps with no predefined purpose that can be used early in the research cycle to explore spatial phenomena (Tukey 1977). Cartography became involved in incorporating findings from the wider visualization community where advances in computing technology and multimedia were allowing the implementation of more dynamic environments and where user customization and interactivity had become possible. By the beginning of the 1990s, cartographers had applied and identified the impact of ViSC techniques to their discipline and understood that they entailed a change in the traditional use and function of maps (DiBiase 1990). Apart from communicating a pre-defined message in an "unambiguous graphic form" (DiBiase *et al.* 1992), maps could also aid individuals in the initial and private stages of research by enabling visual thinking where information is explored interactively in order to form hypotheses about the data under study (MacEachren 1994; MacEachren and Kraak 1997).

The idea of maps with no predefined meaning for aiding information processing has prompted a change in the research paradigm of cartography which has incorporated new definitions of map use:

> *"All map use involves both visualization (defined loosely as the prompting of visual thinking and knowledge construction) and communication (defined loosely as the transfer of information), but map use can differ considerably in which of these activities is emphasized" (MacEachren and Kraak 1997: 337).*

If used privately, maps may be highly interactive, transient and concerned with exploring unknown information. Products of this individual process may be created or intended for specialised users who "need not conform to fixed standards of symbolism, geometry or even scale" (Wood 1994: 16). On the contrary, a map for communication does not afford direct manipulation or interaction as it is typically concerned with transmitting a clear message to

an audience. Thus, this approach has been referred to as (Cartography)[3] as it addresses the type of audience, the degree of interactivity with the map and its purpose given the particular audience to which it is intended (MacEachren 1994: 5).

By allowing fast processing and display of larger volumes of information, technical advances in computing hardware and software have also contributed to the generation of new and innovative map uses. For instance, by removing or lessening the limitations of data collection and processing researchers have been able to become more concerned with issues of representing data quality and error propagation (Buttenfield and Beard 1994; van der Wel *et al.* 1994). Uncertainty may be introduced by a natural variability in the phenomena being measured; by an inability to measure phenomena precisely; or when data that is too coarse or generalised for a particular application is used in spatial analysis (Ehlschlaeger *et al.* 1997). This can be a common problem with the analysis of geospatial data in computer environments because an imprecise reality is given a precise digital representation. Visualising uncertainty can make users aware of its implications in the outcome of spatial analysis as recent research has shown (MacEachren 1992; Evans 1997; Kyriakidis 2001). Therefore, embedding error (deviations from a 'true' value) or accuracy (the discrepancy from an assumed value) diagnostics in a display can be of crucial importance. Innovative representations have been suggested for this purpose. Animation, for instance, has been suggested as a means to communicate the locational uncertainty of attributes (Fisher 1994), as well as to generate information about how data uncertainty affects a particular application (Ehlschlaeger *et al.* 1997). Moreover, the use of other sensory information such as sound have been used to highlight the location of error, while tone or rhythm have been used to provide further details about its relative proportion (Fisher 1994; Krygier 1994).

Estimulating the research concerned with innovative map uses addressed to different audiences and for different purposes was one of the driving forces behind the creation of the International Cartographic Association (ICA) Commission on Visualization in 1995. The ICA aimed at facilitating "connections among cartographic researchers around the world and researchers in other disciplines working on various aspects of scientific visualization" (MacEachren and Kraak 97: 336). Its main purpose was therefore to develop a research agenda that incorporated new advances and trends from the fields leading the development of visual data analysis techniques. Despite these efforts, graphic design in cartography remained heavily influenced by the long tradition of paper map representation.

Post 1999, the (current) ICA Commission on Visualization and Virtual Environments aimed at building upon the efforts of its predecessor by defining a 'new' field, geovisualization, that included a very wide variety of related disciplines and techniques:

> *"Geovisualization integrates approaches from visualization in scientific computing (ViSC), cartography, image analysis, information visualization, exploratory data analysis (EDA), and geographic information systems (GISystems) to provide theory, methods, and tools for visual exploration, analysis, synthesis, and presentation of geospatial data" (MacEachren and Kraak 2001: 1).*

Geovisualization, therefore, typically comprises at least two key components: flexible computer environments and a user as part of the exploration process. By flexible environments, we mean systems that allow the user to represent and interactively manipulate data in multiple graphical displays, which enable the investigation of 'what if' scenarios or questions that prompt the discovery of relations or patterns that are useful (Beshers and Feiner 1993). Software and hardware advances have made computer processing faster and new scripting and prototyping tools allow developers to deliver interactive systems (Dykes 1999) that can increasingly meet this requirement. However, the size and complexity of (spatial) datasets can have an impact on their flexibility and speed. Nevertheless, geovisualization tools must still permit users real-time direct manipulation and interaction with the data on-screen in order to facilitate hypothesis formulation. Some examples of systems that provide a number of the visualization techniques aforementioned for investigating spatial information include: the Cartographic Data Visualizaer or cdv (Dykes 1997), Descartes (Andrienko and Andrienko 1999) and GeoVista Studio (Gahegan *et al.* 2001), to name just a few recent academic efforts (see Table 2.1 for more).

The second critical component of visualization, the engagement of human beings in the exploration process, implies taking advantage of their expertise for identifying relevant patterns in the dataset. Not only can visualization methods make "good use of the human eye's unparalleled ability to recognise inherent structure and relationships" (Gahegan 1996: 1), they can also build upon the user's specialist knowledge to gather information on the data structure. Consequently, visualization "relies on creating an impression of data properties within the mind of the observer" (Gahegan 1999: 289). Hence, results obtained from a visual exploration process may very well depend on the person interacting with and interpreting the data. This cognitive nature of visualization makes it difficult to compare it against other analytical methods and its tools difficult to assess. Furthermore, the wealth of theory and methods that may be embedded in geovisualization tools might be of modest benefit if they are not developed to support the relevant information processing of their human operators or if they are difficult to use and understand. Therefore, apart from the visualization fields that contribute to geovisualization, human computer interaction and cognitive psychology

have come to play a central role in geovisualization. This is reflected in the topic of much current research which is described below.

2.4.1 Issues of Cognition in Visual Representations

The use of animation in cartography provides an interesting example of how visualization techniques have been adopted by a number of disciplines and have changed the way in which we represent and investigate data. In addition, it raises questions regarding the adequacy of visualization methods for supporting information processing. Map animation has been around as a common choice for depicting time series of georeferenced data at least since the 1960s (Campbell and Egbert 1990). However, new techniques and technology have permitted the development of innovative applications that allow greater user control of the animations. As early as the 1980s, systems such as HyperCard or MacChoro II where available and facilitated a common type of animation that used the film/video metaphor to present a succession of maps as a digital movie (Peterson 1999). Although they permitted some user interaction, there was little room for altering the order and speed of frames.

After the 1990s, new media and improved distribution formats have made cartographic animation commonplace. For instance, the Internet is a relatively new medium for the dissemination of geospatial information to a much wider and varied audience that includes many non-cartographers. Formats such as animated GIFs (created by assembling individual JPEG or GIF files), MPEG or QuickTime also permit the easy creation of movies or sequences of images, although they are very much in the spirit of the earlier systems that only allow minimal user interaction with the graphical information. More dramatically though, purpose designed programming languages have enabled the provision of applications served from the web (or otherwise) that use innovative metaphors to provide users with data and information (Cartwright 1999). "The new technologies open to the wide Internet audience give access to spatial information in the form of intelligent objects equipped with properties and behaviour" (Kahkonen *et al.* 1999: 430). Java, for instance, is an object-oriented technology developed by JavaSoft as a networked environment and programming language for developing small applications (applets) that can be downloaded into any browser. This technology has permitted the development of applications that are portable across multiple platforms, are highly adaptable to changing environments and do not need special software to be viewed or used. Moreover, application of similar technologies has enabled cartographic animations that are more than a sequence of maps arranged to depict change over time. Animation can instead prompt from user actions that trigger changes on data presented on-screen. 'Mouse-over' interaction, for example, has been used to alternate between maps

showing different attributes or time periods and reveal spatial patterns in the data.[10] Interactive devices have also been implemented which can be dragged or clicked on by the user for instance to control the display of attribute values through time (MacEachren *et al.* 1998; Andrienko *et al.* 2001). By providing maps that change in response to actions on the part of the users or changes in the data, maps and their animation can be used as tools that support spatial data exploration.

Nevertheless, the question still remains whether animation and dynamic presentation facilitates information processing. Larkin and Simon (1987) suggest that diagrams may reduce cognitive processing because all information is made explicit and perceptually available which reduces the effort for making inferences. In other words, graphic displays remove some of the information processing load from conscious thought to pre-conscious visual processing (Friedhoff and Benzon 1989), which facilitates perceptual inferences. Extrapolating this result, it is commonly stated that animation may be expected to enhance understanding when illustrating dynamic processes because it explicitly depicts motion. However, there is very limited formal evidence of this to date and user studies on the effectiveness of animated versus static displays have produced contending results (Slocum *et al.* 2001).

Narayanan and Hegarty (2000) developed a cognitive model of how people understand multimedia presentations including static diagrams, animations and text in the domains of mechanics and algorithms. They found that mixed-mode presentations designed according to a cognitive model outperformed traditional presentations. This effectiveness was regarded to stem from structure and content which is independent of the medium of presentation. In addition, computer-based presentations which afforded interactivity did not outperform static paper-based presentations. However, the effectiveness of animation via direct manipulation and interaction may be highly dependent on the purpose of using the technique. Narayanan and Hegarty (2000) suggest that it is possible that in some situations a dynamic presentation might be more compatible with a cognitive model than a static presentation. Engaging animations and interactivity may have an impact on motivation. For instance, Morse and Lewis (2000) found that increasingly complex data might be more amenable to visual treatment. Furthermore, that as tasks become more difficult, visual presentations become more valued by users. Nevertheless, Narayanan and Hegarty (2000) argue that static presentations may have other benefits such as promoting active learning and

[10] For a good example developed using JavaScript see

engaging cognitive processes of comprehension. Furthermore, they argue there is also a real risk of the 'hands-on, minds-off' problem when using animations "in which the sensory stimulation and interactivity of multimedia and its externalization of information distract from, and discourage, the cognitive processing that deep learning requires" (Narayanan and Hegarty 2000: 190). Morrison *et al.* (2000) have even suggested that animation is not effective for educational purposes. One possible explanation for this is that "although animation may increase explicit perceptually available information, it may not automatically improve understanding. Visual explicitness itself does not necessarily guarantee accurate perception of specific information, nor does perception of information guarantee comprehension" (Jones and Scaife 2000: 231).

On the other hand, Park (1998) maintains that animation allows attracting and directing subject's attention. Furthermore, that it is well suited for representing domain knowledge that involves movement and explaining complex phenomena such as structural and functional relations. Kaiser *et al.* (1992) also found that when one dimension of the information was being depicted, animation facilitated observation. However, in a geospatial domain, animated depictions require decisions to be made on the spatial resolution of the individual images, the order in which these are presented, as well as on the time intervals between them when the display is not fully controlled by the user (MacEachren *et al.* 1992; DiBiase *et al.* 1992; Ehlschlaeger *et al.* 1997). An animated sequence may be too rapid and complex to allow the processing of information. Not only does it require the processing of each image but also the comparison among images in the dynamic sequence. Thus, Andrienko *et al.* (2000) argue that animation facilitates information processing only if it is under complete user control. Moreover, Monmonier and Gluck (1994) found that users perceived maps of demographic time series as difficult to understand and frustrating to watch when the animations could not be interactively controlled. Nevertheless, "even when under the viewer's control, a rich and rapid sequence of graphics can overwhelm the eye-brain system" (Monmonier 1992: 23). Still, Kousoulakou and Kraak (1992) found that animated maps allowed a more rapid interpretation of spatiotemporal information than static maps.

These issues point to an even deeper debate. Although there is a large amount of literature on graph comprehension, it is unclear how this research can be generalised to the comprehension of complex graphs such as those utilised in geovisualization because little

---

<http://maps.unomaha.edu/AnimArt/ActiveLegend/JavaScript/java.html>.

theoretical or empirical work has been undertaken. Past research on graph comprehension has greatly focused on the encoding and perception of graphs and on the cognitive processes that make a graph easier or harder to interpret (Pinker 1990; Kosslyn 1989; Cleveland and McGill 1986). Most studies have focused on data spaces with only a limited number of variables (typically two or three) in order to retain experimental control (Trafton *et al.* 2000). Furthermore, in these studies, experiment participants were typically undergraduates who were required to extract relatively low-level information from a graph that had been prepared in advance. However, users of complex visualizations generally have relevant prior knowledge, deal with larger amounts of variables over a range of scales and must decide what information to visualize for making some decision or inference (Freedman and Shah 2002). Furthermore, in a geovisualization context, users aim at obtaining insights or deriving hypothesis from manipulating the displays.

One possible way forward is to assess the effectiveness of visualizations for defined types of data, users and sets of tasks. Such was the purpose of the experiments that are described in detail in the following chapters. These use vector data, georeferenced data users, and consider a set of geovisualization tasks. In other words, the evaluations investigate whether the computer-based tools tested allow the identification of patterns in the data (Becker and Cleveland 1991; MacEachren and Ganter 1990), help users to define questions, develop hypothesis and find answers (MacEachren *et al.* 1992; Gahegan 2001). Andrienko *et al.* (2002) point out that, in geovisualisation, the researcher often wishes to evaluate a concept or technique rather than a system but that the idea needs to be implemented before any evaluation is possible. This poses a limitation for evaluating the effectiveness of visualizations for information processing and brings us to consider whether the implementation of a visualization technique in a piece of software is easy to use by its potential users.

Before turning to issues of usability in geovisualization environments, it is important to note that 3D displays, or more precisely, displays which use depth to represent a third (or higher) data dimension, were not considered in the experiments that form the basis to the original research reported here. The main reason for this is that, given the data used in the evaluations and the type of tasks considered, there was scarce documented evidence that using 3D displays would facilitate improved user information processing. On the contrary, a number of studies suggest the opposite. Westerman and Cribbin (2000), for instance, were interested in the differences between 2D and 3D visualizations in terms of their semantic richness and cognitive demands. They conducted a study for information search and found

that the additional semantic information provided by the 3D solution was outweighed by the increased cognitive demands that it generated for its processing. Similarly, Sutcliffe and Patel (1996) showed that 3D images do not appear to provide any significant advantage over 2D displays for information browsing. This is not to say that 3D displays may not be beneficial in certain scenarios, as those discussed in §2.3.1, but rather suggest that the usefulness of graphics is likely to be a function of the task that the presentation aims to support (Casner 1991). For the tasks, data type and user population considered in this research, we could not find compelling evidence that 3D displays provide any additional benefits in terms of facilitating cognition and perception. This however may constitute an interesting topic and future extension of this research.

2.4.2 HCI in Geovisualization and GISc

Large volumes of geospatial data are currently available but may not be used effectively since little has been done to make geographic information handling technologies accessible to non-experts. This problem can be even more dramatic for those users outside the geoinformation field that are unfamiliar with the terminology and concepts embedded in software tools (Traynor and Williams 1995). Hence, user-centred design of geovisualization tools, as well as the assessment of both existing and new methods, present current major challenges. Human computer interaction (HCI) is "a discipline concerned with the design, evaluation and implementation of interactive computing systems for human use and with the study of major phenomena surrounding them" (Hewett *et al.* 2002). HCI is concerned with enhancing the quality of interaction between humans and computer systems within the physical, organisational and social aspects of the user's environment to produce systems that are usable, safe and functional (Preece *et al.* 1994). Research on these issues is based on the assumption that the needs, capabilities and preferences for the way users perform an activity within an environment should influence the design and implementation of a system in order for it to match user requirements. Knowledge about the users and the work they need to accomplish, as well as about the technology needed to support them makes HCI a multidisciplinary field of research.[11]

Interest in HCI is not a new issue in geographical information systems (GIS) and geographical information science (GISc), although it has only recently emerged in geovisualization (Slocum *et al.* 2001). During the late 1980s, cognitive aspects of HCI for geographical information systems (GIS) were discussed at workshops of larger conferences or as sections of books where HCI issues in GIS were not their primary focus. The 1990s,

however, saw a strong international research interest in the topic. Evidence of this can be found in four workshops that were held between 1990 and 1994 in the US and Europe, which explicitly discussed HCI aspects in GIS,[12] as well as in at least two books—Medyckyj-Scott and Hearnshaw (1993) and Nyerges *et al.* (1995)—published solely on the topic.

The interest in cognitive aspects of HCI for GIS can be explained at least by two main trends. On the one hand, the relatively recent increased availability and power of affordable personal computers broadened their user community, which was no longer restricted to "a technical and mathematical priesthood" (Baecker and Buxton 1987: 51). This new user community both desired and demanded interfaces that required technical expertise. On the other hand, GIS had been developing since, at least, the late 1960s and had reached a state where functionality such as map production, display and spatial data analysis was commonly provided. Concerns in the GIS community about increasing processing speed and storage requirements were the main topics on the early agenda, but it expanded at the beginning of the 1990s to include how GIS were used and how they could accommodate the needs of a broader range of users. This had not hitherto been a matter of substantial interest, as the principle focus was upon facilitating advances in GIS functionality to satisfy expert user needs, and the prime movers were computer system designers, developers and the systems' manufacturers (Medyckyj-Scott and Hearnshaw 1993). It was realised, however, that GIS were "more likely to fail on human and organizational grounds... than on technical ones" (Medyckyj-Scott 1992: 106) and that the deficiencies of the systems in terms of human factors could compromise their future success.

GIS evolved out of a number of fields including geography, cartography and database management. Hence it is argued that GIS requires its users to have considerable knowledge of these fields to operate the system (Traynor and Williams 1995). Furthermore, GIS entail the computer literacy of its users and a considerable time investment in order to become familiar with "an interface that reflects the system's architecture" (Traynor and Williams 1997: 288). This is made evident by the limited use of sophisticated functionality of GIS (Craig *et al.* 2002). Nevertheless, GIS users vary in expertise and may use the technology in one of a large number of application areas, and may also demand different functionality and

---

[11] For a brief history of HCI see Haklay and Tobón (2003).

[12] 'Cognitive and Linguistic Aspects of Geographic Space', July 1990, Spain, NATO Advanced Study Institute. 'User Interfaces for GIS', June 1991, USA, NCGIA Initiative 13. 'Task Analysis in Human-Computer Interaction', June 1992, Austria, 11th interdisciplinary workshop on Informatics and Psychology. 'Cognitive Aspects of Human-Computer Interaction for Geographic Information Systems', March 1994, Spain, NATO Advanced Research Workshop.

analytical power. Accommodating such a wide spectrum of needs is a challenge in its own right that must take into account a number of factors—such as the components and requirements of users' work, their capabilities and limitations, the types of support the system can provide, and where it can be provided most effectively.

2.4.2.1 Usability and Geovisualization

Apart from understanding how to improve users' work processes, HCI is also concerned with understanding how people use computer systems in order to develop or improve the system's or interface design. Central to the concept of HCI is usability, which deals with the description of a system's fitness for purpose and its ability to facilitate potential user tasks. Usability studies aim to enhance user performance by helping to design, develop and deploy systems that meet user requirements (Landauer 1995) and allow work to be carried out "safely, effectively and enjoyably" (Preece *et al.* 1994). Usability deals with these aspects and it applies to all aspects of a system's user interface, defined as the medium through which a user interacts and communicates with the computer (Nielsen 1993).

Usability evaluation techniques comprise a good number of methods that are suitable for different types of systems and users and which are also appropriate for different stages in system design (Fuhrmann *et al.* 2003). These methods are aimed at obtaining either quantitative evidence about user or system performance, or qualitative information about user perception and opinion. They are also often combined in order to obtain richer insights into the issues raised by the experiments. In practice, evaluating a system to meet user requirements in a particular application domain may demand the tailoring of available techniques (Fuhrmann and Pike 2003).

Usability testing is one technique, which requires test participants to complete tasks that potential users of the system would wish to accomplish and observing and recording how they work and perform (Dumas and Redish 1999). Performance measurements recorded during these tests may include the time it takes subjects to complete a task, as well as measures of success and error rates. Analysis of these performance measures—either statistical or otherwise—is frequently complemented with information obtained from post-test interviews or questionnaires. The main reason for this is best illustrated with an example. 'Throughput' or 'ease of use' is one metric used in HCI for system evaluation. It can be recorded by measuring the length of time required to perform tasks and documenting the errors made (Preece *et al.* 1994). However, in the context of geovisualization, differences in these times might not only be indicative of ease of use. Shorter times could arise as users

gain confidence and learn how to use the system, or they might very well be suggestive of users operating the software without consideration of the results they obtain. Longer times could signal difficulties with understanding the graphics or using the tool but they may also arise because users feel motivated to investigate some information in detail. Moreover, in a system designed for exploration, there are no defined sequences of steps to follow in order to reach an answer. Thus, part of the user's operation of the software would consist of experimentation with the data and the graphical displays, which can hardly qualify as user errors or actions that do not accomplish desired goals (Nielsen 1993). Hence, performance measurements may depend on a number of interrelated factors, such as the ease of use of the software, users' experiences, the complexity of the tasks performed and the functionality provided to support them. A combination of evaluation methods can help elucidate information about the various aspects of a system's usability, as well as the reasons for diverse user behaviour and experience.

2.5 Towards an Understanding of User Tasks

Usability evaluation methods however require the clear definition of user tasks and goals. If we recall the main purpose of geovisualization—gaining understanding of a problem or dataset by re-expressing data in graphical displays—defining tasks in this context may not be a straightforward undertaking because of their exploratory nature. In addition, geovisualization techniques can be useful and pertinent for a wide range of applications with an equally large variety of potential users and data types. At present, this implies that evaluations must be custom tailored to each system and the domain dependent tasks it is designed to support. This makes it difficult to control for the effects of the geovisualization tool and the particular concept under study.

It has been previously discussed that the other main component of geovisualization is the user whose background knowledge interacts with the data presented on-screen as he interprets their meaning. This may be seen as a limitation in the sense that the results of a geovisualization session can depend on the individual's interests and abilities. It is commonly accepted that maps (or more generally visual displays) are capable of many interpretations (Unwin 1996) and thus even marginal changes to a representation can potentially prompt different meanings. In addition, the information extracted from the representations may itself be a function of the understanding that the users have regarding the data or the phenomena being mapped. Nevertheless, it is our perspective that this is one of the strengths of geovisualization as a method of exploring data in the initial stages of analysis as it may prompt insights and hypothesis that were not anticipated and which can be then

examined using confirmatory analysis techniques such as statistics. The value of geovisualization is that it may allow the generation of ideas regarding the plausible relations within a dataset without imposing structure to the data before they are better understood. In addition, it puts the user at the centre of the data analysis process and thus forces us to take into account the work process that must be supported by a computer-based tool in order to facilitate the unlocking of the advantages that geovisualization has to offer.

Although not all of these issues can be addressed by HCI or usability evaluation alone, the use of formal HCI models and techniques can lead us to an improved appreciation of user tasks, and can also facilitate the design of systems that are fit for purpose and which support user requirements. Combined with a definition of the tasks users might attempt with geovisualization tools, the analysis of user performance may be conducive to an understanding of cognitive processes. Therefore, the combination of HCI methods and visualization task taxonomies for evaluating geovisualizations "can make a significant contribution to increasing the sophistication and rationality of visualization design based on cognitive work requirements" (Wood and Broodlie 1994: 7).

The software systems described in Table 2.1 and in the following chapters, implement some of the graphical methods described earlier. Apart from supporting user interaction, the key characteristic of these environments is that they each provide a visual depiction of queries made about data and the continuous reformulation of goals, hypotheses or insights through the manipulation of various data views (Kaushik and Rundensteiner 2001). However, it still remains an open research question as to what type of functionality these systems should provide and the type of tasks they are each best suited to solve. These issues are explored in depth in the following chapters in environments designed for GIS users and for the exploration of area-based georeferenced information. It is in this context that the geovisualization tools and techniques presented are evaluated to assess whether they do facilitate the investigation of spatial data.

Table 2.1 Example Software for Geovisualization

Software	Description
DecisionSite Map IS DecisionSite Map Interaction Services	DecisionSite is a web-enabled software aimed at visually exploring and analysing point data that are non-spatial. It is a very flexible tool in terms of how different views can be modified by the user. It provides an environment for the rapid identification of trends, anomalies, outliers and patterns using 2D and 3D scatter plots, histograms, bar charts, pie charts, and parallel coordinate plots, trellis plots, heat maps, line charts. **Salient features** include: • Interactively filtering (excluding or including) data using a graphic device on the GUI that performs SQL queries in real-time (up to +/- 10,000 records). • Histograms can be binned as determined by the user using a wizard. • DecisionSite was extended to provide spatial context to data analysis using a link to ArcExplorer. The add-on product, Map Interaction Services (Map IS), enables very limited GIS-type functionality and display of spatial data and limited linking to the visualization tool. *Developed by Spotfire (http://www.spotfire.com/).*
DataDesk	Interactive software for EDA uses linking and various forms of brushing to connect information provided by various views. Although not spatially enabled, one of its **salient features** is supporting large amounts of data (the vendor argues that it can smoothly handle visualizations with more than a million records). DataDesk provides statistical functionality such as linear regression, PCA, cluster analysis, ANOVA, and different forms of correlation. Also, it provides a drag-and-drop interface for building customized templates where combinations of analysis and plots can be designed. Finally, it provides boxplots *a la* Tukey (1977) which can be linked to other visualizations. *Commercially available from http://www.datadesk.com/DataDesk/.*
cdv Cartographic Data Visualiser	CDV implements several re-expression and brushing techniques that are very effective for exploring relations among variables in an area-based dataset. It provides alternative views (dotplots, boxplots, circle maps, cartograms, scatterplots and parallel coordinate plots) of one, two or three variables simultaneously. **Salient features** include: transient symbolism, multiple views, thematic map colour coding that can be defined by the user, symbol scaling, an interface that is interactive and tailored to a specific set of users. *Developed by J. Dykes as part of his doctoral thesis (University of Leicester) on cartographic visualisation and exploratory spatial analysis. Available from http://www.geog.le.ac.uk /jad7/.*
Descartes	Descartes' main concern is providing traditional cartographic presentation methods, which are interactive and dynamic for the exploration of areal data. It uses multiple representations of the information, each being best suited for certain types of analysis, to explore the data. Dynamic choropleth maps, pie charts, bar diagrams, dotplots, boxplots, and scatterplots are available to the user. **Salient features** include: • Use of sliders to represent the range of a numeric variable displayed on the map. They can be moved to interactively change map classifications and colour coding. • Use of a colour matrix created from dividing the range of two variables into intervals. "$M \times N$ classes, [are defined] where M and N are the numbers of intervals for the first and for the second variable, respectively" (Adrienko and Adrienko 1999). A colour scale is then selected for each variable. This generates a colour matrix on top of which a scatterplot is laid. Colours in the matrix thus create a visual classification of the data by value ranges. *Developed at the German National Research Centre for Information Technology (GMD) by G. Adrienko and N. Adrienko. Functionality is described in Adrienko and Adrienko (1999) and a web supplement of this paper with demos in http://allanon.gmd.de/and/java/iris.*

CommonGIS Common Access to Geographically Referenced Data	CommonGIS is a WWW-based GIS with specific functions for the automatic generation of thematic maps. Its purpose is the dissemination of georeferenced data to a broad cross-section of the public (Andrienko and Andrienko 1999a). It is an open, object-oriented, distributed system. **Salient features** include the automated creation of maps which are customised for different user types. *Developed by a consortium, Common GIS (http://commongis.jrc.it/index.html) brings together three previous developments (Descartes, Lava/Magma by PGS (http://www.pgs.nl) and Vizard by IGD (http://www.igd.fhg.de/igd-a5/) as the basis for development and integration.*
GeoVista Studio	A Java-based experimental programming environment designed to allow the integration of components (Java beans) for the analysis of geographic data using visualization techniques. A **Salient feature** of this alternative is the possibility of potentially linking any component that has been built using JavaBeans technology. The parallel coordinate plot shows a good example of the benefits of interaction and flexibility in the analysis of the plot. Variables can be assigned to each axis as well as the variable used to classify data into ranges and colour code the lines in the plot accordingly. *Developed at the GeoVista Centre, Penn State University (http://www.geovistastudio.psu.edu/).*
Density Dial	Its most **salient feature** is providing users with a tool to interactively vary the statistical value at which intervals between two classes occur (Ferreira and Wiggins 1990).
RADVIZ	Its most **salient feature** is the use of a projection to map multidimensional points into 2D space. Points are equally spaced around the circumference of a circle or along the vertices of a regular polygon. Any point inside the polygon is a projection or representation in 2D space of a point in *m*-D space. This method is particularly useful for the identification of outliers, extreme values and correlation among variables (Ankerst and Keim 1996).
Regard Formerly Spider	Allows the exploration of the distributions of data for any subset of locations selected from a map. Uses brushing to link various views of the data including an areal map, bar chart, histogram, trellis plot, scatterplot, and attribute table. *Developed by Haslett, J., Wills, G., and Unwin, A.*
MANET Missing Are Not Equally Treated	Designed for area-based spatial data to implement interactive graphics tools for data sets with missing values. **Salient features** include allowing users to experiment with the sensitivity of a histogram to the width of its bins (Unwin *et al.* 1996). By dragging their edges, the data are reclassified into new bins and the user can see the effect on the resulting distribution. *Developed by Unwin, A and Hofmann, H. Available from http://www1.math.uni-augsburg.de/Manet/.*
SpaceStat	Developed for the analysis of spatial lattice data through the use of spatial statistics and spatial econometrics. Non-linear optimisation routines are implemented in the system to obtain maximum likelihood estimates for spatial regression models. Its most **salient feature** is facilitating testing for the presence of spatial autocorrelation and the estimation of models that incorporate this form of dependence. A link to ArcView GIS from ESRI permits the visualization of results. *Developed by Anselin, L. Commercially available from http://www.spacestat.com/.*
INFO-MAP	This tool allows the creation of proportional symbols or choropleth maps from data values; interpolation of point data as contour or kernel maps; scatterplots, frequency distributions, variograms, K-functions and correlograms can be created. Areas, perimeters and adjacency of polygons can be calculated as well as nearest neighbour distances. Written to run on MS-DOS, it supports small databases. *Developed by Bailey and Gatrell (1995).*

3 Study 1: Usability evaluation of a geovisualization tool

In a preliminary version of this thesis, I tried to change the look of footnotes. A bounding box shaded in light grey had magically appeared around the little line that separates the footnotes from the text. After reading the help files and trying all the menu options, I gave up. I had to copy the text, paste it with no format in a new document, and start from scratch again. 'People consider a product 'easy to learn and use' in terms of the time it takes to do what they want, the number of steps they go through, and the success they have in predicting the right action to take" (Dumas and Redish 1999: 5). I would not think that this particular word processor is very popular among PhD students writing dissertations.

3.1 Linking plots, linking software

The previous chapter discussed a number of visualization methods developed in various fields and their application to spatial data. A broad range of these techniques is not generally provided in any single software system. On the contrary, there is great variability between the specific features and flexibility to accommodate user tasks that available software tools offer. However, it might not be efficient for a system to offer a comprehensive range of analysis techniques. Not only do systems become larger and heavier to run as they demand more computer resources, but adding functionality may not solve the problem of supporting user requirements. In addition, providing "more choice can equate with more difficulty of use since the user has to remember a greater number of options, commands and procedures" (Medyckyj-Scott 1994: 106). However, different types of data and tasks demand different analysis and visualization methods (Kleinmuntz and Schkade 1993; Casner 1991) that may not be supported by a single environment. In order to overcome this difficulty, linkages between packages are a common solution aiming at supplementing systems' functionality.

GIS applications are no exception to this practice. Goodchild (1987), for instance, suggested a taxonomy of six basic classes of GIS-based spatial analysis operations that he viewed as an 'ideal' set of methods from the spatial analyst's perspective, which were not provided by any single package available in the 1980s. GIS may still not include all the types of analysis that some critics consider appropriate (Haining 1994; Bailey 1994). Therefore, two different approaches known as loose and close coupling (Goodchild *et al.* 1992) have been implemented for linking GIS to other packages and solving this deficiency. In the former case, objects, data and commands are written to intermediate files and passed between systems. In closely coupled implementations, commands in one system can be called by another. More integrated approaches, defined by Anselin and Getis (1992) as encompassing or modular links, imply the development of routines in a system's macro or scripting language. Examples of close, loose and encompassed implementations with ESRI's

(Redlands, CA) products for exploratory data analysis (ESDA) are described in Anselin (1998) and some examples are given in Table 3.1. The first column contains the names of the coupled systems; the second a short description of the link and the functionality it provides; the last column describes the degree of integration between the systems.

Table 3.1 GIS and ESDA implementations with ESRI's products

Software and Link	Characteristics	Coupling
ARC/INFO	Kriging and geostatistics are provided as part of the in-built functionality.	Encompassed or modular
ARC/INFO and SAGE	Running in UNIX, linkage is achieved through a client-server architecture where the GIS is the server and communicates to SAGE by Remote Procedural Calls (RPC).	Loose
ArcView	Statistical functionality (written in Avenue) has been implemented as an extension. Later versions (ArcGIS) include geostatistics as part of the toolbox.	Encompassed or modular
ArcView and S-Plus	Running on Windows, the link uses S-Plus' Application Programming Interface (API) to implement a client-server architecture where S+ acts as the server and ArcView as the client. Functionality includes the estimation of local Moran's I and Geary's c and spatial weighted regression which is added as an extension in ArcView.	Loose
ArcView and SpaceStat	Uses Dynamic Link Libraries (DLL) written in C++ which are called from within ArcView using Avenue scripts. Statistical graphs provided include histograms, box plots, scatterplots, Moran scatterplots. There may be dynamic links between them but they are linked to a single view in ArcView.	Combination of close and loose
ArcView and XGobi / XploRe	Running on UNIX, the link uses RPC in a client-server architecture whereby any system can be the client or the server but only one link can be opened at a time. Focus on geostatistics (variogram cloud plots, spatially lagged scatterplots) for point data only.	Close

At least three of the basic operations suggested by Goodchild (1987) were related to statistical and spatial statistical analysis. All of the examples in Table 3.1 illustrate efforts to augment the GIS spatial analysis 'toolbox' to include statistical and geostatistical analysis and graphical displays. However, the implementations described in Table 3.1 are typically "slow in execution, due to the deficiencies of the scripting language in terms of algorithms and data structures necessary to carry out statistical computations. Moreover, the size of datasets that can be analysed with these approaches is severely limited and precludes realistic applications" (Anselin 1998: 84). Some examples illustrate the complexity and limitations of this type of integration. The first is the link between ArcView and XGobi (a visualization package by

Buja *et al.* 1996) and XploRe (an interactive statistical environment by Hardle *et al.* 1995). For instance, to dynamically link point data in a view in ArcView and scatterplots in XGobi, (x, y) co-ordinates and attribute values of a point have to be passed from ArcView to XGobi. A proprietary data structure to the link is used for this purpose. If points are brushed, then XGobi passes their identifiers back to ArcView using the same proprietary structure. If the selection of points is done in ArcView, the selection is first passed to XGobi where the relevant points are highlighted and then back as described earlier. This form of communication between the systems severely restricts the amount of records that can be analysed.

The second example is the ArcView and S-Plus link. The user must interact with a menu in ArcView that offers only a few options and little or no control of the parameters in the estimation and definition of weights matrices. Although an experienced user of S-Plus can alter them from within the statistical package, the process is far from straightforward and prone to errors. In addition, "S+ objects are written to intermediate files before they are linked to a table in ArcView" (Anselin 1998: 88). Information passed in this form cannot be controlled or monitored and the link unreliable as a unique identifier in both systems is not included in the transfer. As a result, S-Plus is frequently unable to perform the requested operations and the user is left with no feedback on the source of error. Furthermore, dynamic linkage of the systems is not possible. In other words, brushing or the simultaneous highlighting between views in the two packages is not supported. Columns of results computed in S-Plus are appended to tables in ArcView which can then be mapped. Finally, the amount of data that can be analysed is highly constrained both by the process of writing to intermediate files (which makes the communication between systems very slow) and by the use of weights matrices (which become rather large even for relatively small spatial datasets). S-Plus is not designed to support spatial data and calculations become intractable after a few thousand points, a task magnitude which can nevertheless be small for a georeferenced database.

The SAGE (Spatial Analysis in a GIS Environment) to ARC/INFO link (Wise *et al.* 2000) provides a further example. This implementation links classical and spatial statistical displays to a map view. However, it uses graphical facilities of ARC/INFO that are "less powerful and flexible than in other similar systems [...] partly because of the nature of the cartographic tools in a standard GIS like ARC/INFO and partly because of the delay caused by the need to communicate between the client and the server" (Wise *et al.* 2000: 183).

In summary, depending on the type of coupling, disadvantages of this approach to linking software have included: i) limits to the size of the datasets that can be used; ii) inefficient algorithms being implemented because of the differences in data structures between systems; or iii) slow execution not allowing real-time dynamic interaction because of the form of the communication between systems. In addition, some of these links have been developed as academic research efforts and have not been updated to communicate to newer versions of the GIS to which they were originally linked. Others have been commercial efforts but have not been developed further or maintained or are not available for newer versions of the systems involved in the linkage. Despite the limitations, linking packages to existing GIS software may make sense if potential users already have spatial data in a GIS and need to undertake some analysis without having to export that data to and fro. It may also make sense to use a GIS for part of the analysis as such systems have a number of facilities including their visual display which may be of use in the analysis undertaken (Wise *et al.* 2000). In addition, new technologies are making the development of links between packages much easier to program, more reliable and in some ways faster than in the examples detailed in Table 3.1. Component object modelling (COM) is one of such technologies which is discussed in detail at the end of this chapter.

A second approach has been the development of systems supporting visual analysis of spatially referenced data such as cdv (Dykes 1997), Descartes (Andrienko and Andrienko 1999), CommonGIS (Andrienko and Andrienko 1999a) and GeoVISTA Studio (Takatsuka and Gahegan 2002), which were reviewed in the previous chapter. The first two systems are designed for exploring lattice data and are not intended to communicate or provide a visualization environment that can add—at least interactively—to GIS functionality. These tools, however, are either proof-of-concept systems that are not easily extensible or customizable, do not easily connect to existing databases, or tend to support the analysis of rather small datasets. CommonGIS was still under development and was not available at the time when the experiments reported here were designed and conducted. GeoVISTA Studio is an open software development environment that allows users to build applications for geocomputation and geographic visualization. This solution is also still under development and demands a great deal of customisation and programming skills if it is to be tailored to particular applications. Nevertheless, it is intended to provide an experimental environment to build up systems and "seek deeper understanding of the kinds of tools required for effective knowledge construction" (Takatsuka and Gahegan 2002).

This chapter considers another tool, DecisionSite Map Interaction Services (Map IS) by Spotfire, Inc. (Boston, USA), which is more in the spirit of the developments described in Table 3.1. DecisionSite is an interactive system for the visual and dynamic exploration of data designed to support decision-making. It provides a number of graphical displays which include scatterplots, parallel coordinate plots (PCP), histograms, line graphs, data tables and details-on-demand, among others. This product was enhanced by linkage to ArcExplorer—a stand-alone software application built using ESRI's MapObjects embeddable mapping and GIS components—which provides limited GIS functionality (simple user interface, basic tools and data management). Thus, using a 'loose' approach, the functionality of visualization and GIS software was combined as a means to develop an environment that supports spatial data users.

This chapter describes a study aimed at providing some evidence on the adequacy of tools built as a coupling of two systems for supporting geovisualization tasks. Although potential disadvantages of this approach have been discussed, new technologies offered the possibility of building better links between software products. The following sections discuss an evaluation where attention was given to participant behaviour while using the system. Furthermore, we discuss some usability problems with the system and its functionality which hindered users' visual exploration of spatial data. Therefore, concepts of usability and usability evaluation are briefly revised next to better explain the purpose and scope of the study.

3.2 Usability

It has been previously discussed (Chapter 2) that the field of human-computer interaction (HCI) is concerned with how to improve users' work processes. It also deals with understanding how people use computer systems in order to develop or improve their design. The aim is to meet user requirements so that they can carry out their tasks "safely, effectively and enjoyably" (Preece *et al.* 1993; ISO 9241). Usability is a sub-field of HCI that deals with these aspects and the issue of how to provide functionality that can be used effectively and with ease (Nielsen 1993).

Usability describes a system's fitness for use and its effectiveness for supporting the interaction between human operators and their computer system. Whatever the aim of the system, it should facilitate users' tasks or enhance their performance by providing appropriate analysis tools. Therefore, it can be specified in terms of how well intended users can master and perform tasks using the system (Butler 1996). From a user-centred

perspective, the aim of a software application is to satisfy users by providing relevant functionality which they find fit for purpose and can enhance their performance. "Testing usability means making sure that people can find and work with [a system's] functions to meet their needs" (Dumas and Redish 1999: 4). Usability is therefore concerned with establishing whether people can use a computer-based tool to complete their tasks effectively and with sufficient ease.

Usability engineering provides systematic methods to measure a system's usability in terms of its learnability, efficiency, effectiveness, memorability, error rate and user satisfaction (Butler 1996; Nielsen 1993). This applies to all aspects of a system's user interface, which can be defined as the medium through which a user interacts and communicates with the computer (Preece *et al.* 1993, Nielsen 1993). The ease of learning a system is measured as the time it takes a potential user to "reach a specified level of proficiency in using it" (Nielsen 1993: 28). Efficiency refers to the level of productivity that the user must achieve once the system has been learned. Effectiveness on the other hand is "the accuracy and completeness with which users achieve specified goals" (Thomas 1996: 109). Memorability measures how easily a system is remembered either after a period of not using it or by casual users. An error in this context is defined as any action that does not accomplish the desired goal (Nielsen 1993), which can be objectively observed and counted to obtain a measure of a system's error rate. Satisfaction refers to how pleasant the system is to use. Preece *et al.* (1994) also consider throughput, flexibility and user attitude towards the system as aspects that must be addressed in assessing a system's usability. Ease of use or throughput is comparable to the concept of efficiency and error rate as it is defined as "the tasks accomplished by experienced users, the speed of task execution and the errors made" (Preece *et al.* 1994: 401). Flexibility refers to the extent to which the system can accommodate tasks it was not originally planned for. Attitude is comparable to user satisfaction or the "comfort and acceptability of use" (Thomas 1996: 109). Ease of use, satisfaction and learnability were the usability aspects investigated with the study discussed below.

Usability evaluation involves recording both objective data on user interaction with the system during performance of task scenarios and subjective data for diagnosing problems. Objective data may include the time spent on tasks, patterns of operations during the test sessions, or errors. Empirical data can be gathered either in a usability lab or in user workplaces. Data can be recorded using multiple cameras and programs for recording and playing back on-screen user interaction. Subjective data include verbal reports that can be recorded either during the evaluation or following the session. These techniques allow

monitoring of user thoughts as well as the problems they are aware of having (Butler 1996). Similarly, qualitative interviews and questionnaires using rating scales can aid in eliciting user satisfaction with the software and perceptions on ease of use and usefulness of a system.

3.3 Usability Study

Usability evaluation normally runs in parallel with a system's development cycle. In research, it can help to assess the suitability of a concept that is being implemented using a computer-based tool, as well as the suitability of the design for supporting some task. This section discusses results from the first study conducted regarding the use of GIS-based tools for exploring spatial data visually and interactively. Its purpose was to provide a qualitative description of the problems that participants had while performing the tasks as well as identifying functionality that enhanced their work process. In other words, the discussion that follows points at strengths and weaknesses of the system and concepts tested. These findings were used in a subsequent development iteration to build an environment that overcomes some of the limitations discussed next.

3.3.1 Goals

The main aim of this study was to evaluate whether the functionality provided by the coupling between the lightweight GIS and DecisionSite served to create an environment that supported the exploration of spatial data. Ultimately, the aim was to identify the potential of linking systems to support geovisualization tasks for spatial data users. In addition, the study was designed so that it would be possible to review participants' work processes, how they thought and behaved when using the system.

It has been argued that dynamic displays facilitate perception (Peterson 1994; Ware 2000). Dynamic graphics are usually interpreted as supporting interaction or direct manipulation of graphical elements on a computer screen that provides an almost instantaneous change of elements (Becker and Cleveland 1987). Those that promote the experimentation with data attribute combinations and transformations, see dynamic displays as essential for exploratory analysis (DiBiase *et al.* 1994; Asche and Herrmann 1994; MacEachren 1994; Openshaw *et al.* 1994). Thus, it was hypothesised that the main shortcoming of DecisionSite Map IS (Figure 3.1) was the lack of a real-time link between the two software components. This refers to the value of providing linked multiple graphical displays that support simultaneous or dynamic brushing as an important feature of systems that allow the exploration of data (Spence 2001; MacEachren 1995; Monmonier 1989).

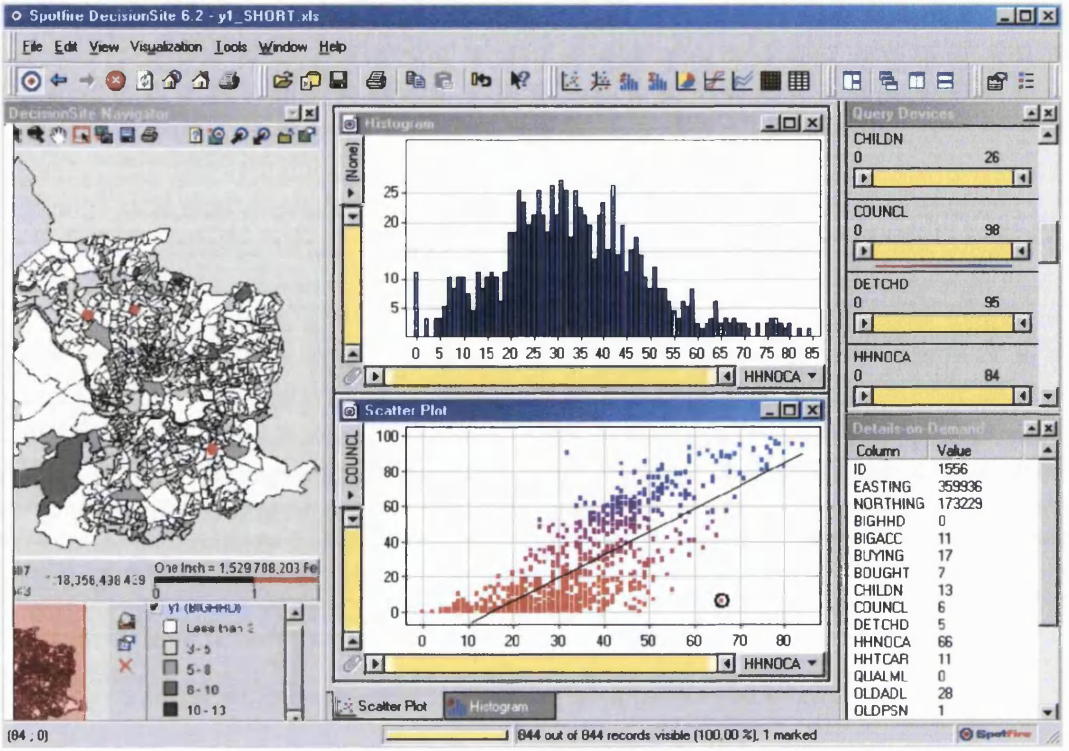

Figure 3.1 DecisionSite Map IS (Spotfire, Boston). To the left of the image is ArcExplorer and DecisionSite appears to the right. The points on the map correspond to a set of records selected in DecisionSite which have been 'transferred' to the GIS as an image and laid on top of the map.

In DecisionSite Map IS, however, there is no simultaneous highlighting between the aspatial graphical displays (such as a scatterplot, PCP, histogram or line graph) and the map. In order to see how data in the former views relate to a map, the user must *transfer* the data from the particular view in DecisionSite to the map in ArcExplorer. In other words, if the user wants to associate attribute values represented on a scatterplot in DecisionSite to their actual location in space, the data must be selected on the scatterplot and then transferred from DecisionSite to the GIS. The selection is then represented on the map in the form of an image file that plots a point on the corresponding geographical coordinates (see the red points on the map in Figure 3.1). Not only is this form of communication lacking a dynamic link between the two components, it cannot be customised differently as ArcExplorer only supports data browsing and does not provide a development environment. It was therefore of interest to find whether the available functionality sufficed for enabling the system as a spatial decision-making tool. In other words, to test the importance of having a dynamic link for attempting geovisualization tasks.

In addition, the identification of potential problems with the current implementation was a main goal since this study was a preliminary enquiry into the characteristics that similar software tools should provide. In this study, particular attention was given to measuring user performance with the system in terms of their ability to generate hypotheses about the data during the tasks provided. It was not expected that participants would follow a strict series of steps to solve the tasks but rather that they obtained plausible answers. The test design and choice of participants aimed at facilitating the evaluation of these goals.

3.3.2 Participants

Ideally, profiles of potential users of a system are developed before the time for a usability test (Dumas and Redish 1999). However, this test was performed on an available product but no previous usability tests, surveys or interviews were available to define user profiles. The main strategy to overcome this difficulty was working with one of Spotfire's usability engineers to understand the aims of providing Map IS as an add-on to the core software. The integration was intended to enhance DecisionSite's decision-making environment to include analysis of data with a geographical component where variables or attributes of interest could be potentially located in space. Using this rough guide as the main purpose of providing the environment, a user profile was prepared to select the participants for the first test.

Participants for this preliminary study were chosen from the GIS industry and academia, as well as from other professions that share the characteristic of using geographical data regularly. They were expected to have experience with interacting with spatial data on-screen and with systems that rely heavily on the use of graphics to provide context to data manipulation. This selection would control for users that had to first learn about spatial data and rather allow us to concentrate in investigating if providing limited GIS functionality to an aspatial visualization system could enhance the process of georeferenced data exploration and decision-making. Finally, this selection of participants aimed at evaluating the benefits that GIS or advanced users of spatial data would obtain from the provision of visualization tools enabled to investigate spatial data.

It was of interest to recruit participants from a range of fields or application areas as it was hypothesised that similar tools could potentially support a broad spectrum of tasks and work processes. Therefore, participants included professionals, masters and doctorate students with a variety of interests but who shared the characteristic of using geographical data regularly. Table 3.2 shows that two GIS professionals, one master's and six PhD students

took part in this study and that their interests or expertise ranged from transport modelling to environmental conservation. They were between 24 and 30 years of age of both genders (four females and five males) and various nationalities. Two other participants (listed at the end of the table) took part in pilot tests designed to assess the adequacy of the tasks that would be used in the main experiments (for the relevance of pilot studies see §3.4).

Table 3.2 Description of participants

User	Gender	Background
1	Male	Architectural Engineer and PhD Student. European (Greek).
2	Male	Environmental Scientist and PhD Student. Asian (Chinese).
3	Female	Geologist and GIS Consultant. North European (Swedish).
4	Male	Geographer and PhD Student. European (Irish).
5	Female	Transport modeller and PhD Student. European (Maltese).
6	Female	Architect and PhD student. European (Greek).
7	Male	Geologist and PhD Student. Asian (Indian).
8	Male	Database specialist and GIS Consultant. European (English).
9	Female	Geologist and MSc in GIS Student. Eastern European (Polish).
Pilot	Male	Archeologist and PhD Student. European (English).
Pilot	Male	Cartographer and Lecturer. European (English).

3.3.3 Tailoring Usability Evaluation Techniques to Geovisualization

The number of participants in the main test was carefully considered and mostly followed the guidelines found to be successful for heuristic evaluation (Nielsen 1994). This is a usability engineering method that involves a small set of evaluators who examine the interface to detect usability problems (Nielsen 1992). Each evaluator typically inspects the interface and their results are either reported at the end or recorded by having each evaluator verbalise their comments to an observer as they go through the interface.

However, the nine participants were not usability or geovisualization experts but rather familiar with spatial data. Therefore a combination of traditional user testing and heuristic evaluation techniques was used for this study. In a user test, the observer typically interprets users' actions and infers how these actions relate to usability limitations of the interface. This allows users who are not knowledgeable of interface design to participate in a usability test. Heuristic evaluation, on the other hand, requires the evaluator to analyse the user interface

and the observer to take notes of the verbalised comments. In traditional user testing, the experimenter does not normally answer questions from the participants during the session. Essentially, the interest is on finding mistakes users make when using the interface and therefore the observers only provide help when it is absolutely necessary. For a heuristic evaluation this is not the case. On the contrary, answering questions from the evaluators is seen as facilitating their better assessment of the usability of the interface with respect to the characteristics of the domain (Nielsen 1994). Therefore, when an evaluator is clearly struggling with the interface and has commented on the problem, they can be given help on how to proceed.

For this study, the role of the experimenter or observer was thus twofold. On the one hand, the experimenter recorded participants comments about the interface and later interpreted their actions aided by the recordings of each user on-screen interactions (§3.3.5). The observer also explained aspects of the interface that were difficult for someone with limited visualization domain expertise (Nielsen 1994). In this manner, participants workload was reduced. Given that not all participants had GIS expertise, this was desirable. However, they all had experience with spatial data and hence could reasonably be expected to be familiar with interfaces for handling such data. Hence, during the test, the experimenter assisted the users in operating the interface only in case of severe problems that restricted or impeded the continuation of the evaluation.

In addition, in a heuristic evaluation, evaluators decide how to proceed with the assessment of the interface to describe it in relation to common usability properties (Nielsen 1994). In this study, however, users solved tasks or scenarios as in traditional usability testing (§3.4). Nevertheless, users were encouraged to verbalise their comments as they went through the process of solving the task and were allowed to engage in discussions with the observer where necessary.

This combination of techniques was appropriate for evaluating the environment for at least three reasons. First, the main interest of this evaluation was to elicit user opinions on how the environment supported the exploration of georeferenced data. Their expertise with different data types and systems served to bring forth a myriad of observations about usability problems in the interface and limitations of the environment. Second, this being the first study and first prototype tested, it was reasonable to keep the evaluation as simple as possible and to concentrate on finding usability problems with the interface. In effect, the test facilitated the observation of this and other issues such as user behaviour as it is

discussed in §3.6. Finally, following the guidelines of heuristic evaluation allowed for the number of users to be kept to a minimum. According to Nielsen and Molich (1990), 80% of the usability problems of an interface can be detected with between 4 and 5 participants during a heuristic evaluation. In the context of usability testing, Virzi (1992) found that 90% of the usability problems can be detected with 10 participants and Dumas and Redish (1999) argue that additional participants are less likely to reveal any new information. Thus, nine participants were considered to be an appropriate number for this evaluation.

In summary (Table 3.3), GIS users and users of spatial data were invited to participate with the purpose of evaluating the usability of the environment for visually exploring spatial data. In addition, starting to understand user behaviour with this type of software was an important issue as it may aid in the design of tools that better support users work processes. It was hypothesised that the benefits for potential users would include the rapid exploration of spatial datasets in an environment that provided GIS-like functionality and spatial context to the investigation of geographical patterns. Nevertheless it was expected that users would have to invest in time to fully understand the workings of the system as well as the information provided by the various data views.

Table 3.3 Summary of goals, benefits and foreseen costs (Dumas and Redish 1999).

User group	Tasks goals	Benefits	Costs
GIS users and users with knowledge of spatial datasets	To find whether the GIS enhances DecisionSite's capabilities for facilitating the exploration of spatial data. To understand user behaviour when utilising an interactive visualization environment.	To explore spatial data visually, provide spatial context and GIS-like functionality. To improve design by incorporating functionality that facilitates users' process of visual data exploration.	Time to learn to use the system in terms of understanding how to use the visualization tools provided.

3.3.4 Invitation to Participate and Motivation

Users were emailed an invitation to participate in the test, which included all the relevant data about the evaluation. The invitation briefly described the system and why potential users' participation was relevant. It detailed the place where the test was going to be held and how it was going to be structured (demonstration, recorded test and taped interview). It also indicated the estimated duration of the test (about one hour and a quarter), as well as the purpose of the test. Its potential interest and benefit to the users as a session on methods for

data analysis was highlighted as users were recruited through networking and no payment was offered for their participation.

Once users had confirmed their willingness to participate, they were given a short description of the dataset (see Chapter 6 for a discussion about the data used). The purpose of this description was two-fold. On the one hand, it provided users with information about the dataset to minimise the time spent during the experiment on understanding the meaning of the variables in the dataset. It also aimed at motivating participants to perform the tasks thoroughly by detailing the particular characteristics and challenges of the dataset. Careful consideration was given as to whether users should have a priori information about the dataset and how this would affect their performance and results. Three arguments can explain choosing to provide this information. First, data users typically approach a dataset with a purpose (Trafton et al. 2000). Be it decision-making, research, confirmation of hypotheses or exploration, investigating a dataset follows from posing a question or aiming to obtain some information. Hence, providing a description about the data was not seen as impeding the discovery process but rather as a means to motivate users to explore it in depth. Second, the information given was mostly descriptive and included a discussion on the relevance and interest of the dataset for participants. Finally, there was a practical concern with limiting the time users would spend in the test as all users were expected to have a restrictive time schedule. Providing a description of the dataset beforehand was expected to allow them to concentrate on learning to use the software and solving the tasks during the test.[1]

3.3.5 Test Design

The usability study was designed to test a limited set of functionality (in particular the limited link between some of the graphic displays) and to take about an hour and a quarter of each participant's time. It was also planned to be conducted during a three-week time span. Within that time frame, two pilot tests were carried out to help refine the tasks, as well as to detect any possible problems with the software or hardware required to conduct and record the sessions. The final two weeks were allocated for the nine participants to select a convenient time to do the test.

Each session had three phases. In the first, users were given a demonstration of the software and an explanation of the data that they were going to use prior to the test. None of the participants had used the software before so the system's demonstration was aimed at

[1] For the invitation to participate and details provided to the users see Appendix A.

allowing them to learn and get comfortable with the functionality that was going to be tested. Also, any general questions about the dataset were discussed beforehand to assure participants understood, at least theoretically, the dataset they would encounter. The second phase included the actual testing, where users where given a set of tasks or scenarios that they had to solve. Their interaction on-screen as well as any comments they made during the tasks were recorded using Matchware's ScreenCorder (http://www.matchware.net/). The third and final phase included a post-test interview to elicit participant views on the appropriateness of the software tool for exploring spatial data visually. The interview was taped for later analysis (see Table 3.4 for a summary of all resources).

Table 3.4 Summary of Test's Resources

Time	1 hour 15 minutes per test
Hardware	1 PC connected to the internet Dictaphone
Software and data files	Spotfire DecisionSite and Map IS ScreenCorder Shapefiles (map and attribute data) Excel file with data subset
Instructions	Detecting relations, outliers and mistakes Choosing candidate locations

3.4 The Tasks

An essential requirement of usability tests is that the tasks are realistic. In other words, that they pose scenarios that potential users of the system would need or want to address (Dumas and Redish 1999: 160). This implied a very careful design of the tasks to be performed especially since the time available for testing was very limited. The tasks were divided into two parts (Table 3.5). The first part was deliberately ill-defined so as to leave room for users to explore the dataset at will. Some guidelines or 'information to look for' were provided only to limit the time users were to spend on the first part of the test to about 20 minutes. However, there was no unique answer to the problem and the interest was more on analysing how participants interacted with the system and used the visualization tools to obtain information from a spatial dataset that contained unknown information. Nevertheless, in terms of accuracy, it was expected that users would be able to discover relations between variables as well as anomalies or potential mistakes in the dataset.

Table 3.5 Tasks

Part 1
1. Find whether variables in this dataset are related in some way and whether their relations occur in any particular areas of Bristol.
2. Can you detect 'outliers' (unusual combinations of attribute values)?
3. Can you detect possible mistakes in the dataset entries?

Part 2
Having explored the data, can you choose candidate locations for (please take note of the polygon ID):
4. A secondary school
5. A gourmet shop

The second part was aimed at testing whether the software did provide an environment suitable for decision-making or problem solving with geospatial data. They implied identifying candidate locations that would be potentially suitable for an activity. Since visualization tools are not intended to be confirmatory but rather support a process of hypotheses formation, there were no unique answers to the questions posed. On the contrary, their outcome was a function of users' criteria for addressing the task. For instance for the task where they had to choose a location for a secondary school, they could choose the characteristics of the catchment area (such as the households income level) where they would potentially have it built. The interest here was on observing how users would define a set of criteria to make a decision with the data provided and whether the system easily enabled this process. Table 3.6 summarises the two parts of the evaluation and their aims.

Table 3.6 Tasks aims and description.

	Part 1
Concern	Testing whether the present form in which Map IS and DecisionSite are coupled helps users to acquire information about or insight into a spatial dataset.
Test Setup	Users were asked to use all the demoed visualizations.
Task Description	Finding relations between variables and their location in space. Detecting anomalies and/or mistakes in the dataset.
	Part 2
Concern	Detecting whether there were usability problems with the interaction between Map IS and DecisionSite for the purpose of supporting spatial data exploration.
Test Setup	Users were asked to choose locations.
Task Description	Deciding for a candidate location(s) for a school and a gourmet shop.

3.5 Results

This section discusses some of the usability problems found with DecisionSite Map IS (for a detailed description see Tobón 2002). In particular, features that facilitated or inhibited participant visual exploration of the spatial dataset are mentioned as well as subject behaviour while using the tool. This is because, as Dix *et al.* (1998: 407) suggest:

"In addition to evaluating the system design in terms of its functional capabilities it is important to be able to measure the impact of the design on the user. This includes considering aspects such as how easy the system is to learn, its usability and the user's attitude to it. In addition, it is important to identify areas of the design which overload the user in some way, perhaps by requiring an excessive amount of information to be remembered."

3.5.1 User Observations about the software

3.5.1.1 Observations of Satisfaction

DecisionSite provides dynamic query devices (Chapter2) for performing structured queries on the fly (Ahlberg and Shneiderman 1994). Query devices can be implemented graphically in a number of ways. Figure 3.2 shows their implementation as two-way sliders that can be dragged by the user to reduce the number of items in a view. An alternative is to use check boxes to select a number of possible values, or radio buttons to select just one value of an attribute. Under the name of alpha sliders (Ahlberg and Shneiderman 1994a), they have also been used to perform queries on text data. Users found the ability to dynamically filter data from the screen using the query devices and the possibility to plot the result onto a map to be the most innovative and salient functionality of the environment. Most of them were not familiar with visualization packages but immediately recognized the value of investigating spatial data interactively: *"You can find [spatial] relationships directly using the map... This is really cool!"* (User 7; see Table 3.2 for a reference to the users). The query devices were perceived as powerful tools that enabled users to focus on subsets of data and quickly find the information they were interested in. They hide the structured query from the users very effectively, allowing them to concentrate on the exploration: *"You do not need to build the query by yourself"* (User 8).

The evaluation therefore highlighted the importance of enabling the user to focus upon subsets of data. The dynamic query filters serve this purpose. Providing similar functionality might be necessary not only when the dataset is large but also when the user needs to narrow down the data to those pertaining to a particular insight or hypothesis. Although a training period may be needed to master the use of dynamic query filters, the devices have been shown to be quick, powerful (Tanin *et al.* 1997; Ahlberg and Shneiderman 1994a),

significantly faster and more pleasant to use than other database query interfaces. It has also been argued that they are an easier method to master for less expert users than SQL (Ahlberg *et al.* 1992).

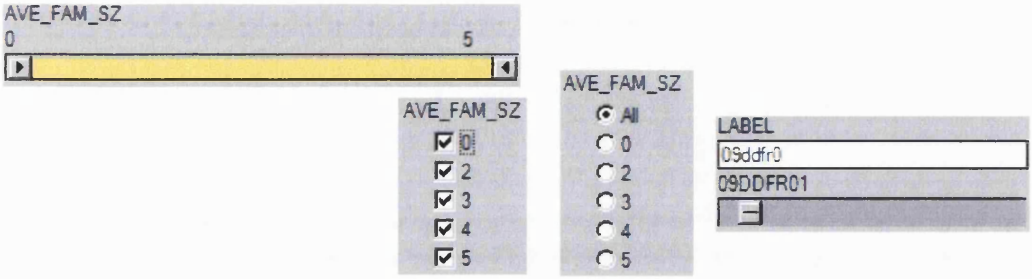

Figure 3.2 Four different implementations of Query Devices in DecisionSite (Boston, MA). From left to right: Far Left) Range Sliders: users move the end of the sliders or the slider itself to select a data range they want to appear in the graphical displays. B) Checkboxes: allow users to select more than one value of an attribute to show in the graphical displays. C) Radio Buttons: allow users to select only one value of an attribute. D) Alpha Sliders: allow the querying of text data.

Manipulating all elements of a visualization (such as the colour coding of objects on a scatterplot, their size and orientation) from a single properties window was also seen as an advantage, especially over doing the same operations using menus: *"You don't have to go to look in menus to find what you want to do which I think is really good because menus are only good when you remember where things are"* (User 3). Its use was rapidly learned and preferred to manipulating each item of a visualization separately. Users testified that this functionality was the most useful when substantial changes were needed. For instance in the PCP, the properties window was preferred when adding or deleting large numbers of variables from the display.

Since the environment evaluated is windows based, more advanced users expect the system to comply with windows standards such as shortcuts. Furthermore, these users would expect some commands to be reserved for particular operations, i.e. control key plus movement of the scroll wheel or button in the mouse to zoom in and out of a view; or control key plus letter 's' to save. Meeting these standards allows the user to recognise familiar functionality with ease and also to learn new contexts in which such functionality can be applied by analogy to systems that they already know. For instance, properties windows are increasingly common in windows based systems. For this reason, they were seen as an advantage of the system as it is a familiar feature with which to customise a view. Even when users did not expect or had not encountered this functionality before, properties windows were perceived

as a more efficient way for customising a view than manipulating it directly when many changes had to be made.

Table 3.7 Frequently of feature acceptance

Advantages	Frequency	User Comments
Ease of learning the system	8/9	*"The graphics' capabilities resemble S-Plus"; "Similar things I've seen before in Data Desk" (User 4); "Looks like Excel" (User 2).*
Ease of using the system	8/9	*"It is really quick and very effective" (User 6).*
Interaction between the GIS and DecisionSite	7/9	*"The ability to dynamically select points and to be able to see them on the map I think is very useful, because generally otherwise [...] you really have to juggle between applications to do that" (User 7).*
Ease of using/understanding the Scatterplot	9/9	*"The Scatterplot is probably the most useful [tool] for finding relationships. It is also the one I'm most familiar with [...] It shows outliers and it shows not only the relationship but the strength of the relationship at the same time" (User 4).*
Ease of using/understanding the parallel coordinate plot	4/9	*"It first looked like it was just loads of lines but once you learn how to use it; it was probably the most useful" (User 1). "It's quite a useful graph as you have more variables" (User 8).*
Ease of using/understanding line chart	4/9	*"Looks like Excel so it's very easy to get used to" (User 6).*
Ease of using/understanding the properties window	8/9	*"All I need to control the look of the windows is right here" (User 1).*
Ease of using/understanding the query devices	7/9	*"Filtering is the best [tool] because you can immediately take away the data you are not interested in" (User 5).*

In addition, the users had previously encountered some of the visualizations used such as the scatterplots, histograms and line charts. Their familiarity made users more comfortable and accounted for some of the user satisfaction with the tool. This in turn may have contributed to the understanding of the system and in learning its functionality. Users however also commented on the speed of the tool. Combined with their perception that the tool was easy to manipulate, this was recognised as a crucial feature for supporting the rapid investigation of the ideas that the exploration process prompted. The details-on-demand window, which provides expanded or contextual information about a feature (Chapter 2) was also a valuable resource for test participants. The contextual information provided in the window depends on the type of operation the user performs. For instance, if the user demands details for a set of geographical areas or graphic objects, then the information is presented in table form as

opposed to a list which is provided when particulars about a single record are required. This context-sensitive and mainly user-driven behaviour allowed the window to provide the details that were relevant to the operation performde. A summary of these and other comments about salient features are summarised in Table 3.7. The table describes the frequency with which a feature was cited as an advantage by users. Excerpts from user comments from the interviews or the recorded test sessions are included as examples.

3.5.1.2 Observations of Frustration

The main downsides of DecisionSite Map IS were: i) not being able to brush (or dynamically link) the map view with the rest of the visualizations in DecisionSite; and ii) not being able to explore the attributes of data plotted into the GIS. In terms of the first disadvantage, the map and the rest of the visualizations cannot be updated interactively. Users found having to load their queries into ArcExplorer or DecisionSite to be problematic since they would easily lose track of the initial idea they were exploring: *"You tend to think of things on the fly and if you don't satisfy your query within a couple of seconds you lose interest and go on to something else especially when you have such a huge data set"* (User 4). In terms of the second disadvantage, points can be plotted on the map only as an image. This does not permit users to change the properties or inspect attribute values of a point plotted onto the map as you can normally do with a layer of spatial information in a GIS. Furthermore, users could not easily relate the data on the visualizations to that on the map. This was due in part to their different representations. For instance, attributes in the dataset were represented as points in a scatterplot or lines in the PCP. On the map, what users could see was areal data. Hence, because there was no visual feedback which cued users that all graphical displays where different re-expression of the same data, they could hardly make the inference: *"I found it very difficult to relate the data points [in the aspatial visualizations] to the map"* (User 6).

Therefore, the evaluation suggested the need for an environment that enables dynamic linkage of all graphic displays which represent different views of the data and allows brushing following user on-screen interaction with such features. Multiple coordinated visualizations can alleviate the difficulties that users experience in relating and exploring complex information by taking advantage of the strengths of each visualization for revealing different aspects of the data (North and Schneiderman 2000). The lack of brushing and linking in particular was found to hinder user understanding that each view in the environment was a different representation of the same dataset. Hence, without this form of coordination, the perception that geographical features or areas on the map had associated attribute values which were represented in scatterplots, histograms, and PCPs was difficult. Thus the potential of the graphical displays for facilitating inference was not efficiently exploited. A

number of studies have indicated similar results in terms of the advantage of coordination between specific displays, which would indicate its advantage over an uncoordinated alternative (Chimera and Shneiderman 1994). Display coordination was found to improve user performance in terms of the time taken to obtain the same information with and without the coordination, as well as user satisfaction (North and Schneiderman 2000).

A further disadvantage with the environment evaluated is that the only way to save a particular data selection plotted in the map is to save an image of the display. This of course would not permit any further interaction with the selected data, i.e. with the recorded scenario. This was seen as a further constrain in the exploration process as users needed to compare various queries to make decisions about the suitability of their criteria for choosing candidate locations. Since the transfer of data was cumbersome, it made the comparison of maps (with data selections plotted on them) cognitively too demanding. This translated into subjects not utilising the map as frequently in the data exploration process and most of the searches being performed in the aspatial graphical displays.

A final problem with the transfer of data onto the map is that every time a transfer is made, the user must set the columns that should be used as latitude and longitude in a dialogue window. This was seen to cause serious problems to users when they entered the wrong variable as the latitude or longitude. For instance, one user mistakenly selected the Eastings to be taken as the longitude and the Northings to be taken as the latitude. Hence, his data selection appeared outside the study area. However, the user could not understand how this would happen since his choice of Latitude and Longitude was accepted by the dialog window and no feedback or warning was given to cue that a mistake like this could occur. This intermediate step should not be present as it interrupts subjects train of thought and is prone to errors.

An alternative to plotting points on the map is to load the map as an image into one of the aspatial views of DecisionSite. This requires the use of a scatterplot, the axes of which are set to latitude (x axis) and longitude (y axis). Although the map may provide a visual aid when used in this manner, relations between other attribute values cannot be explored in this same window as the axes must be fixed to latitude and longitude. Other graphical displays can be used simultaneously to support the exploration of the data through the dynamic brushing that occurs in all the views in DecisionSite. However, this approach excludes any possibility of manipulating and interacting with the map, i.e. changing the colour coding of areas or

obtaining information about the data from the map itself. In other words, the map is left to provide context but cannot be used as an exploratory tool in its own right.

The environment offers one more possibility which does not solve the problem of integrating the map as part of the tools for exploring the data. Subsets of spatial information can be selected from ArcExplorer and transferred into DecisionSite. In theory, this enables users to concentrate in a smaller area. However, if a dataset was already opened, this solution obliges the user to start a new session with the subset of data selected. Whatever dataset the subject was exploring must be closed in order to complete the transfer of the new data selection. Any customisation that the user might have given to the displays is also lost. In addition, as the size of the data to transfer increases, the operation becomes extremely slow.

Apart from these limitations, some users also found the interaction with the query devices difficult. They have been tested for their ease of use and said to be adopted and understood quickly with limited training (Ahlberg 1996). However, even after a demonstration, some users scarcely made use of the query devices or could not understand the effect they had on the dataset. In effect, once a range slider is moved to filter for instance a high value of an attribute, the whole record containing that particular value for that variable is excluded from all the views and the analysis. However, some users did not realise that the effect was on a whole record but thought it was only on the particular attribute from which the value was being filtered. Other users opted for actually deleting the unwanted graphical objects from a view, which was also possible. Although deleting a record serves the same purpose as the query devices, i.e. narrowing the amount of data seen on-screen and inspected, it has a permanent effect on the database. This indicates that the effect of moving the position of a graphical widget may not be as easy to grasp without training or practice, although the operation necessary to achieve it is simple.

In summary, DecisionSite MapIS is intended to permit "users to visualize and interact with [GIS] data within the Spotfire DecisionSite environment" (http://www.spotfire.com/products/mapper.asp). However, in the opinion of one of the respondents, the environment *"allows you to visualise results spatially but is not enabled to really act as a tool for exploration"* (User 4). Queries plotted in the form of points from DecisionSite onto the map cannot be further explored or their attribute values inspected. They are added as an image on top of the map (as opposed to add as a layer or shapefile) and thus their size, shape or colour cannot be modified. In addition, the data transferring was cumbersome and inhibited the exploration process. Table 3.8 summarises the main problems participants

faced when using DecisionSite MapIS, the frequency with which users reported the same problem and excerpts from users' comments to exemplify their views.

Table 3.8 Users' frequent problems

Problems	Frequency	User Comments
Dynamic link: Trying to relate information in all views	8/9	*"[Brushing] would be very useful now because I selected 5 points and plotted them on the map but now I want to know where they are on this scatterplot" (User 6).* *"The software is not good at relating spatial information. It is good at relating trends between datasets" (User 4).*
Use of the map as an exploration tool	7/9	*"You don't really use the map to find things" (User 5).*
Viewing previous queries	7/9	*"I would like to have more interaction with the map, because if I need to find an area where there are certain characteristics, I would like to perhaps choose [some points first], then I would like to display them on the map, and then select some of them or then perhaps investigate the selection that I've done from the map. So, a kind of two step investigation, forwards and backwards, which I would find a lot easier" (User 3).*
Using/understanding the query devices	7/9	*"I didn't really remember I could filter" (User 3).* *"I need to move all the sliders to get rid of these points, right?" (User 2).*

3.6 User Behaviour

Apart from identifying flaws in the interface of a visualization system for improving its design, the purpose of this study was also to better comprehend how users perceive and understand information when inspected visually. Spence (2001) suggests that navigating in a graphical display to investigate data implies a four stage cycle that loops as many times as users deem necessary (Figure 3.3). In such processes, users look at data displayed in some form, searching for content or some recognisable pattern (Unwin 1996). The user must then remember this information and attempt to build an internal model with which to interpret the data. If the observed information can be explained, the user must decide where to look next in order either to test or refine the cognitive model. Otherwise, if the data interpretation does not support the initial hypothesis, it has to be modified and users must decide where to look next to find information that may aid in its improvement.

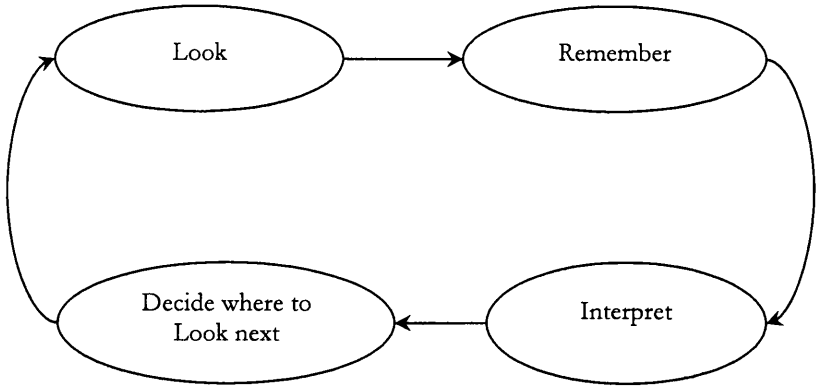

Figure 3.3 Visualization Process (Spence 2001)

The fact that a subject must remember a visual result when using DecisionSite Map IS before they can move to further explorations partly explains why the map was used infrequently to perform the tasks using the environment evaluated. The transfer of data of interest onto the map was slow and demanded thinking on other issues such as selecting the correct data attributes as latitude and longitude. This put a strain on the cognitive load of subjects who usually opted to find answers to the tasks without taking the spatial information that the map provided into account. Only towards the end of the task would users resort to the map as a means of checking their answer.

Klatzky (1975) suggests a model of information processing where a set of three memory stores serve to obtain information from graphical displays. The first store is known as the sensory register or iconic memory which holds information in sensory form long enough for it to be recognised (Peterson 1994). Through a matching process with knowledge previously acquired and stored in long term memory, the image sensed is sent to short term memory where information is temporarily stored. The contents of short term memory must be rehearsed in order to be maintained (Matlin 2002). If the subject is required to do a task which interferes with the rehearsal, then the information in short term memory decays and is not transferred into long term memory. When a visual display is used, the information may not need to be rehearsed as it is directly available to the user. Hence the display acts as an extension or memory aid. However, if the display prompts an idea, it needs to be explored immediately as it may not be rehearsed into long term memory and remembered. If we take Klatzky's model to be a plausible explation of the dynamics of information processing, it then further stresses the need to support rapid feedback from views in a visualization system. Hence, a link between views that requires intermediate steps to relate the information they portray will not be an effective means to explore and retain items in short term memory.

Figure 3.3 suggests that users loop through the process of looking, remembering, interpreting and deciding where to look next in order to find some information by navigating a visual display. Figure 3.4 builds upon this basic model to include the user behaviour observed during the tasks and also inferred from later conversations with the subjects. User exploration of the information may start by browsing through the dataset to look for content that can prompt ideas. As they browse, they gather information in support of their hypotheses. It was seen that users typically had a goal in mind such as corroborating some intuition that they had about the relation between two attribute values. In a visualization system, this can be done by trying different combinations of variables in a display that may suggest a particular relation. Changing the appearance of the graphical objects may also serve this purpose for instance by changing glyph shape or colour according to the values of some other data attribute. With the gathered information users interpret the data and evaluate their initial hypotheses, if any. If necessary, they may then amend or completely discard their initial hypothesis. This would include an evaluation phase that is not explicitly considered in the model of Figure 3.3 but which was repeatedly seen in users work processes. Loops around these steps could occur either because subjects needed to obtain more information about a particular task but also as a means to evaluate a plausible answer.

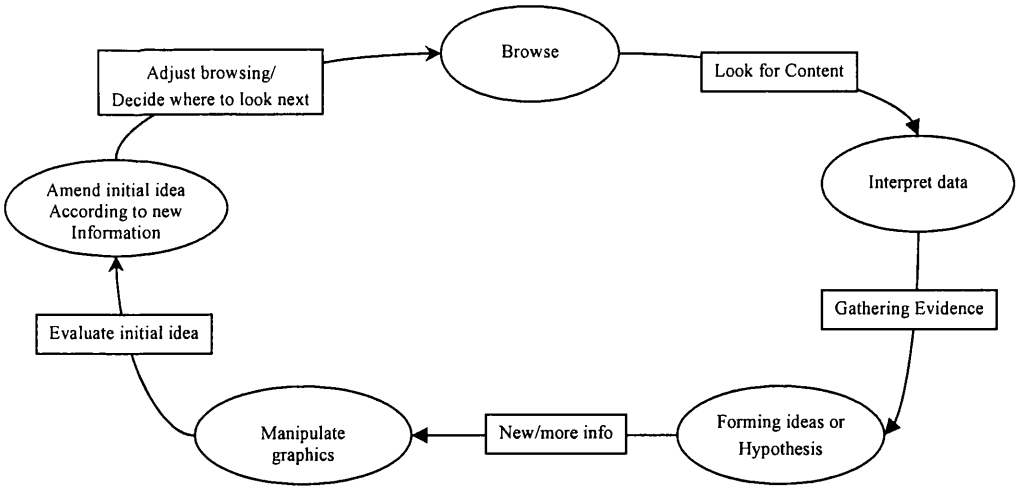

Figure 3.4 Visualization Process II

However, the looped sequence of events in Figure 3.4 does not always occur smoothly. For instance, at least one of the nine users surveyed here was not able to obtain any ideas or form any hypothesis from browsing through and looking at data on-screen. This user could not see the benefit of having a number of tools for re-expressing and investigating the data and hence was unable to use the package to her advantage. This may suggest that investigating

data visually is not necessarily as simple as it is frequently implied. In fact, data visualizations entail a process in which information is encoded in a display (Cleveland 1993: 1) and users may need to learn how to decode the information it portrays. For instance, all users had initial difficulties in understanding the information provided by the PCP as they had not used one before. Once explained, at least three out of the nine participants found it to be the most useful tool for exploring multi-dimensional data. Nevertheless the sole manipulation of the visualization cannot always hint at all the information it can provide. This suggests that effective use of visualisations and graphic devices may be conditional on training. Larkin and Simon (1987) suggest that this is because graphic displays can be extremely useful and efficient for drawing inferences due to the way in which they index information. Hence they are useful if the user is aware of how information is encoded and can take advantage of its properties. Therefore, extracting relevant information from data views promptly and easily requires knowledge about the display and possibly practice in its use or application. Even with easy to operate graphic devices such as a scroll bar for focusing upon a subset of the data, the meaning of any operation needs to be learned and its consequences understood.

The evaluation suggested that participants were able to attain a high level of proficiency with the software rather quickly. A plausible explanation for this is that the tool is based mostly on direct manipulation of interfaces in order to permit users to concentrate on the task in hand rather than on mastering the use of the tool (Rutkowski 1982). Direct manipulation allows the "continuous visual representation of objects and actions of interest, [as well as] rapid, incremental, reversible operations whose impact on the object of interest is immediately visible" (Ahlberg *et al.* 1992: 619). Most operations with the system provided constant feedback. This may allow novice users to start working with little training while still being powerful for advanced users (Tanin *et al.* 1997; Ahlberg *et al.* 1992). Nevertheless, it was sometimes the case that users went a few times around the loop in Figure 3.4 without any indication that they had obtained any insight on the data. They were also reluctant to advance ideas when the exploration process was in progress. Even though they knew the purpose of the environment and the tasks was the generation of hypotheses, this behaviour might be a by-product of users feeling that they were being tested and hence had to perform to some standard. At least in one case, however, this was due to the user being unable to interpret the visualizations or understand the purpose of the environment.

Finally, the loop in Figure 3.4 suggests that users incrementally improve their mental models as they browse through and re-express the data. However, most users often investigated data or relations that caught their attention and were far from organised in their exploration.

Nevertheless, they managed to obtain rich insights following different routes. Figure 3.5 illustrates two-way relations between the stages in the loop that may better describe individual ways of forming hypotheses. It shows how users may go backwards and forwards between steps before proceeding to higher levels of abstraction. Furthermore, it shows how some 'steps' can be skipped or simply different operations sequences. For instance not all hypotheses may be prompted after a display has been interpreted by the user. Similarly, users may not amend or evaluate their hypotheses as they build them but may wait until more are gathered to compare them or opt for one.

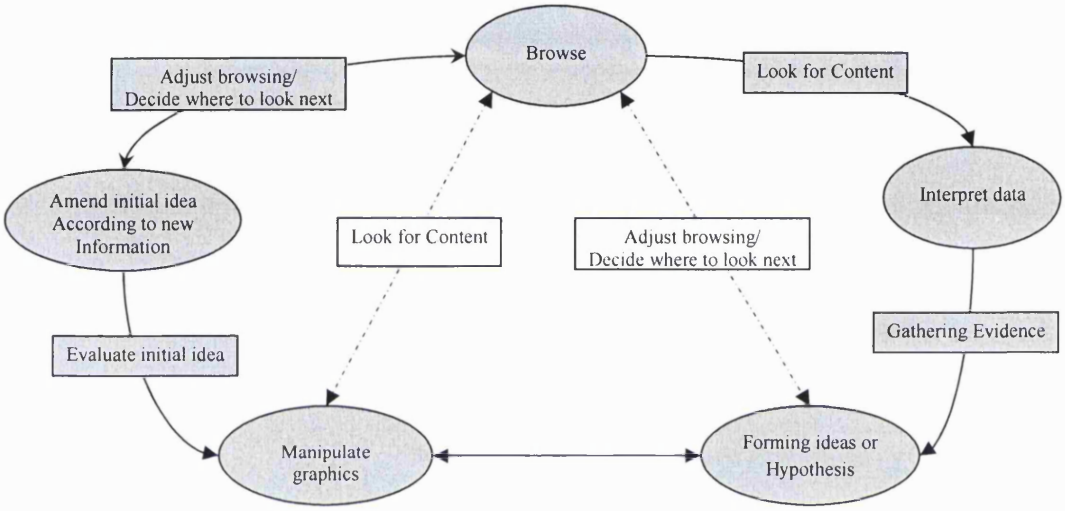

Figure 3.5 Visualization Process III

3.7 Next iteration: Improving the tool

The environment evaluated provides a powerful tool for the visual exploration of data. However, this chapter has discussed various limitations in terms of usability and functionality that make the investigation of georeferenced data difficult even for the study participants who were all spatial data users. In particular, the evaluation showed the need for a dynamic integration of the displays, which implies that data moves between the systems as a response to user actions (Anselin *et al.* 1993), as opposed to a link where data need to be passed between the components of the software environment. Nevertheless, the study has shown the potential of a modular approach to systems integration for providing an environment that supports a particular set of users and tasks.

The development of a second environment for spatial data exploration was thus undertaken (Figure 3.6), which aimed at creating a tool that banked on the features found to facilitate

user data investigation during this study and address those issues which hindered participant understanding of the data and the information they contained. The environment was built by coupling a commercially available GIS (ArcGIS from ESRI) and DecisionSite (from Spotfire) for supporting the dynamic and interactive exploration of vector data (point, lines and polygons and associated attribute values). Although the potential limitations of this approach were discussed at the beginning of this chapter, the discussion that follows explains why it was considered and developed further.

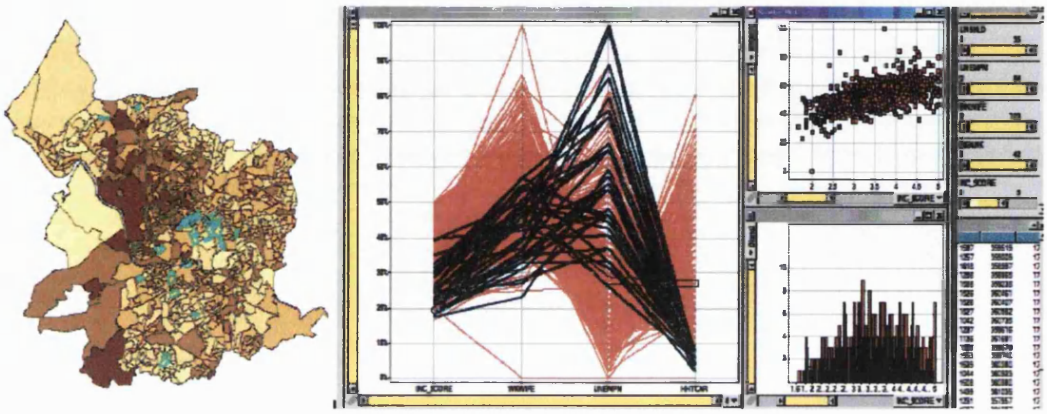

Figure 3.6 Environment's data views for the visual exploration of spatial data

3.7.1 The COM approach

There is currently no standard package that supplies and integrates a range of functionality for supporting the use of visualization throughout the scientific cycle (Gahegan *et al.* 2001). Instead users often resort to a myriad of programs over which the desired functionality is scattered, which may be difficult to connect or that require effort on the part of the user in terms of transferring data and results from one system to the other. This is partly because packages will tend to support specific data structures and formats and will be designed to perform particular forms of analysis. GIS for instance, are less strong in data exploration as they lack the flexibility or functionality to re-express data in ways that support brushing and direct manipulation of some displays. In contrast, they are designed to store information about the topology of the data which is not a necessity in statistical packages. Since providing a broad host of analysis features in a single system may not only be implausible but also not necessarily an efficient means of carrying out the data analysis, various forms of links between GIS and other data analysis packages have been attempted. However, many of these impose constrains on the size of the datasets that can be investigated or on the performance of the system as aforementioned.

Furthermore, little is known about how visualizations should be combined to support an effective process of knowledge construction via visual data exploration, or whether the sequence of graphical display use affects user processes of information discovery (Gahegan *et al.* 2001). Added to the wide range of applications to which visualization techniques are applicable even within the realm of geographic information, as well as the exploratory nature of geovisualization (Slocum *et al.* 2001), these aspects make the specification of a system based on functionality requirements or task analysis problematic.

Yet, combining the functionality in one or at least integrated and seamless environment would be desirable. Takatsuka and Gahegan (2002) suggest that scientific research implies a set of stages such as observation, data exploration, model construction, simulation, verification and communication of results that would benefit from an environment that supports these activities in the integrated form. In addition, the increasing complexity of spatial data, both in terms of number of observations and number of variables, demands new techniques for spatial data analysis of both an exploratory and a confirmatory nature (Anselin 2000, 1995).

Although visualization techniques may aid in the understanding of data, quantitative analysis may be needed to validate or confirm the hypothesis generated through visual data exploration. GIS are "inherently visual and much of their power lies in the potential to display [large amounts of] information visually" (Goodchild 2000: 8). They also provide some spatial data analysis and processing techniques as part of their off-the-shelf functionality. However, they treat graphical display as an output or endpoint of an analysis process and do not facilitate its use for visualization. Hence, they lack the flexibility needed to support visual data exploration and thinking. Some visualization packages provide statistical functionality that may be used for confirming hypotheses. They however may not be designed to support data structures common to spatial information. However, a standard feature is their ability to link different graphical displays representing a variety of aspects of a dataset through highlighting, which has been suggested to enhance the data analysis process. The strengths and limitations of each family of data analysis systems starts to suggest the potential for their coupling in terms of facilitating user tasks.

Despite the technical difficulties encountered in coupling systems, GIS appear to be the logical meeting point of modules for (visual) spatial data analysis. On the one hand, they are designed for displaying geographical data and capable of handling large datasets. On the other, they are increasingly open environments in the sense that they allow more

customisation and their functionality can be extended either by programming or by "developing mechanisms that allow the analyst to mix and match components in order to accomplish specific tasks" (Anselin 2000: 15).

Hence, a tight coupling approach is explored in the following chapters with the development of the environment shown in Figure 3.6, which combines the functionality of DecisionSite and ArcGIS software. This was achieved by coupling the two pieces of software using a COM software architecture that allows the components (or pieces of compiled code that provide some service to the rest of the system) made by different software vendors to be combined into applications (Williams and Kindel 1994). In other words, the coupling can be said to be 'close' (Table 3.1) as the functionality is written in the GIS programming language. The development of the tool was achieved using VisualBasic for Applications (VBA) in a Windows 2000 platform. VBA is the COM-compliant programming language provided with ArcGIS. Nevertheless, COM is a binary standard for component interoperability and function calling between components, which is not dependent on a particular programming language and is available on multiple platforms including Windows, Windows NT, Apple Macintosh and UNIX. Hence, other software systems could be integrated in a similar manner using other programming languages that can create structures of pointers and explicitly or implicitly call functions through pointers. This is because:

"components are manipulated indirectly through one or more interfaces, which are specifications or contracts to perform a related set of functions. Interfaces are pointers to virtual tables in memory or 'vtables', which themselves are pointers to different parts of the actual object code. More precisely, an interface is a pointer to a pointer to a vtable, which allow multiple instances of an object to access the same vtable if necessary" (Ungerer and Goodchild 2002: 44).

In other words, this not only means that the environment reported here could have been created using other languages, but also that components written in different languages can communicate and be reused.

COM allows the two applications to connect and communicate directly without using intermediate or temporary files for exchanging data or functions. Hence when coupled, they behave seamlessly like one environment to the user. The communication in this case was between two systems residing in the same computer, but it can also occur over a network. Furthermore, it allows the analyst or researcher to write her own functions in any other COM compliant programming languages. These can not only be more robust and powerful than proprietary macro languages such as Avenue, but the latter are being abandoned and replaced for industry standard object-oriented macro languages. This situation is desirable as

the researcher may not have to learn new programming languages and their syntax but rather the software objects which allow her to control the application.

Although functions could have been written, for instance, in C++ or JavaScript for developing this environment, VBA was used for a number of reasons. First, VBA has become the macro language for multiple software packages not only because of its degree of support, but also because it provides features like code completion and code debugging. Second, although C and C++ are more robust than VBA, this gap is rapidly closing (Lomax 1998). Finally, if the routines are not compiled, their details can be inspected and functions and subroutines recycled hence avoiding the 'black-box' approach of some of the coupling strategies of Table 3.1 (such as the S-Plus and SpaceStat link to ESRI's products) and even of the functionality provided in most commercial GIS packages. This may be desirable for research purposes as it allows programs to be shared as well as the inspection, adaptation and further development of routines.

Other recent examples of linkages using COM technology include Ungerer and Goodchild (2002) and Zhang and Griffith (2000). The former uses COM for developing a coupling strategy to perform spatial analysis (using Microsoft Excel) within a GIS (ESRI's ArcGIS). The latter implements a link between a database (Microsoft Access) and ESRI's MapObjects in order to use GIS components from within the database. Similarly to the first approach, the environment developed and discussed in the following chapters aimed to create an application using components from different sources. Since the GIS that was used as one of the components does not rely on a proprietary programming language, analysis functionality that is not part of the GIS 'tool box' could be embedded into the GIS with relative ease to expand the capabilities of the GIS by providing the user with the functionality available in the visualization software.

The coupling between ArcGIS and DecisionSite uses a single module of code which utilises components and objects from both the GIS and the visualization software. Hence, the components from the two packages access the same data. This avoids the problems of data transfers that where a major drawback of modular approaches to linking software. In addition, this approach overcomes the limited flexibility of the GIS for visually exploring spatial data by coupling it to a visualization package. Still, it retains the GIS capabilities for input, output and display of geographically referenced data.

Providing users with visualization functionality from within the GIS is convenient for at least two reasons. On the one hand, it permits use and integration of spatial data from any source into the exploration process. On the other hand, if multiple or large databases have to be analysed, this approach relieves the user from unnecessary work and effort in manipulating data between packages, thus freeing their concentration and cognitive processing resources for addressing the data analysis task itself.

3.8 Challenges

Although frameworks for the design of visualizations that match user tasks have been proposed (Card *et al.* 1999; Casner 1991), many questions remain about the effectiveness of various visualizations on particular tasks. The focus of research in visualization design has been on the development of techniques and not until recently has there been an interest and coordinated effort to empirically evaluate visualizations.[2] Nevertheless, reports of user-centred design of information visualizations in the literature are uncommon and still in an early stage in the sense that most studies have been conducted in the past 5-7 years (Chen and Yu 2000). The widespread proliferation of the use of visualization techniques in many domains has therefore made evident the significance of empirical evaluation of the principles and methodologies these techniques embed (Chen and Czerwinski 2000).

The following chapters draw on a number of techniques that provide a structured approach to systems evaluation. More importantly, they propose a means of designing the tasks with which to assess geovisualization tools and suggest topics of investigation that would lead us to a better understanding of how these tools are used so we can design them accordingly. They also discuss a number of limitations and challenges for evaluating geovisualizations.

For instance, time constraints limit the amount of visualizations or functionality that can be presented to users and hence evaluated for some purpose. Evaluations usually need to be as short as possible not only because participants tend to have restricted time schedules but also because tiredness may affect their results. During the evaluation discussed in this chapter, all available functionality was demonstrated to participants. Users either spent a long time over and above the training period to understand their meaning or they restricted their interaction to a few selected graphic displays. Therefore, the functionality and visualizations tested in the following experiments had to be severely restricted.

Similarly, complex functionality can hardly be evaluated if many other elements have to be learned. For instance, visual variables such as the shape, rotation and colour of scatterplot glyphs in the environment evaluated can be customised. This can be used to encode more variables apart from those used on the two axes and hence explore associations between multiple attribute values. These visual variables were perceived by users as being complex to master and were time-consuming to learn and interpret. Hence, users preferred using a number of coordinated visualizations, or a display such as the PCP to explore multiple data attributes. This hardly means that the techniques described are not appropriate, but that learning all the functionality provided by the environment and the use of all the visualizations with a short time span was too demanding for participants.

This chapter has shown the relevance of usability evaluation for assessing computer-based tools as well as its utility for understanding users work processes. This is an issue that requires careful attention as the main difficulty for evaluating geovisualization tools thus far has been the difficulty in defining user tasks. Careful planning of usability evaluations, even when carried out with a small number of participants, can provide the researcher with a wealth of empirical information regarding the effectiveness of a tool or a technique for exploring data visually. The remainder of this dissertation therefore further investigates user tasks in a geovisualization setting.

[2] Some examples include: the special issue of the International Journal of Human Studies on empirical evaluation of information visualizations in 2000; the inclusion of cognitive and usability issues of geovisualization in the agenda of the ICA commission (1999) and the inclusion of cognitive issue of GIS in the NCGIA research agenda (1996).

4 Study 2: Geovisualization Task Taxonomy[1]

"Studies of people using graphics for the purpose of solving problems or performing information-processing tasks find little evidence to support the claim that graphics are superior to other types of presentations such as tables, lists, outlines, or text. This is in striking contrast to the popular belief that graphics are the preferred medium for information presentation. These studies do not reject our intuitions about the usefulness of graphics but rather find that the utility of any information presentation is a function of the **tasks** *that the presentation is being used to support" (Casner 1991: 112; original emphasis).*

4.1 Lessons from the first evaluation: About task definition

If we recall the broad objective of gaining insight and constructing knowledge about spatial data as the goal of using a geovisualization tool, then tasks to achieve this goal should reflect the exploratory nature of geovisualization. In other words, they should allow users some degree of freedom to investigate the data using the computer-based tool. Closed questions, though often simpler to answer and straightforward to analyse, did not appear to capture the purpose of the tools being evaluated. Therefore, the tasks in all experiments were designed so that users would decide on the criteria employed in the problem solving. Such criteria could include the variables they chose to solve a question, the values and value ranges considered, as well as the area explored.

In the first study, in order to assure that the tasks were exploratory, they were purposefully ill-defined. From the five tasks that were completed by all users (see Table 3.5), the initial three were open ended to allow participants to explore the dataset at will. The remaining two implied choosing candidate locations for different facilities, such as one for a grocery shop or a school. Although the open ended questions did instigate participants to investigate relations between different attributes using the graphic displays provided by the computer environment, the goal of these tasks was too broad. Hence, users spent very long trying to investigate as many relations as they possibly could in order to feel that they had addressed the task thoroughly. Alternatively, they would only explore the relation between a few attributes and manifest that, if they would have more time, they would explore all combinations of variables systematically. Participants were more comfortable and confident when addressing the last two questions, which not only demanded more familiar spatial operations, but also had a more specific goal. This allowed them to better assess or evaluate their own answers.

[1] This chapter is based on Tobón (2003).

Designing tasks that approximate the type of problems with which potential users may be confronted was a main concern in the second experiment. Tasks were also defined so as to allow users to decide the criteria for reaching a solution including the combination of graphic displays, variables and operations with which to obtain relevant information, as it is discussed in the next section. In addition, analysing user behaviour in the first study via the recorded on-screen interaction was very time consuming. This highlighted the importance of defining tasks that not only reflect what potential users might want to accomplish with the software, but also facilitating the recording and interpretation of user performance. Thus, the following section details the goals of the evaluation and its design. In §4.3 and §4.4, a methodology for defining tasks that are realistic and still allow more systematic and speedier analysis is discussed. Furthermore, §4.7 and §4.8 suggest that a combination of methods for recording not only user performance but also subjective views on user satisfaction with the software environment can be a very rich and informative means for detecting limitations of the tool, as well as understanding user needs and thought processes that need to be supported.

4.2 Usability Evaluation

4.2.1 Goals

The understanding of user tasks was at the centre of the experiment as stated in Table 4.1. Furthermore, deepening our understanding of how these tasks are accomplished and how they can be better supported was expected to shed light on what geovisualization tools should provide. Hence, the evaluation aimed at obtaining information about the ability of the environment to support the tasks defined in §4.3, as well as learning about task definition. Users' perceptions on ease of use and usefulness of the software was also of interest, as well as the flexibility of the environment to re-express and manipulate data and the extent to which it supported users' work processes.

Table 4.1 Goals, benefits and foreseen costs (Dumas and Redish 1999).

User group	Task goals	Benefits	Costs
GIS users of varying expertise	To increase the understanding of geovisualization tasks. To find whether the environment facilitated the exploration of spatial data. To find out about user's perceptions on ease of use and usefulness of the environment.	To extend GIS functionality by providing geovisualization tools while making use of the GIS capacity to store and display large amounts of spatial information.	Time taken to learn to use the system in terms of understanding how to use the visualization tools provided.

4.2.2 Participants

The environment discussed at the end of Chapter 3 was designed for GIS users as a prototype system that would seamlessly embed geovisualization functionality within a GIS. The profile considered for this evaluation mainly consisted of users with some GIS experience or who used GIS to support their work processes. This selection was of crucial importance because the aim of the evaluation was not to find information about usability problems of the GIS as that might constitute a whole research project in its own right. GIS have been shown to be difficult to use by novices because GIS require users to have or acquire considerable technical knowledge before they can operate the system (Traynor and Williams 1995). This presents major obstacles to novice users in terms of navigating an interface that embeds a language, world view and concepts that support the system's architecture rather than the user's work view (Tobón and Haklay 2003). Therefore, participants were expected to be familiar with the basic GIS functionality that is part of most commercial packages. This would allow them to operate the GIS as a component of the developed environment with very little training.

Nevertheless, it was of interest to recruit participants from a range of fields or application areas that could potentially use the tools evaluated in their own work. Therefore, participants included professionals, masters and doctorate students with a variety of interests but who shared the characteristic of using digital geographical data regularly. Table 4.2 shows that, for this study, 7 GIS professionals, 8 master's and 5 PhD students took part and that their interests or expertise ranged from location based services to issues of data accuracy and error propagation in GIS. They were between 24 and 36 years of age of both genders (8 females and 12 males). One more participant, listed at the end of Table 4.2, was recruited for a pilot test aimed at assessing both the tasks and the environment before the main test was conducted.

The aim of the evaluation was also not simply to perform a usability study of DecisionSite. The results from the first evaluation discussed in Chapter 3, had already identified a number of shortcomings with the initial interface and functionality, with some of the data views and with the modus operandi of some of the graphic devices. Functionality or graphic displays, which could not be amended and needed to be included in the environment were carefully explained during a pre-evaluation demonstration. This was to prevent users from having difficulties which sources could be traced to known usability problems with DecisionSite.

Apart from their profile, the number of participants is a crucial issue for any usability study and experimental design as its selection is closely related to the purpose of the evaluation. For the first study described in Chapter 3, the number of participants followed HCI and usability evaluation guidelines for giving sufficient evidence and depth to a usability study (Nielsen and Molich 1990; Virzi 1992). For this evaluation however, the purpose was to further our understanding of the tasks that should be supported by geovisualization tools. Hence, the evaluation was designed to provide statistical evidence of the significance of the factors used for the task definition. For this reason, the number of participants followed considerations on the desired power of the experiment as is discussed in detail in §4.4.

4.2.3 Invitation to Participate and Motivation

Potential users were invited to take part by letter (sent via email). The invitation broadly described the environment they would use (a seamless software environment which coupled GIS and a visualization tool), the technical knowledge they needed to have (basic knowledge of any GIS package), and the purpose of the test (facilitating the visual exploration of spatial data in a GIS environment). Details such as the duration of the evaluation, the place it would be held and, more crucially, the importance of each person attending to two sessions if they committed to participate, were also included.

Participants were not paid but rather explained the interest they could potentially have in the geovisualization system as a tool for data investigation. As with the first evaluation, this proved effective with the user sample invited to take part in the study as only one recruitment wave was necessary and the response rate was close to 50%. This rate is high considering the extent of the commitment required to take part in the evaluation. Participants who replied to the invitation and confirmed their acceptance were emailed a more detailed explanation of the evaluation and its purpose, as well as a description of the dataset and the variables they would explore during the sessions (see Appendix A for details on all correspondence to the users).

Table 4.2 Users' Description

User ID	Occupation	Research Interests and Expertise
1	MSc GISc	Retail. Weekly use of GIS (intermediate). European (British).
2	MSc GI	Web-based information systems and GIS. Weekly use of GIS (beginner). Asian (Chinese).
3	MSc GI	Water Distribution Systems. Daily use of GIS (intermediate). Indian.
4	GIS Technician	Digital data production and processing. Daily use of GIS (intermediate). European (British).
5	MSc Remote Sensing	Measuring biodiversity. Daily use of GIS (intermediate). European (Spanish).
6	MSc GI	Coastal Management. Weekly use of GIS (beginner). Asian (Japanese).
7	PhD Geography	Retail geography; geodemographic profiling and marketing. Daily use of GIS (expert). European (British).
8	PhD Geography	Spatial data accuracy; errors and error propagation in GIS Daily use of GIS (expert). European (British).
9	PhD Geography	Dynamic location strategies for emergency services. Daily use of GIS (expert). European (German).
10	PhD Geography	Location Based Services; urban way-finding. Daily use of GIS (intermediate). Asian (Chinese).
11	MSc GI	Transport and congestion. Weekly use of GIS (intermediate). African.
12	GIS Consultant	Logistics. Daily use of GIS (expert). Northern European (Swedish).
13	PhD student	Multi-agent simulation; geodemographics. Daily use of GIS (intermediate). South American (Brazilian).
14	GIS Consultant	Gravity modelling. Daily use of GIS (expert). European (Spanish).
15	GIS Consultant	Local government and planning. Weekly use of GIS (beginner). European (British).
16	MSc GIs	Projections in diagrams. Daily use of GIS (intermediate). European (British).
17	GIS Consultant	Utilities and telecommunications. Daily use of GIS (expert). Northern European (Norwegian).
18	IT manager	Cartographer. IT manager. Infrequent use of GIS (Beginner). European (British).
19	MSc Conservation	Ecological patches. Daily use of GIS (beginner). European (British).
20	GIS Consultant	Geographer. GIS Consultant. Daily use of GIS (expert). European (Irish).
Pilot	GIS Lecturer	Public Participation and Usability in GIS. Daily use (expert). Israeli.

4.2.4 Evaluation Design

The main goals of the experiment outlined in §4.2.1 were to be undertaken using the following evaluation methods. In terms of the first goal, obtaining information about the ability of the environment to support user tasks, the study included a usability evaluation aimed at measuring user performance in terms of the time it took participants to perform each task, as well as their success in carrying them out. These measures were to be analysed statistically as is discussed in §4.4. The second goal, that of obtaining information from the users about their perceptions on usefulness and ease of use of the software were measured using a Likert scale questionnaire, as discussed in §4.7. Interviews provided further information about users' work processes, their views on the supported functionality and limitations of the environment.

4.2.4.1 Evaluation Phases

The evaluation was divided into four phases each of which wase conducted by the researcher in a neutral venue. In an initial pre-test phase, participants were explained that the environment was designed to support data exploration. It was also clarified how to interpret each one of the graphic displays and were given a demonstration of the functionality they would be asked to use. In order to avoid biases, i.e. giving more information to some users, a 'script' with instructions was followed throughout this phase. Subjects also had a chance to discuss the dataset before commencing the evaluation. Issues such as the area of the UK they would explore (Bristol, England), the data format (vector data), the type of variables (numerical values as proportions of total households in an enumeration district) and their source (mostly from the 1991 UK Census of population) were made clear to all participants before the evaluation started.[2] Afterwards, participants were given a set of training or warm-up tasks to complete. The first study highlighted the importance of giving users an introductory demonstration as it cannot be expected that they will be able to grasp the ideas of map interactivity and manipulation just from looking at the map, data views and controls. However, a brief period of training is usually found to be sufficient for understanding the purpose of interactive geovisualization tools (Andrienko *et al.* 2002). Users were then encouraged to ask any questions about the system, as they were advised that no help in solving the tasks would be provided during the evaluation phase. Most of their questions regarded the use and purpose of the graphical displays and the meaning of variables in the dataset. Finally, users were told that they would be interviewed and given a questionnaire at the end of the two sessions regarding their perceptions about the software. Giving this

[2] For a detailed account of the dataset used in this experiment see Chapter 6.

information was of importance as "without a prior set towards thinking of the system with a view to evaluating it, people might not categoryze and remember their experiences" (Root and Draper 1983: 85) with the tool. This pre-test stage took an average of 30 minutes.

In the second or main testing phase, users completed a set of 16 tasks on their own. Usability metrics such as time taken to complete tasks and error rates were recorded by the researcher. This stage took users an average of one hour. In a third post-test phase, users where first asked to complete a questionnaire on perceived ease of use and usefulness of the environment. Then, the researcher conducted semi-structured interviews to obtain information about participants' perceptions of the environment and details about how they accomplished some of the tasks (Figure 4.1). The interviews were taped and were subsequently analysed (see Table 4.3 for a summary of resources). This stage took an average of 30 minutes. Results from phases 2 and 3 of the evaluation are discussed in greater detail in the sections to follow.

Table 4.3 Summary of Test Resources

User Time	1 hour 30 minutes per test
Hardware	1 PC connected to the internet 2 Monitors Time watch Dictaphone
Software	ArcGIS DecisionSite
Data files	Shapefiles with vector geometry of EDs for Bristol .dbf and Excel file with Census SAS data for the study area

Every user committed to completing two sessions within a seven day period. Therefore the second phase was completed twice by each participant. Tasks for each session were similar in type or structure but differed slightly in the wording and the data attributes analysed. The purpose of repeating the experiment was twofold. On the one hand, it would facilitate comparison of possible differences in user performance between the sessions to investigate learning effects in terms of understanding the environment and the data. On the other, this research design would make it possible to increase the precision of the experimental measures as discussed in the following section.

With respect to the data explored, the first evaluation discussed in Chapter 3 suggested that a limited number of variables should be provided for users to investigate during the experiments. The main reason for this was that users spent a great deal of time trying to

assimilate the meaning of the variables and the ways in which they were scaled and not all participants used the full range of variables provided. Providing an unnecessarily large set of variables placed a strain on the rather limited time that users had available to become used to the environment and complete all tasks. Therefore, in this second experiment the number of variables was reduced from 24 to 16, which was still thought to provide sufficient data to explore and potentially formulate hypotheses about various socio-economic characteristics of the study area. It is relevant to note that during this second evaluation, all users asked whether there were other variables or data sources available to address some of the tasks. This was seen as an endorsement of the software environment as it efficiently allowed users to explore the dataset provided and formulate ideas that were complex enough to include a good number of variables and still lead them to demand more data in order to refine the knowledge construction process. It thus appears that this environment was better suited than the first evaluated to explore spatial datasets. This contention is explored in more depth in the sections to follow.

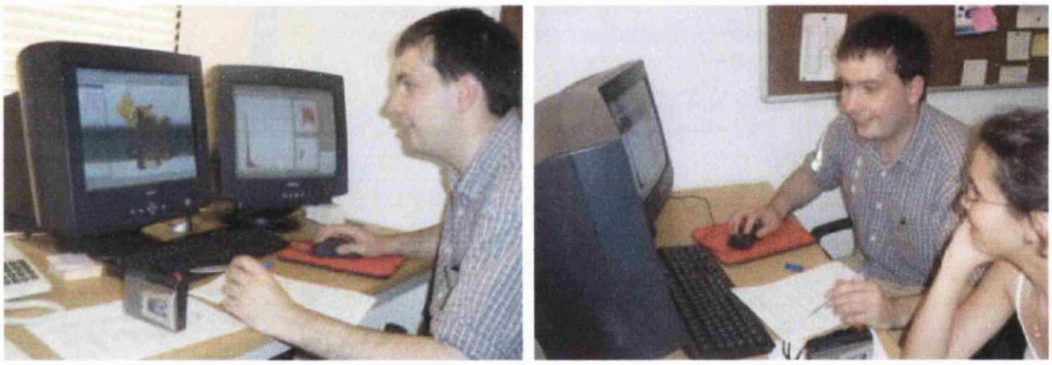

Figure 4.1 Experiment setting: Left) User completing the test. Right) User being interviewed by the observer.

4.3 Task design

Tasks can be of varying types that depend, for instance, on the number of attributes involved, the geographical extent of the area in which analysis is to take place, or the operation that needs to be performed in order to obtain some information from the data. In turn, the evaluation of an environment that allegedly supports one or many task types should take these various aspects into account. This would allow detecting whether an environment better supports some task types than others and to improve the system accordingly. In the evaluation discussed in Chapter 3, for example, it was found that the environment was better suited to identify anomalies in data attribute relations than to summarise spatial patterns. The main reason for this was the lack of dynamic linking between the map and all other data

views, which hindered the completion of tasks that demanded inferences on variables' spatial trends.

Factors that may intervene in the visual exploration of data must therefore be considered. Some work has been done outside the geovisualization field to appraise one important aspect, namely the visual operators or cognitive visual actions, necessary to accomplish a task's goal. Casner (1991) and Zhou and Feiner (1998, 1996) have worked on characterising both visual tasks and data types for the purpose of automating the design of connected visual displays. Casner (1991) in particular is concerned with characterising the task that a graphic is intended to support in order to automate the design of displays that are effective for their users and needs. Wehrend and Lewis (1990) explore a classification of visualization techniques that is independent of particular application domains but that provides a conceptual framework for describing techniques and graphics relevant to particular problems. In Schneiderman's (1998) data-structure-oriented taxonomy, information visualizations can be classified according to the data types they use.Keim (2002) further suggests a classification of visualization techniques that takes into account both the data type and the interaction and distortion techniques used in visualization. A similar approach for geovisualizations has been suggested by Crampton (2002). He characterises interactivity types and argues that they can be used and combined for designing digital environments.

Using a task analysis methodology, Knapp (1995) selected four visual operators from those considered by Wehrend and Lewis (1990) which she found to be used the most in work which utilised visual designs. The four cognitive visual actions are 'identifying', 'locating', 'comparing' and 'associating'. Following Wehrend (1990), they can be defined as follows. Identification refers to establishing a characteristic by which an individual object can be distinctively recognised. Identifying a scalar or height of a bar in a chart are good examples. Locating is determining the specific position of an object (such as a particular scalar in a bar chart) or a spatial feature. Comparing implies examining objects or features to note their similarities and differences. Associating objects or features is to link them in a relation. For instance, one may associate the income with particular areas in a region. Knapp (1995) describes these four operators as representing a logical progression in the visual process.

In addition to the visual operator, the task that the user attempts can have more components, as aforementioned. Take for instance the following example: 'Do areas with a high proportion of children tend to be located towards the periphery of the study region?' Apart from locating such areas, this task can also be broken down into the number of spatial

units explored or the extent of the area of interest and the number of variables or attributes that need to be interrogated to reach an answer. Therefore, the approach taken here was to break down tasks into three factors or components. Each of the three components or factors—visual operator, number of spatial units and number of attributes—can have more than one value or 'level'. Tasks can hence be defined as the various combinations of these three factors at their various levels, thus providing a task taxonomy.

In order to limit the number of tasks that users would attempt in the evaluation and to keep the experiment design as simple as possible, the levels for each factor were restricted a limited number. For instance, there are a number of visual operators that could be considered but only four—identify, locate, compare and associate—were used to define this factor as having four levels. Similarly, the number of spatial units could comprise any continuous quantity. However, the geographic extent factor was defined as having only two levels, namely one (or two) and many. Finally, the number of attributes was also restricted to having two possible levels: one (or two) and many. These same two factors (number of spatial units and attributes) also with two levels are considered by Edsall *et al.* (2001) for defining a task typology. In our case, the typology is thus formed by all the possible combinations of their levels with those of the visual operator. All combinations of the levels of these three factors define the 4*2*2 = 16 tasks that where considered in the evaluation (Figure 4.2).

Different combinations of factor levels define task types as well as their complexity, if the levels go from minimum to maximum, low to high, or correspond to a logical progression of operations. Figure 4.2 illustrates how the first factor, the visual cognitive operation performed, can range from identifying a feature or attribute value to associating spatial patterns of multiple variables in disjoint geographical areas (Knapp 1995). Similarly, the geographical extent of the area can include one (or two) spatial units or many. Finally, single or multiple attributes may be explored. Thus, the simplest type of task users attempted was identifying one attribute value in one particular area of Bristol. The task complexity increases as higher levels or the factors are combined. For instance, more complex tasks involved comparing the spatial pattern of two variables in a portion of the study area. Therefore, defining the tasks as combinations of factors allows investigating the effect of different visual operators, the size of the geographical area and the amount of attributes explored on user performance. This is highly appropriate as it allows the design of experimental studies that can be easily adapted to test different environments and their ability to support different tasks (Chen and Yu 2000). Defining tasks based on a visual taxonomy which is domain-

independent would be a useful guide for developing visual interfaces in general (Morse and Lewis 2000) as it would allow testing the visualization in isolation of its application.

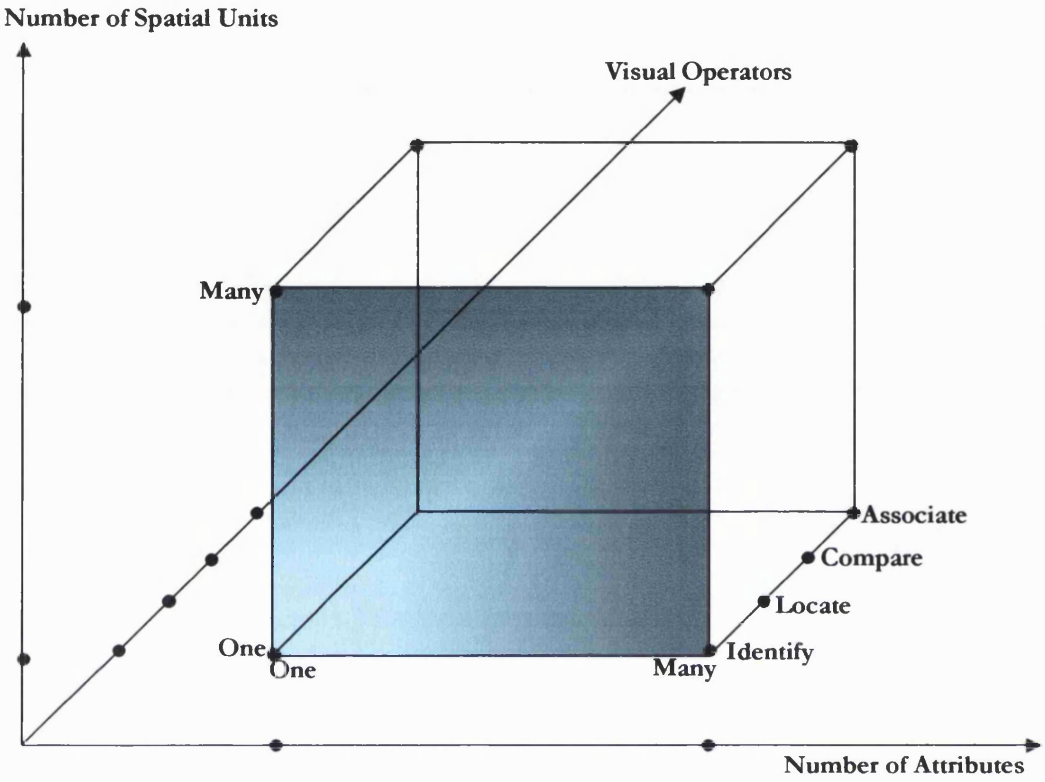

Figure 4.2 Framework for formulating user tasks as combinations of levels of three factors: The visual operation performed, the number of spatial units or the extent of the geographical area and the number of attributes investigated.

Table 4.4 provides some examples of the 16 tasks that users attempted.[3] The first column gives a reference number to the type of task. The numbers go from 1 to 16, the tally of tasks that each participant completed in any one session. The next three columns describe the task type or, in other words, the combination of different levels of the three factors aforementioned (Figure 4.2). For instance in task 1, users had to identify the value for one attribute in one area, i.e. they had to identify the lowest income score for an enumeration district (ED) in Bristol. The third column contains examples of the actual tasks that users attempted.

[3] The complete set of tasks can be found in Appendix A.

Table 4.4 Task Typology

	Visual Operator	Number of Spatial Units	Number of Attributes	Examples
1	Identify	1	1	Can you identify the area with the lowest household income?
2	Locate	1	1	
3	Compare	1	1	
4	Associate	1	1	
5	Identify	1	Many	
6	Locate	1	Many	Areas with a high proportion of working wives tend to have low proportions of children. Can you locate an area where this is not the case?
7	Compare	1	Many	
8	Associate	1	Many	
9	Identify	Many	Many	
10	Locate	Many	Many	
11	Compare	Many	Many	Is the relation between aggregate household income and being a male resident with a qualification different in the north west than in the south east of the study area?
12	Associate	Many	Many	
13	Identify	Many	1	
14	Locate	Many	1	
15	Compare	Many	1	
16	Associate	Many	1	You are an estate agent with a client interested in an area where she could locate a Gourmet Shop. Can you recommend 3 locations within the study area?

4.4 Experimental design

An experiment where we are interested in all or nearly all factor combinations is called a factorial design. It allows estimation of the effect of any of the factors upon the outcome or response variable (such as user time performance), as well as their interactions. In other words, it allows measurement of the average effect of each factor as well as differences of the effect at varying levels of the other factors. The tasks in the study were defined as combinations of factors (spatial extent, number of attributes and visual operators) and factor levels (four for the latter and two for the remaining factors). As such, this experimental design can assist in the understanding of effects such as that of the visual operation on the response variable. Thus, it allows to understand *inter alia* what is the effect of the visual operator on the time taken by subjects to complete a task and how does it vary for different amounts of variables or spatial units explored. Main effects of factors and the interactions

between them are discussed in §4.5 but we turn first to discuss the advantages and power of using experimental designs.

4.4.1 Advantages of Factorial Experiments

There are at least four reasons that account for the power of factorial experiments to estimate main effects and the interactions between them. On the one hand, all factors can be allowed to co-vary, obviating the need for a design in which a starting point for each factor is selected and then successively varied while holding the rest constant. The latter are frequently described as 'one-factor-at-a-time' experiments, which are used extensively in practice (Cox 1958). However, if the effect of one factor on the outcome of a task depends on the levels of other factors or, in other words, if there is an interaction between the factors, the 'one-factor-at-a-time' experimental design will produce poor results (see Montgomery (1997) for a detailed discussion). For instance, if the effect of the number of spatial units explored is found to be different when users 'compare' spatial phenomena rather than when they 'locate' geographical features, consideration may be given to what distinguishes tasks as defined by different visual operators. More generally, factorial experiments may shed light on how different factor level combinations affect the outcome of geovisualization tasks.

Factorial designs are also more efficient than one-factor-at-a-time experiments when studying the effect of two or more factors as fewer observations have to be made. An observation in this context is the outcome or response of a subject to a task, which makes this property highly desirable. A common problem of experiments with human subjects is that the results of lengthy tests tend to be affected by respondent fatigue (Keppel 1982). In addition, participants tend to have restrictive time schedules that do not allow them to take part in multiple or very long evaluations. This has a direct effect on the design of the evaluation or experiment and must be carefully considered.

Factorial designs allow many factors and levels of each factor to be defined. Furthermore, the gain of using factorial experiments increases with the number of factors considered, without compromising the experiment's power (Cox 1958). Although adding factors and levels may increase the complexity of the results, the task definition illustrated in Figure 4.2 can, in principle, be adapted to incorporate as many factors and levels as may be deemed necessary to generate a task taxonomy for evaluating geovisualization tools or concepts. In addition, if the experimental design is a simple one, such as that discussed here where all users complete all tasks and there are no missing values, the interpretation of the results is

straightforward. The substantial amount of literature concerning the application of this mature technique to a myriad of cases can provide further guidelines with which to design more complex experiments as necessary.

A fourth advantage of factorial experiments is that their precision can be increased by fine tuning the experimental design rather than taking more observations or including more participants in the evaluation. This is an important point because some aspects of the evaluation, such as the individual differences in participant backgrounds and levels of expertise, may lie outside the experimenter's control and could create possible sources of variability in the experiment. However, if the errors are not systematic, the estimate of each factor's effect will only differ from their true value by a random error (Cox 1958).[4] For instance, participants must be carefully selected to represent the potential users of the environment in order to avoid biases in subject performance arising out of inappropriate sample selection.

4.4.2 Standard Error and Standard Deviation

The size of the experiment's errors can be assessed using the standard error estimate, the magnitude of which provides an indication of the precision of the experiment. It is defined here as:

$$\text{Standard Error} = \text{Standard Deviation} * [\,2/(\text{number of users per task})]^{1/2} \qquad \textbf{(3.1)}$$

In other words, the standard error estimate refers to the random variation of an estimate from an experiment, while the standard deviation refers to the uncontrolled variation of individual observations. Note that in order to reduce the standard error by one half, the number of users that would need to take part in the evaluation would have to increase fourfold. Hence, reducing the experiment's standard deviation can increase precision more effectively. This property also works to our advantage since a common limitation of usability evaluations is finding participants that fulfil a profile of potential system users who have the time to participate (Dumas and Redish 1999). However, this requires that we know something about the standard deviation so that we can anticipate the precision that will be attained with a given number of participants.

[4] Random is used here in the statistical sense meaning that variations do not show any reproducible pattern.

In cases were previous data are available, we can make an estimate of the standard deviation and hence decide on the appropriate sample size to use in the experiment. Although the first evaluation examined a somewhat different piece of software—especially in terms of the dynamic interaction afforded between the graphical displays—the tasks that users performed were similar in nature. They were also exploratory and consisted of scenarios that participants had to solve. These similarities allow using participant time performances of the first evaluation for acquiring an idea of the standard deviation in this second experiment or an estimation of the study's precision. A worked example to illustrate this follows.

Table 4.5 User log of times (in seconds) during first evaluation

	Task 1	Task 2	Task 3	Task 4	Total (U)	Mean (U)
User 1	6.6	6.1	6.0	5.7	24.4	6.1
User 2	6.3	5.8	5.9	5.2	23.2	5.8
User 3	6.5	5.6	6.8	6.6	25.5	6.4
User 4	6.8	5.6	5.2	5.6	23.2	5.8
User 5	6.0	6.3	6.5	5.4	24.2	6.1
User 6	6.9	5.0	6.3	6.1	24.2	6.1
User 7	6.4	6.5	5.9	6.3	25.1	6.3
Total (T)	45.5	40.9	42.6	40.9	169.9	**6.1**
Mean (T)	6.5	5.8	6.1	5.8	**6.1**	

In order to estimate the standard deviation, we need to "measure that part of the variation that is not due to real [task] effects and that cannot be disregarded as systematic variation between the [users]" (Cox 1958: 28). Table 4.5 shows the natural logarithm of the time it took each user (in seconds) to complete each task during the first evaluation,[5] the log sum and log mean by user (Total (U) and Mean (U), respectively), and the log sum and log mean by task (Total (T) and Mean (T), respectively). To compute the standard deviation we first need to calculate the square of the residuals defined by the following equation:

$$\text{Residual}^2 = (\text{Observation - mean observation for the user - mean observation for the task + overall mean})^2 \quad \textbf{(3.2)}$$

Table 4.6 gives the squared residual for each observation. For instance, the residual for the first observation is $(6.6 - 6.5 - 6.1 + 6.1)^2 = 0.01$.

The standard deviation measures the magnitude of the residuals divided by the residual degrees of freedom, i.e. the number of users minus one multiplied by the number of tasks minus one:

[5] For an explanation on the data transformation into logs see §4.5.1.

$$\text{Std Dev} = \left[\left(\frac{1}{(users-1)(tasks-1)}\right)\sum_i residual_i^2\right]^{1/2} \quad \textbf{(3.3)}$$

In this case, the residual standard deviation is 0.48 and replacing this value in Equation (3.1) we obtain an estimate of the standard error:

$$\text{Std Error} = 0.48 * (2/7)^{1/2} = 0.26 \quad \textbf{(3.4)}$$

Table 4.6 Squared Residuals

	Task 1	Task 2	Task 3	Task 4
User 1	0.00	0.04	0.01	0.02
User 2	0.00	0.05	0.00	0.11
User 3	0.10	0.27	0.17	0.17
User 4	0.37	0.00	0.37	0.00
User 5	0.28	0.23	0.20	0.17
User 6	0.21	0.76	0.04	0.05
User 7	0.12	0.24	0.15	0.06

We can interpret this result by saying that the true value of the difference between two of the estimated means in Table 4.5 is in error in more than $\pm 1.96*0.26 = \pm 0.51$ (the confidence limits) with a probability of 95% (the confidence level). For example, if we take the difference of means for tasks 1 and 2 in Table 4.5 (6.5 - 5.8 = 0.7), the result will lie between [0.7 – 0.51, 0.7 + 0.51] or [0.19, 1.21]. By taking the exponential of these values to convert the range to seconds, [1.21, 3.35], we see that it is rather small.

It was anticipated that the precision of the measurements of the second experiment would be increased (i.e. the sample standard deviation would decrease) by inviting subjects to participate who could be more closely defined as potential users of the system, namely people with at least some basic knowledge of GIS functionality. Furthermore, to increase the precision of the experiment, the number of users as well as the number of observations for each task was increased. Twenty participants took part in the second experiment and completed two whole sessions of sixteen equivalent tasks each. Assuming the same standard deviation, if we replace the number of users by 20 in Equation (3.1) and conduct the experiment with one repetition, the standard error would be reduced to 0.11:

$$\text{Std Error} = 0.48 * (2/20*2)^{1/2} = 0.11 \quad \textbf{(3.5)}$$

By comparing the results from Equations (3.4) and (3.5) we obtain a standard ratio of 0.26/0.11 = 2.36 which indicates that together with increasing the number of subjects, the experimental design would more than double the precision of the first experiment. Consequently, before performing a follow-up experiment, we can establish an approximate confidence interval for the standard error at any level of probability. In this case, our new estimate is that the results of the experiment would be subject to a margin of error of $\pm 1.96*0.11 = \pm 0.22$ with a confidence level of 95%, which is considerably lower than the random variation estimated for the first experiment.

In cases where previous data are not available, the number of observations (or in this case the amount of participants performing a number of tasks) would have to be determined by the experimenter's judgment of the likely variability of the population's characteristics, perhaps with reference to ancillary sources of information. Alternatively, Cox (1958) suggests that a smaller pilot experiment is conducted first or that the study is conducted in stages so the observations from the first stage can be used to estimate the standard deviation and hence the appropriate sample size for the next stages. In our case, the number of participants in the first evaluation (Chapter 3) followed suggestions posed in the HCI literature. That evaluation in turn served to provide a broad measure of precision for this second set of experiments.

An estimate of the standard deviation as well as of main effects and interactions of the factors considered in the task definition can be obtained from most statistical packages which provide an analysis of variance (ANOVA). We turn to this technique next and discuss the models estimated for explaining the performance measures recorded. We also discuss some assumptions of ANOVA concerning the distribution of scores which must be met or at least understood in order to perform correct inferences from the results.

4.5 Analysis of Variance

A factorial experiment where all participants complete all tasks is known as a within-subjects or repeated measures design. Repeated measures ANOVA is the most common technique with which to study the changes in performance that result from successive experience with a task (Keppel 1987). Repeated measures or within-subjects ANOVA is used to estimate factors' main effects and interactions when all members of the sample (all test participants) are measured under different conditions (all tasks) for the same characteristic (time to complete a task). In this case, all 20 users finished a set of 16 tasks in each of the two sessions they all attended. Since the 16 tasks can be interpreted as the conditions, it is said

that the measurement of the dependent variable (time performance) is repeated. The time taken to complete each task was therefore recorded separately for each individual and each session. This provided two observations (one per experiment) per user and per task.

The main reason for using repeated measures ANOVA is that a great deal of variation between participants is to be expected because of differences in their individual characteristics, which causes the estimates from standard ANOVA to be large. We saw an example of this with the performance of some of the users in the study of Chapter 3 (Table 4.5). This occurs because individuals respond differently to the tasks for reasons such as changes in attention, motivation and learning (Keppel 1982). Repeated measures ANOVA is a powerful way of accounting for this heterogeneity because it takes into account the fact that there is more than one observation per subject, which can be used to supply information on the main effect of each factor. Furthermore, when sample members are difficult to recruit, this design is economical because each member is measured under all conditions while the power of the test is increased due to the reduction in the error variance (ITS 1997; Keppel 1982).

The model underlying the analysis of variance is specified by expressing the dependent variable, such as time performance, as an equation of the form:

$$y_{ijk} = \mu + \beta_1 NU_i + \beta_2 NA_j + \beta_3 VO_k + \beta_4 (NU * NA)_{ij} + \beta_5 (NU * VO)_{ik} + \beta_6 (NA * VO)_{jk} + \beta_7 (NU * NA * VO)_{ijk} + \varepsilon_{ijk} \quad \textbf{(3.6)}$$

where y_{ijk} is the response time for each task; μ is the overall mean of the population; NU_i corresponds to the number of spatial units factor (with 2 levels $i = 1,2$); NA_j is the number of attributes explored factor (at $j = 1,2$ levels); VO_k is the visual operator factor (with four levels $k = 1,2,3,4$); the composite terms denote the interactions between the factors; the β's are the corresponding parameters for each of the explaining variables and interaction terms; and ε_{ij} is the experimental error term. Using this model, we wish to estimate its parameters $(\mu, \beta_1, ..., \beta_n, \sigma^2)$ and test the hypothesis of whether different tasks, or combinations of the three factors at various levels, have different effects on user responses. This is equivalent to testing whether the task means are equal:

$$\begin{aligned} H_0 &: \mu_1 = \mu_2 = ... = \mu_n = 0 \\ H_a &: \mu_n \neq 0 \end{aligned} \quad \textbf{(3.7)}$$

for at least one n which is the number of all possible factor combinations. This is also equivalent to testing the hypothesis that all coefficients are zero:

$$\begin{aligned} H_0 &: \beta_1 = \beta_2 = \ldots = \beta_7 = 0 \\ H_a &: \beta_n \neq 0 \end{aligned} \tag{3.8}$$

ANOVA compares differences among task means (also called between-groups differences) and differences among subjects receiving the same task (or within-group differences). In the within-subjects ANOVA, differences among the conditions (between-groups differences) are due to the inconsistency with which participants perform under the various tasks (Keppel 1982). The within-groups source of variability is largely refined in the within-subjects ANOVA estimation by subtracting the between-groups differences to reflect the behaviour of each subject in the context of the experiment. This accounts for instance for the observation that, although participant behaviour may vary form task to task, subjects that are fast in performing one particular task will tend to be fast in all of them and participants who are slow will tend to be slow in all the tasks. By removing the between-subject differences from the within-group estimates of unsystematic variation, an increased gain of using a within-subjects design is achieved as the estimates take the individual differences into account (Keppel 1982). By partitioning the variability in the observations into these two component parts, the between and within-group variability, an F ratio may be calculated between them to assess whether they are significantly different.

4.5.1 A note on the assumptions of ANOVA and data transformations

We have discussed that the deviation of each score form the grand mean of the population contains two components in a within-subjects ANOVA design: a between- and a within-group deviation. The latter forms the basis of the estimate of the experimental error and ANOVA assumes that the errors are independent both between- and within-groups. Thus, the tasks in the experiments had to be administered in a random order that was different for each participant. This implies that users were randomly assigned to task sequences which were themselves organised in different random orders. This was necessary to avoid biases in the statistical analysis of the results that could arise as a consequence of presenting the tasks in a particular order.

Table 4.7 Normality of the response variable

Distribution moments	Time	LnTime
Variance	18119.13	0.56
Standard Deviation	134.61	0.75
Mean	190.79	5.00
Kustosis	3.92	0.24
Skewness	1.61	-0.50

ANOVA also assumes that the dependent variable is normally distributed, that the population variances are equal, and that the dependent variable is continuous (Keppel 1982). If the population from which a sample is drawn violate one or more of the ANOVA assumptions, the results of the analysis may be incorrect or misleading. For instance, if the assumption of normality is violated, ANOVA may not be the most powerful test available and our conclusions based on the F-statistic may be affected. Thus, the data analysis performed in the previous section as well as that in the reminder of this chapter and the next, uses the *log* of the time (in seconds) taken to complete a task as the response variable. This was necessary to attain Normality. Note the moments of the distributions of the response variable Time for this evaluation and that of its log, LnTime, in Table 4.7. As the skewness and especially the extreme kurtosis show, Time is not Normally distributed. This problem can be addressed either by using a nonparametric test or by employing a simple transformation such as taking the log.

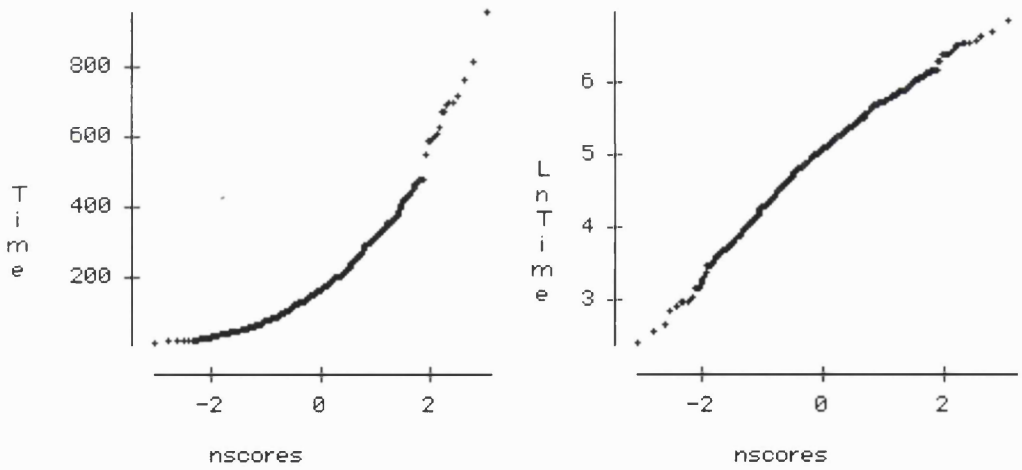

Figure 4.3 Normal QQ Plots: Left) Time. Right) LnTime.

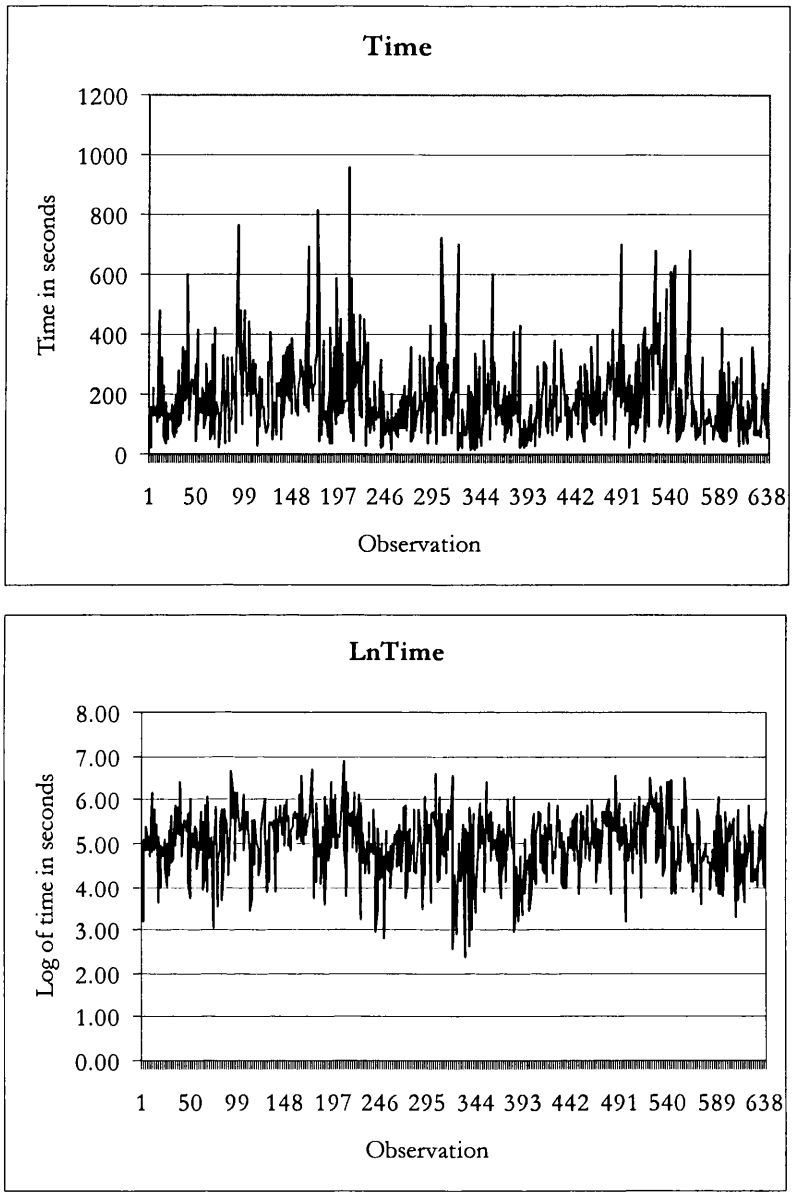

Figure 4.4 Line graph showing the volatility of the variance for: Top) Time; Bottom) LnTime.

Transformations, or a single function applied to each data value, are used to correct problems of non Normality or unequal variances. In our case, the logarithms of the sample values where taken to reduce skewness to the right in the response variable Time. Figure 4.3 shows the Normal QQ plots (§2.3.3.2) for Time and LnTime. Note that LnTime is slightly skewed but is not kurtotic. Transforming the values of Time also reduced or stabilised the variance as shown in Figure 4.4 and Table 4.7. This was also crucial as the third and final assumption of ANOVA is the homogeneity of the error variance. Unless a specific transformation suggested by scientific theory is available, the log or a power—such as the reciprocal square (-2), the reciprocal (-1), the reciprocal square root (-0.5), the square root

(0.5) or the square (2)—should normally be used (Tukey 1977). The main issue with transformations is that because they involve changing the metric or scale in which the data are analysed, care should be taken in the interpretation of the results.

4.5.2 Main Effects and Interactions

Table 4.8 presents a summary of the ANOVA results of the experiment.[6] Note that apart from the number of spatial units, attributes explored and cognitive visual action which the task involved, these results include a fourth factor, Day, which corresponds to the session in which the tasks were performed. This factor was added to the ANOVA model specified in Equation (3.6), along with its possible interactions with the other three factors, because it was of interest to see the effect of learning through practice with the software and the data on time performance.

4.5.3 Main Effects

The results for the main effects in Table 4.8 are analysed first. For the number of spatial units (NU) factor, the null hypothesis tested is that the size of the area does not have an effect on the time performance. However, since the F ratio for this hypothesis is large [$F(1,19) = 55.95$; $p < 0.00$], we must reject the null hypothesis and conclude that time performance does change with the size of the area explored in the population from which the sample was drawn. Similarly for the number of attributes (NA) factor, the null hypothesis of no effect of the number of attributes explored in the time taken by participants to complete a task was rejected, since the associated F ratio for this effect is large and significant [$F(1,19) = 86.96$; $p < 0.00$]. Also, the visual operation (VO) performed had a significant effect on time performance since [$F(3,57) = 28.59$; $p < 0.00$]. Finally, performance on the tasks also varied significantly during the two sessions—although less strongly so as the F statistic for the D factor shows [$F(1,19) = 6.65$; $p < 0.018$].

Figure 4.5 illustrates these main effects. The dotted line represents the grand mean of the response variable, i.e. the mean time to complete all of the tasks across all users. Thus, the main effect of each level of each factor is the deviation from the grand mean. For the number of spatial units explored it is clear that as the area to explore increases, the time to investigate it also augments. Table 4.9 indicates that the effects of the levels of this factor are significant. The same can be said of the number of attributes to explore factor. The visual operator shows a more interesting pattern. 'Identify' and 'locate' tasks had a similar

[6] All estimations were done in Minitab version 13 and SAS.

significant effect on users response, which were also below the overall mean (Figure 4.5), thus the negative values of their coefficients in Table 4.9. In addition, not only did the 'compare' and 'associate' tasks took significantly longer, but the compare tasks appear to be more complex given their deviation from the overall mean time. Finally, the last panel in Figure 4.5 shows the main effect of the session or day in which the tasks were completed. In other words, it indicates that with practice, users took significantly lower times to complete the tasks posed during the second session.

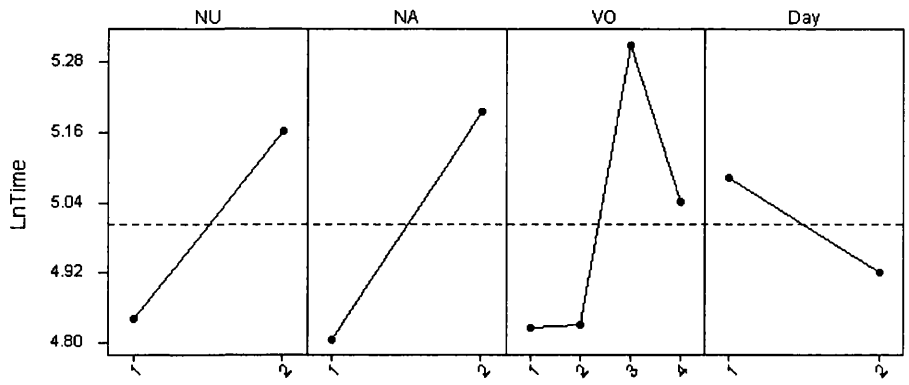

Figure 4.5 Main effects of the response means for each factor level. From left to right: Main effects of the number of spatial units explored; Main effect of the number of attributes investigated; Main effect of the visual operator; Main effect of the day in which the tasks were completed which reflects user practice with the environment.

4.5.4 Interactions

The presence of interactions between these four factors was also of interest and most were found to be significant (Table 4.8). For instance, a significant interaction was found for both NU*VO [$F(3,57) = 14.25$; $p < 0.00$] and NA*VO [$F(3,57) = 9.57$; $p < 0.00$] but not for NU*NA [$F(3,57) = 0.07$; $p < 0.80$]. From the latter result we can conclude that any differences in time performance levels did not reliably depend on the spatial extent of the area being explored in conjunction with the amount of variables investigated.[7] In addition, Table 4.9 further indicates that the NU*VO interaction was not significant when the task involved identifying one area ($p > 0.28$). Also, the NA*VO interaction was not significant when one attribute value was being identified or located ($p > 0.07$ and $p > 0.53$, respectfully).

[7] All 2-way interactions with the session or day (D) variable (D*NU [$F(1,19) = 16.76$; $p < 0.00$]; D*NA [$F(1,19) = 6.82$; $p < 0.017$]; D*VO [$F(1,19) = 19.92$; $p < 0.00$]) were significant or marginally so.

Before discussing this results further, it is necessary to note that one 3-way (NU*NA*VO) and the 4-way interaction (NU*NA*VO*D) were also found to be significant—[$F(3,57) = 12.49$; $p < 0.00$] for the 3-way and [$F(3,57) = 5.99$; $p < 0.00$] for the 4-way interaction. The 4-way interaction indicates that the 3-way one occurred differently during the two sessions. We now turn to analyse the 3-way interaction for each of the two sessions as this reveals some interesting issues regarding user behaviour.

Table 4.8 ANOVA results summary

	DF	Sum of Squares	Mean Square	F Ratio	P Value
Main Effects					
NU	1	16.5	16.5	56.0	0.00
Error (NU)	19	5.6	0.3		
NA	1	24.6	24.6	87.0	0.00
Error (NA)	19	5.4	0.3		
VO	3	25.2	8.4	28.6	0.00
Error (VO)	57	16.7	0.3		
D	1	4.2	4.2	6.7	0.02
Error (Day)	19	12.1	0.6		
2-way Interactions					
NU*NA	1	0.0	0.0	0.1	0.80
Error (NU*NA)	19	3.6	0.2		
NU*VO	3	15.1	5.0	14.3	0.00
Error (NU*VO)	57	20.1	0.4		
NA*VO	3	7.8	2.6	9.6	0.00
Error (NA*VO)	57	15.6	0.3		
D*NU	1	4.5	4.5	16.8	0.00
Error (D*NU)	19	5.1	0.3		
D*NA	1	2.1	2.1	6.8	0.2
Error (D*NA)	19	5.8	0.3		
D*VO	3	14.4	4.8	19.9	0.00
Error (D*VO)	57	13.7	0.2		
3-way Interactions					
NU*NA*VO	3	8.7	2.9	12.5	0.00
Error (NU*NA*VO)	57	13.2	0.2		
D*NU*NA	1	0.1	0.1	0.5	0.50
Error (D*NU*NA)	19	4.7	0.3		
D*NU*VO	3	0.8	0.3	0.9	0.44
Error (D*NU*VO)	57	16.8	0.3		
D*NA*VO	3	0.1	0.0	0.1	0.94
Error (D*NA*VO)	57	14.3	0.3		
4-way Interactions					
NU*NA*VO*D	3	4.6	1.5	6.0	0.00
Error (NU*NA*VO*D)	57	14.4	0.3		
Total	620	295.6			

Table 4.9 General Linear Model: LnTime versus NU, NA, VO, D

Term		Coefficient	S.E.	T	P Value
Constant		5.00	0.03	203.23	0.00
NU					
1		-0.16	0.03	-6.49	0.00
NA					
1		-0.19	0.03	-7.92	0.00
VO					
1		-0.18	0.04	-4.14	0.00
2		-0.17	0.04	-3.99	0.00
3		0.31	0.04	7.20	0.00
D					
1		0.08	0.03	3.28	0.00
NU*NA					
1	1	-0.01	0.03	-0.18	0.86
NU*VO					
1	1	-0.05	0.04	-1.08	0.28
1	2	-0.19	0.04	-4.42	0.00
1	3	0.24	0.04	5.54	0.00
NU*D					
1	1	-0.08	0.03	-3.38	0.00
NA*VO					
1	1	-0.08	0.04	-1.85	0.07
1	2	0.03	0.04	0.62	0.53
1	4	0.17	0.04	3.94	0.00
NA*D					
1	1	0.06	0.03	2.31	0.02
VO*D					
1	1	0.25	0.04	5.81	0.00
2	1	-0.15	0.04	-3.42	0.00
3	1	-0.07	0.04	-1.71	0.09

4.5.5 Graphical Analysis of Main Effects and Interactions

To interpret the 3-way significant interaction using a 2D display, it is necessary to fix one factor and interpret the interactions of the other two. Figure 4.6 to Figure 4.9 plot the mean time response for tasks with different level combinations of NU and NA, for each level of VO. For instance, each point in Figure 4.6 shows the mean response of all users in tasks where the visual operation is that of identifying an attribute value or feature. In order to understand how the interactions differed in each session, it is necessary to further split the data into the two days. Therefore, the graphs on the left of all figures plot the means of response times for all participants which were gathered during the first session and the graphs on the right those of the second session. In each graph, pairs of points that represent tasks where NU has the same level is joined by a dotted or solid line labelled 'one spatial unit' or 'many spatial units', respectively. This is to facilitate the comparison between the levels of NU. The y axis in all graphs indicates the (log of) time in seconds. The x axis codes the two levels of the NA factor, namely, how many attributes where explored. The

graphs can be interpreted as follows: Parallel lines indicate no interaction between NU and NA at a particular level of VO. Positive or negative slopes indicate a main effect of NA of the same sign on the response variable. Thus, horizontal lines indicate no average effect of NA. Coincident lines, on the contrary, are an indication of no main effects of NU.

For instance, Figure 4.6 (Left) shows the mean time performance averages for all users during the first session, where the visual operation of the tasks attempted was to identify an attribute value or feature. Hence, the dotted line shows the mean time averages for tasks 1 (first point) and 9 (second point) in Table 4.4, whereas the solid line represents the mean performance times for tasks 5 and 13, during the first session. The graph shows an interaction between NU and NA which we know is significant from the ANOVA analysis.

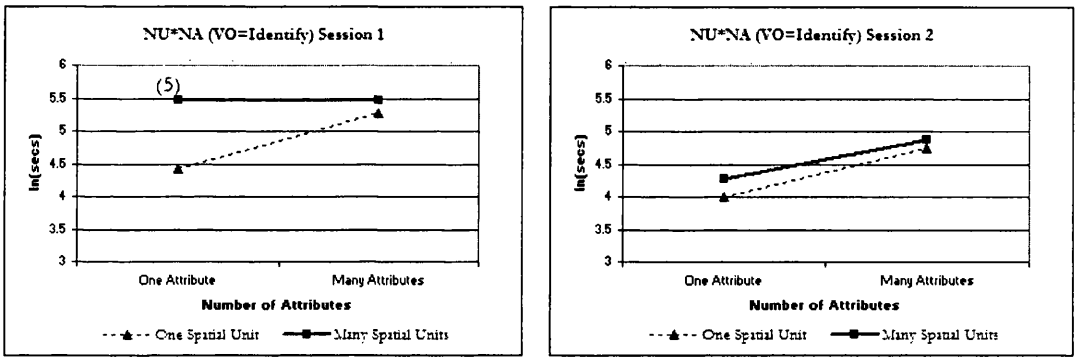

Figure 4.6 (Left) NU*AE interaction for session 1 when VO is fixed to its first level; (Right) NU*AE interaction for session 2 when VO is fixed to identify

Compare this graph to that in Figure 4.6 (Right) where the mean times represented are for the second session.[8] The first notable difference is the lower levels in the means for all tasks performed during the second session indicating that participants got faster with practice. Second, most of the lines in Figure 4.6 show a positive slope which indicates the (significant) main effect of the NA factor in time performance. The fact that the lines are almost coincident in the right graph is evidence that there was no significant main effect of NU during the second sessions when the task implied identifying a feature or value. Finally, the most interesting point in the graph is that corresponding to task 5 (Table 4.4) as it is causing a significant NU*NA interaction in the first session that is not present during the second session. As opposed to all other tasks where users had to identify a value or a feature, in task 5 (session 1) they had to perform the identification *relative* to a geographic entity. The same point is also causing the non-significant main effect of the NU factor when set to its higher level, as the nearly flat slope of the solid line in Figure 4.6 (Left) indicates.

Figure 4.7 should be interpreted similarly as above, although the visual operator in both graphs is now fixed to the second level 'locate'.[9] Note that in Figure 4.7 (Left) the lines are parallel and both have positive slopes. This is an indication of the main effects of both NA and NU, but that there is no interaction between them when VO is set to 'locate'. Also note the point in the right graph labelled as 6, which corresponds to that task in Table 4.4. Participants were required in task 6 (second session) to locate a value *relative* to a spatial entity as opposed to finding the location of a particular value or feature as was the case with all other tasks represented in Figure 4.7. Again that result is causing the NU*NA interaction in the second session which is not present in session 1. We will return to this issue shortly.

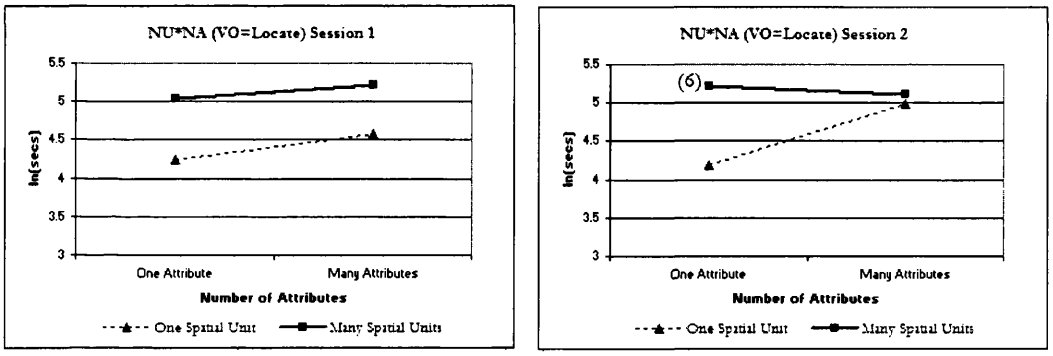

Figure 4.7 (Left) NU*NA interaction for session 1 when VO is fixed to its second level; (Right) NU*NA interaction for session 2 when VO is fixed to locate

Figure 4.8 shows the NU*NA interactions for sessions 1 and 2 when the visual operation was 'compare'.[10] In most tasks users had to compare values or patterns. The point labelled with (3) however, indicates the mean times for a task where users had to compare two features and answer why the expected *relation* between them was not present. It is easily seen in the graphs how these tasks are responsible for the NU*NA interaction in both sessions and for the non-significant main effect of NU in session 2.

[8] Points show tasks 1 and 9 for the dotted line and 5 and 13 for the solid line but for the second session.

[9] Points are the means for tasks 2 and 10 for the dotted line and 6 and 14 for the solid line. The graph to the left corresponds to tasks completed during the fist day while the graph to the right shows the means of those completed during the second day.

[10] Points in the graphs are the means for tasks 3 and 11 for the dotted line and 7 and 15 for the solid line in the graphs Figure 4.8. The graph to the left shows the means for the tasks in session 1. The graph to the right shows the means for the tasks in session 2.

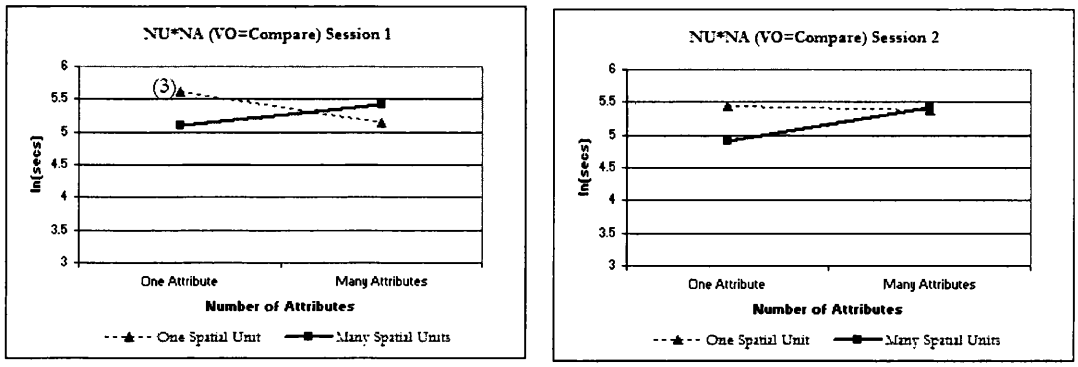

Figure 4.8 (Left) NU*NA interaction for session 1 when VO is fixed to its third level; (Right) NU*NA interaction for session 2 when VO is fixed to compare

The last pair of graphs in Figure 4.9 show the NU*NA interaction for both sessions for tasks were the visual operation was to 'associate' values or patterns. Note that points in the graphs are the means for tasks 4 and 12 for the dotted lines and 8 and 16 for the solid lines. Two main effects are evident in the left graph whereas only a main effect for NA is present in the graph to the right. All associate tasks implied doing operations relative to spatial entities which appears to be the reason for the absence of interactions.

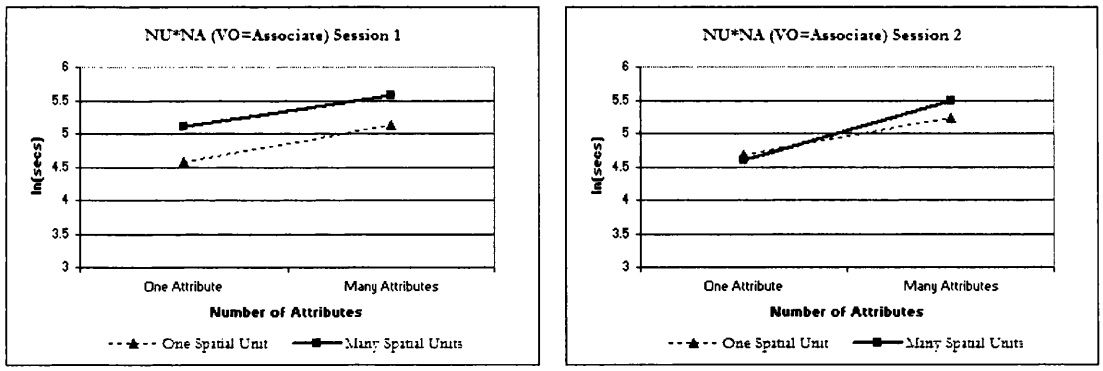

Figure 4.9 (Left) NU*NA interaction for session 1 when VO is fixed to its highest level; (Right) NU*NA interaction for session 2 when VO is fixed to associate

The aforementioned results suggest that there may be fundamental differences between the cognitive processes required to participate in the visualization of geographic spaces and those required for visualization of information spaces that are not spatial. The reason for this is not only the different visual operations that must be carried out but also the information processing tasks that need to be accomplished with geospatial data. In our case, relative comparisons of features in space are partly responsible for the interactions between the factors considered. If we recall that the visual operators used here were taken from research in Information Visualization, this study suggests that we need task taxonomies that are much

more refined to characterise geovisualization. This is an important new finding in the sense that it may imply that the processes that must be supported in geographic visualization are not only different but possibly more complex.

Figure 4.10 and Figure 4.11 show the NU*NA*VO interactions for each level of the spatial extent factor. Figure 4.10 shows the mean times for all tasks where the spatial extent was at the lower level (1 or 2 units), while Figure 4.11 shows the mean times for all tasks where the spatial extent was at the higher level (many units). Figure 4.10 shows two lines labelled 'One Attribute' and 'Many Attributes' which correspond to the number of attributes explored. The former represents the results for tasks 1-4 in Table 4.4 and the later the results for tasks 9-12. Hence, the first point on the 'One Attribute' line in Figure 4.10 is an average of the response times for task 1 in the two sessions and for all users.

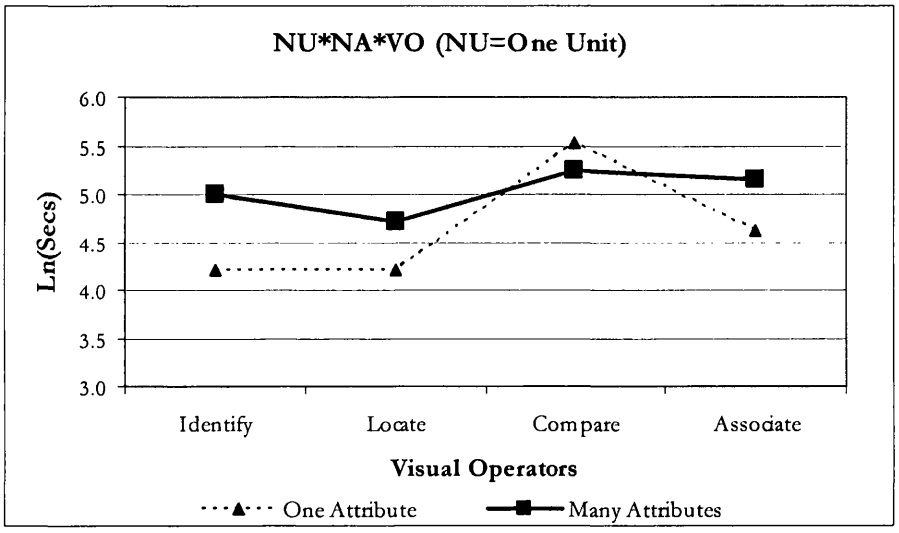

Figure 4.10 NU*NA*VO: Interactions of Spatial extent, Attribute extent and Visual Operators with Spatial extent set to the low level (results are averages of both sessions for all users).

Figure 4.10 shows that, in general, tasks where users had to investigate many attributes took longer. This is also the case when the number of spatial units was higher as shown by Figure 4.11. In particular, the tasks where the VO was to 'compare' or to 'associate' patterns or trends, took appreciably longer than those where users had to 'identify' or 'locate' values. The interaction of the three factors (NU, NA and VO) is also made evident by the shape of the lines. Note that if the lines were parallel, then there would not have been an indication of significant systematic differences between tasks implying different visual operators and numbers of attributes explored. These interactions, however, are slightly more complex to

interpret because they occur differently for each level of VO. We know this not only because the lines in the graphs are not straight but also from the analysis in the paragraphs immediately prior. In Figure 4.10, for instance, there is an unexpected high value of the 'compare one attribute value in two geographical locations' tasks (such as 3 in Table 4.4). The reason for this can be found in the way users approached task 3 during the first session. Most participants tried to solve the scenario by inspecting a large number of variables apart from the one inquired about in the question. Hence, the task resembled a multiple attribute exploration one and the average time for task 3 appears higher than that for the tasks where many attribute values had to be actually compared. This points to a difference in the strategies taken by subjects which are discussed in depth in the following chapter.

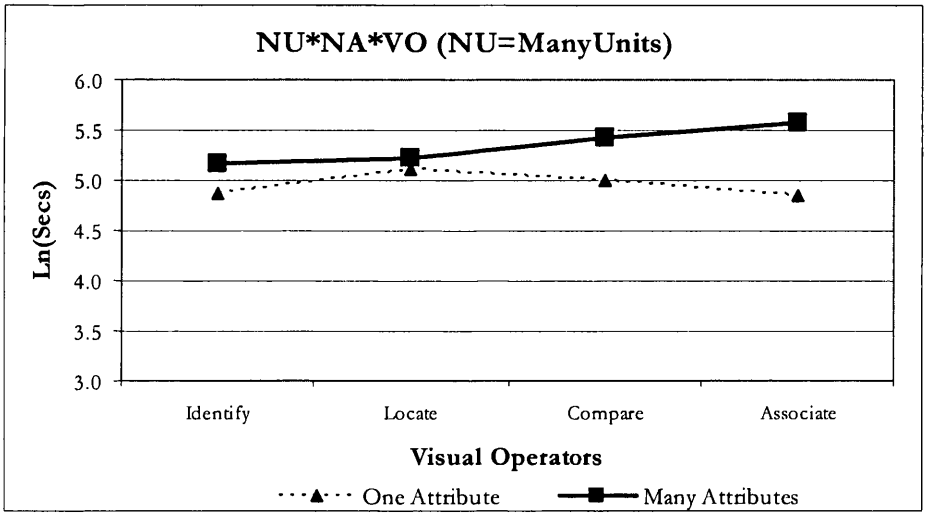

Figure 4.11 NU*NA*VO: Interactions of Spatial extent, Attribute extent and Visual Operators with Spatial extent set to many units (results are averages of both sessions for all users).

Figure 4.11 shows that exploring larger areas takes increasingly longer as the visual operation augments. However, the 'locate' tasks took appreciably longer when only one attribute was explored. This was explained previously when Figure 4.7 was analysed and arises because of the slightly different nature of task 6 in both sessions. Nevertheless, it is also of interest that there is little difference between the 'identify' and 'locate' tasks when one or many attributes are explored. In addition, there is a clear difference shown both Figure 4.10 and Figure 4.11, between the times for the 'identify' and 'locate' levels of the visual operators when compared to the 'compare' and 'associate' ones. This points at the fact that the 'compare' and 'associate' tasks are clearly of a different and higher level of complexity. Knapp (1995: 366) suggests that the 'identify' and 'locate' operators "work together to characterize and position singular or multiple objects relative to both a spatial and a temporal dimension", which this results may also imply. Furthermore, the figures above also show that, in general, tasks where

the VO was set to 'associate' take less time than those where the VO was set to 'compare'. According to Knapp (1995), however, the four VO levels are a logical progression of analysis of geographical data which would imply that in order to make an association, the previous steps of identifying, locating and comparing would need to be taken. This would entail that associating tasks should take longer but the results do not support this assertion. This issue will be explored in the following chapter, along with some of the interesting interactions found and discussed above.

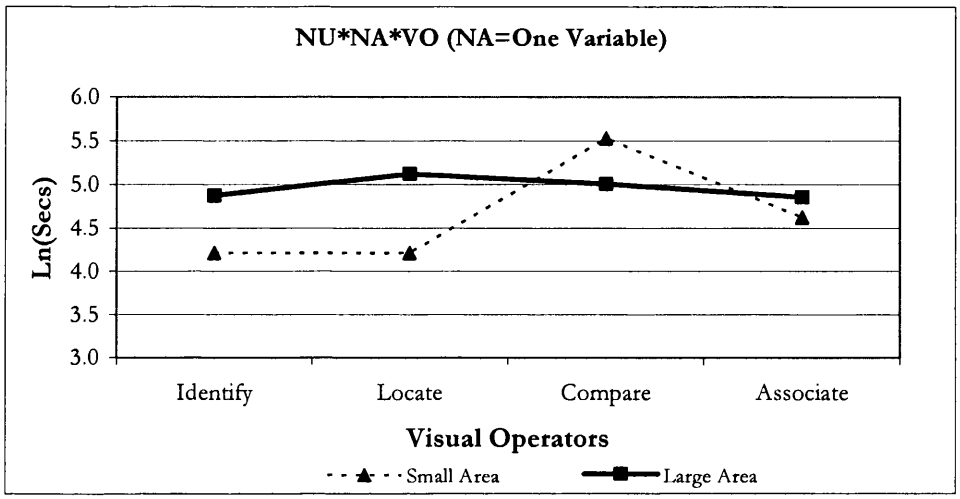

Figure 4.12 NU*NA*VO: Interactions of Spatial extent, Attribute extent and Visual Operators with Attribute extent set to its lower level (results are averages of both sessions for all users).

If we fix the number of attributes explored as opposed to the number of spatial units, we obtain the graphs in Figure 4.12 and Figure 4.13. Note that in Figure 4.12, the point on the 'Small Area' curve which surpasses its 'Large Area' counterpart corresponds to task 3. Once again it indicates that performing tasks relative to spatial features added to the cognitive processing of individuals and reflected on the time they took to complete tasks. Figure 4.13 shows an interesting result as it would appear that in tasks where more attributes were explored, there are relatively small differences in time to complete them, as the proximity and resemblance of the 'Small Area' and 'Large Area' lines suggest. However, there is still evidence of interaction between the factors, especially at the two higher levels of VO. Note that we have shown with ANOVA that differences in times are statistically significant. Therefore, the overall tendency of more complex tasks taking longer to complete can also be inferred from the slight positive slope of the two lines.

In summary, all main effects of the three factors considered were significant. Therefore, the size of the area to explore and the amount of variables considered have a positive effect on

the time taken to complete a task. Similarly, its outcome will depend on the visual operation performed. However, users performance is dependent on the combination of the levels of these three factors. For this reason, the factorial analysis used is appropriate and allows to inspect the interactions in detail. Of importance is the finding that tasks where the visual operation has to be performed relative to spatial features or attribute values, the task appers to be cognitively more demanding.

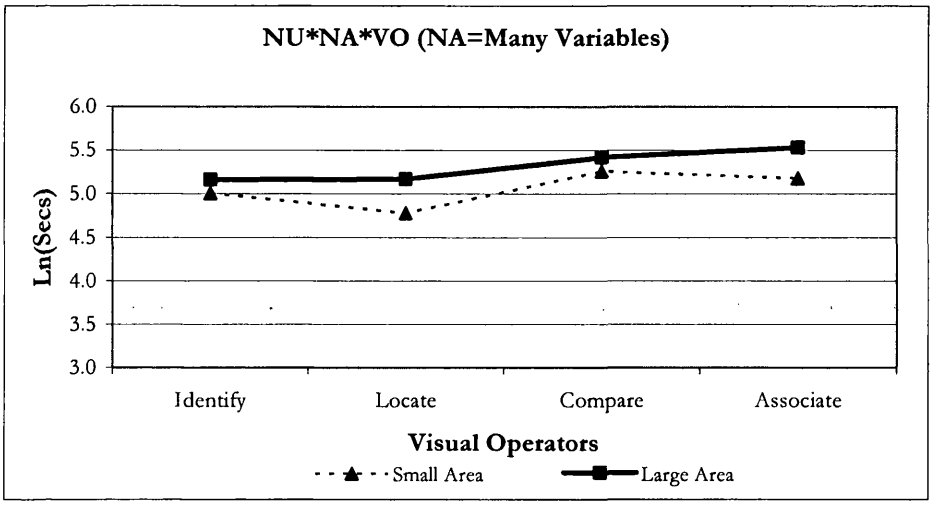

Figure 4.13 NU*NA*VO: Interactions of Spatial extent, Attribute extent and Visual Operators with Attribute extent set to many units (results are averages of both sessions for all users).

4.6 Binary Logistic Regression

The previous section provides a quantitative or numerical study of the performance measurement taken during the experiment. It therefore offers an analysis of causal relations between variables. However, it does not provide an understanding of the processes (Popper 1980) that the participants went through in order to arrive at an answer to the tasks posed. Given the observation that user satisfaction with software will affect the willingness to use it, analysing participants' experience with the tool was necessary to understand whether the environment could support the development of their work on particular tasks. In addition, this understanding may shed more light on their cognitive processes. Therefore, the analysis of the qualitative data gathered from the study as well as users' subjective appreciation of the experience is pursued in the following sections to complement the quantitative findings described before.

We start by analysing participants' success in attempting the study tasks. Being successful in a task was characterised as arriving at the correct conclusions given the criteria that users had

employed in order to respond to a task. The purpose of coding user success was to extend our understanding of their performance since not only the speed of completion is of interest but also the accuracy with which users can reach an answer. Therefore, a success binary variable was recorded indicating whether the task was completed successfully (1) or unsuccessfully (0). Logistic or logit regression can be used to analyse this type of variables and is typically given by either one of the following equivalent equations:

$$[p/(1-p)] = \exp(a + \beta x + \varepsilon) \tag{3.9}$$

$$\ln[p/(1-p)] = a + \beta x + \varepsilon \tag{3.10}$$

where p is the probability that an event occurs such as the user completing the task successfully; $1-p$ is the probability of finishing the task unsuccessfully or not finishing it; $p/(1-p)$ is the "odds" or the probability of the event divided by the probability of the non-event;[11] $\ln[p/(1-p)]$ is the "log odds" or "logit of p"; β is the vector of coefficients of the explanatory variable x or a vector of coefficients if x is a matrix of explaining variables; ε is the error term; and exp = 2.71828... is the base of the natural log.

Equation (3.10) makes it clear that the logistic regression model is a non-linear transformation of the linear regression used to constrain the estimated probabilities between 0 and 1, which are the only possible values that the dependent binary variable can take. This gives the logistic distribution an S-shape bounded by 0 and 1. The interpretation of β, however, must be done with care. Although β can be interpreted as the effect on the dependent variable of a unit increase in x in either Equations (3.9) or (3.10), since the dependent variable is a ratio, this effect depends on the relation between p, $p/(1-p)$ and $\ln[p/(1-p)]$. In other words, instead of the slope coefficient β being the rate of change in the dependent variable as x changes, β is interpreted as the rate of change in the "log odds" as x changes. Given that $p/(1-p)$ is a monotonically increasing function of p, $\ln[p/(1-p)]$ is a monotonically increasing function. Hence if p increases (decreases) $p/(1-p)$ and $\ln[p/(1-p)]$ also increase (decrease) (Retherford and Choe 1993). Similarly, if $\ln[p/(1-p)]$ increases (decreases), then $p/(1-p)$ and p also increase (decrease).

[11] For instance if p = 0.75 and (1-p) = 0.25, the odds of the event occurring is 0.75:0.25 or 3:1.

Fitting a logit model to the evaluation data, the log odds of completing a task successfully (logit p) was found to depend on the number of attributes explored, the visual operation performed and the day of the session. Table 4.10 shows the results where β is the estimated logit coefficient, S.E. is the standard error of the coefficient, significance gives the significance level of the coefficient (for instance, the coefficient on NA is significant at the .01 or 99% confidence level) and $\exp(\beta)$ or e^{β}, henceforth is the "odds ratio" of the individual coefficient. Hence, increasing the level of the NA factor, which means augmenting the number of attributes explored, has the effect of decreasing the log odds of being successful, $\ln[p/(1-p)]$, in -0.57 which, because of the monotonic property, also decreases $p/(1-p)$ and p or the odds and the probability of successfully completing a task. On the contrary, taking the test for the second time increased the odds of being successful by a factor of 0.53 which given the monotonic property implies that $p/(1-p)$ and p also increase.

Table 4.10 Logit Regression Results

	β	S.E.	DF	Significance	$\exp(\beta)$	95.0% C.I. for $\exp(\beta)$	
						Lower	Upper
NA(1)	-0.57	0.21	1	0.01	0.57	0.38	0.85
VO			3	0.01			
VO(1)	0.39	0.29	1	0.18	1.48	0.84	2.59
VO(2)	-0.55	0.27	1	0.04	0.58	0.34	0.98
VO(3)	0.18	0.28	1	0.54	1.19	0.68	2.10
D(1)	0.53	0.20	1	0.01	1.70	1.16	2.50
LnTime	0.28	0.05	1	0.00	1.33	1.21	1.46

To understand the magnitude of these effect take Equation (3.9) and, following Rutherford and Choe (1993), let:

$$\frac{p}{1-p} \equiv \Omega = \exp(a + \beta x) \tag{3.12}$$

Now assume an increase in x of one unit so that Equation (3.12) becomes:

$$\Omega' = \exp(a + \beta(x+1)) = \exp(a + \beta x + \beta) = \exp(a + \beta x)\exp(\beta) = \Omega \exp(\beta) \tag{3.13}$$

which can be written as:

$$\frac{\Omega'}{\Omega} = \exp(\beta) \tag{3.14}$$

Therefore, from Equation (3.13), a unit increase of x multiplies the odds by a factor of e^{β}, henceforth $\exp(\beta)$, which is called the "odds ratio" due to the result of Equation (3.14). Furthermore, taking the natural log of Equation (3.14) we obtain:

$$\ln\Omega' = \ln\Omega + \beta \tag{3.15}$$

and therefore, the coefficient β represents the *additive* effect of a unit change in x on the log odds of completing a task successfully and the odds ratio exp(β) represents the *multiplicative* effect of a unit change in x on the odds of completing a task successfully. Hence, the effect of increasing the level of the NA factor is *ceteris paribus* exp(-0.57) = 0.565, or decreasing the odds of completing a task in 39%. On the other hand attempting the tasks during the second day multiplied by exp(0.55) = 1.733 the odds of completing a task successfully or increasing the odds by 73.3%. The results for the visual operator are rather interesting. Note that increasing the level of this factor from 'identify' to 'locate' slightly increases the odds of completing a task successfully in exp(0.39) = 1.477 or 47.7%. Increasing the level of VO from 'locate' to 'compare' decreases the odds of completing a task successfully in exp(-0.55) = 0.577 and increasing once more the level of VO from 'compare' to 'associate' increases the odds of completing a task successfully by a factor of exp(0.18) = 1.197 or 19.7%. Note however that the parameters for the VO levels are not significant. Nevertheless, including all the independent variables was found to improve the model or significantly increase the ability to predict participants' success ($\chi^2_6 = 258.8, p < 0.001$).

In summary, the findings of the logistic analysis are as follows. First, practice with the environment increases the likelihood of success. This was expected due to the known importance of training in mastering graphical displays (Sutcliffe *et al.* 2000; Burin *et al.* 2000; Kyllonen *et al.* 1984a). Second, exploring more attributes decreased the likelihood of success. On the contrary, the spatial extent of the area being explored had no effect on the likelihood of completing a task. In addition, participants where more likely to be successful in locating features or values than in identifying them. They were also more likely to be successful in making associations than comparisons. These two findings are in direct contrast to Knapp (1995) assertion about the four operators being a logical progression in the processing of

information. Nevertheless, compare and associate operations had a lower likelihood of being completed successfully than identify and locate operations. This would indicate that the former two are of a higher level of difficulty.

4.7 Questionnaires

The usability of a computer system is said to be defined according to the effectiveness, efficiency and satisfaction that its use entails (Preece *et al.* 1994). The efficiency of the environment developed was measured and analysed in §4.5 in terms of user time performance and its effectiveness was assessed in terms of user success rates as discussed §4.6. It was of interest, however, to learn about user perceptions about the environment, namely whether participants found it enabled them to explore spatial data and whether they found it useful and easy to use. These aspects were inspected using a questionnaire originally proposed by Davies (1989) to characterise and evaluate information systems from a user perspective. Hence, the purpose of the questionnaire was to function as a tool with which to extract this information from the users, as well as to record and collect it systematically.

From the definition of the word 'useful'—"capable of being used advantageously, beneficially, or for several purposes" (Collins Dictionary 1987: 1106)—it follows that perceived usefulness is the degree to which a person believes there exists a benefit in using a particular tool to enhance her work process (Davis 1989). Likewise, the word 'ease' means "freedom from discomfort, worry or anxiety; lack of difficulty, labour or anxiety" (Collins Dictionary 1987: 312) from which it follows that perceived ease of use is the degree to which a person believes that using a particular tool would be free of effort or stress (Davis 1989). Since usefulness and ease of use are fundamental determinants of potential user acceptance and usage of information systems (Preece *et al.* 1994; Davis 1989) then, if the tool is to be used by a set of people larger than that made by its developers, obtaining information about these aspects and how to improve them is of enormous importance.

Nevertheless, there is no objective way of measuring perceptions in general or perceptions of usefulness and ease of use in particular. Davis' questionnaire provides a subjective measure using a so-called Likert or summative scale. It is an example of a System Usability Scale (SUS) which gives a "global view of subjective assessments of usability" (Brooke 1996: 1991). Likert scale questionnaires are composed of a set of statements or items to which subjects provide answers according to their degree of agreement. For instance, in a seven point scale like that used in the questionnaire for this study, participants may: strongly agree, agree, possibly agree, not know whether they do or do not agree, possibly disagree, disagree,

or strongly disagree with each statement. Hence, for each item, respondents were instructed to choose one of these seven possible answers or anchors. If they felt they could not respond to a particular item, they marked the centre point of the scale ('don't know'). Possible responses could be limited to five anchors or expanded to more than seven if there might be a reason for wanting more gradations of detail in the answers. The number of anchors could also be a factor of two as opposed to being an odd number if there was reason for eliminating the middle 'don't know' answer and force users to either agree or disagree with a statement. Table 4.11 shows the 12 questions presented to the 20 participants in the study.

Table 4.11 Questionnaire's Items and Corresponding Constructs being Measured

	Perceived Usefulness[12]	**Construct Measured**
1	Using the software would help me investigate spatial data more quickly.	Speed
2	Using the software would enable me to improve my exploration of spatial data.	Performance
3	Using the software would increase my productivity in investigating spatial data.	Productivity
4	Using the software would help me find more information from the exploration of spatial data.	Effectiveness
5	Using the software would make it easier to investigate spatial data.	Facilitate task
6	I found the software useful for investigating spatial data.	Useful
	Perceived Ease of use[13]	**Construct Measured**
7	Learning to operate the software was easy for me.	Easy to learn
8	I found it easy to get the software to do what I wanted it to do.	Reliability
9	My interaction with the software was clear and understandable.	Clarity
10	It was easy to become skilful using the software.	Skill
11	It was easy to remember how to use the software.	Remember
12	I found the software easy to use.	Ease of use
	GIS Usage[14]	
1	Do you have regular access to a computer?	
2	How often do you use a computer?	
3	How often do you use a GIS?	
4	As a GIS user, would you describe yourself as a beginner, intermediate or expert user?	

Each response to an item was given a score from 1 to 7. The main attribute of Likert scales is that all the scores for each item can be added so that the perceptions about the software

[12] The seven possible answers to this question were: strongly agree, agree, slightly agree, don't know, slightly disagree, disagree, and strongly disagree.
[13] The seven possible answers to this question were: strongly agree, agree, slightly agree, don't know, slightly disagree, disagree, and strongly disagree.
[14] These questions were only used to characterize participants but where not part of the Likert scale.

system can be quantified. To achieve this, it must be proven that each item of the questionnaire does in fact make a statement about the construct it attempts to measure (such as effectiveness or ease of use) or provide a manifestation of the degree to which the questionnaire actually measures them (Kirakowski 1994). Similarly, the reliability of the questionnaire as a measurement instrument must be determined. This is the degree of accuracy of the questionnaire or the extent to which the same answers would be obtained if it were administered again to the same person under the same conditions (Laitenberger and Dreyer 1998). Davis' experiments have been replicated by others who have confirmed that the questionnaire is a reliable and valid measurement instrument of user perception of ease of use and usefulness (Dennis *et al.* 1992; Laitenberger and Dreyer 1998). In addition, Laitenberger and Dreyer (1998) successfully applied the rating scale to a domain different to that for which it was designed. This is because information obtained from a questionnaire tends to be independent of the software system, tasks and users who answer it if it is reliable and discriminates between the constructs (Kirakowski 1994; DeVellis 1991). Thus, they permit a general assessment of usability or bear cross-system comparison if that was an issue of interest (Brooke 1996).

For these reasons, Davis' questionnaire was adopted to obtain information about participants subjective appreciations with which to complement the statistical results discussed previously. In order to reflect the environment being evaluated, however, some minor changes to the wording of the items in the original questionnaire were made. However, if such modifications do not change the overall meaning of the construct that is being measured, then the properties of the questionnaire should remain unchanged.[15] In addition, the questionnaire included factual questions to ascertain user frequency of GIS use and expertise with GIS usage. They however were not part of the summative scale but where used to characterise users. The items of the questionnaire administered are shown in Table 4.11 along with the construct being measured by each statement. The questions focus on people's attitudes to the software system and it is an example of the so-called satisfaction questionnaires used in HCI to find information about users attitudes concerning their perceptions of efficiency and helpfulness of a system, its learnability and ease to control (Kirakowski 1994).

Each one of the first 12 questions in Table 4.11 aimed at measuring an aspect of the usefulness or the ease of use of the system. This included the perceived speed with which

[15] Personal communication with faculty and staff at the Departments of Computer Science at UCL and the UCL Interaction Centre (UCLIC).

people could arrive to an answer, as well as the effectiveness and the ease with which the process was done. Although Table 4.11 presents all the questions in order according to the usability aspect being measured, they were given to participants in a random order. The main reason for randomizing the questions was to avoid some response pattern especially between questions that are similar.

4.7.1 Ratings for perceived usefulness

Given the properties of the Likert scale, all responses for each item in the questionnaire were summed to obtain a rating for each construct being measured (Table 4.12). The ratings of the summative results for the perceived usefulness of the tool range between 34 and 41. Given that the maximum rating is 42, we can conclude that most of the subjects considered the tool to be useful. Table 4.12 shows the mean, minimum and maximum scores given to each item. Out of a maximum score of 7, five out of six aspects of the environment's usefulness received an average score above 6. This implies that users agreed or agreed strongly with the statements provided. The highest average score (6.4) was given to the statements inquiring about the effectiveness of the environment and whether users thought it increased their performance. The lowest (6.0) was given to that inquiring about whether users perceived the tool made them more productive in exploring the data. Although to a lesser degree than the other aspects, users still believed the tool had a positive impact on their productivity. The fact that the environment was very flexible and permitted the re-expression of data in real time was reflected in user assertions that the environment was not only useful but also facilitated and made the completion of their tasks quicker possibly making their work process more effective.

4.7.2 Ratings for perceived ease of use

The ratings of the summative results for the ease of use of the tool range between 28 and 38. Given that the maximum rating is 42, we can conclude that most of the subjects considered the tool to be easy to use although this aspect was marked slightly lower than the tool's perceived usefulness. Somewhat surprisingly, the item with the highest score was the ease of remembering how to use the tool (6.1 on average). It was expected that the complexity of the environment would hinder this score. A possible explanation for this is the fact that the tool is composed by a GIS the main functions of which were known to participants either from using the same or other commercial packages. This might have had a positive impact on the recall score. On the other hand, users completed the two sessions within a seven day period. This might have been too short a span of time to evaluate whether users could easily remember the tool after a period of disuse. However, users rated the ease of becoming skilful

in operating the environment the lowest (5.4 on average). In the interviews most users expressed that they thought they would be able to do more sophisticated analysis once they had more time to master all the functionality provided. Nevertheless, on average, users agreed that the tool was easy to use (5.9) and that it was easy to learn (5.8).

Table 4.12 Questionnaire Results

	Mean	Min	Max
Usefulness	6.2	4	7
Quick	6.3	4	7
Performance	6.4	6	7
Productivity	6.0	4	7
Effectiveness	6.4	6	7
Facilitate	6.1	5	7
Useful	6.4	6	7
Ease of use	5.7	3	7
Easy to learn	5.8	4	7
Clear	5.7	4	7
Controllable	5.6	3	7
Skilful	5.4	4	7
Ease to remember	6.1	5	7
Easy to use	5.9	5	7

4.8 Interviews

Qualitative methods of data collection include non-structured procedures from observation to interviewing, self reports or written narratives. Their focus is within the situation for which they are used to elicit information and on individuals' perceptions. Therefore, the information collected is not established *a priori* as the gathering of information is open-ended in nature. Qualitative analysis typically does not have a numeric translation of data beyond deriving percentages or frequencies. The focus adopted here is therefore descriptive (Ericsson and Simon 1984) and for shedding light onto the meaning of the numerical results previously obtained.

Semi-structured interviews are one of the most frequently used qualitative methods. They allow a focused, conversational, two-way communication while still being rather open. Relevant topics were initially identified and some questions prepared but these tend to be very general and only serve the purpose of providing a framework for the interview. Therefore, many of the questions were improvised during the interview, allowing both the interviewer and the person being interviewed the flexibility to probe for details or discuss issues of interest to individuals. Thus, in semi-structured interviews the relationship between the identified topics and the issues confronted by each subject become the basis for more

specific questions which do not need to be prepared in advance.[16] Consequently, every interview is slightly different as they try to capture as much as possible the subject's thinking about a particular topic or a practical task (del Barrio 1999). Although interviews may allow deepening and clarifying subjects' conceptions or particular actions while interacting with the software, they have at least two limitations. First, they can be time consuming not only to carry out but also to analyse. Second, some degree of training is needed in order to avoid leading the subject into giving an answer or, on the contrary, not asking enough questions for fear of leading a subject. However, it is an adequate tool to capture how a person thinks of a particular domain.

Table 4.13 summarises some of the comments users made regarding the environment and gives the frequency with which they talked about nine of its positive aspects. Most users commented on how visualising data in multiple representations aided their understanding of the dataset because it made patterns and relations more obvious and immediately available. They also highly appreciated the easy selection of records that was possible in all the views, especially in DecisionSite. In addition, most users commented on the ability of the environment to facilitate focusing on subsets of data and reducing visual clutter by using a graphical device such as the query sliders. Furthermore, the speed of this tool was seen as an advantage and a main reason for making the exploration effective. One user even double-checked most of his answers because he felt they were too easy and quick to obtain for them to be correct. All participants thought of possible applications of the environment to their line of work, especially for comparing scenarios and distilling information from datasets.

All users expressed that accomplishing the same tasks using only a GIS would either take too long or be too difficult to be a practical means of obtaining the same information. Moreover, all participants found the experience enjoyable. They believed the tool was fast to learn and that they improved greatly in their analysis as they had more time to spend using the environment. They also perceived, although less strongly, that the tool was easy to learn, control and master. Nevertheless, at least one user—the one with less GIS expertise—thought the tool was very complex. Still, all users were especially interested in the flexibility of the visual displays, the speed of the environment, and the easy selection of data that could be seen and inspected in any of the visualizations.

[16] Appendix A shows an outline of the topics that were discussed in the interviews.

Table 4.13 User's Perceptions

Advantages	Frequency	User Comments
User satisfaction	20/20	*"I'm used to querying data from tables in a GIS so I enjoyed being able to see the information I was interested in, on thescreen...It made it more intuitive for me" (User 20).* *"I can't believe I was doing this for an hour and a half! I was really having fun!" (User 6).*
Speed of analysis	19/20	*"Things happen in real time!" (User 12).* *"Some of the analysis could have been done using only the GIS but it would have been torturously slow!" (User 20)*
User performance	19/20	*"In the second session, I was definitely quicker and I think I was more confident and much more comfortable with the data and with the system" (User 20).*
Environment's effectiveness	19/20	*"The environment is very useful for seeing spatial distribution of attributes, pick out trends and outliers which can be very difficult to do using only a map" (User 19).* *"Using this environment is much better than looking for attribute values in tables" (User 5).*
Environment's usefulness	20/20	*"It's very useful to see all the data all at once" (User 13).*
Ease of learning the product	18/20	*"It was very fast to learn" (User 13).*
Ease of understanding the environment	15/20	*"In the first day, there was a lot to take in" (User 19).* *"The idea that you can perform something that is actually quite complex in a simple way was the most pleasing aspect of [the environment]" (User 20).*
Ease to remember	18/20	*"During the first session everything was new so I was trying to answer the questions and learn the environment at the same time. The second time around, I could simply go to the point" (User 13)*
Ease of using the product	18/20	*"Once I had a go at it, I found it reasonably easy to explore the data" (User 19).* *The environment is really easy to use" (User 10)*

4.9 Problems and Possibilities for Further Development

A number of limitations of the system became evident during the second set of experiments. Most subjects comments about shortcomings related to the limited facility of the system for comparing selections or scenarios (Table 4.14). One possible way of addressing this has been implemented in a number of software packages that allow the user to temporarily choose or select data objects of interest with different colours. DataDesk (DataDescription Inc., Ithaca)

provides a good example of this functionality (see Table 2.1 for a description of the system). This software not only supports brushing and linking, but also the colour coding of selections made by the user (Figure 4.14). Unfortunately, nor ArcGIS or DecisionSite allow the storing of multiple selections which would facilitate the development of this feature. Hence it was not possible to assign different colours to different data subsets as in Figure 4.14. This was therefore a major limitation of the approach of linking the two pieces of software that had no simple solution apart from low-level programming.

Table 4.14 Users' frequent problems

Problems	Frequency	User Comments
Highlighting	2/20	*"Highlighting doesn't stay up so it is difficult to see" (User 4).*
Multiple selections with different colours	7/20	*"Selection would have been facilitated if you could use colours to compare the selections" (User 5).*
Easy de-selection	6/20	*"If you're exploring options and you find you have selected too many things, it is difficult to undo" (User 4).*
Roll-back	10/20	*"Selection would have also been easier if you could save and compare your current selection to previous queries" (User 5).*
Other data views	2/20	*"Are there any other graphics available?" (User 1).*

An alternative possibility to this problem would be to allow users to store the query that produces a particular subset selection. This is an interesting approach as it would allow users to roll-back to previous selections of interest in order to compare scenarios. This possibility should be further explored because it would resemble the roll-back metaphor that is now available in most desktop packages and which takes the user to a previous state of their working environment, file or browser. Its wide availability makes us anticipate that users may be able to use it to their advantage as there is a high likelihood that they would be familiar with similar software behaviour. In order to implement the roll-back functionality in a geovisualization environment, a structured query containing the data selection would need to be stored either automatically or on user demand. Users could even assign names to the most satisfying scenarios and have them listed in a window for easy display, comparison or roll-back. In addition, the data subset containing records that may have been filtered out from the views would also need to be stored in a structured query. This would require minimum memory space and programming in our case, as the same procedures and routines that have already been developed would be used and only the SQL would be stored and passed to the appropriate routines.

A means of easily comparing selections either by facilitating different colour coding or selection roll-back is seen as necessary as most users commented that they needed to compare their answers at different stages of their analysis. In addition, the results in §4.5 showed that comparison tasks took longer to accomplish and §4.6 suggested that they had a reduced likelihood of being completed successfully. It is possible that the lack of a means to compare selections contributed to these results. Without it, the only way in which users could accomplish the comparison of scenarios was to make a selection, remember it, make another selection and either compare it mentally or take notes and repeat this process a number of times until participants were satisfied that they had reached the desired appraisal of the different scenarios. Two users tried to create layers of information in the GIS that contained only the desired selection and used them for comparison with the original map. However, they only attempted this once or twice because they thought it took too long. In the experiment debriefing, two other users contended that they thought about this solution but that they did not attempt it because they thought they did not have enough time or it would become too cumbersome an operation. Easy comparison of selections or data subsets would therefore likely improve user performance and satisfaction, and would also greatly facilitate potential users work.

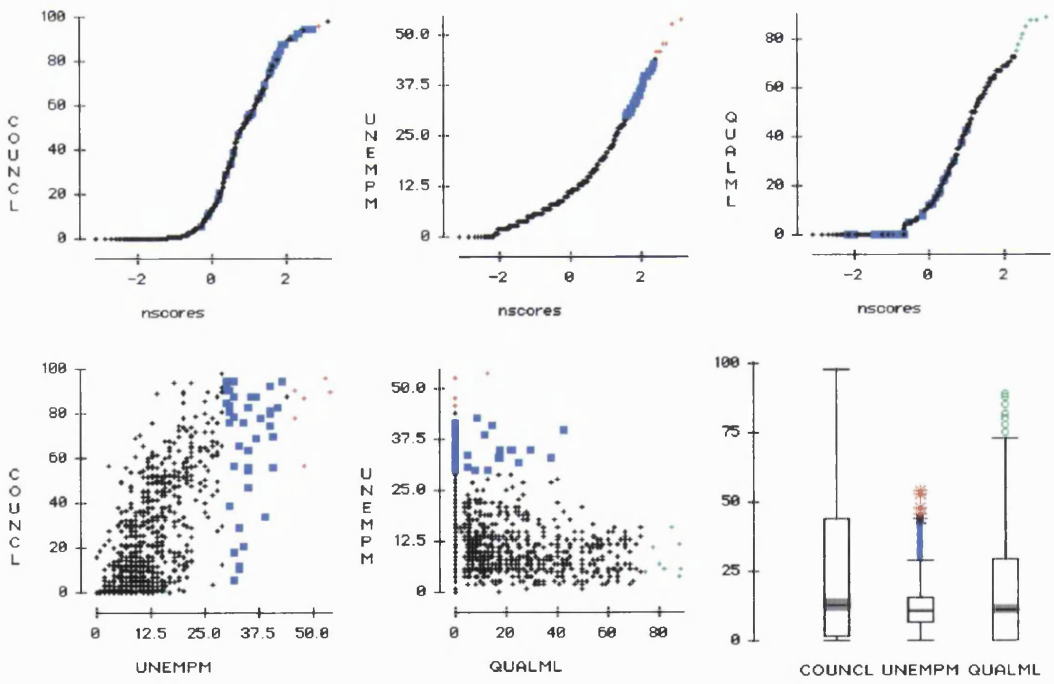

Figure 4.14 Colour coding multiple selections in six linked data view DataDesk (DataDescription Inc., Ithaca): The top three graphic displays show Normal QQ plots for three variables. The lower three displays show two scatterplots and boxplots for the same variables. The colour coding (green, red and blue) show were the selected records (outliers as shown by the boxplots) are in all data representations.

Two other problems identified by users were addressed. The first, facilitating the de-selection of single or sets of features, was fairly straightforward to fix programmatically. The second entailed coordinating the filtering of records that occurred on the graphic displays of DecisionSite with a similar behaviour of the features on the map. In other words, Chapter 2 discussed the concept of dynamic filtering where a scrollbar is used to perform structured queries on the fly, in order to facilitate the exploration of a subset of the data. In DecisionSite, this is accomplished by eliminating unwanted objects from the views. This behaviour was coordinated with the map so that the user would have an indication of the areas that were not being included in a particular subset of records of interest. Since it might be undesirable to make areas or features on the map disappear from view, other visual cues could be given to flag the areas that were not being considered in some scenario because their corresponding records had been filtered out in some other view. For instance, the transparency of the colour by which an area is coded could be changed, or the thickness of its boundary or some other aspect of map symbology could be altered. Such a visual cue would synchronise an event such as filtering between the two pieces of software. Changing the transparency of an individual polygon is not supported by ArcGIS. Thus, hatching was used as an alternative. This functionality was not tested in subsequent experiments but it is worth exploring in future research.

It is of importance to note that highlighting or the temporal change in object symbology, was a functionality that was available but was not appropriately used. When users hover over an object in any view it is highlighted, as well as objects in all other views which correspond to the same database record. The purpose of this visual cue is to facilitate the identification of one database record from a selection set or simply from a collection of objects of interest without having to select it (by drawing a bounding polygon around it). There are two reasons why users may not have taken advantage of this functionality. On the one hand, the highlighting occurred rapidly and it might have been difficult to see it in views that were already cluttered with objects. On the other hand, when the areas in the map are small, the visual cue is difficult to detect. This could be improved, for instance, by allowing the highlighting to remain as long as the user is over an object in any view.

4.10 Conclusion

This chapter has shown how experimental design can be effectively used to assess the adequacy of both geovisualization tools for supporting user tasks and task typology definitions. On the one hand, it allowed the detailed investigation of how well each task was

supported by the computer environment and thus facilitated its fine tuning. On the other, the statistical analysis provided detail insight into user performance for each factor combination. This allowed a better understanding of the factors and their interactions and how they affect user work processes. These findings are of relevance as it suggests the potential value of learning about geovisualization tools and tasks through experimental design. Conducting evaluations in a similar fashion not only permits a systematic yet rich analysis, but would also contribute to accumulating a body of empirical evidence and comparable evidence that would permit a deeper consideration of the suitability of geovisualization tools and concepts for supporting potential users exploratory data analysis. This in turn, is argued, would contribute to build more solid theoretical foundations for geovisualization.

A second important finding of this experiment is that there may be fundamental differences between the cognitive processes required to participate in the visualization of geographic spaces and those required for visualization of information spaces that are aspatial. It is argued that the reason for this is not merely a difference between the visual operations considered but, more fundamentally, because of the information processing that must be accomplished with geospatial data. This is an important new finding in the sense that it may imply that the processes that must be supported in geographic visualization are not only different but more complex than those needed to investigate aspatial data. This result is also of interest as it contributes to the findings reported both in the Information Visualization and spatial cognition literature from which geovisualization feeds. It therefore suggests an exciting possibility of cross-fertilisation between these areas of research where geovisualization can start to contribute to the debate, for instance, by providing evidence on the cognitive demands of geospatial tasks.

Finally, it was also observed during the experiment and during discussions with the participants that many attempted the same tasks in a different manner. In addition, many used or changed strategies for solving a single task. Possible differences in the strategies taken by participants are of interest as they may help explain the disparity between the time taken by subjects to complete tasks. In addition, they may point towards different behaviours or individual differences for solving geovisualization tasks. The following chapter considers a task taxonomy of finer granularity with which to examine these issues further as well as that of user behaviour when performing geovisualization tasks.

5 Study 3: Refining the characterisation of geovisulization tasks

"Graph comprehension is constrained by the goals of the cognitive system that processes the graph and by the context in which the graph appears" (Feeney et al. 2000: 149).

5.1 Lessons from the second evaluation

In the previous chapter, three factors were used for defining geovisualization tasks—the cognitive visual operators, the spatial extent of the area explored and the number of attributes investigated. Combinations of levels of these three factors defined a set of tasks that, if supported by the environment developed, were hypothesised to enable users to accomplish geovisualization goals, namely hypothesis generation and knowledge construction. The experiment conducted indicated that the environment did support the taxonomy of tasks tested and that it aided users in obtaining information from a spatial dataset. Nevertheless, some limitations of the environment were identified. For instance, improving the ability of the environment to facilitate the comparison of different scenarios is likely to enhance user performance and possibly speed the data exploration process.

Of great interest was the finding that there may be fundamental differences between the cognitive processes required to participate in the visualization of geographic spaces and those required for visualization of information spaces that are not spatial. The experiment revealed that the four visual operators considered (i.e. identify, locate, compare and associate) are too coarse to characterise spatial operations. Tasks where the aim was to find the relation to or relative position from one spatial entity to another, exhibited distinct patterns of user behaviour in the results. In other words, the four visual operators were insufficient to typify some of the tasks that users needed to accomplish when dealing with geospatial data.

Graphic representations can allow inferences to be made quickly and easily. Larkin and Simon (1987) suggest two principal reasons for this. First, these representations permit rapid searches of information based on the location of the components on a diagram, including information about relations with adjacent components. Together with the human visual system, these diagrams facilitate perceptual inferences, such as detecting cues from a graphical display in order to elicit information relevant to the problem at hand. Larkin and Simon (1987) find that this inference is largely independent of the representation, if the information content of the representations is equivalent between graphic displays. It is in the recognition process, or the matching of elements in the representation to elements in the

data, where graphics are more advantageous, particularly if the dataset is large. This is because diagrams or graphic representations make relations in the data explicit and directly perceivable. They can be understood as a form of external memory, which places information that may be needed for future inference together. Hence, they appear to "improve recall for presented data, [and] reduce short-term memory loads during problem solving" (Casner 1991: 118). In other words, "visualization can act as an extension of cognitive processes, augmenting working memory by providing visual markers for concepts and by revealing structural relationships between problem components" (Ware 2000: 335).

However, even when a display successfully encodes the relation between elements, it might still not help to solve the problem at hand. It has been previously mentioned that, in the field of Information Visualization, some work has been carried out to characterise the tasks supported by techniques for data investigation (Casner 1991; Zhou and Feiner 1998, 1996; Wehrend 1990; Wehrend and Lewis 1990; Knapp 1995). These taxonomies attempt to break up problems into tasks that can be supported by visual representations according to the type of data that they represent and, in some cases, also consider the work processes their potential users might need to support. These taxonomies "interface high-level presentation intents" (Zhou and Feiner 1998: 392), such as informing the user about a relation between attributes, with low-level visual techniques, such as linking and brushing between views. They also proved useful in a geovisualization context for defining user tasks that need support when exploring unknown data with a vague goal. Furthermore, they enabled the evaluation of the tool used for such purpose.

However, these typologies are incomplete or at least too general to describe geovisualization tasks. In Information Visualization, maps have not been considered as representations that can be interacted with and used as exploratory tools for reducing the search for information and facilitating knowledge construction. On the contrary, they are mostly seen as providing the context in which the analysis of some attributes—other than (x, y) coordinates—may be meaningful. The fact that "spatially extended region[s] or object[s]" (Wehrend 1990: 16) are little more than a backdrop that gives context to analysis is not surprising since the data with which Information Visualization is concerned are typically abstract and non-geographical. Thus, position or location of elements in a graph representing abstract data may be relevant—along with their arrangement or position relative to other elements in the view—for depicting similarities between data items. However, relative position and arrangement are interpreted as attributes from which a display is constructed (Wehrend 1990).

These 'attributes' may be used to improve the design of visual displays by making them more effective in the sense of facilitating the perception of the information they attempt to convey, as well as increasing their efficiency by speeding user searches for information. In contrast, when representing georeferenced information, the space of the display is usually reserved for depiction of geographic space (MacEachren *et al.* 1998). In addition, it is well known that in geography "everything is related to everything else, but near things are more related than distant things" (Tobler 1970: 236) and that attribute patterns and relations will tend to manifest in space accordingly. Hence, spatial relations and relative positions between geographic entities are not attributable to the semantics of a visual graphic designed to facilitate perception as in the case of Information Visualization displays. Rather, they are attributable to spatial processes that may exist in the actual physical space that the map represents. Hence, maps as visualizations of spatial information can indeed portray information that gives access to details about attributes in a dataset as in aspatial visualizations, but relative position or location of geographical entities has a meaning in itself over and above facilitating perception.

This chapter will therefore describe the third and final evaluation which is aimed at providing evidence in support of the need for a task typology that reflects the specificity of spatial tasks. The evaluation was conducted using the same environment used for the second study (Chapter 4), but was slightly improved to take into account some of the usability problems and identified limitations. The experimental design and the purpose of the evaluation are discussed in §5.2. The tasks considered are detailed in §5.2.4 and the results are presented in §5.2.5. The data used for this experiment is from the 2000 U.S. Census of Population for the area of Los Angeles, as described in more detail in Chapter 6.

5.2 Experimental design

5.2.1 Experiment Goals

There were two main goals for this third set of experiments. The first was to investigate the differences between tasks found in the previous chapter. It has been previously shown that tasks where a cognitive visual operation was performed relative to a spatial feature took distinctively longer than similar tasks where the relative spatial comparison was not required from users. In order to inspect this difference in more detail, each participant completed two sets of four questions. Therefore, eight tasks were examined to consider whether a finer

granularity in the task definition is relevant for describing geovisualization tasks. The second goal of the experiment was to investigate different user approaches to solving the tasks. It was hypothesised that any differences would occur not only because of the diverse complexity of the tasks but also because of the cognitive processes that might be invoked when addressing each type. For this reason, a second aim of this third experiment was to inspect subjects' strategies. It was hypothesised from observing user behaviour in the second evaluation that participants used different strategies for solving tasks and sometimes even within tasks. This is considered in §5.3 which details and discusses how the data recorded served to explore this issue.

5.2.2 Profile and Number of Participants

The profile of subjects in an evaluation should be carefully considered to reflect the potential users who would be likely to use the system. The environment evaluated was the same as that discussed in Chapter 4 (Figure 5.1), which was designed for GIS users. The nature of the environment required that subjects for this third study had some basic knowledge of off-the-shelf GIS functionality but were not expected to have any skill in using visualization tools. Eight participants from a range of fields or application areas who could potentially use the environment to support their own work processes were recruited. Table 5.1 shows that four participants were doctorate students with interests ranging from three-dimensional (3D) modelling to location based services. Three other participants were professionals working in the GIS industry. The final participant, although not in the GIS business, had received graduate level training in GIS. Subject ages ranged from 26 to 35 years of age. Three females and five males of different nationalities (although mostly European) took part. Potential users were invited to participate by letter (sent via email) and, as in previous experiments, were not paid for their time. Only one recruitment attempt was necessary and the response rate was more than 70%.

The number of users was chosen to satisfy the requirements of the experiment design explained below and to attain the desired standard error or test power. Hence, the standard deviation estimate from the second experiment (0.36), calculated as in §4.4.2, was used to obtain an indication of the number of participants that would be necessary to achieve the desired power. The standard error for this third experiment was estimated at $0.36*[2/(2*4)]^{1/2} = 0.18$. We can derive a confidence interval from this calculation to suggest that the estimates from the experiment would be subject to a margin of error of no more than $\pm$ 1.96*0.18 = $\pm$ 0.35 with a probability or confidence level of 95%.

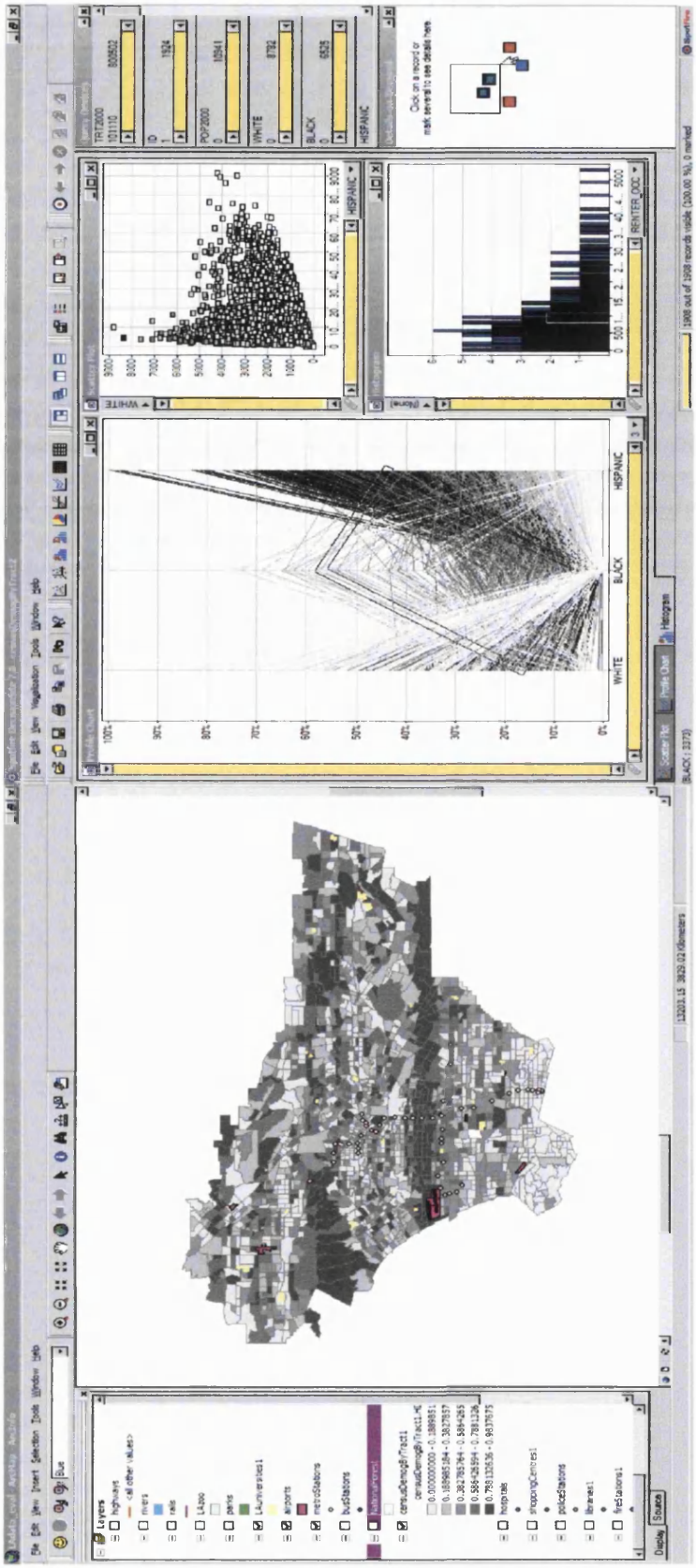

Figure 5.1 The environment to be evaluated closely coupled ArcGIS (ESRI, Redlands) and DecisionSite (Spotfire, Boston).

Table 5.1 Users' description

User ID	Occupation	Research interests or Expertise
1	GIS Consultant	Transport. Daily use of GIS (intermediate). European (Spanish).
2	PhD Student	Location Based Services. Monthly use of GIS (Intermediate). Asian (Chinese).
3	GIS Analyst	Utilities. Daily use of GIS (expert). Northern European (Norwegian).
4	PhD Student	Spreading of infectious diseases. Weekly use of GIS (Intermediate). European (Portuguese).
5	PhD Student	Urban Planning. Daily use of GIS (Intermediate). Near Eastern (Turkish).
6	IT support	Databases. Rarely uses a GIS. European (British).
7	PhD Student	3D Urban Modelling. Rarely uses a GIS. European (British).
8	Environmental Protection	Conservation. Daily use of GIS (intermediate). European (British).

5.2.3 Evaluation Phases

The experiment was divided into five phases each of which was conducted by the researcher in a usability laboratory (Figure 5.2). In the first phase, participants were given an explanation of the environment's functionality and a detailed description of the dataset as in the previous two experiments. A 'script' was followed as in the previous two experiments to give all users a similar demonstration of the environment and avoid biases due to providing participants with different information. Afterwards, subjects were given a set of training or warm-up tasks to complete. This training phase is not only useful to establish familiarity with the environment but is also of crucial importance to enhance the effectiveness of visual interfaces for information retrieval (Tobón 2003; Andrienko *et al.* 2002; Sutcliffe *et al.* 2000), as well as to improve subject performance (Burin *et al.* 2000; Kyllonen *et al.* 1984a). Users were then encouraged to ask any questions about the system or the data, as they were advised that no help in solving the tasks would be provided once the actual experiment started. Most of their questions regarded the use and purpose of the graphical displays and the meaning of variables in the dataset. This pre-test stage took an average of 30 minutes.

The second was a testing phase, in which users completed four tasks on their own. Their on-screen interaction with the software was captured and recorded in the usability laboratory (Figure 5.2). Also, a script was prepared so that every interaction of the users with the software was automatically captured, coded and written to a log file. This would later

facilitate the inspection of the frequencies with which participants used each display on each task, as well as the sequences in which the data views were used.

Figure 5.2 Usability Lab: Left) Room where participants completed the tasks using the computer to the left. They used the computer to the right to transcribe the answers to the tasks. In the middle is the camera that was used to record participants face in order to trace the direction of their gaze. Right) This room was separate from the one in which users completed the evaluation and was used to monitor their progress. The monitors on the top show the video streamed by the camera, as well as the computer screen used by participants. These two video sources were mixed using the equipment above and recorded onto VHS tape for subsequent analysis.

In addition, a video recording of the session was made using a camera to ascertain the direction of the participants gaze. According to the eye-mind assumption, which was empirically validated by Just and Carpenter (1976), "when a person is performing a cognitive task while looking at a visual display, the location of his or her gaze on the display corresponds to the symbol that he/she is currently processing in working memory" (Narayanan and Hegarty 2000: 182). Therefore, the analysis of the videos would further allow characterisation of how subjects addressed each task. To facilitate the analysis, the two video streams (that of the on-screen interaction and that recording participants gaze) were recorded and mixed into VHS tapes. Together with the logs, the on-screen interaction and video recordings would permit the analysis of user strategies as is discussed in detail in later sections. Any comments made by participants while doing the tasks where also captured by microphones installed in the lab. All participants were advised that they were being recorded, assured that all results would be kept anonymous, and asked for their consent to record the session. This stage took users between 30 minutes and one hour.

This first set of four tasks was followed by a short semi-structured interview conducted by the researcher and a break. The purpose of the interview was to gather information about user processes for accomplishing some of the tasks as it is discussed in a later section. This stage took an average of 20 minutes. The fourth phase was similar to the second in that participants completed a second set of four tasks. It is important to note that the tasks administered in phases two and four were of different nature as they evaluated different sets of visual operators and actually constituted two separate experiments. The evaluation ended with a final semi-structured interview which also aimed at eliciting participant thoughts and to learn about their understanding of the dataset. Table 5.2 summarises the experiment's resources.

Table 5.2 Summary of Experiment's Resources

User Time	1 hour 30 minutes per test
Hardware	PC connected to the Internet Monitor Camera Video streaming equipment Video Mixer VHS Time watch
Software	ArcGIS DecisionSite
Data files	Shapefiles with vector geometry of tracts for Los Angeles,CA. .dbf and Excel file with U.S. Census data for the study area.

5.2.4 Tasks

Two sets of cognitive visual operators were considered in this experiment. The first set was concerned with 'identify' and 'distinguish' operations, while the second investigated 'compare' and 'associate' ones. From the previous experiment it was found that performing visual operations *relative* to a spatial feature had significant effects on users' cognitive load and behaviour. Therefore phase two considered not only 'identify' (I) and 'distinguish' (D) but also 'identify relative to a spatial feature' (IS) and 'distinguish relative to a spatial feature' (DS). The difference is that the first two only imply the comparison of values, which could be accomplished using any graphical display. For that reason, they are referred to as 'non-spatial' (identify or distinguish) tasks in later sections. Tasks characterised in this manner break up the cognitive visual operation into finer levels in order to investigate the information processing during the data exploration. Phase four, on the other hand,

considered 'compare' (C) and 'associate' (A) as well as 'compare relative to a spatial feature' (CS) and 'associate relative to a spatial feature' (AS). The first two are generally also referred to 'non-spatial' (compare or associate) tasks in later analysis. Each set of four operators in each phase was used as a factor with four levels. Using the experimental design terminology, each set of four visual operators define four treatments that each subject encountered in the corresponding testing phase.

Two issues must be clarified at this point. The first regards the use of the visual operator 'distinguish' as opposed to 'locate' in this experiment. This consideration obeyed two reasons. First, the second experiment showed that 'identify' and 'locate' were very similar in terms of their effect on user performance. Only one was considered here so that the suitability of the technique to include other visual operators of interest could be appreciated. Second, 'distinguish' was included because it has not only been considered by the Information Visualization literature and thus could facilitate the comparison of some results, but is also one of the operators considered by Bertin (1983). It was of interest to ascertain whether, when presented with a map plus other views, respondents would prefer using the maps' colour coding or feature symbolism in order to distinguish between values or whether they would use other graphic displays for the same purpose.

The second issue to point out is the number of factors used in this experiment. The size of the area to be explored (NU) as well as the number of attributes (NA) used to solve a task were factors found to have significant main effects on subject performance (Chapter 4). Their effects were as expected: as their levels increased, the time taken by subjects to solve a task augmented because a larger number of spatial units or variables to consider increased the complexity of the task. The interaction effect between these two factors and the cognitive visual operator, however, depended on the visual operator. Generally, the interaction effect with the visual operator also increased the performance time when many attributes and spatial units where being explored. However, on average, there was no significant interaction between the NU and NA factors. Thus, since the effect of the NU and NA factors was somewhat established in the second experiment, these were not considered in this third set of experiments. There are two main reasons for this. The first is that the aim of the third set of experiments was to disaggregate the visual operator in order to inspect it in detail and better characterise geographic visualization tasks. Second, it was found in the previous experiment that there was a significant learning effect on performance as the subjects were exposed to the environment, the tasks and the data. Thus, it was reasonable to think that the order in which treatments were administered could have had a significant

effect in user performance as the user would be more experienced towards the end of the evaluation. Alternatively, users can also get increasingly tired which may also affect their performance. In order to test these hypotheses, a Latin Square experimental design was required. This technique only permitted one task dimension (such as the visual operator) to be investigated at a time. Since our main aim was to focus wholly on the differences of the visual operator, only this factor was considered. Thus, the words task and treatment are used henceforth interchangeably because the only difference between tasks administered was the treatment itself, i.e. the visual operator.

5.2.5 Latin Squares Design and Analysis

There are at least two major disadvantages with the repeated measures design used for the experiment in Chapter 4 where each user completed all tasks. The first is that subjects may show a positive practice effect in the form of a general improvement during the course of the evaluation or, alternatively, a negative practice effect in the form of fatigue that builds up as the session progresses (Keppel 1982). This implies that the performance measure may reflect both the effect of the task the participant is tackling and the effect of prior experience in the experiment. In other words, the two effects may be confounded. Although tasks were administered in a random order which was different for each participant, it was necessary to corroborate that the findings of Chapter 4 pertaining the visual operator were not a result of the order in which tasks were administered. A possible solution to this problem is administering the tasks to participants in sufficient testing orders "to ensure the equal occurrence of each experimental [task] at each stage of practice in the experiment" (Keppel 1982: 371). This is known as counterbalancing and may be achieved using a Latin Squares experimental design.

Practice effects are assumed to affect all tasks equally. Nevertheless, the effect of administering a task first may be different on subject's performance on subsequent tasks in a specific way. In other words, the effect may be differential and dependent on the sequence in which it was given relative to the other tasks. For instance, if there is a task that is particularly onerous, subjects receiving this task first may not behave in the same way as subjects receiving it last. In this case, the order in which tasks are administered may affect later conditions in different ways and counterbalancing will not control for the differential effects that may carry over to subsequent tasks. Although important in some experiments, a differential carry over effect was not a major concern in this study for at least two reasons. The first is that tasks of similar difficulty were administered as two separate experiments. It had already been found in a previous experiment that tasks such as 'identify' and 'locate'

were less complex than 'compare' and 'associate' operators. The second phase of this study concentrated on the less complex tasks while the fourth phase of the experiment tested the differences between the more complicated tasks. Thus, differential effects within the second or the fourth testing phases were not deemed to be of great importance. This was acceptable as the results of each phase were to be analysed separately. Second, enough time was left between the two testing phases which included an interview to elicit subject explanations of how they arrived at particular answers. Any carry over effects between the two phases were therefore regarded as being attributable to practice with the tasks and the dataset. The variation of the order in which the treatments were administered using the Latin Squares design accounts for this effect, as aforementioned (Keppel 1982).

If we wish to study n tasks and decide that each participant will attempt each task at least once, then we have n choices for the first task to administer, $n-1$ choices for the second, and so on. In total, there would be $n! = n(n-1)(n-2)...(2)(1)$ possible sequences. For four tasks, there are $4! = 24$ possible sequences in which the tasks can be administered. Latin Square designs are combinations of these sequences which preserve certain properties while significantly reducing the number of tasks to test. The one used for this experiment is shown in Table 5.3. A, B, C and D denote four visual operators and P1, P2, P3 and P4 the position in which the treatments were given. Two participants where allocated to each sequence which means that two users received the tasks in the order A, C, B and D while two users were administered the treatments in the order B, D, A and C, and so on. It is important to note that the actual visual operators were assigned randomly to the A, B, C, D designations and that the users were also assigned randomly and in equal numbers to the sequences. Finally, the same Latin Square design was used for both phases two and four for simplicity.

Table 5.3 Latin Squares (Sequence by position)

	Testing position			
Sequence	**P1**	**P2**	**P3**	**P4**
I	A	C	B	D
II	B	D	A	C
III	C	A	D	B
IV	D	B	C	A
Sum	PP1	PP2	PP3	PP4

The Latin Square in Table 5.3 is certainly not the only one available as other sequences out of the possible 24 could have been combined. Furthermore, if we combine all the possible sequences we would obtain 24 possible Latin Squares arrangements. The mathematical

calculations that determine what sequences conform to a Latin Square design are beyond the scope of this chapter. However, the main property of the design in Table 5.3 is that any pair of visual operators in a given sequence would appear in the reversed order in another sequence. For instance in sequence I, the pair (A, C) is given in that order and before the pair (B, D). In sequence II, on the other hand, (B, D) is administered before (A, C). Also, if one treatment precedes another, the reverse order is found in a different sequence in the Latin Square. For instance in sequences I and II, A is before C whereas in sequences II and IV the opposite holds. The selection of the Latin Square may frequently be done at random as there were no good reasons in our case for choosing any one in favour of another (see Senn 2002 for a detailed discussion). More importantly though, in order to counterbalance the effects of practice, each task must be presented an equal amount of times in the first, second, third or fourth positions. Note that in each column and row in Table 5.3, each treatment appears only once. This means that it is administered once in each of the four possible positions.

It is important to mention one major disadvantage of Latin Squares designs, namely, that all observations are needed. In other words, participants must complete all tasks and all recorded outcomes must be used in the statistical analysis. Experiments where subjects may 'drop-out' and not finish all the tasks are not good candidates for this type of experimental design. This does not imply that the tasks attempted must be correct, but simply that they must be finished and the time to completion recorded and used for the analysis. In this experiment, users had no time restrictions for completing tasks but had to state when they had finished each one by pressing a button on the screen.

Table 5.4 Latin Squares (Sequence by visual operator)

	Factor Levels				
Sequence	A	B	C	D	**Sum**
I	P1	P3	P2	P4	S_1
II	P3	P1	P4	P2	S_2
III	P2	P4	P1	P3	S_3
IV	P4	P2	P3	P1	S_4
Sum	V_1	V_2	V_3	V_4	T

In order to perform a within-subjects ANOVA analysis, we can rearrange the data according to the treatments as in Table 5.4. This is so that the calculations illustrated in Table 5.5 can be understood more easily. Table 5.5 illustrates the estimations for our experiment where the visual operator (V) and the subjects (S) are treated as factors (Keppel 1982). All subjects receive all levels of the V factor as treatments. Therefore, V is a within-subjects or repeated measures manipulation. Summing and squaring the scores in Table 5.4 by column, and

dividing this amount by the number of subjects we obtain [V] = $[(V_1+V_2+V_3+V_4)^2/8]$. Summing the scores in Table 5.4 by row, we obtain the total score for each subject at every level of factor V. Squaring those amounts and dividing them by the total number of levels of the visual operator (v) we obtain [S] = $[(S_1+S_2+S_3+S_4)^2/4]$. Finally, by squaring and summing each one of the scores obtained by each individual (each of the Ps in Table 5.4), we obtain [VS]. Dividing this result by (v) multiplied by the number of subjects (s), we obtain [T]. Note that uncontrolled sources of variation in this design

"cannot be due to differences ***between*** *subjects, since each subject serves in all of the conditions, but rather to the* ***inconsistency*** *with which subjects perform under the different treatments. This inconsistency is reflected by the Treatment * Subject interaction" (Keppel 1982: 385, original emphasis).*

Table 5.5 Within Subjects ANOVA: (V*S) design

Source	DF	Sums of Squares	MS	F
V	v-1	[V] - [T]	SS_V/DF_V	MS_V/MS_{V*S}
S	s-1	[S] - [T]	SS_S/DF_S	
V*S	(v-1)(s-1)	[VS] - [V] - [S] + [T]	MS_{V*S}/DF_{V*S}	
Total	(v)(s) - 1	[VS] - [T]		

The above analysis, however, does not reflect that the order in which the tasks were administered was designed to counterbalance effects of practice. Hence, when a practice effect is present, a main effect would be confounded with the interaction between the visual operator and the subjects. This can be easily removed in practice using a Latin Squares analysis, which takes into consideration the position in which the treatments were administered. This in turn has an effect on the size of the error term, MS_{V*S}. In order to remove such effect, we sum all the columns in Table 5.3, square and sum the values and divide them by the total number of subjects to obtain [PP] = $[(PP_1+PP_2+PP_3+PP_4)^2/8]$ and include this term in the calculations, as shown in Table 5.6. The residual variation, $MS_{LS\,Res}$, therefore has the practice effect removed.

Table 5.6 Latin Squares

Source	DF	Sums of Squares	MS	F
V	v-1	[V] - [T]	SS_V/DF_A	$MS_V/MS_{LS\,Res}$
S	s-1	[S] - [T]	SS_S/DF_S	$MS_S/MS_{LS\,Res}$
PP	p-1	[PP] - [T]	SS_P/DF_P	$MS_P/MS_{LS\,Res}$
LS Residual	(v-1)(s-1) – (p-1)	$SS_{V*S} - SS_P$	$MS_{LS\,Res}/DF_{LS\,Res}$	

Similar to the discussion in Chapter 4 (§4.5), the model underlying the analysis of variance when a Latin Squares design is used is specified by expressing the dependent variable, such as time performance, as an equation of the form:

$$y_{ijk} = \mu + \beta_1 V_i + \beta_2 S_j + \beta_3 PP_k + \varepsilon_{ijk} \tag{4.1}$$

where y_{ijk} is the response time for each task; μ is the overall mean of the population; V_i corresponds to the visual operator effect at each level i (for $i = 1,2,3,4$); S_j is the subject effect for the jth subject (in this case we had 8 participants so $j = 1,2,...,8$); PP_k is the effect of the position in which the task was administered ($k = 1,2,3,4$); and ε_{ij} is the experimental error term. Using this model, we wish to estimate its parameters $(\mu, \beta_1, ..., \beta_3, \sigma^2)$ and test the hypothesis of whether different tasks have different effects on user responses, whether the subjects have an effect, and whether there is any trace of a position effect after this has been taken into account by the Latin Squares design for administering tasks to subjects. This is equivalent to testing whether the task means are equal:

$$\begin{aligned} H_0 &: \mu_1 = \mu_2 = \mu_3 = 0 \\ H_a &: \mu_n \neq 0 \end{aligned} \tag{4.2}$$

for at least one n which is the number of all possible factor combinations. This is also equivalent to testing the hypothesis that all coefficients are zero:

$$\begin{aligned} H_0 &: \beta_1 = \beta_2 = \beta_3 = 0 \\ H_a &: \beta_n \neq 0 \end{aligned} \tag{4.3}$$

5.2.5.1 Results of the second phase

The data in Table 5.7 show the results for each individual in every task of the second phase of the experiment. The columns in the table correspond to each one of the tasks or levels of the cognitive visual operator factor (I, D, IS, DS). The rows are the observations for each subject. Since pairs of subjects were administered the treatments in a particular sequence, there are two rows for each of the four possible sequences in the Latin Square. Note that the natural logarithm of the performance measurement was used in both cases to stabilise the variability of performance times between subjects. The effect of the transformation is shown

by the two graphs in Figure 5.3 where the raw times and their logs are plotted. For a further discussion of why the logs were used throughout the experiments see also §4.5.1.

Table 5.7 Results for the visual operators equal to Identify and Distinguish

	Factor Levels				
Sequence	**A = DS**	**B = IS**	**C = I**	**D = D**	**Sum**
I	6.23	4.60	3.97	6.35	21.14
	5.81	6.31	5.30	6.58	23.99
II	6.15	4.88	5.62	6.17	22.82
	6.31	5.25	5.18	6.48	23.22
III	6.59	6.32	6.27	6.45	25.63
	6.02	5.17	4.91	5.99	22.10
IV	5.30	4.36	4.88	5.80	20.34
	4.54	4.19	4.06	4.84	17.64
Sum (S)	46.95	41.07	40.20	48.67	176.87
Mean	5.87	5.13	5.02	6.08	

When the values are arranged by the testing position in which they were administered, we obtain Table 5.8. Bear in mind that scores in any given column do not correspond to the same task but to the position in which they were administered. If we now apply the formulas in Table 5.6 or estimate the model in Equation 4.1 using the log of the time (in seconds) that it took each participant to respond the 'identify' or 'distinguish' tasks as the independent variable, we obtain the results in Table 5.9.[1] These are the results for the ANOVA using a Latin Squares design. The effect of the visual operation was found to be significant as it was hypothesised ([$F(3,18) = 14.1$; $p<0.00$]). Note that the PP factor, which measures the effect of the position in which the task was administered, is not significant ([$F(3,18) = 0.58$; $p<0.65$]). This is what we aim at when using this technique as the Latin Squares design should control for the effect that the treatment's position might have on user performance.

Table 5.8 Results organised by task position (Phase 2)

	Task position			
Sequence	**P1**	**P2**	**P3**	**P4**
I	6.23	3.97	4.60	6.35
	5.81	5.30	6.31	6.58
II	4.88	6.17	6.15	5.62
	5.25	6.48	6.31	5.18
III	6.27	6.59	6.45	6.32
	4.91	6.02	5.99	5.17
IV	5.80	4.36	4.88	5.30
	4.84	4.19	4.06	4.54
Sum	43.99	43.09	44.74	45.05
Mean	5.50	5.39	5.59	5.63

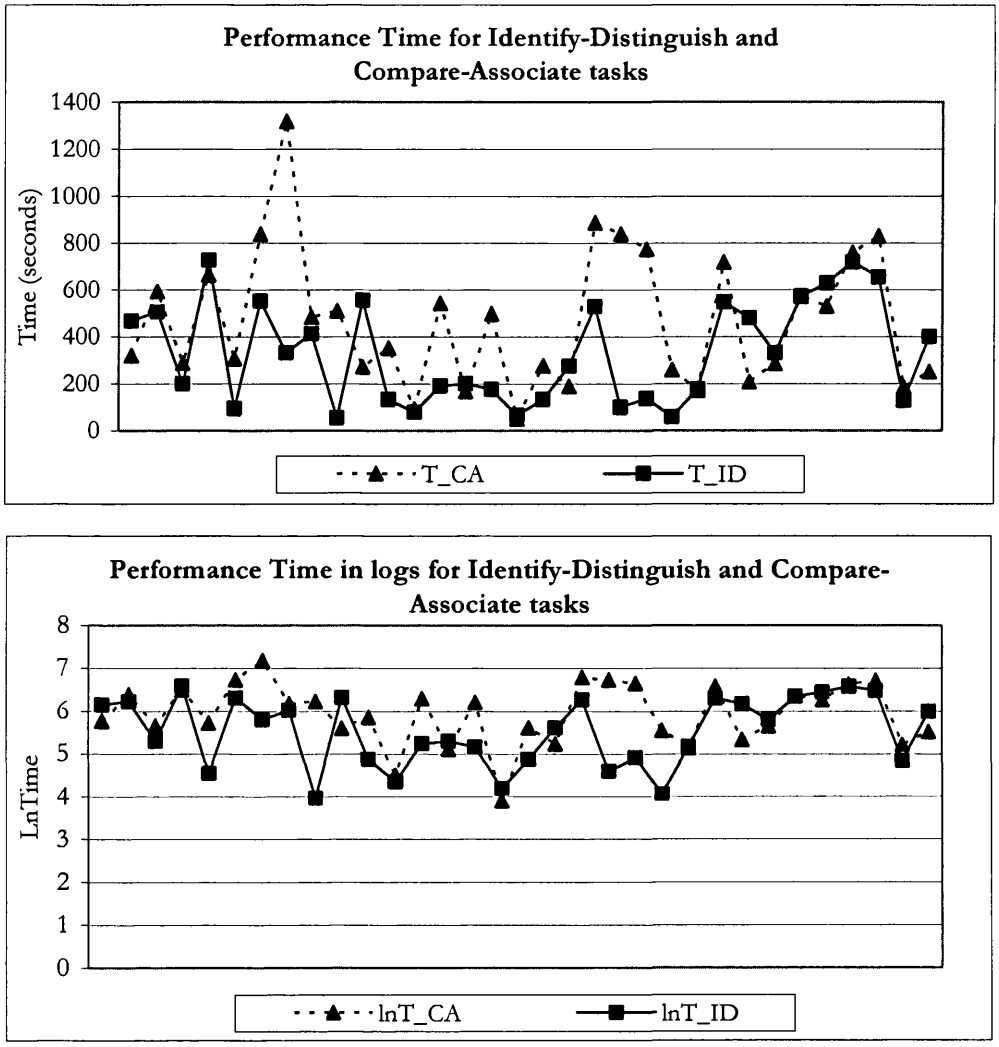

Figure 5.3 Time performance for all tasks and users: Top) Raw times in seconds for all tasks. Bottom) Log of times taken to complete all tasks.

We would expect the average time taken on the task administered first to be longer. Indeed, the tasks administered in the first position took slightly more than those given in the second as the mean times estimated by column in Table 5.8 show. However, tasks administered in positions three and four took slightly longer. Nevertheless, Figure 5.4 (Centre) shows that the main effects of the levels of the position factor did not deviate considerably from the grand average (dotted line). The Latin Squares design was nevertheless appropriate as it is discussed in a later section.

[1] All estimations were done in Minitab version 13 and S-Plus version 4.5

Table 5.9 Latin Squares Analysis for operators Identify and Distinguish

Source	DF	Sums of Squares	MS	F	p-value
V	3	6.67	2.23	12.22	0.00
S	7	10.43	1.49	8.19	0.00
PP	3	0.29	0.10	0.53	0.67
LS Residual	18	3.28	0.18		

Table 5.10 shows the estimates of all coefficients for the linear regression. The sign of the coefficients indicate whether the effect was below (negative) or above (positive) the overall mean response time. The results indicate that the two distinguish tasks took considerably and significantly longer than the grand average for all tasks. In addition, the two 'identify' tasks, spatial and aspatial, took less time than the 'distinguish' tasks and than the overall mean. Figure 5.4 clearly shows the greater main effects of the 'distinguish' tasks (first and last points on the curve) as deviations of the grand mean. It was expected that the spatial tasks would take longer than their non-spatial counterparts but this was not observed for the 'distinguish' tasks. They however show the greatest deviation from the overall mean.

Table 5.10 Estimates of effects for 'identify' and 'distinguish' tasks

Term	Estimate	Standard Error	t	p-value
Mean	5.53	0.07	73.28	0.00
Visual Operator				
V (DS)	0.34	0.13	2.61	0.02
V(IS)	-0.39	0.13	-3.01	0.01
V(I)	-0.50	0.13	-3.86	0.00
Subject				
S(1)	-0.24	0.19	-1.27	0.22
S(2)	0.47	0.19	2.47	0.02
S(3)	0.18	0.19	0.93	0.36
S(4)	0.28	0.19	1.46	0.16
S(5)	0.88	0.19	4.62	0.00
S(6)	-0.00	0.19	-0.01	0.99
S(7)	-0.44	0.19	-2.33	0.03
Position				
P(1)	-0.03	0.13	-0.22	0.82
P(2)	-0.14	0.13	-1.08	0.29
P(3)	0.07	0.13	0.5	0.62

The fact that the 'distinguish non-spatial' task took longer might have obeyed at this task demanding the exploration of more variables (at least four) than its spatial counterpart (which required the investigation of at least two attributes). As it was shown in Chapter 4, the number of attributes to explore has a positive effect on time performance. In addition, the 'identify' spatial task took on average much less than it might have been expected. This is partly due to the very short times taken by users 7 and 8 in this task. This points at the influence that individual differences may have on the results. Figure 5.4 (Right) illustrates the

effects of the eight users on the response mean. Users 1, 7 and 8 were significantly faster than the average, whereas 2 and 5 where particularly slow (see also Table 5.10). The range for each participant's average task completion is particularly large: 1.5 minutes for User 8 and 11 minutes for User 5.

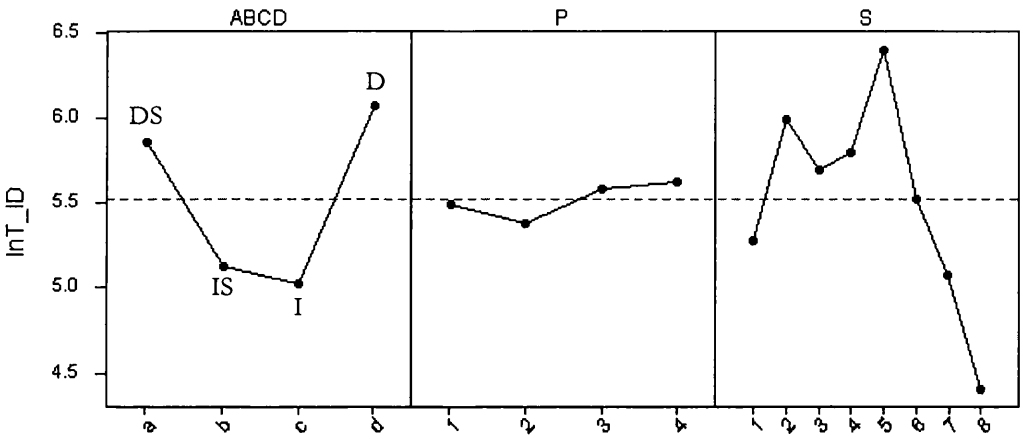

Figure 5.4 Main effects of factors on the performance variable (log of time to complete 'identify or 'distinguish' tasks): Left) Main effects of the visual operator. Centre) Main effects of the position in which tasks were administered. Right) Main effect of subjects.

Before analysing the results for phase 4, we must check whether the Latin Squares design was more efficient when compared to alternative designs, i.e. completely randomised or block designs. In other words, whether there is a significant difference between the subjects or the positions in which the tasks were administered. First we compare the efficiency of the Latin Squares to the completely randomised design. In the latter design, the tasks would be assigned to the subjects at random. This would imply that there is no link between the subjects and the response. In order to evaluate whether a complete randomised design would have been more efficient, we must first estimate its linear model:

$$y_i = \mu + \beta_1 V_i + \varepsilon_i \qquad \textbf{(4.4)}$$

Now we may compare the residual standard error of the complete randomised model (0.4861) to that of the Latin Squares (0.2797) using the following equation:

$$[\text{Standard Error (alternative design)/ Standard Error (Latin Squares design)}]^2 \qquad \textbf{(4.5)}$$

which gives us a 73.8% improvement when we use the Latin Squares over the complete randomised design. We can repeat this procedure to compare whether the Latin Squares is

more efficient than a block design. This latter design implies that the experimental units (the subjects) are not homogeneous. Thus, we would group the subjects into blocks that minimise the variation within the group and maximise the variation between the blocks. For instance, if we had recruited males and females with the purpose of investigating gender differences, we would block the subjects by sex. In our case, however, we considered each subject to be a block. All subjects selected met two characteristics—some knowledge of GIS and no knowledge of visualization tools—so any other differences between them can only be considered to be because of their individual qualities. Hence, this would imply that they can be regarded as a heterogeneous experimental material because of their individual differences. In order to compare whether a block design is more efficient than the Latin Squares we first estimate its linear model:

$$y_{ij} = \mu + \beta_1 V_i + \beta_2 S_j + \varepsilon_{ij} \tag{4.6}$$

We then compare the residual standard error of the block design model when each subject is taken to be a block (0.29) to that of the Latin Squares design (0.30). Replacing these numbers in Equation 4.5, we obtain 4.6% which is the increase in efficiency when we use Latin Squares. Alternatively, we could take the position in which the tasks were administered to be the block variable. It is reasonable to expect that the variability of subjects' performance would be different for tasks administered first than for those administered towards the end of the test because of learning effects. We therefore estimate the following block design model:

$$y_{ik} = \mu + \beta_1 V_i + \beta_2 PP_k + \varepsilon_{ik} \tag{4.7}$$

and compare its residual standard error (0.50) to that of the Latin Squares design (0.30). Replacing these numbers in Equation 4.5, we obtain a 78.8% gain in efficiency by using a Latin Squares design. These results therefore indicate that using Latin Squares was a good design choice.

5.2.5.2 Results of the fourth phase

The data in Table 5.11 show the results for each individual in every task of the fourth phase of the experiment. Once again, the scores correspond to the natural logarithm of subject time performance (Figure 5.3 shows how this transformation stabilises the variance). The columns in the table correspond to each one of the levels of the cognitive visual operator

factor (C, A, CS, AS). The rows are the observations for each subject. Pairs of subjects were thus given the same sequence as in the second phase.

Table 5.11 Results for visual operators equal to Compare and Associate

	Factor Levels				
Sequence	**A = A**	**B = AS**	**C = C**	**D = CS**	**Sum**
I	6.39	6.23	6.73	6.35	25.70
	7.19	5.12	6.58	6.63	25.51
II	5.77	5.86	5.24	5.34	22.21
	6.73	6.30	5.15	6.72	24.90
III	6.50	5.61	6.79	6.27	25.17
	6.18	6.21	6.65	5.53	24.57
IV	5.66	4.51	5.62	5.66	21.45
	5.73	3.89	5.55	5.22	20.39
Sum (S)	50.14	43.73	48.31	47.71	189.90
Mean	6.27	5.47	6.04	5.96	

When the values are arranged by the testing position in which they were administered, we obtain Table 5.12. Note again that scores in one column do not correspond to the same task. Applying the formulas in Table 5.6 or estimating the model in Equation 4.1 for the log of the time it took each participant to respond the 'compare' or 'associate' tasks, we obtain the results in Table 5.13. They show the output of the ANOVA using a Latin Squares experimental design. The effect of the visual operation was also found to be significant ([$F(3,18) = 3.26$; $p<0.04$]), as well as the effect of subjects on the response variable ([$F(3,18) = 3.81$; $p<0.01$]). Note that the position is not significant ([$F(3,18) = 1.32$; $p<0.30$]), which is the desired result.

Table 5.12 Results organised by task position (Phase 4)

	Testing position			
Sequence	**P1**	**P2**	**P3**	**P4**
I	6.39	6.73	6.23	6.35
	7.19	6.58	5.12	6.63
II	5.86	5.34	5.77	5.24
	6.30	6.72	6.73	5.15
III	6.79	6.50	6.27	5.61
	6.65	6.18	5.53	6.21
IV	5.66	4.51	5.62	5.66
	5.22	3.89	5.55	5.73
Sum	50.04	46.45	46.82	46.57
Mean	6.26	5.81	5.85	5.82

The average time taken on the tasks is closer to what we would expect in this case. Table 5.12 shows that the mean time for the tasks completed first is larger than all the rest. This is also clear in Figure 5.5 (Centre) which shows the main effect of the position. The first task took clearly longest and also more than the grand average (dotted line). All other tasks took

less and have similar mean times among them, which deviate only slightly from the overall average time to complete the tasks.

Table 5.14 shows the estimates of all coefficients for the linear regression. The results indicate that the 'associate' task deviate significantly more from the average response time than the 'compare' tasks. However, the AS task took less than its aspatial counter part and less than the CS task. In addition, the aspatial 'compare' tasks took slightly longer than its spatial counterpart. This can also be easily seen in Figure 5.5 (Left) which shows the main effects of the visual operator.

Table 5.13 Latin Squares Analysis for operators Compare and Associate

Source	DF	Sums of Squares	MS	F	p-value
V	3	2.74	0.91	3.26	0.04
S	7	7.47	1.07	3.81	0.01
P	3	1.11	0.37	1.32	0.30
LS Residual	18	5.03	0.28		

Note again in Figure 5.5 (Right) that subjects varied greatly in the mean times they took for solving tasks. Users 7 and 8 were again fast whereas users 2 and 5 where again slow when compared to the whole group. Nevertheless, all users were, on average, slower in completing these tasks as the overall mean (dotted line) indicates. This was to be expected as we know from Chapter 4 that 'compare' and 'associate' operators are of higher complexity and this reflects on user performances. User 1 however was much slower during this set of tasks. This user was the slowest on average but also provided some of the best answers to the tasks as §5.3.2 discusses. Although User 3 maintained the same average task completion time as compared to Phase 2 of the experiment, his mean time appears relatively low when compared to the group of participants. In other words, he was in fact much faster during Phase 4 of the evaluation. More importantly, it was expected that the spatial tasks would take longer. However, this was not observed (see Figure 5.5, Left). The reason for this is rather interesting. It lies in the very long completion times that subjects 2 and 5 took in the task administered first—which were non-spatial associate and compare treatments, respectively—and in the very short times that subjects took in the task administered second—which was a spatial compare treatment (see the values shaded in Table 5.12).

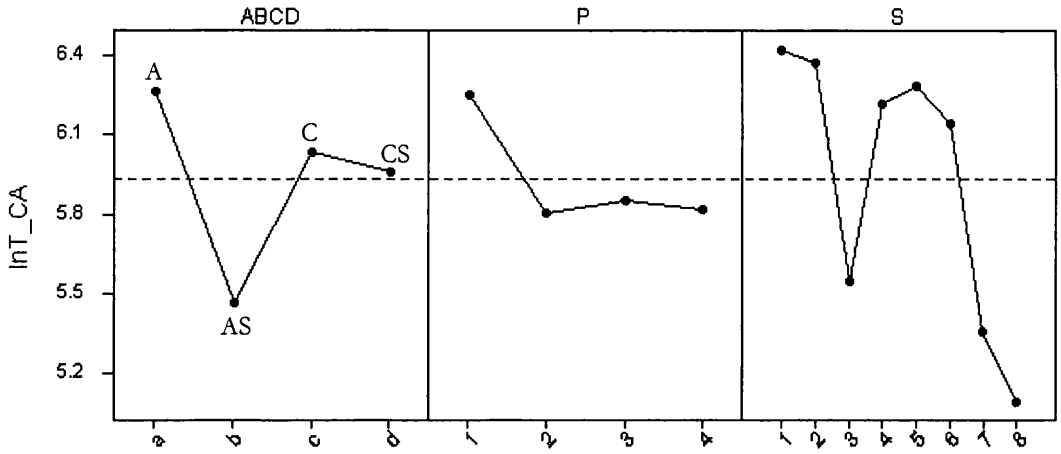

Figure 5.5 Main effects of factors on the performance variable (log of time to complete 'identify or 'distinguish' tasks): Left) Main effects of the visual operator. Centre) Main effects of the position in which tasks were administered. Right) Main effect of subjects.

Before we analyse this last point in detail, we must check how much more efficient was the use of the Latin Squares compared to completely randomised and block designs. We first estimate the same models as in Equations 4.4, 4.6 and 4.7 but for y, the performance variable, being the time taken by users to complete the 'compare' and 'associate' tasks. First, we compare the residual standard error of the completely randomised and block designs (0.4890) to that of the Latin Squares deign (0.1656). Replacing these values in Equation 4.5, we obtain that the Latin Squares represented an improvement compared to the complete randomised design of 295.3%. Second, we compare the residual standard error of the block design when the position in which the tasks were administered is used as the block variable (0.5365) to that of the Latin Squares design (0.1656). Replacing these values in Equation 4.5, we obtain an increase in efficiency of 323.97%. Third, we compare the residual standard error of the block design when the subjects were used as the block variable (0.1552) to that of the Latin Squares design (0.1656). Replacing these values in Equation 4.5, we find that the Latin Squares was slightly less efficient than this block design in 6%. Still, it was considered that the Latin Squares was an appropriate choice of design. This highlights the relevance of taking into account the position in which tasks are administered in these experiments.

Table 5.14 Estimates of effects for 'compare' and 'associate' tasks

	Estimate	Standard Error	t	p-value
Mean	5.93	0.09	63.47	0.00
Visual Operator				
V (A)	0.33	0.16	2.06	0.05
V(AS)	-0.46	0.16	-2.89	0.01
V(C)	0.11	0.16	0.65	0.52
Subject				
S(1)	0.49	0.25	1.99	0.06
S(2)	0.44	0.25	1.80	0.09
S(3)	-0.38	0.25	-1.54	0.14
S(4)	0.28	0.25	1.17	0.26
S(5)	0.36	0.25	1.44	0.17
S(6)	0.21	0.25	0.84	0.41
S(7)	-0.57	0.25	-2.31	0.03
Position				
P(1)	0.32	0.16	1.98	0.06
P(2)	-0.13	0.16	-0.79	0.44
P(3)	-0.08	0.16	-0.50	0.62

In summary, these results are point towards an important finding, namely that individual differences have an significant role on experiments where visual spatial operations must be accomplished. Similar results have been obtained in the field of cognitive psychology as is discussed in detail in §5.3. Therefore, we turn now to analyse user strategies as a means to understand some of these differences.

5.3 User Strategies

In the experiment discussed in Chapter 4, it was found from interviews and a Likert scale questionnaire that the system evaluated was perceived by participants as being both useful and easy to use for accomplishing visual exploratory tasks. However, Sutcliffe *et al.* (2000a) found that while visual user interfaces for information searching might be usable, they may not help users achieve improved performance. User performance in the second experiment was analysed using recorded measures of the time taken by users to complete each task and the number of successful answers during the two testing sessions. It was found that user performance improved during the second session because of the previous experience with the software environment. It was also found that, across all users and tasks, 80% of all answers were correct, or more precisely, within the range of plausible solutions. 77% of respondent answers during the first day were correct and 83% during the second, suggesting an improvement. These findings are in accordance with those of Kyllonen *et al.* (1984a) who suggest that subjects who have a high degree of proficiency in a particular domain perform best if their strategies and skills are sharpened through practice and feedback. By high proficiency in a domain, we mean in this case that they were all average to expert users of GIS and data. In terms of feedback, users had an introductory or training period before each

experiment started, during which they could ask any questions including those concerning the identification of ways of finding particular information. Practice, on the other hand, was gained over the course of the two experiments

In addition, following observation of user behaviour and from the interviews carried out during the experiment reported in Chapter 4, it was hypothesised that participants were employing different strategies in order to reach their answers. A number of studies show that subjects differ in the strategies they use and often shift strategies between and within spatial tests to accommodate different aspects of tasks that they might find difficult (Kyllonen *et al.* 1984a; Kyllonen *et al.* 1984; Snow 1978; Cooper 1980). Kyllonen *et al.* (1984: 1325) cite Snow (1978) who points to three possible sources of individual differences in task performance which are loosely referred to as strategy differences: "sequence differences, where subjects differ in the steps that are executed; route differences, where subjects differ in the steps that are included; and summation or strategy differences, where the entire processing program differs from subject to subject or differs within a subject for different items of a task". These three possible sources of individual differences were investigated for the third and last set of experiments.

5.3.1 User Behaviour

In an attempt to understand the strategies used by subjects in the context of geovisualization, respondents' mental and physical operations were studied (Sutcliffe *et al.* 2000; 2000a). Users were video taped and the tapes analysed to investigate their behaviour through the operations they performed with the software. These operations are defined in Table 5.15. In order to solve a task, users had to perform a sequence or combination of low-level actions such as searching and selecting information. With the functionality provided by the environment, this could be done either using the GIS (map) or any of the other non-spatial graphical displays. Off-the-shelf GIS functionality offers selection by attribute and by location, as well as colour coding of areas for representing attribute values. Both on the map and on the rest of the visualizations, users could interactively select objects by enclosing them in a bounding polygon. In addition, data filtering using query devices was available for concentrating on subsets of data and emptying all views (apart from the map) from unwanted clutter.

Table 5.15 User Behaviours

Physical Operations (Observed)	
1. Execute a search by attribute	Query the database using the search by attribute dialogue window in ArcGIS
2. Execute a search by location	Send a spatial query to the database using the search by location dialogue window in ArcGIS
3. Filtering	Send an SQL query to the database using the filtering devices (Query Devices) in DecisionSite
4. Highlight or navigate records	Browse or hover records or spatial features
5. Select records interactively	Draw a bounding box around records to be selected in any of the graphic displays
6. Evaluate content	Scan a record's attributes using the Details-on-Demand window in DecisionSite or the Information Tool in ArcGIS
7. Manipulate map	Change colour coding or layers displayed
Mental Operations (Observed or verbalised in interview)	
8. Decide threshold	Decide the criteria or variable values of relevance for solving a task
9. Select records	Choose subsets of data to investigate

5.3.1.1 Frequencies

Table 5.16 shows the number of times users performed these operations for each task. They were extracted and computed from the recorded logs and from observing the videos. Note that the 'decide threshold' operation has been computed by summing all activities that entailed identifying a minimum or maximum value (or both) when searching information within a subset of the data. This includes user actions such as performing a 'search by attribute' in the GIS and operating the query devices for 'filtering', which are both equivalent to performing structured queries. Similarly, the 'select records' operation has been computed by summing all selection activities such as selecting from any graphical display interactively or 'selecting by location' in the GIS.

Table 5.16 suggests that these two sets of operations, namely 'decide threshold' and 'select records', account for 77.5% of all the user actions. 36.4% of the operations across all tasks consisted of interactive searches and selections while 41.8% consisted of filtering using the query devices. Given that all participants were GIS users and that they had no prior experience with visualization tools, this latter result is a very high proportion of the operations which are not performed using the GIS. In addition, note that searches by attribute values in the GIS and filtering using the query devices are equivalent operations, namely sending a structured query to the database. However, users much preferred using the

query devices. Participants expressed the view that the query devices allowed them to assess the sensitivity of the outcome to the value thresholds imposed by the structured query. In addition, users considered the devices supported their thought process as their effect on the data explored could be seen in near real time. It is through immediacy of the visual feedback generated that the query devices allow users to explore and assess their ideas as they have them. This is therefore an indication of the ease with which data could be explored using visualizations other than the map display, such as the PCP, scatterplot or query devices. It also points to the usefulness and benefit of combining spatial data displays with flexible, interactive and rapid data query, as well as with aspatial data views that represent different relations present in the data. Other studies have also reported the benefit of these features for information search and exploration (Sutcliffe *et al.* 2000; Ahlberg and Shneiderman 1994; Andrienko *et al.* 2002).

Table 5.16 Frequency of user operations from computer logs and videos

	IN	DN	IS	DS	CN	AN	CS	AS	Total	%
Physical Operation (Observed)										
1. Search by attribute	4	4	6	5	1	2	4	0	26	1.6
2. Search by location	0	0	0	0	0	0	0	1	1	0.1
3. Interactive Selection	36	47	53	69	131	104	102	56	598	**36.4**
4. Filtering	34	65	58	166	61	112	142	49	687	**41.8**
5. Highlight records	31	33	27	35	37	28	28	35	254	15.5
6. Evaluate content	14	1	14	2	9	8	3	4	50	3.0
7. Manipulate map	4	1	10	1	0	2	1	9	28	1.7
Mental Operations (Observed)										
8. Decide threshold (1+4)	38	69	64	171	62	114	146	49	713	**43.4**
9. Select records (2+3)	34	45	42	64	130	96	99	50	560	**34.1**
Total (1 to 7)	123	150	168	278	239	256	280	150	1644	

A further breakdown of the data in Table 5.16 is very revealing. Table 5.17 shows the number of times different operations were carried out by users for the 'identify' and 'distinguish' tasks alone. The column labelled 'Total' is the total number of operations performed for all 'identify' and 'distinguish' tasks. We can see that interactive searches and filtering account for most of the operations (28.5% and 44.9%, respectively), with the use of query devices being the most frequent action. The column labelled 'Total Non Spatial' corresponds to all operations performed when completing 'non-spatial' tasks, namely, the 'identify non-spatial' and 'distinguish non-spatial' tasks. For these tasks, the difference in the amount of interactive searches and filtering operations is much smaller (30.4% and 36.3%). The column labelled 'Total Spatial' corresponds to the number of operations performed when addressing 'spatial' tasks, i.e. 'identify' and 'distinguish' relative to a spatial feature.

Contrary to the values for the aspatial tasks, the spatial ones show a much larger proportion of the use of interactive structured queries (27.4% and 50.2%, respectively). The column labelled 'Total Identify' corresponds to total number of operations for all identify tasks, both spatial and not. The column labelled 'Total Distinguish' corresponds to the total number of operations for all distinguish tasks. For the 'identify', the difference between frequencies of interactive searches and filtering is rather small (30.6% and 31.6%). For the 'distinguish' tasks, there is a much larger proportion of the use of query devices (27.1% and 54%, respectively). Therefore, the higher proportion of filtering actions in the grand total arises because the query devices were used more frequently in the 'distinguish' tasks and in tasks with a spatial component.

Table 5.17 Frequency of user operations from computer logs and videos (Compare and Associate tasks)

	Total	%	Total Non Spatial	%	Total Spatial	%	Total Identify	%	Total Distinguish	%
Physical Behaviour (Observed)										
1. Search by attribute	19	2.7	8	2.9	11	2.5	10	3.4	9	2.1
2. Search by location	0	0	0	0	0	0	0	0	0	0.0
3. Search interactively	205	**28.5**	83	**30.4**	122	**27.4**	89	**30.6**	116	**27.1**
4. Filtering	323	**44.9**	99	**36.3**	224	**50.2**	92.0	**31.6**	231	**54.0**
5. Highlight records	126	17.5	64	23.4	62	13.9	58	19.9	68	16.0
6. Evaluate content	30	4.2	14	5.1	16	3.6	28	9.6	2	0.5
7. Manipulate map	16	2.2	5	1.8	11	2.5	14	4.8	2	0.5
Mental Behaviours (Observed)										
8. Decide threshold (1+4)	342	47.6	107	39.2	235	52.7	102	35.1	240	56.1
9. Select records (2+3)	185	25.7	79	28.9	106	23.8	76	26.1	109	25.5
Total (1 to 7)	**719**		273		446		291		428	

Table 5.18 shows the number of operations for the 'compare' and 'associate' tasks only. The column labelled 'Total' is the total number of operations for all 'compare' and 'associate' tasks. We can see that interactive searches and filtering account for most of the observed actions (42.5% and 39.9%, respectively). Contrary to the case of the 'identify' and 'distinguish' tasks, the difference in frequencies between these two operations is less apparent and the most frequent action is the interactive selection and search of information. The column labelled 'Total Non Spatial' corresponds to the total number of operations for 'non-spatial' tasks, i.e. the 'compare non-spatial' and 'associate non-spatial' tasks. For these tasks, interactive searches are slightly higher (47.5% and 35%, respectively). The column labelled 'Total Spatial' corresponds to total number of operations for 'spatial' tasks, i.e.

'compare' and 'associate' relative to one or more spatial features. Note that the previous pattern is reversed (36.7% and 44.4%, respectively), in other words, filtering operations were more frequent. The column labelled 'Total Compare' corresponds to the total number of operations for all compare tasks, spatial and none. Interactive searches were the most common action in this case, accounting for 45% of all total user behaviour. The column labelled 'Total Associate' corresponds to the total number of operations for all associate tasks. If frequencies are summed by tasks where an association is made, then there is little difference between the interactive searching and filtering behaviours (39.4% and 39.7%, respectively). This adds to the finding of Chapter 4 that 'compare' and 'associate' are not necessarily a logical progression in the analysis of spatial data as Knapp (1995) asserts, because the higher amount of operations required to complete the 'compare' tasks may be indicative of a higher complexity. Of interest is the fact that all behaviours are performed 28.6% more times (925 * 100 / 719) than in the 'identify-distinguish' experiment. This also adds to the results of the evaluation in Chapter 4 where it was found that 'compare' and 'associate' tasks are of a different level of difficulty than 'identify', 'locate' or 'distinguish' tasks.

Table 5.18 User Behaviours frequencies for Compare and Associate tasks (Identify and Distinguish tasks)

	Total	%	Total Non Spatial	%	Total Spatial	%	Total Compare	%	Total Associate	%
Physical Behaviour (Observed)										
1.Search by attribute	7	0.8	3	0.6	4	0.9	5	1.0	2	0.5
2. Search by location	1	0.1	0	0.0	1	0.2	0	0.0	1	0.3
3. Search interactively	393	**42.5**	235	**47.5**	158	**36.7**	233	**45.0**	160	**39.4**
4. Filtering	364	**39.4**	173	**35.0**	191	**44.4**	203	**39.1**	161	**39.7**
5. Highlight records	128	13.8	65	13.1	63	14.7	65	12.5	63	15.5
6. Evaluate content	20	2.2	17	3.4	3	0.7	12	2.3	8	2.0
7. Manipulate map	12	1.3	2	0.4	10	2.3	1	0.2	11	2.7
Mental Behaviours (Observed)										
8. Decide threshold (1+4)	371	40.1	176	35.6	195	45.4	208	40.1	163	40.2
9. Select records (2+3)	375	40.5	226	45.7	149	34.7	229	44.1	146	36.0
Total (1 to 7)	**925**		495		430		519		406	

In summary, by inspecting the frequency of graphical display usage and low level operations performed by participants to achieve a goal, it is possible to characterise how subjects addressed certain tasks and analyse how they used the software to solve them. This enables us to make some general comments. The first is that the data provide testimony to the

importance of immediate feedback from the computer tool to user actions when performing an interactive and visual exploration of a dataset. This was illustrated by the preference of the query devices to more standard interfaces at which SQL statement must be typed in full. By sparing the user the effort of writing syntactically correct queries and executing them for different values, query devices allow users to concentrate on addressing subtantive tasks.

In terms of the tasks performed it was found that, in absolute numbers, spatial tasks account for more low level operations than their non-spatial counterparts. This is suggestive not only of the higher degree of complexity that spatial tasks may pose, but also of the different approach that users take when solving them. This is discussed in the following section. Also note that participants made considerable use of the query devices in order to solve spatial tasks. This indicates the benefit that software tools designed for the analysis of spatial data would gain from providing similar devices.

In absolute numbers, 'distinguish' tasks accounted for more low level operations than 'identify' ones and once again, the query devices were the most frequently used graphical widget. In addition, the map was used more for the 'identify' tasks. This is not to imply that visual variables are not used for distinguishing but rather that if users are presented with other graphical devices that facilitate the operation in some way, they may be likely to use them more frequently and even prefer them to using colour coding for assessing, or in this particular case, distinguishing values. Nevertheless, some of the most successful strategies for spatial tasks analysed below included the use of the map's colour coding as a component. Finally, 'compare' tasks required more manipulations of the graphical displays or devices while the map views were more frequently utilised in tasks of association. These remarks are not intended in anyway to assess the validity of Bertin's visual variables but rather to suggest that their efficacy may be mediated by both the graphical display in which they are embedded and the task that needs to be addressed. Although this point is not discussed further in this dissertation, it constitutes a relevant topic which is a good candidate for formal experimental design and usability testing.

5.3.2 Articulate and Evaluate search cycle

When presented with complex visualizations, Trafton *et al.* (2000) found that users would not only interrogate the visualization in the hope of seeing something of interest, but would also directly search for the information that they wanted to extract. Therefore, the evaluation is frequently done against a background of *a priori* expectations. For instance in our experiment, when commenting about a task in which a user had chosen an inappropriate variable to

inspect, he explained: "*When I was doing the task I was not obtaining what I was expecting to see. There was something strange going on. The reason was that I was using the wrong data on the PCP*" (User 1). This emphasises that users may have more or less well-formulated goals in mind when they investigate the data and that their prior knowledge will influence the structure of the search.

Graphical comprehension is thus affected by both the display characteristics and the viewer's prior knowledge. When activated, domain knowledge will guide the processing and encoding of graphical information which in turn may influence the findings made by subjects. Shah (2000) suggests that when viewers have particular expectations, they are likely to try to confirm those relations and may ignore points that are inconsistent with the anticipated trend. This domain knowledge may even lead to systematic biases and errors in interpretation. For instance, in our experiment, User 6 had great difficulty solving a task about public transport in Los Angeles, California. He was expecting areas of higher population to also be high in public transport usage. However, this pattern is not always observable in Los Angeles because of the very high use of private transport. The subject later expressed that he felt lost and confused because he could not confirm his prior expectations. Furthermore, he ignored the other transport variables available in the dataset (such as that quantifying private transport use by area) which could have suggested a different explanation.

This is similar to Chin and Brewer (1992) who found that, in dealing with anomalies, novices may simply ignore them. For instance, in our experiment, not all users noticed that the tracts with higher public transport usage were not only low in population but also badly connected to the public transport network (which included buses, trains and subway). On the contrary, experts focus their attention on anomalous data possibly for the very reason that relevant explanations are activated in experts and not in novices (Freedman and Shah 2002). Users 1, 3 and 8, who were the most experienced in terms of GIS and database use, particularly commented on the oddity of such areas. On the other hand User 5, who works in the field of urban planning, focused on this aspect of the data to hypothesise about the dynamics of the city in terms of the origin-destination flows of residents to and from particular areas.

Nevertheless, prior knowledge will only be activated by the graph if it is accessible and if it is relevant in a specific context (Freedman and Shah 2002). For instance, some of the participants had not used Census data for a long time. They commented that it was difficult for them to understand what records in the dataset meant and hence query the data in search for information. Regarding a question where users had to specify the relation of one variable to three ethnic groups (Whites, Hispanics, and Blacks), User 2 commented: "*I haven't dealt*

with Census data for the last couple of years... I always wanted to choose [data] by ethnic origin but the variables are numbers of the ethnic group in a tract. I can't know all the Whites, or all the Hispanics [that, for instance, use public transport]. All that I can know is that a tract is high in the number of Whites. The data set slowed me down; my understanding of the dataset slowed me down". This user eventually understood the variables and found the information she was looking for but had difficulty remembering how to interpret Census data. By contrast, viewers that do not have particular expectations might tend to describe local maxima and minima. For instance, in a question in which users had to compare the Hispanic and Black populations, User 5 quickly generalised that in all tracts the former ethnicity was always more represented than the latter. This was not necessarily the case in all areas, which could have been noted by inspecting any of the visualizations more closely.

User behaviour for resolving the tasks could generally be described as a cycle between articulating and evaluating a search (Figure 5.6). The articulation phase mostly includes the selection of the views that the subject uses for solving the task, the variables considered and the selection of a subset of the data, for instance using a structured query or interactively selecting records. Once the search is articulated and a set of candidate records is selected, users evaluate their results by inspecting and comparing the views previously chosen and examining values in greater detail. Some of the users would typically loop a greater number of times between the articulation and evaluation of a search while refining their questions. Sutcliffe *et al.* (2000) found that major determinants of search success included users searching in iterative cycles and carefully evaluating the relevance of the records they selected as answers to the questions posed by the experiment. Search strategies were a major determinant of search performance. The following sections show how iterative cycles were a crucial component of more elaborate answers. On the other hand, poorer performances were attributable to difficulty in assessing the results from a search, difficulty in understanding the data provided, or a lack of careful inspection of some graphical displays. Poor motivation, which has been found in some studies to be a cause of poor performance, was ruled out in this case since users participated in the study knowing that it demanded a good deal of time and that there was no monetary remuneration to their taking part. Hence, their willingness to participate was an indicator of their interest in the experiment or the software they would use.

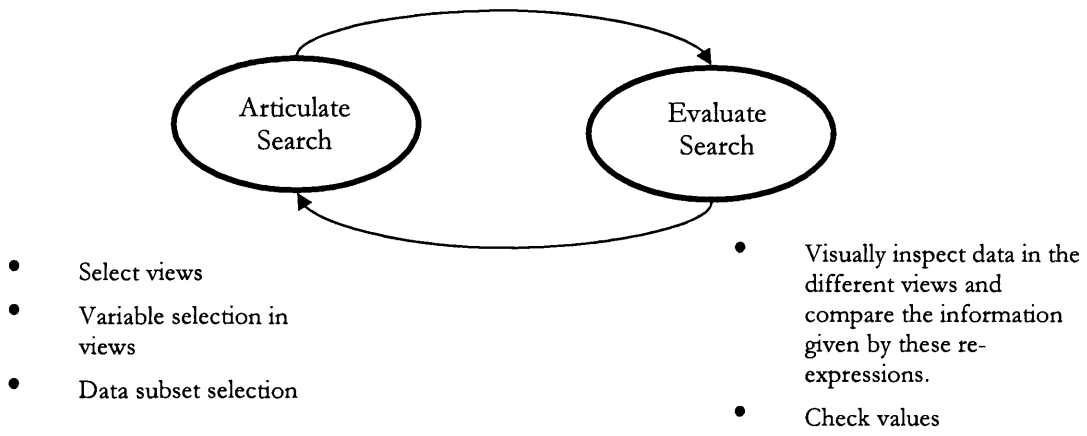

Figure 5.6 Articulate and evaluate search cycle

In order to analyse patterns and search strategies, behaviours were sequentially recorded to investigate transitions between specific visualizations and low level operations with the interface. Thus, in order to evaluate participant strategies it was necessary to code how many times each of the activities in Table 5.15 followed each other. The following sections analyse some of the more elaborate strategies adopted for solving the tasks. Note that the 'identify non-spatial' tasks is not considered below. This task was a fairly simple one and most users adopted very similar strategies. Also, because it was simple, there was no room for users to alternate or try different strategies because they typically found the correct answer on their first attempt. Hence, it is not considered further.

5.3.2.1 IS task

The spatial 'identify' task asked users: *Where are the tracts with the most users of public transport? Please describe their general location.* To solve this task, User 5 first explored the data using structured queries (selecting by attribute in the GIS). She first randomly chose any value from the list provided in the attribute selection window to perform the query but decided the result did not provide a meaningful answer. The second time around the articulate-evaluate search cycle, she carefully selected the value that would determine the subset of data she would explore. In order to further inspect her selection, she used colour coding in the map to easily pick out areas where the level of public transport was high and decided that her selection was appropriate. Figure 5.7 illustrates in a diagram how this user's behaviour is reflected in the sequence of graphic displays utilised. Thus, when a structured search did not prove successful, she rapidly moved to colour coding and visual inspection. It was therefore observed in this and other user behaviours that the articulate-evaluate loop could be associated with changes in data exploration strategy.

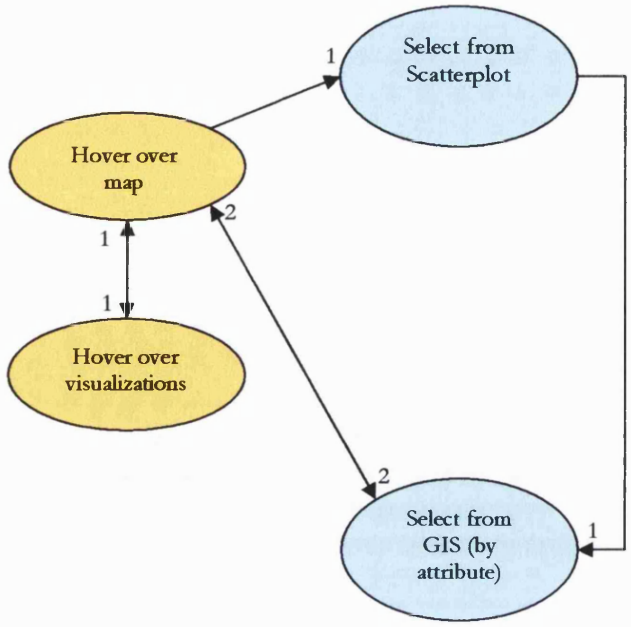

Figure 5.7 User 5 sequence of graphic display use for task IS

Before continuing her exploration, this user reflected on the adequacy of a colour coding scheme—which showed proportions as opposed to counts of the variable explored—and its effect on her attempts at visual corroboration. She then decided to change the variable being portrayed. This was an interesting behaviour as the inspection of variables and the strategy shifting made the user consider the nature of the data and its implications for the answers being reached.

The second part of the question required the user to describe the common locational characteristics of the tracts that she had identified on the basis of attribute values. Because they were not clustered in one area, she was forced to think of a reason that would explain their location. Hence, her strategy was to visually evaluate her answer against other layers of information provided in the GIS. She concluded that she could explain the pattern of public transport use by the location of other geographical features such as universities. Therefore, this user not only looped at least three times between articulating and evaluating a search, but she also corroborated her answers by visually comparing them to spatial features. In addition, she used the map's colour coding as an aid to cross check that she was considering the correct tracts in the analysis

5.3.2.2 DN task

The non-spatial 'distinguish' task asked users: *Is there a difference in the household size of the three ethnic groups?* User 1 gave the most complete and in depth analysis of all answers to this

question. The main reason for this was that he iterated seven times around the articulate-evaluate search cycle, systematically considering the different possible household sizes (1-6 people) for each ethnic group. Hence, the user would first set the view to show the variables of interest, make a selection on that view to concentrate on a subset of the data, consider the selection and then narrow the search if necessary. It must be noted that he only used interactive selections from a non-spatial view to solve this question, and that these were sufficient to provide him with an acceptable answer. Figure 5.8 shows that although the strategy adopted was very simple, the number of iterations compensated for its lack of sophistication and allowed a refined answer to be provided. The large number of articulate-evaluate loops, however, arose partly because the user initially selected the incorrect set of variables to explore. Nevertheless, his systematic consideration of the results obtained at each step allowed him to realise the problem, amend it and construct a good hypothesis.

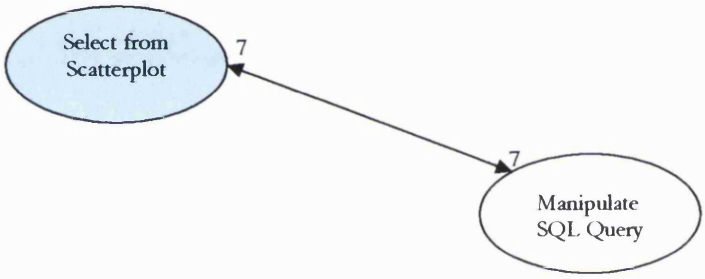

Figure 5.8 User 1 sequence of graphic display use for task DN

5.3.2.3 DS task

The spatial 'distinguish' task asked users: *Where do middle age and older people tend to reside?* Similar to the previous case, User 7 iterated six times around the articulate-evaluate loop. He would mainly focus upon the data of interest using the query devices and then select the data he wanted to explore using a scatterplot. However, in order to corroborate his answer, he would assess it visually on the map. In addition, because the areas that met his criteria for answering the question were not clustered, the subject attempted to find a relation between his chosen tracts and other geographical features such as proximity to the centre, to the sea or to major roads. Thus, this subject frequently used the strategy of selecting data on aspatial visualizations before corroborating the result on the map. However, he did not use the query devices before posing this question. Using this functionality cleared his views from clutter possibly facilitating his searches. Combining it with a visual inspection of the data as well as repeating the strategy provided him with an efficient and accurate way to reach a solution.

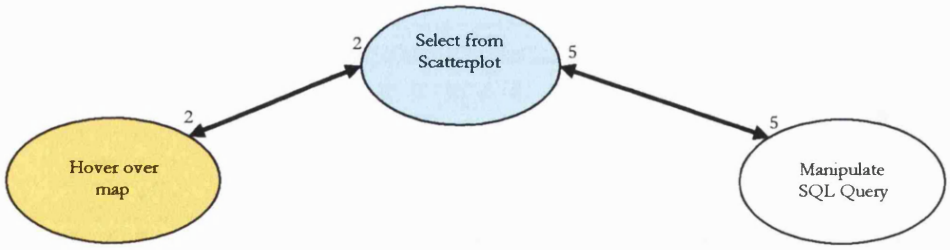

Figure 5.9 User 7 sequence of graphic display use for task DS

5.3.2.4 CN task

The non-spatial 'compare' task asked users: *Can you describe the older population in Los Angeles by ethnic origin and where they reside?* The best answer for this task was again provided by User 4. In this task the subject systematically prepared all the graphical displays he was going to use. In those displays he performed selections followed by narrowing of the data selected. In some cases however, he would select a data subset and then filter out records to narrow the search, only to refresh the views and start again. Hence, this user iterated six times around the articulate-evaluate loop in order either to test an idea or to refine a query. The map was not necessary to answer this question except for the second part of the task which forced to user to look at the map. Therefore, only at the end did the user look at the map display to obtain a picture of the spatial manifestations of his answer. This user spent a long time in this task which partly explains the more elaborate interaction with the environment, as shown in Figure 5.10.

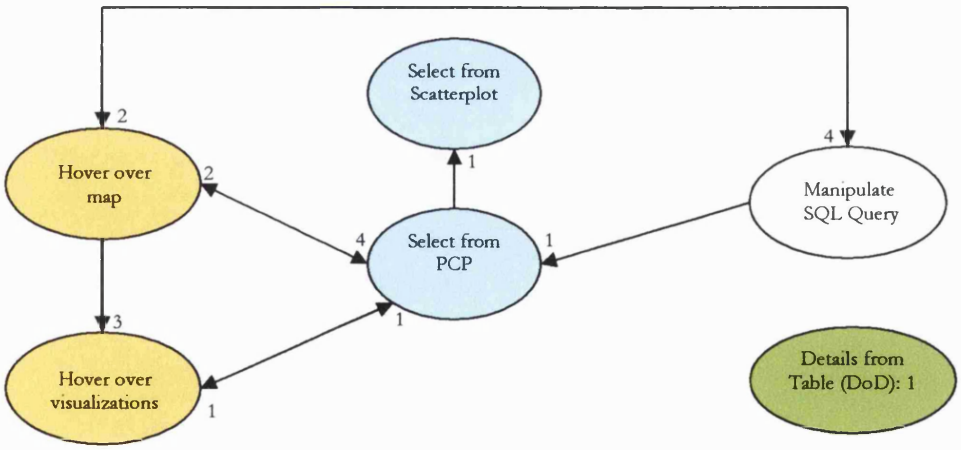

Figure 5.10 User 4 sequence of graphic display use for task CN

5.3.2.5 CS task

The spatial 'compare' task asked users: *Is there a relation between the spatial distribution of tracts with a higher percentage of Hispanic population and the distribution of tracts where the majority of its residents are*

White? User 6 not only correctly identified the location of tracts with high percentages of Hispanic and White populations, but he also investigated other combinations like areas where there was a similar mix of each ethnicity. The reason for his detailed exploration of the data was that he found what he described as a striking pattern, namely interleaving bands where higher White or Hispanic population lived, running east-west across the study area. He therefore tried to understand the ethnic composition in each of the bands he found. This behaviour is shown by the elaborate sequence of graphical displays utilised (Figure 5.11). In addition, he tried to explain their occurrence by investigating other socio-economic variables such as domestic residence ownership, as well as spatial features such as highways.

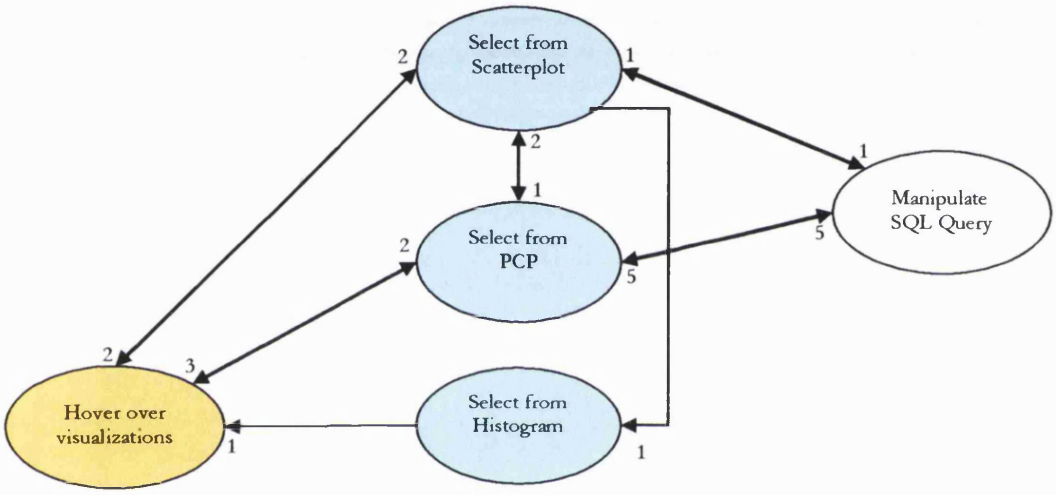

Figure 5.11 User 6 sequence of graphic display use for task CS

User 5, in contrast, mostly manipulated the map's colour coding to answer this question. She only selected data once in the PCP but instead changed the map's appearance three times to provide an answer to this question. She also visually contrasted the colour coding with other layers of geographic features such as roads and trains. She was doing so to test ideas of how the city functioned. This user was looking for evidence of previously acquired knowledge and what she called 'city borders' or geographical features that dissected the location of the different ethnicities. By exploring this idea she obtained a deep insight into the ethnic divides in the city of Los Angeles.

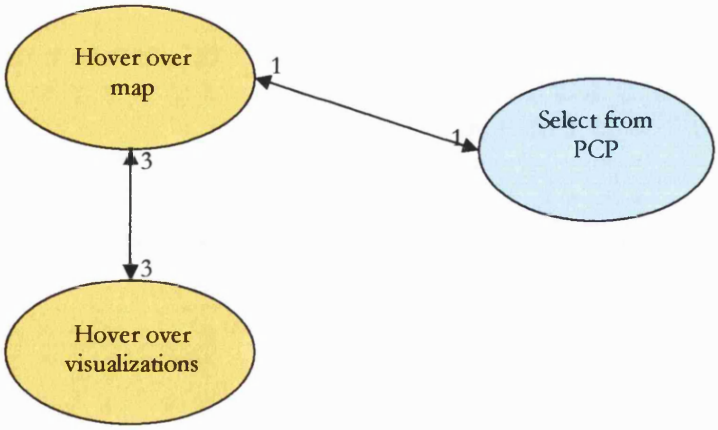

Figure 5.12 User 5 sequence of graphic display use for task CS

5.3.2.6 AN task

The non-spatial 'associate' task asked users: *Is there a relation between owner occupancy and ethnic origin and does it show a pattern in space?* User 4 started to explore this task using the PCP as several variables needed to be explored simultaneously. However, she rapidly shifted strategy and opted to use three scatterplots to address this task. She set them out systematically in order to show the variables that she considered needed to be explored. Thus the subject invested a lot of time setting up all the displays but once this was done, she rapidly solved the task by looking at each of the displays. This task did not require the use of the map except at the end where users were required to report where their selected tracts occurred in space. Hence, only at the end did the subject look at the map to address this part of the question.

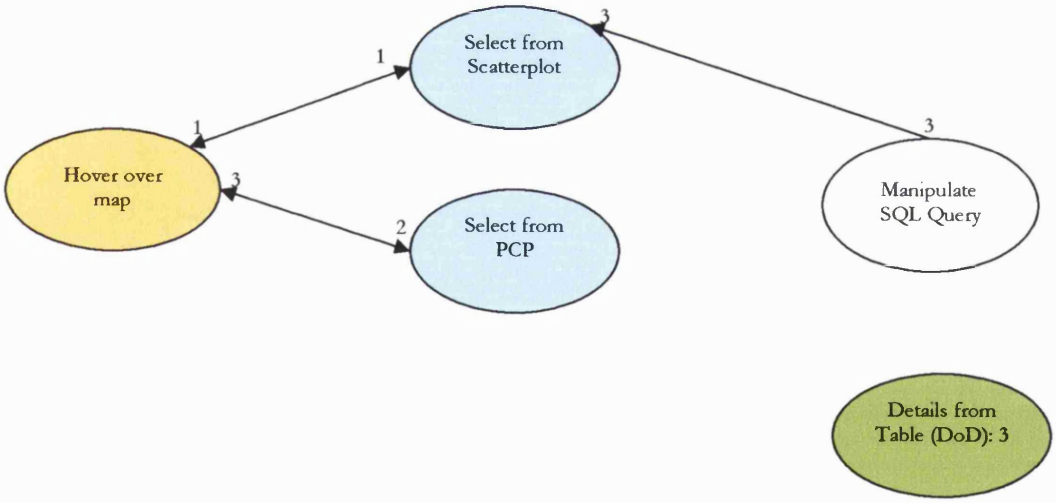

Figure 5.13 User 4 sequence of graphic display use for task AN

5.3.2.7 AS task

The spatial 'associate' task asked users: *Is there a relation between the use of public transport and the location of colleges and universities?* In this question User 7 systematically inspected the relation of

each layer of public transport (trains, subways and buses) to the location of universities in order to reach an answer. He did this visually, only using the map which was colour coded by the value of the public transport variable. On the other hand, User 4 did not make use of the areal information (the census tracts layer) for the visual inspection of the relation between universities and public transport. Instead, he used other layers such as the location of train and bus stations and the road network to address the question. He also used all other aspatial graphical displays to investigate numerical relations between variables.

5.3.3 Summary of Strategies

The analysis of user comments during the post-test interviews as well as their interaction with the software provide good insight into the strategies adopted to solve the geovisualization tasks. Two major approaches can be generalised from this which we will call systematic and holistic. Similar approaches have been detected by other studies (Cooper 1980; Carpenter and Just 1979 cited in Kyllonen *et al.* 1984a). Users who took a systematic approach to the tasks would start by selecting the views they would use and methodically set the variables they would inspect on each. They would only use views not initially set up if they were unsatisfied with the results of the search, in other words, when they changed their strategy to accommodate the lack of information obtained by their original approach. Thus, subjects frequently switched strategies to address variations in task demands (Kyllonen *et al.* 1984) and this was evident on some occasions because of a switch in the graphical display used for exploration. They would also check other graphical displays to corroborate their answer against different data visualizations.

Users who adopted a holistic approach, utilised and set graphical displays as their analysis progressed. Being less systematic, these subjects commonly used views which plotted information that was not relevant for the particular task but which was left from previous analysis. For instance, they would use the map with a colour code that corresponded to a variable investigated in an earlier task and that had no direct relation with what they were investigating later on. It was observed that, in some cases, these users were more prone to not observing relations, making analysis that were irrelevant to the particular task, or making mistakes. They would commonly loop a greater number of times around the articulate-evaluate cycle and eventually set the visualizations to show the necessary variable combinations. This would highlight the importance of automating at least some of the visualization defaults to prevent user mistakes as others have suggested (Andrienko *et al.* 2002; Sutcliffe *et al.* 2000a). This would also benefit the 'systematic' users by saving them

time in setting graphical displays while preventing those subjects who used a 'holistic' type of interaction from using views that do not portray information relevant to the task.

At least four more strategies where observed. Subjects would perform searches by: i) value; ii) location; iii) by value and then by location; iv) by location and then by value. In particular, subjects would tend to use aspatial displays (PCP, scatterplot, histogram, query devices) first for tasks which inquired about attribute values. Having selected a subset of candidate records that met some criteria that answered the question, they would commonly check how the answer corresponded with the map in order to see the manifestation in space of the particular relation found. Very seldom was the map used as an exploration tool. By far the most common strategy entailed selection of attribute values in one of the non-spatial graphical displays and then using the map to verify the result. Nevertheless, it was also observed that users searched the map for information using either its colour coding or selecting a location for detailed analysis. In the latter case they would choose some areas and then inspect their attribute values directly on the map or in one of the other geographical representation.

In most cases subjects would iterate more than once between any of these strategies. Selections of data subsets by value were more frequently performed using the query devices as shown in Table 5.16 but they were also preformed using the structured query dialog window ('select by attribute') available from the GIS. When selecting by value first, users would typically start by narrowing the subset of data that they considered being of interest by either filtering non-relevant data from the views or selecting records in one view and only considering that selection. Narrowing the search involves the implicit definition of a threshold that reflected a numerical criterion which assisted in choosing data of interest. In some cases, such numerical criteria reflected a categorical decision such as selecting high or low values of a particular attribute. When this led to a search by location, the defined selection would be inspected visually on the map in order to detect patterns or evaluate the relation between spatial features. However, the selection could also be inspected in other graphical views to check the relation of the answer to other variables or simply corroborate that the same answer was being shown in other displays. Strategies that began with a location based search were few in number, as aforementioned. They however would be started by either making a spatial selection using the GIS, or by zooming in on one area of the map where the user was expecting to operate.

5.4 The relevance of strategies for information processing

Strategy differences have been studied in cognitive psychology in the context of spatial task performance. The literature identifies at least two kinds of spatial tasks: spatial visualization and spatial relation tasks. The former is defined as "one of several correlated abilities concerned with the encoding, transformation and recognition of spatial information [that involve] the mental manipulation of the entire spatial configurations, often by changing the relation of elements to one another" (Salthouse *et al.* 2000: 188). Spatial relation tasks require judgement as to whether two figures are the same and the comparison typically requires performing a mental rotation of one of the figures. In the context of geovisualization, users manipulate the data via graphical displays or graphical widgets. As Just and Carpenter (1976) suggest, this manipulation provides an indication of user mental processes. Similar to the spatial visualization tasks of the experiments in cognitive psychology, the interactive manipulations of the graphical displays in the experiments reported here also change the relations between objects on the screen. Hence, they are also concerned with encoding, transforming and recognising some spatial pattern. Alternatively, they may help to verify or contradict some idea or hypothesis that the subject might obtain through data exploration. Therefore, results from the cognitive psychology literature may enrich our analysis and provide avenues of further research.

As aforementioned, tests of spatial visualization in cognitive psychology require the execution of a series of mental transformation on stimuli. This is the case, for instance, of paper folding tests where a subject is first given an image of a paper with a hole punched in it. The subject must determine whether the following figures she is shown correspond to folds of the same piece of paper. Thus, the output from intermediate transformation steps must be stored temporarily while some other information is processed. In addition, the products of early transformations must be maintained during the execution of later transformations. The sequences of transformations are necessary so that the subject can tell whether the first and last stimuli are the same by bringing different parts of the figure into correspondence with one another. Similarly, in the tests discussed throughout this dissertation, once a user manipulated any graphical widget, the combination of information the subject had on the screen changed and gave way to new relations between attribute values. Thus, the environments tested also forced users to remember patterns or relations while processing some other information. Unlike the classic cognitive psychology experiments, the user could go back to previous scenarios or combinations of images on the screen. Nevertheless, they still had to remember how they obtained a particular pattern in a view or combination of views.

One approach to explaining individual differences has been to take into account subject spatial abilities when explaining the strategies they adopt for solving spatial tasks. Spatial ability can be defined as the aptitude to generate, retain and transform abstract visual images (Kyllonen *et al.* 1984a). The cognitive study of abilities therefore aims at identifying mental structures, processes, representations and strategies that underlie test performance (Burin *et al.* 2000). This might be a plausible avenue of research for explaining the different strategies adopted by participants in our experiment, as well as the wide range of time performances and the significant main effects of subjects on such measure (see Figure 5.4 and Figure 5.5). However, there are contradictory results in the literature about the choice of strategy being related to the spatial visualization ability level (Chen and Yu 2000; Salthouse *et al.* 2000) or even to gender (Burin *et al.* 2000).

Salthouse *et al.* (2000), for instance, tested three possible sources of individual differences in spatial visualization ability. First, the authors suggest that they might be attributable to variations in the ability to preserve the quality of the spatial information during the aforementioned transformations. This hypothesis would suggest that differences in spatial visualization ability occur because of variations in respondent effectiveness to generate accurate and complete internal representations of spatial information. The second hypothesis suggests that individuals may differ in their efficiency for executing a transformation. This is to say that individuals differ in terms of their speed in carrying out mental manipulations. A third hypothesis suggests that people can vary in terms of their abilities to preserve the representations during the transformations. Salthouse *et al.* (2000), however, could not find any evidence to support any of these three hypotheses. Nevertheless, they did find that individual differences in spatial visualization ability differ when other processing operations must be performed, even when the concurrent operations are not spatial transformations.

These results highlight the discussion in Chapter 2 (§2.4.1) about the merits of graphic displays in problem solving. External representations not only function as memory aids, but also facilitate inference by allowing the direct recognition of features (Larkin and Simon 1987), as well as the access to knowledge and skills (Reisberg 1987). In addition, graphic displays may determine the information that is perceived as well as the structures that are discovered from a particular representation and the information processes that are activated (Zhang 1997). Since subjects had access to the different displays whenever they needed them, individual differences due to the quality (i.e. accuracy and completeness) of the information incorporated in the internal representations of spatial information, the speed of

executing relevant transformations, or people's abilities to preserve a representation during transformations, would be less relevant. However, the operation of interactive systems or tools may demand other processing activities that can interfere with the visual interpretation of data. For instance, users need to decide what attributes to plot, are constantly manipulating the graphical displays and data objects, as well as processing the information they obtain in search for insights or hypothesis. Thus, individual differences in keeping track of a representation while it is transformed, maintaining a representation after it has changed, remembering the changes made to a display, comparing displays after a transformation, retaining a representation while viewing a second display, or maintaining detail and preserving information during following explorations, are of importance when using geovisualization systems. It was shown in Chapter 2 that if the system where these displays are embedded does not facilitate a seamless and dynamic interaction it would force users to allocate their cognitive resources to activities (such as operating the system) that are not necessarily conducive to the understanding of the data.

Salthouse *et al.* (2000) used the desktop analogy of working memory (Broadbent 1971) to reconcile these findings. The metaphor suggests that there can be a trade-off between storage and processing because the greater the 'surface area' devoted to storage of some information, the less is available for working on processing. The space on the desktop is partitioned into separate regions for storage and processing and the same portion of the desktop is available to each of these activities. A similar analogy would explain individual differences when processing and storing operations are performed simultaneously as a problem of "maintenance of information when storage space is restricted" (Salthouse *et al.* 2000: 228). This stresses the importance of good tool design that can facilitate its operation in order to minimise the amount of resources that users have to allocate to activities that may interfere with their goal of deriving hypotheses from the data explored.

Empirical studies that focus on the strategies employed by users when interacting with visualizations for exploring or searching for information may also help address the question of the effectiveness of visualizations for supporting particular tasks. However, there are few such studies (and none to the best of our knowledge for geovisualization) with which to compare the results obtained here as their results have reached very different conclusions. Chen and Yu (2000) undertook a synthesis of all empirical evaluations reported in the information visualization literature, which included an experimental design very much in the spirit of those discussed here. The purpose of their review was to identify invariant relations suggested collectively by the empirical findings of the studies. Chen and Yu (2000)

concentrated their synthesis on information structures visualized as trees and networks which were used for information retrieval tasks. In addition, they concentrated on studies that included measures of accuracy or efficiency as dependent variables. Adopting these selection criteria, they were left with a sample of only six studies out of an original 35 which were concerned with the effects of visual interfaces and user cognitive abilities upon the accuracy or efficiency of information retrieval. By combining the results of the six studies, they were able to confirm two hypotheses suggested in the information visualization literature. First, that cognitive ability has an effect on accuracy when using visualization interfaces. This is in contrast to the finding in the study by Salthouse *et al.* (2000), which suggests that people varying in spatial visualization ability do not differ in the accuracy of decisions involving spatial information. Second, that a user with stronger cognitive abilities performs more efficiently than one with weaker abilities. This is also contrary to the findings of Salthouse *et al.* (2000: 201) that "differences in visualization ability are not associated with variations in the efficiency of executing transformations, or in the duration of processes associated with encoding" spatial information, in the speed of executing relevant cognitive operations or in the precision or amount of information that is remembered.

Finally, the combined evidence from the studies reviewed by Chen and Yu (2000) seems to suggest two further hypotheses. First, that given equal levels of cognitive ability, users tend to perform better when using less sophisticated visualization interfaces. This is not surprising if we take into account the fact that users must be trained in the use of a new system or interface, as has been previously discussed in Chapter 4 (§4.2.4.1 and §4.6). This imposes major restrictions upon evaluations that involve users as they tend to have restricted time schedules, which implies that only limited functionality can be demonstrated and tested. Provided that users have the same experience and limited time to do a test, it would be expected that the simpler the interface the easier it would be to master within the time allocated to the experiment. The second hypothesis is that results regarding whether visual-spatial interfaces enable users to perform better than traditional retrieval interfaces (for instance using command line) in terms of accuracy or efficiency were heterogeneous. This also takes us back to the discussion of Chapter 2 (§2.4.1) regarding the effectiveness of visual and interactive displays to their static counterparts were it was discussed that the suitability of a representation depends not only on the task that must be supported but on the data that the display must portray.

5.5 Conclusion

The experiments reported in this chapter were more concerned with determining whether participants used alternative strategies to accommodate the difficulty of some tasks than with determining how their spatial abilities might have affected their performance. It has been suggested that strategies may affect subjects' performance and that they are related to their aptitudes (Cooper 1980) because "different strategies involve different solution algorithms, so the cognitive components could be different for each strategy" (Burin et al. 2000: 276). In addition, variations in spatial performance which have been previously linked to subject characteristics and abilities have been found to be "linked more directly to differences in solution strategy" (Schultz 1991: 489). Therefore, studying user strategies in a geovisualization context may not only provide a deeper understanding of subjects thought processes when interacting with spatial data but they may also expand on the findings of the existing literature.

It has been suggested here that strategy variation is particularly important in assessing performance in relation to spatial visualization tasks, since these are particularly susceptible to strategy shifting (Lohman and Kyllonen 1983). It was indicated that users not only performed operations in different sequences, but also differed in the steps included or excluded from the strategies. This was shown to influence the time taken to complete different tasks. Therefore the presence of strategies made the analysis of the differences between the so-called spatial and non-spatial visual operators more complex. Nevertheless, we were able to establish the increased difficulty that visual spatial tasks posed through the analysis of frequencies and sequences of interaction with the software used for the testing. These issues make the assessment of strategies in a visualization context an interesting route of research.

However, the role of strategies in spatial performance is as yet not fully understood, and our approach here may contribute to the ongoing debate. Although subject use of different solution strategies and their effect on increasing the variability of spatial visualization performance has been demonstrated, studies have used different categorizations of solution strategies as well as different spatial tasks. This can make the comparison of results or their generalisation to a wider population difficult. Thus, this experiment has provided further evidence with regards to the importance of characterising geovisualization tasks in a manner that supports formal experimental design. This may not only allow the evaluation of software and geovisualization concepts but also facilitate the comparison of studies and the building of a solid body of evidence regarding the characteristics and functionality that

geovisualization systems should provide for facilitating data exploration and hypothesis testing.

Although this chapter highlighted the importance of individual differences in the strategies adopted by individuals when attempting to solve visualization tasks, "studies that have examined the relationship between strategy and ability suggest that a subject's ability profile plays a role in solution strategy choice and efficiency" (Burin *et al.* 2000: 276). This constitutes in our view an important avenue of research that may lead us to a deeper understanding of user cognitive processing and may facilitate the development of guidelines for tool development. Thus, although not attempted here, classifying people by their abilities would be necessary. This, however, demands a battery of ability tests to be administered which have been developed in psychology to address rather specific research questions. Pursuing this thread of analysis would therefore require close cooperation with cognitive psychologists who can contribute with their expertise to the design of experiments that can help us understand whether spatial abilities have any relevance in a geovisualization context.

6 Data Description and Considerations

As discussed in the introduction, this dissertation has been fundamentally concerned with visualization and user interaction. The experiments discussed in the previous three chapters were nevertheless grounded in the real world and the tasks given to participants had substantive implications. Census and associated socio-economic data were used for the three experiments described in Chapters 3 to 5 and the tasks administered were designed to reflect operations or activities that potential users may need to support.

The environments tested were designed to support data in a vector format. Therefore, the nature of the data used embodied an object digital representation of the world where its elements are viewed as a series of entities located in space (O'Sullivan and Unwin 2003). Thus, the nature of the data and the environments evaluated can be expected to have an important effect on the results and insights attained from this research regarding geovisualization and user interaction. The specifics of the datasets used, however, are of no particular relevance to the work described. All of the respondents had a generic familiarity with Census data and had no particular knowledge of the case study areas. Therefore, only a passing reference to the datasets used has been provided thus far. Nevertheless, having described the three sets of experiments conducted, it is appropriate to discuss the general characteristics of the data used and how some of their particular characteristics might have impacted upon the results obtained.

The experiments discussed in Chapters 3 and 4 were grounded in a UK setting and used a hybrid dataset comprising 1991 Census and 1996 lifestyles data for Bristol. It was planned that during the experiment users would explore the relation of income to other socio-economic measures. Hence, data from the lifestyles survey was fused to the Census data set because the former has a direct measure of income which the latter has not (Harris and Longley 2000). U.S. 2000 Census data for Los Angeles, California were used in the last experiment (Chapter 5) which includes a direct measure of income collected by the U.S. Census Bureau. The environments tested, particularly the one discussed in Chapters 4 and 5, were best suited for vector data. Apart from the data type explored, it is possible that the choice of national setting, variable selection and dataset may have impacted upon the results of the experiments. For this reason, the datasets are described in this chapter, along with the implications they might have had in the results of the evaluations discussed.

6.1 UK 1991 Census Small Area Statistics and Lifestyles Data

The 1991 UK Census Small Area Statistics (SAS: the most recent UK Census data available at the time that this research was carried out) provide tabulations of census data at the enumeration district (ED) level. This unit of aggregation is the finest level of granularity permitted for public dissemination, for reasons of privacy and confidentiality. In 1991, "there were [...] 147,644 English EDs, 6,600 Welsh EDs and 38,254 Scottish Output Areas" (Harris 1999: 27), with an average of 148 households resident in each. Although the Census remains the best source of population data at fine spatial scales, it has at least five limitations. Most SAS variables are based upon total counts of the enumerated population per ED. However, 20 out of the 95 SAS tables are based on 10% samples. This poses limitations in the classification of households. For instance, each household can belong to one and only one of 18 socio-economic groups and four types of housing tenure. In a 10% sample, some of the 18*4 = 72 possible sub-group combinations may have a zero count (72 – 10 = 62 at the most) and hence suggest that none of the EDs population belongs to the 'empty' sub-groups when this may in fact be a chance characteristic only of the 10% sample.

A second limitation is that SAS data are subject to data-blurring methods to preserve confidentiality. Barnardization is one of these methods which randomly adds 0, 1 or –1 to the counts (Harris 1999). The effect of this procedure is that some row marginals in the SAS tables will sum to more than 100% of the ED population. Participants in the study spotted these apparent anomalies but they had to be told their source as the effect of Barnardization could not possibly be inferred from the data as provided. In addition, Census data, as of the time of its collection, becomes increasingly out of date and it is only updated every 10 years. This was not much of an issue in our case as the experiments were less concerned with the accuracy of the information derived from the dataset but rather with the usability of the environment, the tasks that users performed and the strategies they adopted for addressing the problem scenarios posed.

Information about household income has never been collected by the UK Census of Population. In order to include a proxy of household income in the dataset used for the experiments, data from a 1996 lifestyles survey were used. Lifestyles questionnaires ask about peoples' interests and activities. They collect data at the individual or household level. Each survey usually consists of a few hundred questions on varied 'lifestyles' topics ranging from income and occupation to shopping habits, holiday and travel preferences. Thus, these datasets are generally very diverse and frequently concatenated from disparate data sources using a common field, the residential address, to form a profile of individual households.

Lifestyles surveys are mostly used for direct marketing to existing or potential customers. Lifestyles data are updated annually and they collect a categorical measure of income at the household or individual level as part of each survey. The lifestyles survey used here asked from respondents to indicate 1 out of 10 possible income ranges to approximate the household income per year. The income variable used for the experiments was adopted from Harris (1999) who calculated an income score for each ED in the Bristol study area by re-coding responses from 1 to 7 as in Table 6.1.

Table 6.1 Income Ranges

Variable	Income Range	Score	Value
X1	< £5,000	1	Yes (1) / No (0)
X2	£5,000 - £9,999	2	Yes (1) / No (0)
X3	£10,000 - £14,999	3	Yes (1) / No (0)
X4	£15,000 - £19,999	4	Yes (1) / No (0)
X5	£20,000 - £29,999	5	Yes (1) / No (0)
X6	£30,000 - £39,999	6	Yes (1) / No (0)
X7	> £40,000	7	Yes (1) / No (0)

Although the frequency of the annual updates and the fine scales at which anonymised records can be made available makes lifestyles datasets a valuable sources of information, they have at least three limitations. First, unlike with the Census, there is no obligation to return lifestyles questionnaires. Hence, respondents present a self-selecting and non-random sample that is unlikely to be wholly representative of the population at large. Furthermore, respondents are not all equally likely to disclose information fully and accurately. There are actually groups that are less responsive and may be less represented. Harris (1999) also suggests that the responsive portion of the population is exclusive of the most and least wealthy. This may occur because some people may not be of interest to the marketers who finance the survey or simply because some sections of the population do not participate. Even though these characteristics of the data were explained to test participants, it was unlikely that they had this more technical aspects present during their exploration or that they considered their possible effects on the analysis made.

These three limitations likely create unknown biases in the income data and this makes their use "fundamentally more biased and unscientific" (Longley 1998). Nevertheless, lifestyles data were an interesting data option to include in the experiments for at least two reasons. First, they are updated every year, as opposed to once every ten years as in the case of the Census data. More relevant to our experiments, however, is the fact that they contain direct questions on income which the Census lacks. Including an income variable was of interest

particularly because participants to the study can identify with it. Thus, the problems with the lifestyles data were accommodated in order to make the experiments more interesting and relevant to participants of the study, particularly given that the objectives of many of the tasks revolved around good interpretation of fine scale measurements.

6.1.1 UK Census data exploration with the environment

The experiment discussed in Chapter 4 included the 19 variables detailed in Table 6.2 for user exploration. These variables were reasonably expected to be meaningful and of interest to a broad based multinational professional user community. Additional variables were considered for the first experiment (Chapter 3) but it has been previously discussed (Chapter 4) that their number was reduced to avoid an overload of information that could not be investigated in the time provided for the experiments. The data include information on the size of the accommodation and the household, the age of the people in the ED, type of property and a proxy variable for income, among others. All variables are provided as percentages of a zonal total. Participants of the experiments found data in proportions to be difficult to comprehend, at least initially. Although the data was explained in an email previous to the experiments and then again explained during the testing sessions, users struggled with data presented in this form. For this reason, the last experiment used raw data or counts instead.

Nevertheless, the environment allowed users to obtain information about the data presented as well as to quickly spot mistakes or anomalies in the data set. An example follows. Figure 6.1 (Left) illustrates the exploration of the relation between working women (*Wkwife*) with unemployment (*Unemp*) and income (*Inc*) at the ED scale (see Table 6.2 for a definition of the variables). By utilising the query devices users could filter data out of view and relieve it of unnecessary objects in order to investigate data subsets. For example, it can be seen in the parallel coordinate plot (PCP) that EDs with high percentages of *Wkwife* are also low in unemployment. In addition, these same EDs are low in the proportion of children (*Child*) and they tend to be located near the city centre. Figure 6.1 (Right) shows instead that EDs with relatively low percentages of *Wkwife* are associated with lower income and higher unemployment rates. In addition, such areas are only associated with high percentages of children in some areas (red circles in Figure 6.1, Right). On further inspection it can be seen that areas of high percentages of *Child* are not apparently associated with any particular income level but that they do tend to be located in the periphery of the city centre (Figure 6.2, Left).

Table 6.2 1991 UK Census SAS variables provided in the experiments

Variable	Definition
Bigacc (Large Accomodation)	% of households in the ED with 7+ rooms
Bighhd (Large Household)	% of households in the ED with 6+ persons
Buying	% of households in the ED in owner-occupied property buying it
Bought	% of households in the ED in owner-occupied property, bought outright
Child (Children)	% of persons in the ED aged 5-14
Youths	% of persons in the ED aged 15-24
Yngadl (Young adult)	% of persons in the ED aged 25-44
Oldadl (Old adult)	% of persons in the ED aged 45-64
Oldpsn (Older person)	% of persons in the ED aged 65-74
Councl (Council property)	% of households in the ED in council or new town rented property
Detached (Detached Property)	% of unshared dwellings in the ED that are detached
Terrce (Terraced property)	% of unshared dwellings in the ED that are terraced
Hhnocar (Household with no car)	% of households in the ED with no car
Hhtcar (Household with 2 cars)	% of households in the ED with 2 cars
Qualm (Qualified males)	% of adult, male residents in the ED who are qualified
Wkwife (Working wife)	% of married females working in the ED
Inc_Score (Income Score)	Income score for the ED, derived from the Lifestyles survey
Unfrnt (Unfurnished property)	% of households in the ED in unfurnished rented property
Unemp (Unemployed)	% of economically active unemployed in the ED

Notably, the graphical displays provided in the environment were very helpful for spotting anomalies or blunders in the data. Note for instance the PCP in Figure 6.2 (Right). It shows some zero values for *Wkwife* and *Unemp*. By highlighting the objects in the data view, the user would obtain the details of the values for every attribute in the corresponding record. For the records selected, it can be seen that most attribute values are zero (red circle in Figure 6.2, Right). This must be associated with missing values in the dataset which are all located in the city centre. As aforementioned, the graphical displays also permitted the fast identification of observations with very high values, such as percentages that were apparently greater than 100%.

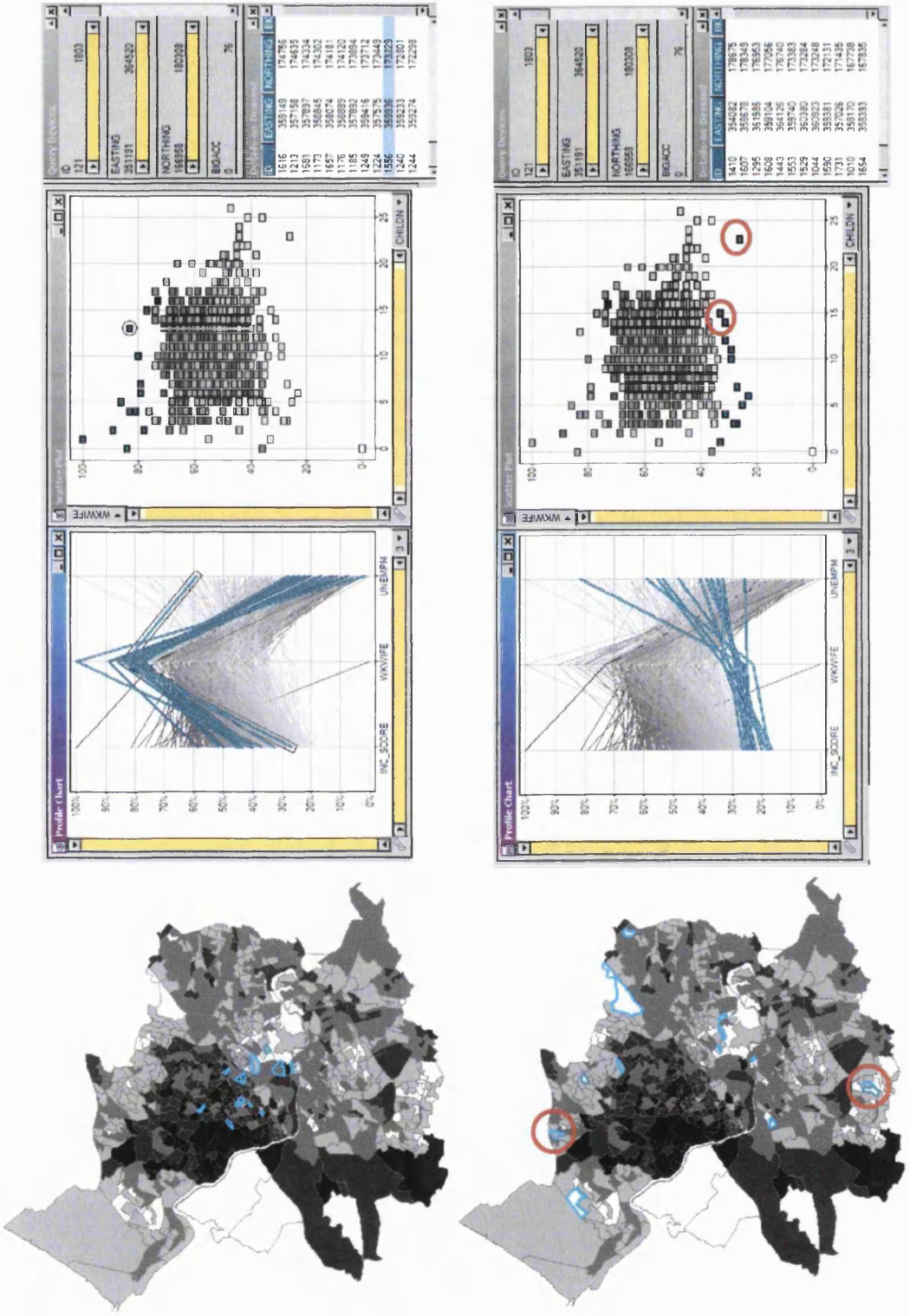

Figure 6.1 Data Exploration of UK 1991 Census data for Bristol: Left) High proportion of *Wkwife* compared to *Inc_Score*, *Unemp* and *Child*. Right) Low proportions of *Wkwife* compared to the same four variables.

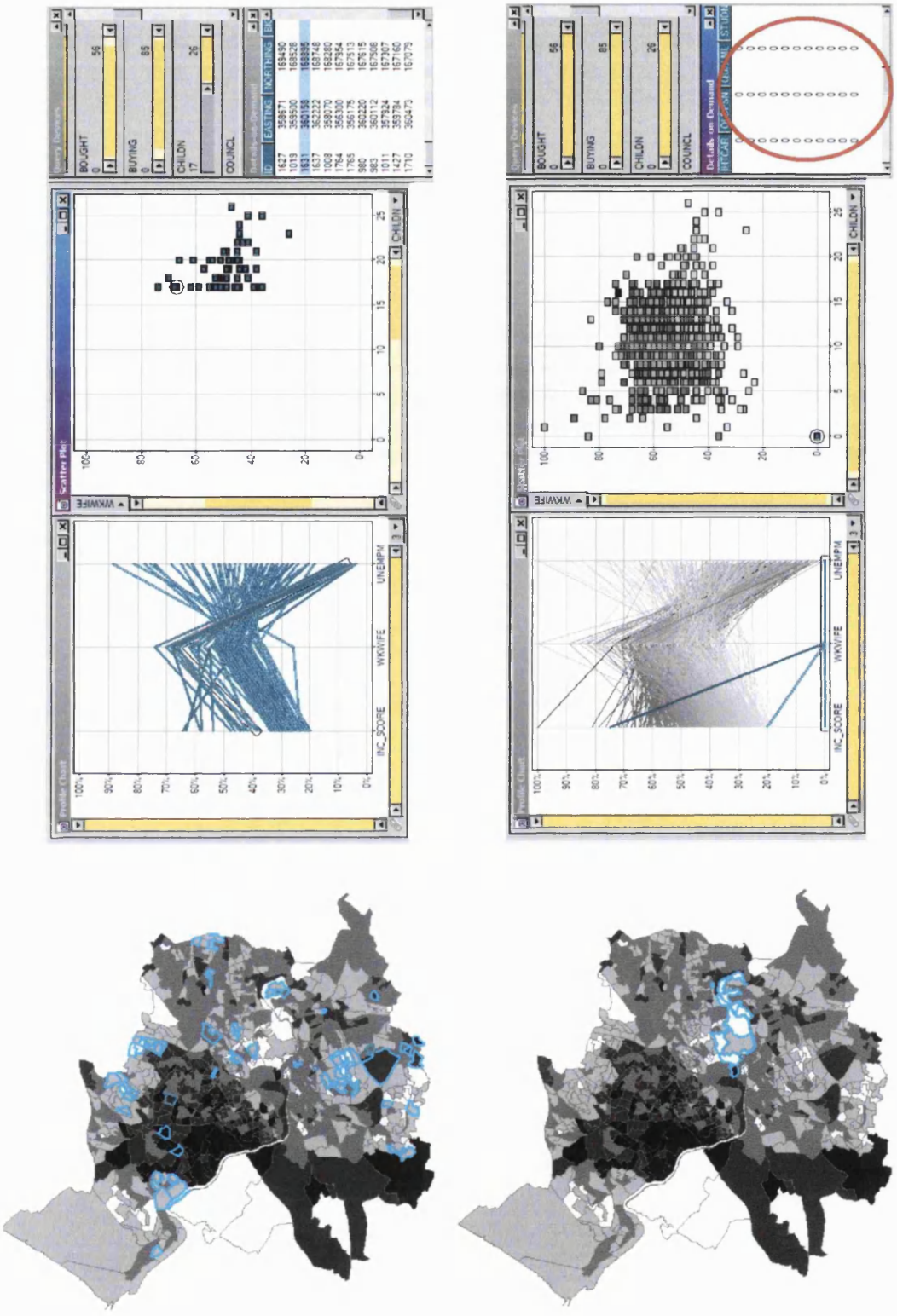

Figure 6.2 Data Exploration of UK 1991 Census data for Bristol: Left) High proportion of *Child* compared to *Inc_Score*, *Unemp* and *Wkwife*. Right) Blunders in the dataset.

It was found in the evaluations that the graphical displays provided within the coupled environment are useful in highlighting values at the extremes of observed distributions. Hence, the first two exploration examples of Figure 6.1 showed selections of high and low percentages of *Wkwife*. Patterns in the mid value ranges of attributes were in most cases less evident. Thus it was possible that interesting observations about relations between mid-range values of variables would only be spotted if the user decided to explore the whole attribute range in more detail. This is partly because of the large number of records in the dataset. As the number of observations increases, so the visual representations rapidly become cluttered with objects. This problem is not specific to the environment tested here, but rather is typical of analysis tools that use graphical displays, as discussed in Chapter 2. Query devices, however, are seen as a useful tool for bypassing this limitation and allowing users to filter in and out records very rapidly.

The environment was also successful in supporting the exploration of the relations between variables both in attribute and geographical space. As it has been previously discussed, participants made use of several displays simultaneously to obtain information on the dataset. By interacting with data objects in any of the representations, corresponding objects in other views would be highlighted or their visual appearance changed in some way (e.g. change colour). In this way, participants were able to use either the map or any other aspatial representation to explore some aspect of the data. Thus, users were not only able to interpret aspatial associations between attribute values but also to inspect their patterns in space.

6.2 U.S. Census 2000

Information about 115.9 million housing units and 281.4 million people across the United States were gathered by the 2000 Census. The census questionnaire asked seven questions of every person and housing unit in the U.S. (also known as 100% characteristics or short form questionnaire) including sex, age, race, Hispanic origin and whether the housing unit was rented or owned. Additional questions were asked of about 17% of households (known as sample characteristics or long form questionnaire) or to about one in every six households in the U.S. These households received the long form questionnaire which, apart from the short form questions, included questions requiring details about ancestry, income, mortgage and size of the housing unit. It was estimated that about 80% of all households in the U.S. were delivered questionnaires. Of these, the national final response rate was 67% (U.S. Census Bureau 2002).

The 2000 Census data are provided at several levels of aggregation: blocks, block groups, tracts, places, counties, metropolitan areas, urban areas, states and voting districts, among others. Blocks are the smallest geographic unit for which the Census Bureau tabulates data. They are generally bounded by streets, legal boundaries or other features. There were approximately 8.5 million blocks identified in the 2000 Census (U.S. Census Bureau 2001). Block groups are a collection of census blocks within a census tract. Tracts are the smallest subdivision of counties averaging around 4000 people, and were first designed to have relatively homogenous demographic characteristics. Data aggregated to the tract level was used for the experiments in Chapter 5.

In order to protect respondent confidentiality, "the Census Bureau withholds data if they represent a population (within a geographic area) so small that the numbers might identify someone" (U.S. Census Bureau 2002: 5) and there is no public access to individual information for a period of 72 years. This introduces randomisation errors to the data similar to those of Barnardization in the 1991 UK Census data. Further sources of error were likely introduced by the automated technology used in data capture. This technology was prone to errors such as the failure to read a field on the form, reading content that was not intended as an input to a field, incorrectly capturing the content on the paper, or correctly capturing content that nevertheless entailed some kind of response error (Conklin 2003).

6.2.1 U.S. Census data exploration with the environment

Most of the variables used for the experiment were extracted from the U.S. Census Bureau Summary File 1, which gives counts and basic cross-tabulations of information collected from all people and housing units. It includes information on age, sex, race, origin, household relationship and whether a household's property is owned or rented. The variables selected for the experiment discussed in Chapter 4 are detailed in Table 6.3. Variables from the short and long form were included in the experiment. They include information on counts of population by ethnic origin and tract, age of the population by tract, income and main mode of commuting to work, among others.

Diverse measures of income were provided in order to allow users to explore differences in their relations to other variables in the dataset. Also, most of the variables were provided as counts except for three: *pWhite*, *pBlack* and *pHisp*, which were the proportion of each of the three ethnicities (White, Black and Hispanic) of the total population in the tract. It was noted in the experiment discussed in Chapter 5 that understanding and using these variables was simpler than when similar ones were used for the Bristol dataset. We believe this is because

the total population (per tract) was provided in the third set of experiments and users could confirm that *pWhite*, *pBlack* and *pHisp* were simply the value of each of the ethnicities divided by the total population in the area. This would indicate that users also utilise the data as presented in table form to investigate the data and the meaning of variables in detail.

Table 6.3 U.S. Census 2000 variables provided in the experiments

Variable	Definition
Pop2000	Total population in the tract in 2000
White	Total White population alone in the tract
Black	Total Black population alone in the tract
Hisp	Total Hispanic population alone in the tract
pWhite	Proportion White population in the tract compared to the total (*White / Pop2000*)
pBlack	Proportion Black population in the tract compared to the total (*Black / Pop2000*)
pHisp	Proportion Hispanic population in the tract compared to the total (*Hisp / Pop2000*)
Age_under5	Male and female under 5
Age_5_17	Male and female 5 to 17 years
Age_18_21	Male and female 18 to 21 years
Age_22_29	Male and female 22 to 29 years
Age_30_39	Male and female 30 to 39 years
Age_40_49	Male and female 40 to 49 years
Age_50_64	Male and female 50 to 64 years
Age_65_up	Male and female 65 years and over
Med _age	Median age, both sexes
Households	Total households
Ave_hh_sz	Average household size
Families	Families
Ave_fam_sz	Average family size
Hse_units	Total housing units
Owner_occ	Housing units, owner occupied
Renter_occ	Housing units, renter occupied
Vacant	Vacant housing units
AggHhdInc	Income of the householder and all other individuals 15 years old and over in the household, whether they are related to the householder or not
medianHhdInc	Median household income based on the distribution for the total number of families including those with no income.
AggFmlyInc	Sum of all incomes of all members 15 years old and over related to the householder
medianFmlyInc	Median family income based on the distribution for the total number of families including those with no income
AggSalaInc	Aggregate salary income which includes total money earnings received for work performed as an employee
PublicTran	People commuting to work using public transport (bus, train, metro)
Car	People commuting to work on private transport (own car or car pool)

Figure 6.3 gives an example of the investigation of the relation between four variables: *White*, *Black*, *Hisp* and aggregate salary income (*AggSalaInc*: see Table 6.3). Note that areas where a large number of *White* population live tend to have very small amounts of *Black* residents. In

addition, areas where the *Black* population is high tend to be low on both *White* and *Hisp* population. This would suggest that each of the ethnicities tend to be located in distinctively different areas of Los Angeles. However, on closer inspection it is possible to find that medium to high numbers of *White* population may live in tracts of medium to high *Hisp* population.

6.3 Conclusion

The geovisualization tasks considered in the preceding chapters and the lessons learned about how users interact with software when visually processing spatial information are thought to be valid beyond the scope of the experiments and the particular dataset used. It was shown, for instance, that the results were found to be robust to issues such as GIS expertise or frequency of computer use (Chapter 4). In other words, these variables were not found to have any effects on the time it took participants to complete the tasks, or to have any interaction with the factors included in the experiments. Although generalising the results in Chapters 4 and 5 to the population from which the sample participants were drawn cannot be achieved without some degree of uncertainty, the power of the experiment designs plus the wide range of conditions investigated through the various combinations of factors and levels, provide some confidence in the appropriateness of extrapolating some of the findings beyond the bounds of the experiments.

Nevertheless, further evaluations along the lines suggested in previous chapters, hopefully conducted by other research groups, would provide results with which to compare those discussed here, enriching the debate about task definition and geovisualization tool design. Furthermore, the use of Census data in all the experiments implies that this dissertation has only considered the exploration of datasets represented in vector format (in particular, that of socio-economic census information) and that of environments suited to their exploration. Investigating the tasks that users of different types of data may need to accomplish constitutes a logical extension of the work commenced in this dissertation and a stimulating avenue of research. Methodologies similar to those used in Chapters 4 and 5 for defining tasks and assessing systems may still be appropriate to evaluate systems that support the information processing of other data types. As it has been stressed throughout this dissertation, the use of experimental design would facilitate both the testing of environments and task definitions, as well as allow the comparison of results gathered by different researchers.

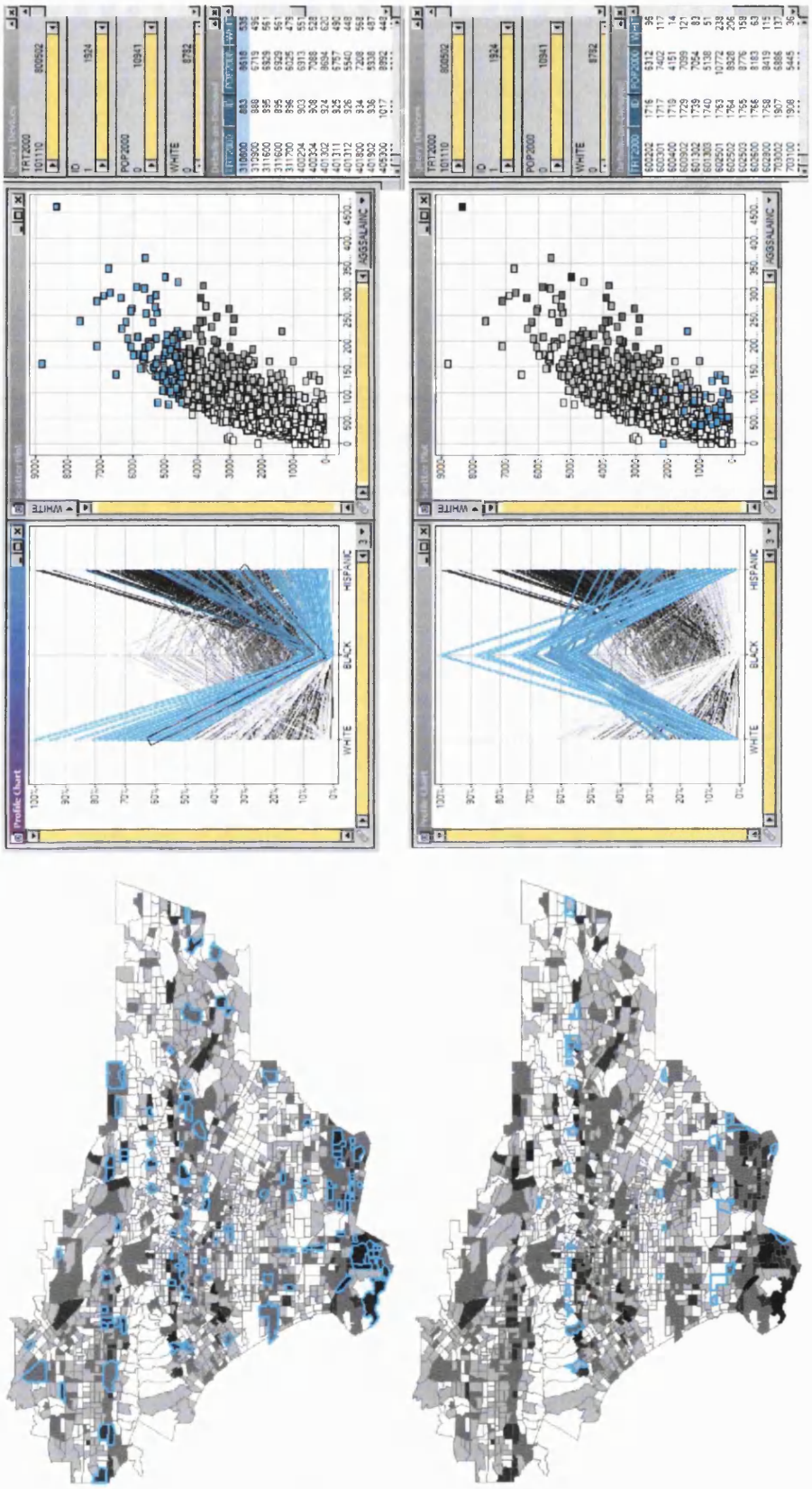

Figure 6.3 Data Exploration of U.S. 2000 Census data for Los Angeles: Left) High proportion of *White* compared to *Black*, *Hisp* and *AggSalaInc*. Right) Blunders in the dataset.

7 Conclusions and Future Research

"It is easy to be dazzled by a display of data, especially if it is rendered with color or depth. Our tendency is to be misled into thinking we are absorbing relevant information when we see a lot. But the success of a visualization tool should be based solely on the amount we learn about the phenomenon under study" (Cleveland 1993: 1).

The work described in this thesis has been presented in a linear manner with an introduction defining the approach to geovisualization as an information processing task. The various sources from which geovisualization draws were explained and a number of techniques were discussed in detail in the context of aspatial and spatial data applications. This in turn led to the consideration of various approaches to computer software development that could provide some of these techniques for exploring georeferenced area data with the goal of deriving hypotheses and building knowledge about the information they may contain.

These considerations were then used to implement and test working software in order to understand the tasks that similar computer-based tools need to support in order to enable geovisualization. Chapters 4 and 5 defined task taxonomies by combining 'levels' of factors, such as the visual operations performed and the number of attributes explored, to characterise work processes that may need to be accomplished in a geovisualization context. A taxonomy of these tasks in turn permitted the creation of formal experimental designs in order to assess not only the task definition but also the capacity of the computer-based tool tested to support the tasks.

In addition, the implementation of the system used in the second and third sets of experiments followed considerations from a usability study of one particular geovisualization environment. Usability evaluation techniques were adapted to the geovisualization domain in order to obtain evidence of software characteristics and functionality which hindered or supported data exploration through geovisualization. This evaluation also provided valuable information regarding the type of tasks that users need to address.

These three aspects of the research—the modular approach to tool development, the usability of the tools and the experimental design approach—are assessed in this final chapter. However, the contentions made throughout this dissertation are reviewed first. Finally, some future avenues of research are discussed.

7.1 Review of assertions

Chapter 3 examined the contention that linked multiple graphical displays are necessary for supporting geovisualization tasks. Although it was found that coupling aspatial data views with a map provides an effective means of representing relations between attributes, the absence of any dynamic link between all of the component graphical displays could in fact hinder the capacity of multiple graphs to provide complementary representations of data.

Chapter 4 in turn evaluated an environment which supported brushing between all graphical displays. It showed that when provided with fast, graphical widgets such as query devices—which enable a quick selection of data subsets—users preferred them to standard GIS-type procedures of querying data using dialog windows. Users found that performing exploratory tasks using only GIS would have been too time consuming to be a practical way of exploring data. This highlighted the limited flexibility of GIS for exploratory analysis in the sense of allowing the user to represent and interactively manipulate data through multiple graphical displays or enabling the investigation of 'what if' scenarios or questions that might prompt the discovery of relations or patterns that are useful. The close coupling of a GIS and a visualization system, however, was shown to overcome this limitation by providing a fast and reliable environment for data exploration. In addition, users found the environment to be fast, fun and easy to use.

Chapters 4 and 5 provided evidence of the benefits that may be gained from devising task taxonomies. They allow for geovisualization to be characterised and for tools to be evaluated. In addition, they can be used to define tasks in experimental designs with which software tools can be formally tested and the behaviour of users better understood. This highlights the importance of conducting empirical studies systematically within a framework—such as using tasks taxonomies and experimental designs—that may allow their comparison and permit a better understanding of potential user cognitive processes when using geovisualization as an exploratory data analysis technique. It is thus argued that methodologies akin to those suggested in the previous chapters should be used more widely in order for geovisualization to routinely build bodies of evidence that allow the better understanding of user work processes and of the cognitive demands of geospatial information processing.

It was also suggested that the tasks designed to test a system must be realistic while also ensuring that they put geovisualization concepts or implementations to the test. This is one of the most challenging issues in geovisualization. Because of its exploratory nature, there is

often no direct comparison with standard usability engineering practice and thus with available techniques for software and interface evaluation. Chapter 3 for instance evaluated the usability of a geovisualization environment which required a number of usability evaluation techniques to be coupled and adapted to geovisualization contexts. This is because tasks used in the evaluations must reflect potential user work processes and at the same time provide sufficient room to explore data in a manner that is conducive to the gaining of insight and knowledge construction, the acknowledged main goals of geovisualization methods and tool use. Therefore, it was made evident that usability evaluation techniques must be tailored to address the specific characteristics of geovisualization.

One important issue that was revealed by Chapters 4 and 5 in particular is that tasks where users had to perform a visual operation that occurred relative to a spatial feature are evidently important and more difficult. This provides some evidence that essentially geovisualization may be different to information visualization as comparison with geographic features is more complex than comparison across other data spaces. Furthermore, Chapter 4 suggested this is a manifestation of the different cognitive processes involved in geovisualization. Chapter 5 also found that user strategies had a major impact on user performance with computer based tools when addressing geovisualization tasks. This suggests that individual differences in the form of abilities and strategies should be investigated further. It also indicates that there may be fundamental differences between the cognitive processes required to participate in visualization of geographic spaces as opposed to information spaces and thus the tools required to support such processes may also need to be different. These findings imply the need to draw from research in other fields such as cognitive science in order to advance in the state-of-the-art of geovisualization.

7.2 Assessment of the modular approach to building geovisualization environments

The approach taken in this dissertation has been that of close coupling a GIS with an aspatial visualization software package, making use of the ArcGIS development platform (ArcObjects™) and the application programming interface (API) of DecisionSite, respectively. ArcObjects is built using Microsoft's COM technology and provides an infrastructure for application customization. Therefore, it was possible to extend ArcObjects by writing COM components using any COM-compliant development language. Thus, ArcGIS and DecisionSite were customised through Visual Basic for Applications (VBA), which is embedded within the GIS software environment.

The API is a set of programming objects that allows users to write software that interacts with an application. DecisionSite, the visualization software used for the close coupling, may be accessed in Java, C++, Visual Basic (VB), VBScript and JavaScript through the API. Note that VBA is the language used to program in VB. Therefore, the DecisionSite API was used as an out-of-process application which means it was launched from the GIS. This allowed the implementation of a geovisualization environment which close coupled two systems that appear seamless to the user. Chapter 3 (§3.1) discussed similar efforts to implement closely coupled or embedded ESDA functionality into GIS. Notably, the approach taken here is different in that it extends the analysis capabilities of the two packages by allowing the user to utilise the whole range of functionality in the two systems. In addition, it does support real time interactivity between all graphical displays in the coupled systems in order to facilitate data exploration.

The environment followed the close coupling approach to geovisualization tool development. It has been shown to be appropriate for data exploration and for supporting successive processes of hypotheses formation and knowledge construction about georeferenced datasets. The results of Chapter 4 (§4.7) suggest that users found the tool to be fast, useful and effective, and these are considered to be crucial features if a system is to gain acceptance by a wide range of users. Furthermore, it was demonstrated that the coupling of a number of views that include a map as well as aspatial representations of a dataset can facilitate not only the exploration of unknown and complex geographical datasets but also the understanding of the data and the relations between attribute values. Thus, it was shown that close coupling can be used rapidly to develop environments that support customisation and that take advantage of available technology. This is important from a research perspective as it may encourage the wider implementation of geovisualizations that can be tried out or tested. By adopting a coupling approach, the analyst or researcher may substantially accelerate the implementation process as many fundamental functionality—such as data input or display—will not need to be developed. This may be desirable particularly if there is interest in the testing of techniques and the fast implementation of concepts that can then allow their testing with potential users.

Coupling such as this can be accomplished with relative ease if an initial investment is made in learning one COM compliant language such as VB. Learning the syntax required to write VB programs and developing applications in VB is a rapid process compared with personal experience using other lower level programming languages. VB allows building and testing of

sophisticated Windows applications without the need to know much about the Windows API.[1] In addition, VB allows the development of bespoke graphical user interfaces with much greater ease than building complete applications in C or C++. Even building C libraries for use in VB can be difficult as the program must accommodate all the details of Windows programming correctly to avoid the application, the developing environment or Windows from crashing. Most importantly, "many algorithms are efficient enough to give good performance" (Stephens 1996: xvii) when written in VB even though VB can be slower than C and C++ on some occasions. This is also an important consideration, as the experiments highlighted the importance of quick system response, feedback and real-time interactivity for geovisualization.

Finally, the development environment for VBA, both in VB or as a hosted language, is very useful. It provides debugging functionality, a complete online help on syntax, function and statement implementations, as well as code completion that greatly assists and accelerates the writing and testing of routines and programs. Therefore, the development through VB of any of the increasing number of systems built on COM technology is facilitated by the large extensive resources that exist to support the VB programmer (such as abundant literature, on-line discussion groups, code snippets and examples).

7.2.1 Limitations and Drawbacks

Nevertheless, some limitations have become evident when using the close coupling of systems as proposed here. Each object in the API of DecisionSite or in the object model of ArcGIS is associated with certain properties, methods and events. Properties are data that describe an object; methods are what we can tell the object to do; and events are things that the object does. When an object in any of the coupled systems does not support a particular property, method or event, the developer is faced with a more difficult programming challenge than aforementioned. In Chapter 4 it was noted that neither ArcGIS nor DecisionSite support multiple selections of data objects from graphical displays. This functionality could potentially allow easier comparison of data subsets or scenarios by users. However, since it was not provided, objects or the properties supporting this behaviour would need to be created.

[1] The Windows Application Programming Interface (API) is a set of predefined functions that are built into the dynamic link libraries (DLLs) and which make up the Windows operating system. End users cannot access these functions. However, programmers can access the code written in the DLLs through the API and use this code in their programs.

Chapter 3 discussed how COM defines an API to allow the creation of components for use when integrating customised applications or when attempting to allow diverse components (such as ArcGIS and DecisionSite) to interact. "However, in order to interact, components must adhere to a binary structure specified by Microsoft. As long as components adhere to this binary structure, components written in different languages can interoperate" (Cornella-Dorda 2001). Because COM is based on a native binary format, components written to its specifications are not platform independent. The system discussed in the previous chapters was developed for a PC-based Windows NT platform. This is the platform where COM is best supported but, from a practical standpoint, happened to be the only one in which both of the coupled systems were available.

Nevertheless, Microsoft is actively supporting COM and developing it further to work in distributed and Web-based scenarios. Distributed COM (DCOM), for instance, is an extension to COM that allows network-based component interaction or processes to be spread across a network. "With DCOM, components operating on a variety of platforms can interact, as long as DCOM is available within the environment" (Cornella-Dorda 2001). This is an exiting possibility for coupling analysis tools that run on different platforms. In addition, COM+ is an ongoing upgrade of COM that aims to provide services such as security and reliable communications that are not integrated in COM.

A major limitation however is that use of the geovisualization environment created in this research requires the user to hold licences for each of the coupled software systems. It is for this reason that this approach may only be a practical and fast method for application development for research purposes. Nevertheless, a program written in VBA can also be adapted to be packaged into a DLL (dynamic link library). A DLL is compiled code that can be accessed only by other programs at the programming level. Functions and subroutines can be put into DLLs in order that they might be shared among many programs running at the same time. The DLL simply needs to be compiled into a callable format, saved in a directory, and registered in the programs that will call it. DLLs are convenient if code is to be shared by various programs or if it will be distributed to numerous people. In this manner, it can be shared to illustrate techniques or tools that may be suitable for testing.

7.3 Assessment of the usability evaluation techniques for evaluating geovisualizations

Chapters 3 and 4 discussed how usability evaluation techniques such as usability testing and heuristic evaluation could be used for detecting problems with the interface of

geovisualization systems that might otherwise hinder the exploration of data. These methods were shown to be useful, both in order to detect usability limitations as well as to better understand the tasks that users may want to attempt in geovisualization. Usability evaluation was also well complemented by other methods of qualitative data gathering such as semi-structured interviews and questionnaires. These techniques helped to elicit user thoughts, greatly clarified how they processed information and made it possible to understand the goals they were trying to reach. In addition, interviews and questionnaires provided us with subjective user views about the tools or techniques tested. This information was important for establishing user satisfaction with the tools which is contended to be essential if the system is to be accepted by a wider user community.

The combination of the aforementioned methods enabled a better understanding of user behaviours and how they accomplished tasks. This in turn was perceived to be invaluable information in the development or fine tuning of the environment and in pinpointing the characteristics or functionality that should be provided to facilitate information processing in a geovisualization domain.

7.3.1 Limitations and drawbacks

As it was discussed in Chapter 4, the exploratory nature of geovisualization makes the clear definition of user tasks and goals difficult to accomplish in practice. This is, however, required in order to define scenarios within which the tool or geovisualization technique might be tested and used. Therefore, it was necessary to adapt the usability evaluation techniques to the geovisualization context. It was shown, however, that methods drawn from the field of HCI can be used to assist the difficult process of defining tasks and building a task taxonomy for geovisualization. Nevertheless, it was necessary to interpret some of the standard usability measures with care when observing user interaction with the coupled software system. For instance, it was not possible to quantify error rates in the more traditional sense of actions that do not accomplish a desired goal because the goal of geovisualization tends to be broad and even difficult to pinpoint. Hence, both quantitative and qualitative information about the coupled system were necessary in order to understand user behaviour and perceptions of the system. It would be illuminating to obtain evidence from other research groups as to how similar evaluation techniques work in the testing of different tools in other application domains. This would provide an indication of the appropriateness of these and other methods for the evaluation of geovisualizations. More importantly, it would provide evaluation techniques that are tailored to the particularities of geovisualization testing.

7.4 Assessment of the experimental design for evaluating geovisualization

In the previous chapters, the first issue explored concerning geovisualization environments was that of their usability and, in particular, the functionality that can sustain or hinder the information processing of data represented on-screen. This in turn informed the development of the environment used in the experiments of Chapters 4 and 5 which explored both task definition and user behaviour.

Thus, experimental design techniques proved to be very powerful in measuring and understanding user performance in a geovisualization context. This was a key aspect not only in identifying how participants used the tools in the experiments but also in understanding their behaviour when solving different tasks. Chapter 4 initially explored a broad range of tasks using a factorial design, which gave invaluable insight into the special characteristics of some tasks. Factorial analysis was seen as an effective method that accommodated some of the limitations of conducting evaluations in a geovisualization domain, such as the difficulty for recruiting potential users to complete the tests and accommodating their restrictive time availability. More importantly, factorial analysis was shown to be a powerful methodology for exploring task taxonomies for a very wide range of data types and applications. This is because it can easily be adapted to include any number of factors or levels of those factors that are deemed necessary for defining tasks.

However, it was shown in Chapter 5 that tasks accomplished in geovisualization experiments can have considerable practice effects which influence user performance. Hence, other experimental design approaches were considered such as Latin Squares. In addition, Chapter 4 indicated that it was necessary to refine the task taxonomy first proposed to account for some characteristics of spatial tasks. Latin Squares was also an appropriate design for investigating tasks primitives in detail. Thus, the experiment of Chapter 5 was successfully designed to learn about such characteristics and investigate user approaches to task resolution, strategies and behaviour. Thus, the evaluations described in this dissertation show that carrying out experiments that go from the most general to the most specific task characterisation are a powerful and relevant approach for an understanding of geovisualization tasks and user behaviour.

7.4.1 Limitations and drawbacks

The major practical limitations of using experimental design techniques were the initial search costs of identifying the appropriate technique for running the desired evaluation, learning the intricacies of the statistical design and interpreting the results of the analysis. For someone who is not familiar with the statistical methods this might require considerable time or guidance, at least initially. However, the power of this analysis tool can be gauged by the sheer size of studies and research fields that use it for hypothesis testing. For instance, factorial designs have an established role in the investigation of a large number of issues in cognitive psychology. This in turn is a valuable source of examples that can provide direction in the design of relevant experiments for visualization. However, the literature is so vast that guidance from researchers in the field is near necessary.

This discussion stresses a point already made in the definition of geovisualization proposed in Chapter 2: geovisualization is a field that draws from advances in other disciplines as its research domain spans from the creation of innovative graphical displays for data exploration to the design of computer systems and their use by humans whose requirements need to be supported. Hence, research in geovisualization can greatly benefit from collaboration with experts from relevant fields. This can be seen as an exciting opportunity for cooperation but it also makes the research demanding and the development of a theory-based field of geovisualization slow on occasion. Not only is the geovisulization researcher required to have at least some familiarity with technological advances that may facilitate the creation, implementation and eventual testing of concepts, she must also be aware of the relevant literature in the disciplines that can so profoundly influence the advancement of a theory of geovisualization. This requires at least some knowledge of the jargon and main debates in such disciplines. Assimilating this knowledge and furthermore applying it to the geovisualization domain is not a simple task.

7.5 Extending the research

7.5.1 Collaborative Geovisualization

Many topics of current interest draw on the combination of techniques and approaches of different disciplines. Such is the case for research on collaborative geovisualization. The environments tested in the previous chapters were designed to be used by users working on their own. Nevertheless, work on many problems is often not done by individuals in

isolation but is approached collaboratively by groups of people (Furhman and Pike 2003). However, geospatial information technologies are not typically designed to support multiple users simultaneously. Designing visualization methods to enable collaborative work has been identified as a research challenge as it entails not only use by individuals, but must consider how the group responds to the system and how its members interact with one another (Slocum *et al.* 2001). The interfaces must support the group work and must account for collaboration that may or may not occur at different places and at different times (Brodlie *et al.* 2002; Brodlie *et al.* 2000; Wright *et al.* 2000). This in itself poses technical problems as data and information must be shared and synchronised in ways that can make its use by people at different locations and even at different times meaningful. In addition, although data exploration may be performed in collaboration, the interface should facilitate the integration of work by individuals, be they in the same location or working remotely, and thus bring different knowledge and experience to the task in hand. Hence, testing these interfaces for their appropriateness and usability poses new challenges in terms of defining the tasks that are relevant to this context, as well as devising the techniques that should be used to record and analyse user performance.

7.5.2 Expertise

Another topic that has attracted much attention is that of customising functionality or the interfaces of exploratory tools according to the expertise of the users. In other words, investigating the characteristics of users to build human-computer interfaces that suit different needs. Many recognise that it is important to know the type of people that will be using the system in terms not only of their expertise with computer systems, but also of their frequency of use and domain knowledge (Slocum *et al.* 2001; Haklay and Tobón 2003). Users can be catalogued, for instance, as naïve, novice, skilled and expert. Some users may use the system only on infrequent occasions, in which case they will likely lose their knowledge and regress to a novice state each time. Others may use it frequently and become expert users. These issues are of interest when designing interfaces that meet different user needs and identifying the level of support that might be appropriate to the greatest number of potential users.

These considerations are relevant to a number of research topics in GIS including one that has received recent attention: Public Participation Geographic Information Systems (PPGIS). Some of the challenges faced by PPGIS are similar to those of geovisualization. PPGIS settings usually call for an open-ended exploration in which users experiment with a GIS to examine various issues that relate to their localities. Therefore, they face similar

problems when defining tasks for testing applications (Haklay and Tobón 2003; Tobón and Haklay 2003). In addition, PPGIS is about empowering GIS users from different backgrounds and levels of expertise and enabling them to use the technology purposefully to capture their local knowledge and advance their goals (Talen 2000). Although the usability of GIS products has improved immensely in recent years, they still require users to have or acquire considerable technical knowledge to operate them (Traynor and Williams 1997). This presents major obstacles to non-expert users in terms of navigating an interface that embeds a language, world view and concepts that support the system's architecture rather than the user's work view (Goodchild 2002; Ghose 2001; Traynor and Williams 1995). The research conducted here addressed the needs of existing GIS users. Therefore, a relevant and timely extension of this work that may greatly benefit from PPGIS research is that of understanding the diversity of potential users that may benefit from exploratory tools. On the other hand, geovisualization can contribute to the design of interfaces that facilitate exploration in PPGIS settings through its accumulated knowledge on the adequacy of graphical displays for supporting different user tasks.

7.5.3 Individual Differences, Strategies and Abilities

The work described in this dissertation has suggested the need to consider differences between individual users, be they in the different strategies adopted when attempting a task, or in the ways that spatial visualization abilities condition the strategies adopted by subjects. Understanding individual differences in the cognitive processes and structures of geovisualization problem solving would facilitate the efforts to adapt techniques and implementations to the characteristics of the cognitive process. Thus, designing environments that optimise the ways in which graphics enhance human performance of information processing is highly desirable.

Our main source for approaching the issue of individual differences and strategies is the literature in cognitive psychology. However, "most psychological experiments tacitly assume that all subjects presented with a task will solve that task in the same way. Lohman and Kyllonen (1983) argue that it is particularly unlikely that subjects use the same strategy on complex tasks. They, however, reference Lucas (1953) who suggests that variations in solution strategy are more a function of the task than of the subject. There are abundant examples that show that there is no consensus in the psychology literature about how individual differences operate on solution strategy. In addition, individual differences are studied by cognitive psychologists using tests that measure specific abilities to then measure their correlation with strategies. It is recognised that the tests can measure different abilities

depending on how the problem is solved and that certain tests are particularly susceptible to alternative solution strategies (Lohman and Kyllonen 1983). Thus, the available literature can only inform the investigation of strategies in a geovisualization context particularly because tasks in geovisualization are somewhat different (as discussed in Chapter 5) and hence the findings and tests used in psychology may be of limited application and use. Therefore, our main research challenge would be to devise a methodology tailored to understand individual differences in strategies in geovisualization.

7.6 Conclusion

This dissertation has investigated the effectiveness of geovisualization as an exploratory method of data analysis. Based on the hypothesis that geovisualization is an information processing activity, its main contention is that users build information or knowledge from the use of a computer based tool and that they differ in their strategies for doing so. These differences may depend on both the tasks users must perform and the data they need to investigate, as well as on their individual differences and prior knowledge. Thus, only by understanding these issues will geovisualization be able to propose a theory of how multiple, dynamic and linked displays facilitate a process of hypothesis formation and knowledge construction.

The possibilities for cross-fertilisation between the evaluation methods reported and the revived interest in cognitive research in geovisualization and GISc thus seem particularly pertinent at present. For instance, research in cognitive psychology on spatial cognition (Richardson *et al.* 1999; Montello and Freundschuh in press), the representation in memory of geographical knowledge and its effects on geographic judgments as well as in biases in geographical knowledge (Friedman *et al.* 2002; Friedman and Brown 2000), have profound implications for building solid theoretical principles for geovisualization (Ware and Plumlee 2003), as well as on the design of computer based tools that aim to support hypothesis formation and the construction of knowledge about spatial data. It is suggested that evaluations similar to those presented here may contribute to such research by supplying information about user tasks and requirements that support their work processes. In addition, the techniques discussed are indeed an effective means to evaluate whether the tools we develop are usable, useful and fit for purpose, which are crucial considerations if geovisualization environments are to support a wider range of users. Those evaluated here had GIS users in mind and were tested for area based georeferenced data. Nevertheless, the task definition and methods suggested for evaluating the systems can be easily adapted or extended to appraise a wide range of tools, data types and user requirements.

References

Ahlberg, C. (1996) "Spotfire: An Information Exploration Environment," *SIGMOD Record*, 25(4): 25-9.

Ahlberg, C. and B. Shneiderman (1994) "The Alphaslider: A Compact and Rapid Selector", *CHI'94 Human Factors in Computing Systems*, Boston, MA, 24-28 April, 365-71.

Ahlberg, C. and B. Shneiderman (1994a) "Visual Information Seeking: Tight Coupling of Dynamic Query Filters with Starfield Displays", *CHI'94 Human Factors in Computing Systems*, Boston, MA, 24-28 April, 313-17.

Ahlberg, C., C. Williamson and B. Shneiderman (1992) "Dynamic Queries for Information Exploration: An Implementation and Evaluation", *CHI'92*, Monterrey, CA, May 3-7, 619-26.

Andrienko, G. L. and N. V. Andrienko (1999) "Interactive Maps for Visual Data Exploration," *International Journal of Geographical Information Systems*, 13(4): 355-74.

Andrienko, G. L. and N. Andrienko (1999a) "Making a GIS Intelligent: CommonGIS Project View", *AGILE'99*, Rome, Italy, April 15-17, 19-24.

Andrienko, N., G. Andrienko, H. Voss, F. Bernardo, J. Hipolito and U. Kretchmer (2002) "Testing the Usability of Interactive Maps in Common GIS," *Cartography and Geographic Information Society (CaGIS)*, 29(4): 325-42.

Andrienko, G. L., N. V. Andrienko and P. Gatalsky (2001) "Mapping Spatio-Temporal Data for Exploratory Analysis", *NTTS & ETK 2001*, Crete, June 18-22, 43-9.

Andrienko, N., G. Andrienko and P. Galatsky (2000) "Supporting visual exploration of object movement", *Proceedings of the working conference on advanced visual interfaces*, Palermo, Italy, 217-20.

Ankerst, M. and D. A. Keim (1996) "Circle segments: A technique for visually exploring lagre dimensional data sets", *IEEE Visualization Conference*, San Francisco, CA, October 27 - November 1.

Anselin, L. (2000) "GIS, Spatial Econometrics and Social Science Research," *Journal of Geographical Systems*, 2(1): 11-15.

Anselin, L. (1998) "Exploratory Spatial Data Analysis in a Geocomputational Environment", in P. A. Longley, S. M. Brooks, R. McDonell and B. Macmillan (eds.), *Geocomputation: A Primer,* Chichester: John Wiley and Sons Ltd, 77-95.

Anselin, L. (1995) "Local Indicators of Spatial Association - LISA," *Geographical Analysis*, 27(2): 93-115.

Anselin, L., R. F. Dodson and S. Hudak (1993) "Linking GIS and Spatial Data Analysis in Practice," *Geographical Systems*, 1: 3-23.

Anselin, L. and A. Getis (1992) "Spatial Statistical Analysis and Geographic Information Systems," *Annals of Regional Science*, 26: 19-33.

Asche, H. and C. M. Herrmann (1994) "Designing Interactive Maps for Planning and Education", in A. M. MacEachern and D. R. F. Taylor (eds.), *Visualization in Modern Cartography,* Oxford: Pergamon, 215-42.

Batty, M., M. Dodge, S. Doyle and A. Smith (1998) "Modelling Virtual Environments", in P. A. Longley, S. M. Brooks, R. McDonell and B. Macmillan (eds.), *GeoComputation: A Primer,* Chichester: John Wiley and Sons, 139-61.

Baecker, R. M. (1987) "Towards an Effective Characterization of Graphical Interaction", in R. M. Baecker and W. A. S. Buxton (ed.), *Readings in Human-Computer Interaction: A Multidiciplinary Approach,* San Mateo: Morgan Kaufmann Publishers, Inc., 471-81.

Baecker, R. M. and W. A. S. Buxton (eds.) (1987) *Readings in Human-Computer Interaction: A Multidiciplinary Approach,* San Mateo: Morgan Kaufmann Publishers, Inc.

Bailey, T. C. (1994) "A review of Statistical Spatial Analysis and Geographic Information Systems", in A. S. Fotheringham and P. Rogerson (eds.), *Spatial Analysis and GIS,* London: Taylor and Francis, 13-44.

Bailey, T. C. and A. C. Gatrell (1995) *Interactive Spatial Analysis*: John Wiley & Sons.

Beard, K. (2000) "Operations on Multiple Representations", *GIScience,* Savannah, Georgia, USA, 28-31 October, 42-43.

Beard, K., B. Buttenfield and S. Clapham (1991) *NCGIA research initiative 7: Visualization of spatial data quality,* Santa Barbara, CA: NCGIA Technical Paper 91-26.

Becker, R. A. and W. S. Cleveland (1987) "Brushing Scatterplots," *Technometrics,* 29(2): 127-42.

Bertin, J. (1983) *Semiology of Graphics: Diagrams, Networks, Maps,* Madison, WI: The University of Wisconsin Press.

Beshers, C. and S. Feiner (1993) "AutoVisual: Rule-Based Design of Interactive Multivariate Visualizations," *IEEE Computer Graphics and Applications,* 13(4): 41-9.

Bishop, I. D. and C. Karadaglis (1996) "Combining GIS-based environmental modelling and visualization: another window on the modeling process." *Third International Conference/Workshop on Integrating GIS and Environmental Modeling,* Santa Fe, New Mexico, January 21-25, CD-ROM.

Bortins, I. and S. Demers (2002) "Cartogram Types", *http://www.ncgia.ucsb.edu/projects/Cartogram_Central/types.html,* accessed July 2003.

Broadbent, D. E. (1971) *Decision and Stress,* London: Academic Press.

Brodlie, K., J. Wood, D. Boyd, L. Sastry, J. Gallop, C. Osland and S. E. Bunn (2002) "Collaborative Visualisation Using Access Grid", *UK e-Science All Hands Conference,* Sheffield, September 2-4.

Brodlie, K., S. Mason, M. Thompson, M. Walkley and J. Wood (2000) "Reacting to a crisis: benefits of collaborative visualization and computational steering in a grid environment", *UK e-Science All Hands Conference,* Sheffield, (*http://www.comp.leeds.ac.uk/markw/Research/Pub/Papers/all-hands.pdf,* accessed January 2003).

Brooke, J. (1996) "SUS: A 'Quick and Dirty' Usability Scale", in P. W. Jordan, B. Thomas, B. A. Weerdmeester and I. L. McClelland (eds.), *Usability Evaluation in Industry,* London: Taylor and Francis, 189-94.

Brown, I. M., D. B. Kidner, A. Lovett, W. Mackaness, D. R. Miller, R. Purves, J. F. Raper, J. M. Ware and J. D. Wood (2002) "Virtual Landscapes: Introduction", in P. Fisher and D. Unwin (ed.), *Virtual Reality in Geography,* London: Taylor and Francis, 95-101.

Brunsdon, C. (2000) "The Comap: Investigating Geographical Pattern via Conditional Spatial Distributions", *GIS Research UK,* University of York, April 5-7, 97-101.

Brunsdon, C., A. S. Fotheringham and M. E. Charlton (2000) "An Investigation of Methods for Visualising Highly Multivariate Datasets", (*http://www.agocg.ac.uk/reports/visual/casestud/brunsdon/contents.htm,* accessed December 2001).

Buja, A., D. Cook and D. F. Swayne (1996) "Interactive High-Dimensional Data Visualization," *Journal of Computational and Graphical Statistics*, 5: 78-99.

Burin, D. I., A. R. Delgado and G. Prieto (2000) "Solution Strategies and Gender Differences in Spatial Visualization Tasks," *Psicológica*, 21: 275-86.

Butler, K. A. (1996) "Usability Engineering Turns 10," *Interactions*, 3(1): 59-75.

Buttenfield, B. P. and K. Beard (1994) "Graphical and Geographical Components of Data Quality", in H. M. Hearnshaw and D. J. Unwin (eds.), *Visualization in Geographical Information Systems,* Chichester: John Wiley & Sons, 150-57.

Campbell, C. S. and S. L. Egbert (1990) "Animated Cartography: Thirty years of Scratching the Surface," *Cartographica*, 27(2): 24-46.

Card, S. K., J. D. Mackinlay and B. Shneiderman (1999) *Readings in Information Visualization: Using Vision to Think*, San Francisco, CA: Morgan Kaufman Publishers Inc.

Carpenter, P. A. and M. A. Just (1979) "Individual Differences in Spatial Ability: An Information Processing Analysis of Psychometric Problems", *Office of Naval Research Annual Constructors Meeting*, Seattle, April.

Cartwright, W. (1999) "Extending the Map Metaphor using Web Delivered Multimedia," *International Journal of Geographical Information Systems*, 13(4): 335-53.

Casner, S. M. (1991) "A Task-Analytic Approach to the Automated Design of Graphic Presentations," *ACM Transactions on Graphics*, 10(2): 111-51.

Chen, C. and M. Czerwinski (2000) "Empirical Evaluation of Information Visualizations: An Introduction," *International Journal of Human-Computer Studies*, 53: 631-35.

Chen, C. and Y. Yu (2000) "Empirical Studies of Information Visualization: A Meta-Analysis," *International Journal of Human-Computer Studies*, 53: 851-66.

Chernoff, H. (1973) "Using faces to represent points in k-dimensional space," *Journal of the American Statistical Association*, 68: 361-68.

Chimera, R. and B. Shneiderman (1994) "An exploratory evaluation of three interfaces for browsing large hierarchical tables of contents," *ACM Transactions on Information Systems*, 12(4): 383-406.

Chin, C. A. and W. F. Brewer (1992) "Psychological responses to anomalous data", *Proceedings of the 14th Annual Conference of the Cognitive Science Society*, 165-70.

Cleveland, W. S. (1994) *The Elements of Graphing Data*, Summit: Hobart Press.

Cleveland, W. S. (1993) *Visualizing Data*, New Jersey: Hobart Press.

Cleveland, W. S. and R. A. Becker (1991) "Visualizing Multidimensional Data," *Pixel*, 2(2): 36-41.

Cleveland, W. S. and R. McGill (1986) "An Experiment in Graphical Perception," *International Journal of Man-Machine Studies*, 25: 491-500.

Collins Dictionary and Thesaurus (1987), London: Collins.

Conklin, J. (2003) "Evaluation of the Quality of the Data Capture System and the Impact of the Data Capture Mode on the Data Quality. U.S. Census Bureau" (*http://www.census.gov/pred/www/rpts/K1b.pdf*, accessed August 2003).

Cook, D., J. Symanzik, J. J. Majure and N. Cressie (1997) "Exploratory Cartographic Visualization: Advancing the Agenda," *Computers and Geosciences*, 23(4): 371-85.

Coombes, M., S. Openshaw, C. Wong and S. Raybould (1993) "Community Boundary Definition: A GIS Design Specification," *Regional Studies*, 27(3): 280.

Cooper, L. A. (1980) "Spatial Information Processing: Strategies for Research", in R. E. Snow, P. A. Federico and W. E. Montague (eds.), *Aptitude, Learning, and Instruction: Cognitive Process Analyses of Aptitude*, Hillsdale, HJ: Earlbaum, 149-76.

Cornella-Dorda, S. (2001) "Component Object Model (COM), DCOM, and Related Capabilities", The Software Engineering Institute (SEI), Carnegie Mellon University, (*http://www.sei.cmu.edu/str/descriptions/com_body.html*, accessed July 2003).

Cox, D. R. (1958) *Planning of Experiments*, Chichester: John Wiley and Sons.

Craig, W. J., T. M. Harris and D. Weiner (2002) "Community Participation and Geographic Information Systems", in W. J. Craig, T. M. Harris and D. Weiner (eds.), *Community Participation and Geographic Information Systems*, London: Taylor and Francis, 3-16.

Crampton, J. W. (2002) "Interactivity Types in Geographic Visualization," *Cartography and Geographic Information Science*, 29: 85-98.

Davis, F. D. (1989) "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology," *MIS Quarterly*, 13(3): 319-40.

DeFanti, T. A., M. D. Brown and B. H. McCormick (1999) "Visualization: Expanding Scientific and Engineering Research Opportunities", in S. K. Card, J. D. Mackinlay and B. Shneiderman (eds.), *Readings in Information Visualization: Using Vision to Think*, San Francisco: Morgan Kaufman Publishers Inc., 37-52.

del Barrio , C. (1999) "The Use of Semistructured Interviews and Qualitative Methods for the Study of Peer Bullying", (*http://www.gold.ac.uk/tmr/reports/aim2_madrid1.html#QUALMETH*, accessed July 2003).

Dennis, P. A. T., A. Adams and R. R. Nelson (1992) "Perceived Usefulness, Ease of Use and Usage of Information Technology: A Replication," *MIS Quarterly*, 6(3): 227-47.

Densham, P. J. and M. P. Armstrong (1995) "Human-computer interaction considerations for visual-interactive locational analysis", in T. Nyerges, D. Mark, R. Laurini and M. J. Egenhofer (eds.), *Cognitive Aspects of Human-Computer Interaction for Geographic Information Systems,* Dordrecht: Kluwer, 179-96.

DETR (1998) *Town Centres: Defining Boundaries for Statistical Monitoring*, London: The Stationery Office.

DeVellis, R. F. (1991) *Scale Development, Theory and Applications*: Sage Publications.

DiBiase, D. (1990) "Visualization in the Earth Sciences," *Bulletin College of Earth and Mineral Sciences*, 59(2): 13-18.

DiBiase, D., A. M. MacEachren, J. B. Krygier and C. Reeves (1992) "Animation and the Role of Map Design in Scientific Visualization," *Cartography and Geographic Information Systems*, 19(4): 204-14.

DiBiase, D., C. Reeves, A. MacEachren, M. von Wyss, J. B. Krygier, J. L. Sloan and M. C. Detweiler (1994) "Multivariate Display of Geographic Data: Applications in Earth System Science", in A. M. MacEachren and D. R. F. Taylor (eds.), *Visualization in Modern Cartography,* Oxford: Pergamon, 287-312.

Dix, A., J. Finlay, G. Abowd and R. Beale (1998) *Human-Computer Interaction*, Hertfordshire: Prentice Hall.

Dorling, D. (1996) *Area Cartograms: Their Use and Creation*: CATMOG 59, Geo Books, Norwich.

Dorling, D. (1994) "Cartograms for Visualizing Human Geography", in H. M. Hearnshaw and D. J. Unwin (eds.), *Visualization in Geographical Information Systems,* Chichester: John Wiley & Sons, 85-102.

Dorling, D. (1992) "Stretching Space and Splicing Time: From Cartographic Animation to Interactive Visualization," *Cartography and Geographic Information Systems*, 19(4): 215-27.

Dorling, D. (1990) *A Cartogram for Visualization*, University of Newcastle upon Tyne: North East Regional Research Laboratory, Research Report 90/4.

Dumas, J. S. and J. C. Redish (1999) *A Practical Guide to Usability Testing*, Exeter: Intellect Books.

Dykes, J. (2002) "Creating information-rich virtual environments with geo-referenced digital panoramic imagery", in P. Fisher and D. Unwin (eds.), *Virtual Reality in Geography,* London: Taylor and Francis, 68-92.

Dykes, J. (1999) "Interactive Maps for Exploratory Spatial Data Analysis: Cartographic Visualization; Approach, Implementation and Application", *PhD Thesis*, University of Leicester, unpublished.

Dykes, J. A., K. E. Moore and J. D. Wood (1999) "Virtual Environments for Student Fieldwork using Networked Components," *International Journal of Geographical Information Science*, 13: 397-416.

Dykes, J. (1997) "Exploring Spatial Data Representation with Dynamic Graphics," *Computers and Geosciences*, 23(4): 261-278.

Edsall, R. M., A. M. MacEachern and L. Pickle (2001) "Case Study: Design and Assessment of an Enhanced Geographic Information System for Exploration of Multivariate Health Statistics", *IEEE Symposium on Information Visualization (INFOVIS'01)*, San Diego, CA, 24-27 July.

Ehlschlaeger, C. R., A. M. Shortridge and M. Goodchild (1997) "Visualizing spatial data uncertainty using animation," *Computers and Geosciences*, 23(4): 387-93.

Erbacher, R. F. and G. G. Grinstein (1995) "Performance issues in a real-time true color data display", *SPIE '95 Conference on Visual Data Exploration and Analysis II*, San Jose, CA, February, 256-62.

Ericsson, K. A. and H. A. Simon (1984) *Protocol analysis: Verbal reports as data*, Cambridge, MA: MIT Press.

Evans, B. J. (1997) "Dynamic display of spatial data reliability: does it benefit the map user?," *Computers & Geosciences*, 23(4): 409-22.

Fabrikant, S. I., M. Ruocco, R. Middleton, D. R. Montello and C. Jorgensen (2002) "The First Law of Cognitive Geography: Distance and Similarity in Semantic Space", *GIScience 2002*, Boulder, CO, 25-28 September, 31-33.

Fairbairn, D., G. Andrienko, N. Andrienko, G. Buziek and J. Dykes (2001) "Representation and its Relationship with Cartographic Visualization: A Research Agenda," *Cartography and Geographic Information Science*, 28(1), 13-28.

Feeney, A., A. K. W. Hola, S. P. Liversedge, J. M. Findlay and R. Metcalf (2000) "How People Extract Information from Graphs: Evidence from a Sentence-Graph Verification Paradigm", in M. Anderson, P. Cheng and V. Haarslev (eds.), *Lecture Notes in Artificial Intelligence 1889*, Berlin: Springer-Verlag, 149-61.

Ferreira, J. J. and L. L. Wiggins (1990) "The Density Dial: A Visualization Tool for Thematic Mapping," *Geo Info Systems*, 10: 69-71.

Fisher, P. F. (1994) "Animation and Sound for the Visualization of Uncertain Spatial Information", in H. M. Hearnshaw and D. J. (eds.), *Visualization in Geographical Information Systems*, Chichester: John Wiley & Sons, 181-85.

Fisher, P. and D. Unwin (2002) "Virtual Reality in Geography: An Introduction", in P. Fisher and D. Unwin (eds.), *Virtual Reality in Geography*, London: Taylor and Francis, 1-4.

Fisher, P. F., J. Dykes and J. Wood (1993) "Map Design and Visualization," *The Cartographic Journal*, 30: 136-42.

Fotheringham, A. S., C. Brunsdon and M. E. Charlton (2000) *Quantitative Geography: Perspectives on Spatial Data Analysis*, London: SAGE Publications Ltd.

Freedman, E. G. and P. Shah (2002) "Toward a Model of Knowledge-Based Graph Comprehension", in M. Hegarty, B. Meyer and N. H. Narayanan (eds.), *Lecture Notes in Artificial Intelligence 2317,* Berlin: Springer-Verlag, 18-30.

Friedhoff, R. M. and W. Benzon (1989) *Visualization: The second computer revolution*, New York, NY: Harry N. Abrams.

Friedman, A., D. D. Kerkman and N. Brown (2002) "Spatial location judgments: A cross-national comparison of estimation bias in subjective North American geography", *Psychonomic Bulletin & Review*, 9, 615-23.

Friedman, A. and N. Brown (2000) "Updating geographical knowledge: Principles of coherence and inertia", *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 26, 900-14.

Fua, Y. H., M. O. Ward and A. Rundensteiner (1999) "Hierarchical Parallel Coordinates for Exploration of Large Datasets", *IEEE Visualization'99*, San Francisco, CA, 43-50.

Fuhrmann, S., P. Ahonen-Rainio, R. M. Edsall, S. I. Fabrikant, E. L. Koua, C. Tobón, C. Ware and S. Wilson (2003) "Making Useful and Useable Geovisualization: Design and Evaluation Issues", in J. Dykes, A. M. MacEachern and M. J. Kraak (eds.), *Exploring Geovisualization,* London: Elsevier, in press.

Fuhrmann, S. and W. Pike (2003) "User-Centered Design of Collaborative Geovisualization Tools", in J. Dykes, A. M. MacEachern and M. J. Kraak (eds.), *Exploring Geovisualization,* London: Elsevier, in press.

Furnas, G. W. (1986) "Generalized Fisheye Views", *CHI'86*, Boston, MA, April 13-17, 16-23.

Gahegan, M. (2001) "Visual Exploration in Geography: Analysis with Light", in H. J. Miller and J. Han (eds.), *Geographic Data Mining and Knowledge Discovery,* London: Taylor and Francis, 260-87.

Gahegan, M. (1999) "Four Barriers to the Development of Effective Exploratory Visualisation Tools for the Geosciences," *International Journal of Geographical Information Systems*, 13(4): 289-309.

Gahegan, M. (1996) "Visualisation Strategies for Exploratory Spatial Analysis", *Third International Conference/Workshop on Integrating GIS and Environmental Modeling*, Santa Fe, New Mexico, January 21-26 (*http://www.ncgia.ucsb.edu/conf/SANTA_FE_CD-ROM/sf_papers/gahegan_mark/gahegan.html*, accessed July 2002).

Gahegan, M., M. Takatsuka, M. Wheeler and F. Hardisty (2001) "Introducing GeoVISTA Studio: A Geocomputational Workbench," *Computers Environments and Urban Systems*, 26(4): 267-92.

Garland, K. (1994) *Mr. Beck's Underground Map*, London: Capital Transport Publishing.

Gershon, N., S. Eick and S. K. Card (1998) "Information Visualization," *Interactions*, 5(2): 9-15.

Ghose, R. (2001) "Use of information technology for community empowerment: transforming Geographic Information Systems into community information systems," *Transactions in GIS*, 5(2): 141-63.

Gibson, J. J. (1979) *An Ecological Approach to Visual Perception*, Boston: Houghton Mifflin.

Goodchild, M. F. (2002) "Foreword", in W. J. Craig, T. M. Harris and D. Weiner (eds.), *Community Participation and Geographic Information Systems,* London: Taylor and Francis, ix-xxiii.

Goodchild, M. F. (2000) "The Current Status of GIS and Spatial Analysis," *Journal of Geographical Systems*, 2(1): 5-10.

Goodchild, M. F. (1991) "The Technological Setting of GIS", in P. A. Longley, M. F. Goodchild, D. J. Maguire and D. W. Rhind (eds.), *Geographical Information Systems: Principles, Techniques, Applications and Management.,* New York, NY: John Wiley and Sons, 45-54.

Goodchild, M. F. (1987) "A Spatial Analytical Perspective on Geographical Information Systems," *International Journal of Geographical Information Science*, 1(4): 327-34.

Goodchild, M. F., R. Haining and S. Wise (1992) "Integrating GIS and Spatial Data Analysis: Problems and Possibilities," *International Journal of Geographical Information Systems*, 6: 407-23.

Haining, R. P. (1994) "Designing Spatial Data Analysis Modules for Geographical Information Systems", in A. S. Fotheringham and P. Rogerson (eds.), *Spatial Analysis and GIS,* London: Taylor and Francis, 45-64.

Haklay, M. and C. Tobón (2003) "Usability Engineering and PPGIS: Towards a User-Centred Design Approach," *International Journal of Geographical Information Science*, 17(6), 577-92.

Haklay, M. (2002) "Virtual Reality in GIS: Applications, Trends and Directions", in P. Fisher and D. Unwin (eds.), *Virtual Reality in Geography,* London: Taylor and Francis, 47-57.

Hardle, W., S. Klinke and B. A. Turlach (1995) *XploRe: An Interactive Statistical Computing Environment*, Berlin: Springer.

Harris, R. J. (1999) "Geodemographics and the Analysis of Urban Lifestyles", *PhD Thesis*, University of Bristol, unpublished.

Harris, R. J. and P. A. Longley (2000) "New data and approaches for urban analysis: models of residential densities", *Transactions in GIS*, 4: 217-34.

Haslett, J., R. Bradley, P. Craig, A. Unwin and G. Wills (1991) "Dynamic Graphics for Exploring Spatial Data with Application to Locating Global and Local Anomalies," *The American Statistician*, 45(3): 234-42.

Hearnshaw, H. M. (1994) "Psychology and Displays in GIS", in H. M. Hearnshaw and D. J. Unwin (eds.), *Visualization in Geographical Information Systems,* Chichester: John Wiley and Sons.

Hewett, T. T., R. M. Baecker, S. K. Card , J. G. Carey, M. M. Gasen, G. Mantei, G. W. Perlman, G. Strong and B. Verplank (2002) "Curricula for Human-Computer Interaction" (*http://www.acm.org/sigchi/cdg/cdg2.html#2_1*, accessed January 2002).

Hinnenburg, A., D. A. Keim and M. Wawryniuk (1999) "HD-Eye: Visual Mining of High-Dimensional Data," *Computer Graphics and Applications*, September: 22-31.

Hoffman, P., G. Grinstein, K. Marx, I. Grosse and E. Stanley (1997) "DNA Visual And Analytic Data Mining", *IEEE Visualization '97*, Phoenix, AZ., October 19-24 (*http://neptune.cs.uml.edu/~phoffman/dna1/*, accessed February 2002).

Hutchins, E. L., J. D. Hollan and D. A. Norman (1987) "Direct Manipulation Interfaces (Excerpt)", in R. M. Baecker and W. A. S. Buxton (eds.), *Readings in Human-Computer Interaction: A Multidiciplinary Approach,* San Mateo: Morgan Kaufmann Publishers Inc., 469-70.

Inselberg, A. (1999) "Don't Panic... Just do it in Parallel!" *Computational Statistics*, 14: 53-77.

ITS (1997) "Repeated Measures ANOVA Using SAS PROC GLM" (*http://www.utexas.edu/cc/docs/stat40.html*, accessed February 2003).

Jacoby, W. G. (1998) *Statistical Graphics for Visualizing Multivariate Data*, Thousand Oaks: Sage Publications.

Jones, S. and M. Scaife (2000) "Animated Diagrams: An Investigation into the Cognitive Effects of using Animation to Illustrate Dynamic Processes", in M. Anderson, P. Cheng and V. Haarslev (eds.), *Lecture Notes in Artificial Intelligence 1889,* Berlin: Springer-Verlag, 231-44.

Just, M. A. and P. A. Carpenter (1976) "Eye Fixations and Cognitive Processes," *Cognitive Psychology*, 8: 441-80.

Kahkonen, J., L. Letto, T. Kilpelainen and T. Sarjakoski (1999) "Interactive Visualisation of Geographical Objects on the Internet," *International Journal of Geographical Information Systems*, 13(4): 429-38.

Kaiser, M., D. Proffitt, S. Whelan and H. Hecht (1992) "Influence of Animation on Dynamical Judgements," *Journal of Experimental Psychology: Human Perception and Performance*, 18: 669-90.

Kaushik, S. R. and E. A. Rundensteiner (2001) "SEE: A Spatial Exploration Environment based on a Direct Manipulation Paradigm.," *IEEE Transactions on Knowledge and Data Engineering*, 13(4): 654-66.

Keim, D. A. (2001) "Visual Exploration of Large Data Sets," *Communications of the ACM*, 44(8): 39-44.

Keim, D. A. (2002) "Information Visualization and Visual Data Mining," *IEEE Transactions on Visualization and Computer Graphics*, 8(1): 1-8.

Keppel, G. (1982) *Design and Analysis: A Researcher's Handbook*, London: Prentice-Hall.

Kuhn, W. (1995) "7+/-2 Questions and Answers about Metaphors for GIS User Interphases", in T. L. Nyerges, D. M. Mark, R. Laurini and M. J. Egenhofer (eds.), *Cognitive Aspects of Human-Computer Interaction for Geographic Information Systems,* Dordrecht: Kluwer Academic, 113-22.

Kirakowski, J. (1994) "The Use of Questionnaire Methods for Usability Assessment" (*http://sumi.ucc.ie/sumipapp.html*, accessed July 2003).

Klatzky, R. L. (1975) *Human Memory: Structures and Processes*, San Francisco, CA: W.H. Freeman & Co.

Kleinmuntz, D. N. and D. A. Schkade (1993) "Information displays and decision processes," *Psychologival Science*, 4(4): 221-27.

Knapp, L. (1995) "A Task Analysis Approach to the Visualization of Geographic Data", in T. L. Nyerges, D. M. Mark, R. Laurini and M. J. Egenhofer (eds.), *Cognitive Aspects of Human-Computer Interaction for Geographic Information Systems,* Dordrecht: Kluwer Academic, 355-71.

Kolácny, A. (1969) *Gestalt Psychology*, New York: Liveright.

Komzak, J. (2000) "Geovisualization", (*http://sgi.felk.cvut.cz/~komzaj1/phd/geovisual/KMi/index.html*, accessed August 2003).

Kosslyn, S. M. (1989) "Understanding Charts and Graphs," *Applied Cognitive Psychology*, 3: 185-226.

Koussoulakou, A. and M. J. Kraak (1992) "Spatio-Temporal Maps and Cartographic Communication," *The Cartographic Journal*, 29(2): 101-8.

Krygier, J. B. (1994) "Sound and Geographic Visualization", in A. M. MacEachren and D. R. F. Taylor (eds.), *Visualization in Modern Cartography,* Oxford: Pergamon, 149-66.

Kyllonen, P. C., D. F. Lohman and R. E. Snow (1984a) "Effects of Aptitudes, Strategy Training and Task Facets on Spatial Task Performance," *Journal of Educational Psychology*, 76(1): 130-45.

Kyllonen, P. C., D. F. Lohman and D. J. Woltz (1984) "Componential Modeling of Alternative Strategies for Performing Spatial Tasks," *Journal of Educational Psychology*, 76(6): 1325-45.

Kyriakidis, P. C. (2001) "Geostatistical Models of Uncertainty for Spatial Data", in C. T. Hunsaker, M. F. Goodchild, M. A. Friedl and T. J. Case (eds.), *Spatial Uncertainty in Ecology: Implications for Remote Sensing and GIS Applications,* New York, NY.: Springer-Verlag, 175-213.

Laitenberger, O. and H. M. Dreyer (1998) "Evaluating the Usefulness and the Ease of Use of a Web-based Inpecstion Data Collection Tool", *Fifth International Software Metrics Symposium*, Bethesda, MD, November 20-21, 122-32.

Landauer, T. K. (1995) *The Trouble with Computers: Usefulness, Usability and Productivity*, Cambridge, MA: MIT Press.

Larkin, J. H. and H. A. Simon (1987) "Why a Diagram is (Sometimes) worth Ten Thousand Words," *Cognitive Science*, 11: 65-99.

Lloyd, D. (2004) "Exploring Error in Socioeconomic GIS: The Case of Defining Town Centres", *PhD Thesis*, University College London, unpublished.

Lloyd, D., M. Haklay, M. Thurstain-Goodwin and C. Tobón (2003) "Visualizing Spatial Structure in Urban Data", in P. A. Longley and M. Batty (eds.), *Advanced Spatial Analysis: The CASA book of GIS,* Redlands: ESRI Press, 267-86.

Lohman, D. F. and P. C. Kyllonen (1983) "Individual Differences in Solution Strategy on Spatial Tasks", in R. F. Dillon and R. R. Schmeck (eds.), *Individual Differences in Cognition,* London: Academic Press, 105-35.

Lohse, G. L., K. Biolsi, N. Walker and H. H. Rueter (1994) "A Classificasion of Visual Representations," *Communications of the ACM*, 37(12): 36-49.

Lomax, P. (1998) *VB and VBA in a Nutshell: The Language*, Sebastopol, CA: O'Reilly.

Longley, P. A. (1998) "Foundations," in P. A. Longley, S. M. Brooks, and R. McDonnell (eds) *Geocomputation: A Primer*, Chichester: John Wiley and Sons, 3-15.

Longley, P. A., M. Goodchild, D. J. Maguire, D. W. Rhind. (2001) *Geographic Information Systems and Science*, Chichester: John Wiley.

Longley, P. A., M. Goodchild, D. J. Maguire and D. W. Rhind (1991) *Geographical Information Systems: Principles, Techniques, Management and Applications*, Cambridge: GeoInformation International.

Lucas, C. M. (1953) *Analysis of the relative movement test by a method of individual interviews*, Princeton: Educational Testing Service.

Lyman, P. and H. R. Varian (2000) "How Much Information" (*http://www.sims.berkeley.edu/how-much-info*, accessed October 2001).

MacDonald, J. A. (1982) *Interactive Graphics for Data Analysis,* Statistics Department, Stanford University, Stanford.

MacEachren, A. M. (1999) "Exploring Geo-Data Spaces: The Search for Meaning", *GeoComputation*, Mary Washington College, USA, July 25-28, 74-5.

MacEachren, A. M. (1995) *How Maps Work: Representation, Visualization and Design*, New York: The Guilford Press.

MacEachren, A. M. (1994) "Visualization in Modern Cartography: Setting the Agenda", in A. M. MacEachren and D. R. F. Taylor (eds.), *Visualization in Modern Cartography,* Oxford: Pergamon, 1-12.

MacEachren, A. M. (1992) "Visualizing Uncertainty Information," *Cartographic Perspectives*, 13: 10-19.

MacEachren, A. M., F. P. Boscoe, D. Haug and L. W. Pickle (1998) "Geographic Visualization: Designing Manipulable Maps for Exploring Temporally Varying Georeferenced Statistics", *IEEE Symposium on Information Visualization*, Research Triangle Park, North Carolina, 87- 94.

MacEachren, A. M., B. P. Buttenfield, J. B. Campbell, D. W. DiBiase and M. Monmonier (1992) "Visualization", in R. F. Abler, M. G. Marcus and J. M. Olson (eds.), *Geography's Inner World: Pervasive Themes in Contemporary American Geography,* New Brunswick: Rutgers University Press, 99-137.

MacEachren, A. M. and J. H. Ganter (1990) "A Pattern Identification Approach to Cartographic Visualization," *Cartographica*, 27(2): 64-81.

MacEachren, A. M. and M. J. Kraak (1997) "Exploratory Cartographic Visualization: Advancing the Agenda," *Computers and Geosciences*, 23(4): 335-43.

MacEachren, A. M. and M. J. Kraak (2001) "Research Challenges in Geovisualization," *Cartography and Geographic Information Science*, 28(1): 3-12.

MacEachren, A. M., A. M. Wachowicz, R. Edsall, D. Haug and R. Masters (1999) "Constructing Knowledge from Multivariate Spatio-Temporal Data: Integrating Geographical Visualization with Knowledge Discovery in Database Methods," *International Journal of Geographical Information Systems*, 13(4): 311-34.

Mackinlay, J. D. (1986) "Automating the Design of Graphical Presentations of Relational Information," *ACM Transactions and Graphics*, 5(2): 110-41.

Mackinlay, J. D., G. G. Robertson and S. K. Card (1991) "Perspective Wall: Detail and Context Smoothly Integrated", *CHI'91*, New Orleans, LA, April 27 - May 2, 173-79.

Marr, D. (1982) *Vision: A Computational Investigation into the Human Representation and Procession of Visual Information*, San Francisco: W. H. Freeman.

Marr, D. (1985) "Vision: The Philosophy and the Approach", in A. M. Aitkenhead and J. M. Slack (eds.), *Issues in Cognitive Modeling*, London: Erlbaum, 103-26.

Marr, D. and H. K. Nishihara (1978) "Representation and recognition of the spatial organization of tree-dimensional shapes", *Proceedings of the Royal Society*, 269-94.

Matlin, M. W. (2002) *Cognition*, Fort Worth, TX: Harcourt College Publishers.

McCormick, B. H., T. A. DeFanti and M. D. Brown (1987) "Visualization in Scientific Computing," *Computer Graphics*, 21(6), entire issue.

Medyckyj-Scott, D. (1992) "GIS and the Concept of Usability", in D. M. Mark and A. U. Frank (eds.), *User Interfaces for Geographic Information Systems: Report on a Specialist Meeting*. NCGIA Report 92-3, 105-10.

Medyckyj-Scott, D. (1994) "Visualization and Human-Computer Interaction", in H. M. Hearnshaw and D. J. Unwin (eds.), *Visualization in Geographical Information Systems*, Chichester: John Wiley and Sons, 200-11.

Medyckyj-Scott, D. and H. M. Hearnshaw (1993) *Human Factors in Geographical Information Systems*, London: Belhaven Press.

Miller, D. R., R. D. Dunham and J. M. Ware (2002) "The application of VR modelling in assessing potential visual impacts of rural development", in P. Fisher and D. Unwin (ed.), *Virtual Reality in Geography*, London: Taylor and Francis, 144-62.

Miller, G. A. (1956) "The Magical Number Seven, Plus or Minus Two: Some Limits to our Capacity for Processing Information," *The Psychological Review*, 63: 81-97.

Mitas, L., W. M. Brown and H. Mitasova (1997) "Simulations of landscape processes on multivariate fields," *Computers and Geosciences*, 23(4): 436-46.

Monmonier, M. (1992) "Summary Graphics for Integrated Visualization in Dynamic Cartography," *Cartography Geographic Information Systems*, 19(1): 23-36.

Monmonier, M. (1989) "Geographic Brushing: Enhancing Exploratory Analysis of the Scatterplot Matrix," *Geographical Analysis*, 21(1): 81-4.

Monmonier, M. and M. Gluck (1994) "Focus Groups for Design Improvement in Dynamic Cartography," *Cartography and Geographic Information Systems*, 21(1): 37-47.

Montello, D. R.., and S. M. Freundschuh (in press) "Cognition of geographic information", in B. McMaster (ed.), *Research agenda for the UCGIS.*

Montgomery, D. C. (1997) *Design and Analysis of Experiments*, Chichester: John Wiley and Sons.

Morrison, J. B., B. Tversky and M. Betancourt (2000) "Animation: Does it facilitate learning?" *AAAI Spring Symposium on Smart Graphics*, Stanford, CA, 53-9.

Morse, E. and M. Lewis (2000) "Evaluating Visualizations: Using a Taxonomic Guide," *International Journal of Human-Computer Studies*, 53: 637-62.

Narayanan, H. N. and M. Hegarty (2000) "Communicating Dynamic Behaviours: Are interactive Multimedia Presentations better than Static Mixed-Mode Presentations?" in M. Anderson, P. Cheng and V. Haarslev (eds.), *Lecture Notes in Artificial Intelligence 1889,* Berlin: Springer-Verlag, 178-93.

Newton, C. (1978) "Graphica: From Alpha to Omega in Data Analysis", in P. C. C. Wang (ed.), *Graphical Representation of Multivariate Data,* New York: Academic Press, 59-92.

Nielsen, J. (1992) "Finding usability problems through heuristic evaluation", *CHI'92*, Monterey, CA, May 3-7, 373-80.

Nielsen, J. (1993) *Usability Engineering*, San Diego, CA: Morgan Kaufmann.

Nielsen, J. (1994) "Heuristic Evaluation", in J. Nielsen and R. L. Mack (eds.), *Usability Inspection Methods,* New York, NY: John Wiley and Sons, 25-62.

Nielsen, J. and R. Molich (1990) "Heuristic Evaluation of User Interfaces", *CHI'90 Human Factors in Computing Systems*, Seattle, WA, April 1-5, 249-56.

Norman, D. A. (1988) *The Psychology of Everyday Things*, New York: Basic Books.

North, C. and B. Shneiderman (2000) "Snap-Together Visualization: Can Users Construct and Operate Coordinated Visualizations?" *International Journal of Human-Computer Studies*, 53: 715-39.

Nyerges, T. L., D. M. Mark, R. Laurini and M. J. Egenhofer (eds.) (1995) *Cognitive Aspects of Human-Computer Interaction for Geographic Information Systems*, Dordrecht: Kluwer Academic.

Office of the Deputy Prime Minister (2002) *Producing Boundaries and Statistics for Town Centres: London Pilot Sttudy Summary Report*, Norwich: TSO.

Openshaw, S., D. Waugh and A. Cross (1994) "Some Ideas About the Use of Map Animation as a Spatial Analysis Tool", in H. M. Hearnshaw and D. J. Unwin (eds.), *Visualization in Geographical Information Systems,* Chichester: John Wiley & Sons, 131-38.

Orford, S. D., D. Dorling and R. J. Harris (1998) "Review of Visualization in the Social Sciences: A State of the Art Survey Report", Advisory Group on Computer Graphics Technical Report Series, 41 (*http://www.agocg.ac.uk/train/review/cover.htm*, accessed January 2003).

O'Sullivan, D. and D. Unwin (2003) *Geographic Information Analysis*, New Jersey: John Wiley and Sons.

Park, O. (1998) "Visual Displays and Contextual Presentations in Computer Based Instruction," *Educational Technology Research and Development*, 46(3): 37-50.

Peterson, M. P. (1994) "Cognitive Issues in Cartographic Visualization", in A. M. MacEachren and D. R. F. Taylor (eds.), *Visualization in Modern Cartography*, Oxford: Pergamon, 27-43.

Peterson, M. P. (1999) "Active Legends for Interactive Cartographic Exploration," *International Journal of Geographical Information Systems*, 13(4): 375-83.

Peuquet, D. J. (2002) *Representations of Space and Time*, New York: Guilford.

Pinker, S. (1984) "Visual Cognition: An Introduction," *Cognition*, 18: 1-63.

Pinker, S. (1990) "A Theory of Graph Comprehension", in R. Freedle (ed.), *Artificial Intelligence and the Future of Testing*, Hillsdale, NJ: Lawrence Erlbaum Associates, Inc., 73-126.

Popper, K. R. (1980) *The Logic of Scientific Discovery*, New York: Routledge.

Preece, J., D. Benyon, G. Davies, L. Keller and Y. Rogers (eds.) (1993) *A Guide to Usability: Human Factors in Computing*, Dorset: Addison-Wesley.

Preece, J., Y. Rogers, H. Sharp, D. Benyon, S. Holland and T. Carey (1994) *Human-Computer Interaction*, Harlow: Addison-Wesley.

Reisberg, D. (1987) "External representations and the advantages of externalizing one's thoughts", *8th Annual Conference of the Cognitive Science Society*, Seattle, WA.

Risden, K., M. Czerwinski, T. Munzner and D. Cook (2000) "An Initial Examination of Ease of Use for 2D and 3D Information Visualizations of Web Content," *International Journal of Human-Computer Studies: Special Issue on Empirical Evaluations of Information Visualizations*, 53(5): 695-714.

Richardson, A. E., D. Montello and M. Hegarty (1999) "Spatial knowledge acquisition from maps, and from navigation in real and virtual environments," *Memory and Cognition*, 27: 741-750.

Robinson, A. H. (1952) *The Look of Maps*, Madison: University of Wisconsin Press.

Root, R. W. and S. Draper (1983) "Questionnaires as a Software Evaluation Tool", *CHI'83*, Boston, MA, December 12-15, 83-7.

Roth, S. F. and J. Matis (1990) "Data Characterization for Intelligent Graphics Presentation", *CHI'90*, Seattle, Washington, April 1-5, 193-200.

Rutkowski, C. (1982) "An introduction to the Human Applications Standard Computer Interface, Part I: Theory and Principles," *Byte 7*, October, 10: 291-310.

Ryan, T. A. J. and B. L. Joiner (1976) "Normal Probability Plots and Tests for Normality" (*http://www.minitab.com/resources/whitepapers/normprob.htm*, accessed December 2001).

Salthouse, T. A., R. L. Babcock, D. R. D. Mitchell, R. Palmon and E. Skovronek (2000) "Sources of Individual Differences in Spatial Visualization Ability," *Intelligence*, 14: 187-230.

Schultz, K. (1991) "The Contribution of Solution Strategy to Spatial Performance," *Canadian Journal of Psychology*, 45: 474-91.

Senn, S. (2002) *Cross-over trials in clinical research*, Chichester: John Wiley and Sons.

Shah, P. (2000) "Graph Comprehension: The Role of Format, Content and Individual Differences", in M. Anderson, M. B. Meyer and O. P. (eds.), *Diagrammatic Representation and Reasoning*, Berlin: Springer Verlag.

Shepherd, I. D. H. (1994) "Multisensory GIS: Mapping Out the Research Frontier", in T. C. Waugh and R. G. Healey (eds.), *Advances in GIS research*, London: Taylor and Francis, 169-87.

Shneiderman, B. (1998) *Designing the User Interface: Strategies for Effective Human Computer Interaction*, Reading, Mass: Addison-Wesley.

Slocum, T. A., C. Blok, B. Jiang, A. Koussoulakou, D. R. Montello, S. Fuhrmann and N. R. Hedley (2001) "Cognitive and Usability Issues in Geovisualization," *Cartography and Geographic Information Science*, 28(1): 61-75.

Snow, R. E. (1978) "Theory and method for research on aptitude processes," *Intelligence*, 2: 225-78.

Spence, R. (2001) *Information Visualization*, Essex: Pearson Education Limited.

Spence, R. and M. D. Apperley (1982) "Data Base Navigation: An Office Environment for the Professional," *Behaviour and Information Technology*, 1(1): 43-54.

Stephens, R. (1996) *Visual Basic Algorithims: A Developer's Source Book of Ready-To-Run Code*, New York: John Wiley and Sons.

Stone, M. C., K. Fishkin and E. A. Bier (1994) "The Movable Filter as a User Interface Tool", *CHI '94*, Boston, MA, 306-12.

Sutcliffe, A. G., M. Ennis and J. Hu (2000) "Evaluating the Effectiveness of Visual User Interfaces for Information Retrieval," *International Journal of Human-Computer Studies*, 53: 741-63.

Sutcliffe, A. G., M. Ryan, M. V. Springett and A. Doubleday (2000a) "Model Mismatch Analysis: Towards a Deeper Explanation of Users' Usability Problems," *Behaviour and Information Technology*, 19: 43-55.

Sutcliffe, A. G. and U. Patel (1996) "3D or not 3D: is it nobler in the mind?" *People and Computers XI: Proceedings of HCI 96*, London, 79-93.

Talen, E. (2000) "Bottom-Up GIS: A New Tool for Individual and Group Expression in Participatory Planning," *Journal of the American Planning Association*, 66(3): 279-94.

Takatsuka, M. and M. Gahegan (2002) "GeoVista *Studio*: A Codeless Visual Programming Environment for Geoscientific Data Analysis and Visualization," *Computers and Geosciences*, 28(10): 1131-44.

Tanin, E., R. Beigel and B. Shneiderman (1997) "Design and evaluation of incremental data structures and algorithms for dynamic query interfaces", *IEEE Symposium on Information Visualization*, Phoenix, AZ, October 20-21, 81-6.

Thomas, B. (1996) "'Quick and Dirty' Usability Tests", in P. W. Jordan, B. Thomas, B. A. Weerdmeester and I. L. McClelland (eds.), *Usability Evaluation in Industry*, London: Taylor and Francis, 107-14.

Thomas, P. and R. D. Macredie (2002) "Introduction to the New Usability," *ACM Transactions on Computer-Human Interaction*, 9(2): 69-73.

Thurstain-Goodwin, M. and Unwin, D. (2000) "Defining and Delineating the Central Areas of Towns for Statistical Monitoring Using Continuous Surface Representations", Transactions in GIS, 4(4), 305-18.

Tobler, W. (1970) "A Computer Movie Simulating Urban Growth in the Detroit Region," *Economic Geography*, 46(2): 234-40.

Tobón, C. (2003) "Evaluating Geographic Visualization Tools and Methods: An Approach and Experiment Based Upon User Tasks", in J. Dykes, A. M. MacEachren and M. J. Kraak (eds.), *Exploring Geovisualization*, London: Elsevier, in press.

Tobón, C. (2002). Usability Testing for Improving Interactive GeoVisualization Techniques. CASA Working Paper Series 45.

Tobón, C. and M. Haklay (2003) "Human-Computer Interaction and Usability Evaluation in GIScience", in P. A. Longley and M. Batty (eds.), *Advanced Spatial Analysis: The CASA book of GIS*, Redlands: ESRI Press, 391-406.

Trafton, J. G., S. S. Kirschenbaum, T. L. Tsui, R. T. Miyamoto, J. A. Ballas and P. D. Raymond (2000) "Turning Pictures into Numbers: Extracting and Generating Information from Complex Visualizations," *International Journal of Human-Computer Studies*, 53: 827-50.

Traynor, C. and M. G. Williams (1995) "Why are Geographic Information Systems Hard to Use?" *CHI'95 Mossaic of Creativity*, Denver, CO, May 7-11, 288-9.

Traynor, C. and M. G. Williams (1997) "A Study of End-User Programming for Geographic Information Systems", *Empirical Studies of Programmers: Seventh Workshop, ESP7*, Alexandria, VA, October 24-6,

Tufte, E. R. (1983) *The Visual Display of Quantitative Information*, Cheshire: Graphics Press.

Tufte, E. R. (1990) *Envisioning Information*, Cheshire: Graphics Press.

Tufte, E. R. (1997) *Visual Explanations: Images and Quantities, Evidence and Narrative*, Cheshire: Graphic Press.

Tukey, J. W. (1977) *Exploratory Data Analysis*, Ontario: Addison-Wesley Publishing Company, Inc.

Ungerer, M. J. and M. F. Goodchild (2002) "Integrating Spatial Data Analysis and GIS: A New Implementation using the Component Object Model (COM)," *International Journal of Geographical Information Science*, 16(1): 41-53.

U.S. Census Bureau (2002) *Census 2000 Basics.* (*http://www.census.gov/mso/www/c2000basics/00Basics.pdf*, accessed March 2003).

U.S. Census Bureau (2001) *Introduction to Census 2000 Data Products.* (*http://www.census.gov/prod/2001pubs/mso-01icdp.pdf*, accessed March 2003).

Unwin, A. (1998) "Visualisation Techniques for Statistics: The Current State of Play", *NTTS 98*, Sorrento, 111-18.

Unwin, D. (1996) "GIS, Spatial Analysis and Spatial Statistics," *Progress in Human Geography*, 20(4): 540-51.

Unwin, D. (1994) "ViSc, GIS and Cartography," *Progress in Human Geography*, 18: 516-22.

Unwin, A. R. and H. Hofmann (1999) "GUI and Command-Line: Conflict or Synergy?" *Proceedings of the 31st Symposium on the Interface*, Chicago, 246-53.

Unwin A., G. Hawkins, H. Hofmann, B. Siegl (1996) "Interactive Graphics for Data Sets with Missing Values – MANET", *Journal of Computational and Graphical Statistics*, 5(2), 113-22.

van der Wel, F. J. M., R. M. Hootsmans and F. Ormeling (1994) "Visualization of Data Quality", in A. M. MacEachren and D. R. F. Taylor (eds.), *Visualization in Modern Cartography*, Oxford: Pergamon, 313-31.

Virzi, R. (1992) "Refining the Test Face of Usability Evaluation: How Many Subjects is Enough?" *Human Factors*, 34: 457-68.

Wachowicz, M. (2001) "GeoInsight: an approach for developing a knowledge construction processbased on the integration of GVis and KDD methods", in H. J. Miller and J. Han (eds.), *Geographic Knowledge Discovery and Spatial Data Mining*, London: Taylor & Francis., 239-59.

Ware, C. (2000) *Information Visualization: Perception for Design*: Morgan Kaufmann.

Ware, C. and M. Plumlee (2003) "3D Geovisualization and the Structure of Visual Space", in J. Dykes, A. M. MacEachren and M. J. Kraak (eds.), *Exploring Geovisualization*, London: Elsevier, in press.

Wehrend, S. (1990) "A Categorization of Scientific Visualization Techniques", *PhD Thesis*, University of Colorado, unpublished.

Wehrend, S. and C. Lewis (1990) "A Problem-oriented Classification of Visualization Techniques", *First IEEE Conference on Visualization*, Los Alamitos, CA, October, 139-43.

Westerman, S. J. and T. Cribbin (2000) "Mapping semantic information in virtual space: Dimensions, variance, and individual differences," *International Journal of Human-Computer Studies*, 53: 765-87.

Wickens, C. D. (1992) *Engineering Psychology and Human Performance*, New York: Harper Collins.

Williams, S. and C. Kindel (1994) "The Component Object Model: A Technical Overview" (*http://msdn.microsoft.com/library/default.asp?url=/library/en-us/dncomg/html/msdn_comppr.asp*, accessed January 2002).

Williamson, C. and B. Shneiderman (1992) "The Dynamic Homefinder: Evaluating dynamic queries in a real estate information exploration system", *SIGIR'92*, Copenhagen, 339-46.

Wise, S., R. Haining and J. Ma (2000) "Integrating Spatial Data Analysis with GIS: Lessons Learned from SAGE", *GIS Research UK*, University of York, 181-83.

Wood, J. D. (2002) "Visualizing the Structure and Scale Dependency of Landscapes", in P. Fisher and D. Unwin (ed.), *Virtual Reality in Geography*, London: Taylor and Francis, 163-74.

Wood, J. D., P. F. Fisher, J. A. Dykes, D. J. Unwin and K. Stynes (1999) "The Use of the Landscape Metaphor in Understanding Population Data," *Environment and Planning B: Planning and Design*, 26: 281-95.

Wood, M. (1994) "The Traditional Map as a Visualization Tool", in H. M. Hearnshaw and D. J. Unwin (eds.), *Visualization in Geographical Information Systems*, Chichester: John Wiley & Sons, 9-17.

Wood, M. and K. Broodlie (1994) "ViSC and GIS: Some Fundamental Considerations", in H. M. Hearnshaw and D. J. Unwin (eds.), *Visualization in Geographical Information Systems*, Chichester: John Wiley & Sons, 3-8.

Wright, H., K. Brodlie, J. Wood and J. Proctor (2000) "Problem Solving Environments: Extending the Role of Visualization Systems", *Euro-Par 2000 Parallel Processing, Lecture Notes in Computer Science 1900*, 1323-31.

Zhang, J. (1997) "The Nature of External Representations is Problem Solving," *Cognitive Science*, 21(2): 179-217.

Zhang, Z. and D. A. Griffith (2000) "Integrating GIS Components and Spatial Statistical Analysis in DBMSs," *International Journal of Geographical Information Science*, 14(6): 543-66.

Zhou, M. X. and S. K. Feiner (1996) "Data Characterization for Automatically Visualizing Heterogeneous Information", *IEEE Symposium on Information Visualization*, San Francisco, CA, October 28-29, 13-20.

Zhou, M. X. and S. K. Feiner (1998) "Visual Task Characterization for Automated Visual Discourse Synthesis," *CHI'98*, Los Angeles, CA, April 18-23, 392-99.

Appendix A: Experiments Material

This appendix contains the invitation to participate, the instructions provided before the test and the tasks administered to users in each experiment. The guidelines for the semi-structured interviews conducted after each experiment are also included. Finally, for the second experiment only, the questionnaire is also included.

A.1 Experiment 1

As discussed in Chapter 2, the main aim of this study was to evaluate whether the functionality provided by the coupling between the lightweight GIS and DecisionSite served to create an environment that supported the exploration of spatial data. Ultimately, the aim was to identify the potential of linking systems to support geovisualization tasks for spatial data users. In addition, the study was designed so that it would be possible to review participants' work processes, how they thought and behaved when using the system. The following subsections describe the materials used throughout the experiments' life-cycle: from the invitation to participants to take part, through the testing session where users completed task scenarios and were then interviewed. These materials are provided here for reference for those wishing to replicate or conduct similar studies.

A.1.1 Invitation to participate

Dear Sir/Madam,

We would like to invite you to participate in a software evaluation exercise of a commercial product designed for the visual exploration of spatial data. Your participation would be most valuable as we are investigating the suitability of the environment for this purpose as well as its potential limitations.

If you would like to participate, you will be invited to a session to the Department of Geography at UCL. You will be given a demonstration of how the software environment works and you will have enough time to get used to its functionality. You will then be asked to solve a few problem scenarios.

To participate, it is not essential that you have any previous experience with other visualisation software. Please email Carolina Tobón (c.tobon@ucl.ac.uk) if you would like to take part. Otherwise, if you have any queries or would like to receive further information about the testing session, please let us know.

Many thanks for your help and interest and we hope to see you soon.

A.1.2 Instructions and Data

Thank you for participating in this usability test. I am sure you will find it interesting and it will help us determine whether the software you will be using can assist users in the interactive and visual exploration of highly multivariate spatial datasets. Before starting, please answer the following general information questions:

1. Name
2. Age
3. Sex
4. Occupation
5. Have you used any of the following types of Software? (Indicate as many as appropriate)
    a. Word Processors
    b. Databases
    c. Spreadsheets
    d. Statistical packages
    e. GIS
    f. Data exploration packages
    g. Graphical packages
    h. Other

The following text is the same you received in a previous email about the dataset you will be exploring. It is included here if you need it for reference during your session.

Lifestyles and Census Small Area Statistics

Lifestyle databases contain individual-level data (or household data) about consumer habits and behaviours. They have been predominantly used for direct marketing, and are mostly obtained from responses to postal consumer surveys or questionnaires. Each survey usually consists of a few hundred questions on varied "lifestyles" topics ranging from income and occupation to shopping habits, holiday and travel preferences. These datasets are generally very diverse and frequently concatenated from disparate data sources using a common field, the residential address, to form a profile of individual households.

However, the lifestyles datasets have various limitations. For instance, since the data are collected for marketing purposes, answers to each question must be chosen from a set of pre-defined responses that are not suitable for all types of analysis. Also, there is no obligation to complete the questionnaires. Hence, the respondents are self-selecting and constitute a non-random sampling of the population (Harris, 1998). These limitations make lifestyles data "fundamentally more biased and unscientific" (Longley, 1998). Why then use this dataset as opposed to the more traditional and widely used UK Census of population? Without entering in the discussion of the limitations of the Census data, 2 reasons make the Lifestyles data an interesting option: 1) they are updated every year (as opposed to once every ten years as in the case of the Census data), and 2) they contain direct questions on income which the Census lacks. Hence, the Lifestyles increased geographic coverage, detail and periodicity offer an alternative source of information that can challenge the traditional use of Census-based geodemographic data both for marketing applications and academic research.

The dataset you will be exploring contains an Income Score variable (Y) derived from a Lifestyles dataset and a set of variables chosen from the 1991 Census' Small Area Statistics (SAS). Hence, the data are aggregated to the Enumeration District (ED) level since that is the finest spatial resolution for publicly available Census data. Therefore, Y is an average score for each ED, where the average is across the 'respondent' households in that ED. The SAS variables are expressed as a percentage of households enumerated by the Census in each ED. Y ranges from 0-10 (where each score corresponds to an income band) and SAS data range from 0-100.

You will be looking at 19 variables available for 844 EDs in Bristol. Briefly, the variables are:

- Bigacc: % Households with 7+ rooms
- Bighhld: % Households with 6+ persons
- Buying: % Households in owner-occupied property buying it
- Bought: % Households in owner-occupied property, bought outright
- Childn: % Persons aged 5-14
- Youths: % Persons aged 15-24
- Yngadl: % Persons aged 25-44
- Oldadl: % Persons aged 45-64
- Oldpsn: % Persons aged 65-74

- Councl: % Households in council or new town rented property
- Detached: % Unshared dwellings detached
- Terrce: % Unshared dwellings terraced
- Hhnoca: % Households with no car
- Hhtcar: % Households with 2 cars
- Qualm: % Adult, male residents qualified
- Wkwife: % Married females working
- Inc_Score: Income score variable
- Unfrnt: % Households in unfurnished rented property
- Unemp: % Economically active unemployed

A.1.3 The Tasks

You have two tasks to solve. They should take around 20 minutes each. We will spend about 10 minutes at the end of the session discussing your results, observations, and general experience using the software.

Part 1

1. Discover whether variables in this dataset are related in some way and whether their relations (if any) occur in any particular areas of Bristol. (You may want to use some of the various tools you were presented in the demo to explore the data).
2. Can you detect 'outlier' or unusual combinations of attribute values?
3. Can you detect possible mistakes in the dataset entries?

Part 2

Having explored the data, can you to choose candidate locations for (please take note of the area ID):

4. A secondary school
5. A gourmet shop

A.1.4 Interview Schedule

Participants were told that the interview was a discussion of the environment for spatial data exploration they had just used. The purpose of the interview was to find about their thoughts and impressions about using the environment with the aim of investigating spatial data at the initial stages of data exploration. They were asked for their permission to tape the interview and were advised that the discussion and their results would be kept anonymous. The topics

discussed are enumerated below. The interview however was semi-structured meaning that it did not have to abide to any particular order. In addition, comments made by the users which were of particular interest were pursued in more depth. Hence, all the interviews differed slightly as they deviated from these guidelines wherever it was considered relevant or necessary.

- How would you describe your general experience with the software?
- Did you need help with the system after the introduction?
- Once your question was solved, did you find other problems?
- Can you recall the variables you explored in the first exercise?
- Was there a reason for choosing those and not others?
- If browsing was initially at random, what was the first interesting observation you gathered from exploring the dataset?
- What caught your attention?
- Was it an outlier, a trend, a cluster? Was it an insight? Was it a confirmation of a hypothesis?
- Can you recall how you came to having that insight?
- When you found some piece of information, did you try to find more about it?
- Was any tool useful for acquiring that piece of information?
- Was any tool difficult to use?
- Was any tool not helpful in giving you information about the data?
- In the second task, why did you choose those candidate locations?
- Did you feel you needed more information to solve the tasks or to answer to your own questions?

A.2 Experiment 2

As it was discussed in Chapter 3, the understanding of user tasks was at the centre of this experiment. Furthermore, the understanding of how these tasks were accomplished and how they can be better supported was expected to shed light on what geovisualization tools should provide was also of interest. Users' perceptions on ease of use and usefulness of the software thus a main concern, as well as the flexibility of the environment to re-express and manipulate data and the extent to which it supported users' work processes. The invitation to participate in the study as well as information provided to participants for completed the

experiment are detailed below. In addition, an outline of the topics explored in the interview and the questionnaire completed by participants are also included for future reference.

A.2.1 Invitation to participate

Dear Sir/Madam,

We would like to invite you to participate in a software testing exercise of a software environment designed for the visual exploration of spatial data. The environment was created by linking ArcGIS to DecisionSite, a powerful visualisation software, to support the investigation of spatial data at the initial stages of data analysis.

Your participation would be most valuable. You will also have the opportunity to use an environment created using COM technology that can support the effective exploration of complex, large and multivariate spatial datasets.

If you would like to participate, you will be invited to come to two sessions to the Department of Geography at UCL. During the first session, you will be given a demonstration of how the software environment works and you will have enough time to get used to its functionality. Then you will be asked to solve a series of tasks or problem scenarios. During the second session, you will again solve a series of problem scenarios and answer a questionnaire in which you will be able to evaluate your experience with the software.

Because of the nature of the test, the two sessions must be completed in two different days. Each should not take longer than one and a half hours and you can choose the time slots that are most convenient for you to come to the sessions.

The sessions will take place from July 29 to August 9 from:

9am-11am; 11am-1pm; 1pm-3pm; 3pm-5pm; 5pm-7pm

Please bear in mind that one person will be conducting the tests so the earlier you reply, the easiest it will be to accommodate the times that suit you best.

To participate, it is not essential that you have used ArcGIS before although knowledge of a GIS package is needed. However, you do not need any previous experience with other visualisation software to participate.

If you would like to participate please email Carolina Tobón (c.tobon@ucl.ac.uk). Otherwise, if you have any queries or would like to receive further information about the testing sessions, please send an email to the same address.

Many thanks for your help and interest and we hope to see you soon.

A.2.2 Data Description: Lifestyles and 1991 UK Census data

Lifestyles Data

Variable Name	Values	Description
INC_SCORE	1.2 – 1.99	Indicates that the average income band for the ED is **under £5,000**
	2 – 2.99	Indicates that the average income band for the ED is **between £5,000-£9,999**
	3 – 3.99	Indicates that the average income band for the ED is **between £10k - £14,999**
	4 – 5.49	Indicates that the average income band for the ED is **between £15k - £19,999**
	5.5 – 7.49	Indicates that the average income band for the ED is **between £20k - £29,999**
	7.5 -9.99	Indicates that the average income band for the ED is **between £30k - £39,999**
	10	Indicates that the average income band for the ED is **over £40k**

1991 Census SAS Data

	Min and Max	
CHILDN	0 - 26%	Percentage of children aged 5-14 in the ED
YOUTHS	0 - 39%	Percentage of people aged 15-24 in the ED
OLDPSN	0 - 27%	Percentage of persons aged 75 + in the ED
2534NK	0 - 42%	Percentage of residents in households, aged 25-34, with no children aged 0-15
BIGHHD	0 - 12%	Percentage of households with 6+ persons
WKWIFE	0 - 100%	Percentage of married females working
STUDNT	0 - 106%	Percentage of residents aged 16+ who are students
UNEMPM	0 - 54%	Percentage of economically active unemployed people in the ED
BIGACC	0 - 76%	Percentage of households with 7+ rooms
BOUGHT	0 - 56%	Percentage of households in owner-occupied property which has been bought outright
BUYING	0 - 85%	Percentage of households where occupants are the owners and are buying the property
COUNCL	0 - 98%	Percentage of households in council rented property in the ED
HHTCAR	0 - 52%	Percentage of households with two cars
UNSKLD	0 - 35%	Percentage of unskilled workers in the ED
QUALML	0 - 89%	Percentage of adult male residents who are qualified

A.2.3 The Tasks

Tasks by factor levels

Session 1

1 or 2 Spatial Units, 1 Attribute

1.1	1 Spatial Unit, 1 Attribute, Identify	Can you identify the ED with the lowest score for income (INC)? Write down the ID.
2.1	1 Spatial Unit, 1 Attribute, Locate	Where is the ED with the highest score for income (INC)? Write down the ID.
3.1	2 Spatial Units, 1 Attributes, Compare	Areas of high income tend to be clustered. By exploring the data, can you think of why the two EDs with the highest score for income (INC) are not close together?
4.1	2 Spatial Unit, 1 Attribute, Associate	Are the two areas with the highest proportion of council accommodation (COUNCIL) close to the city centre?

Many Spatial Units, 1 Attribute

5.1	Many Spatial Units, 1 Attribute, Identify	Identify the EDs with the lowest income score (INC<2) that are close to the river. Write down the IDs.
6.1	Many Spatial Units, 1 Attribute, Locate	Where in Bristol are a large proportion of households buying the property they live in (BUYING)?
7.1	Many Spatial Units, 1 Attribute, Compare	Are there any apparent clusters of EDs with a high proportion of large households (BIGHLD)?
8.1	Many Spatial Units, 1 Attribute, Associate	Are areas of high unemployment close to the historic centre?

1 or 2 Spatial Units, Many Attributes

9.1	1 Spatial Unit, Many Attributes, Identify	Typically, EDs with a high proportion of working wives (WKWIFE>60) have a relatively low proportion of households living in council accommodation. Can you identify an ED where this is not true?
10.1	1 Spatial Unit, Many Attributes, Locate	Are any banks located close to the area with the highest proportion of qualified males?
11.1	2 Spatial Units, Many Attributes, Compare	Find the ED with the largest proportion of young couples with no kids (2534NK) and write down the ID. How does it compare to the level of income (INC) and qualified males (QUALML) in that ED?
12.1	2 Spatial Unit, Many Attributes, Associate	Is the area with the highest proportion of young couples with no kids (2534NK) close to facilities such as banks and supermarkets?

Many Spatial Units, Many Attributes

13.1	Many Spatial Units, Many Attributes, Identify	Young working couples (2534NK) tend to live in affluent areas in Bristol. Are there EDs with a large proportion of 2534NKNK residents that also have high percentages of council accommodation (COUNCIL)

		and unemployment (UNEMP)? If so, write down its ID.
14.1	Many Spatial Units, Many Attributes, Locate	Are there any EDs with unexpected combinations of income (INC) and unemployment levels (UNEMP)?
15.1	Many Spatial Units, Many Attributes, Compare	Is the relation between income (INC) and being a male resident with a qualification (QUALML) different in the NW than in the SE?
16.1	Many Spatial Units, Many Attributes, Associate	You are an estate agent with a client interested in an area where she could locate a Gourmet Shop close to a road. Can you recommend 3 locations? Write down the EDs' IDs.

Session 2

1 or 2 Spatial Units, 1 Attribute

1.2	1 Spatial Unit, 1 Variable, Identify	Can you identify the ED with an unusually high proportion of working wives (WKWIFE)? Write down the ID.
2.2	1 Spatial Unit, 1 Attribute, Locate	Can you locate an ED with an unusually high proportion of students (STUDNT)? Write down the ID.
3.2	2 Spatial Unit, 1 Attribute, Compare	Are EDs with a large proportion of children (CHILDN) close to secondary schools?
4.2	2 Spatial Unit, 1 Attribute, Associate	Apart from where there are no residents aged 75+ (OLDPSN), are the EDs with the lowest proportion for this variable close to each other or to a particular area?

Many Spatial Units, 1 Attribute

5.2	Many Spatial Units, 1 Attribute, Identify	Identify areas with a high proportion of student population. Write down the IDs.
6.2	Many Spatial Units, 1 Attribute, Locate	Close to the river, which are the EDs where most properties have been bought by their tenants (BOUGHT>90)? Write down their IDs.
7.2	Many Spatial Units, 1 Attribute, Compare	Are there any apparent clusters of EDs with low proportions of large houses (BIGACC)?
8.2	Many Spatial Units, 1 Attribute, Associate	Are areas with a low proportion of children close to each other?

1 or 2 Spatial Units, Many Attributes

9.2	1 Spatial Unit, Many Attributes, Identify	Students tend to live in less richer areas. Is there an ED with a high income score (INC>6) which also has an unusually high proportion of students?
10.2	1 Spatial Unit, Many Attributes, Locate	Typically, EDs with a high proportion of working wives (WKWIFE) also have low proportions of children (CHILDN). Can you locate an ED where this is not the case?
11.2	2 Spatial Units, Many Attributes, Compare	In most areas in Bristol, EDs where households live in large accommodations (BIGACC) also have large

		proportions of children (CHILDN) and YOUTHS. Can you find at least one location where BIGACC is not associated to high levels of CHILDN and YOUTHS and explain why this might be?
12.2	2 Spatial Unit, Many Attributes, Associate	Is the area with the highest proportion of teenagers (YOUTHS) close to secondary and high schools?
Many Spatial Units, Many Attributes		
13.2	Many Spatial Units, Many Attributes, Identify	Can you find any EDs with a high proportion of young children (CHILDN) and working wives (WKWIFE) that also has a high level of INC? If so, write down its ID.
14.2	Many Spatial Units, Many Attributes, Locate	Do EDs with a high proportion of children (CHILDN) or large percentage of older residents (OLDPSN) tend to be located towards the periphery of the study area?
15.2	Many Spatial Units, Many Attributes, Compare	Where do EDs with a high proportion of residents aged 75+ (OLDPSN) tend to be located and how do they relate with the income score (INC) of adjacent areas?
16.2	Many Spatial Units, Many Attributes, Associate	You have been hired by Bristol council to recommend 1 location for building a high school for deprived teenagers. Where would you build it? Write down the ED's ID.

A.2.3.1 Tasks administered: Day 1

1. Can you identify the ED with the lowest score for income (INC_SCORE), which is different from 0?
2. Can you locate an ED with an unusually high proportion of students (STUDNT)?
3. Areas of high income tend to be clustered. By exploring the data, can you think of why the two EDs with the highest score for income (INC_SCORE) are not close together?
4. Are the two areas with the highest proportion of council accommodation (COUNCL) adjacent to the city centre?
5. Can you identify the EDs with the lowest income score (INC_SCORE) that are adjacent to the river?
6. Where in Bristol are a large proportion of households investing in buying the property they live in (BUYING)?
7. Are there any apparent clusters of EDs with a high proportion of large households (BIGHLD)?
8. Where are areas of high unemployment (UNEMPM) located relative to the historic centre?

9. Typically, EDs with a high proportion of working wives (WKWIFE) have a relatively low proportion of households living in council accommodation (COUNCL). Can you identify an ED where this is not true?
10. Are any banks near the area with the highest proportion of qualified males (QUALML)?
11. Find the ED with the largest proportion of young couples with no kids (2534NK). How does it compare to the level of income (INC_SCORE) and qualified males (QUALML) in that ED?
12. Is there a relation between the location of some banks and that of young couples with no kids (2534NK)?
13. Young working couples with no kids (2534NK) tend to live in affluent areas in Bristol. Are there EDs with a large proportion of 2534NK residents that also have high percentages of council properties (COUNCIL) and unemployment (UNEMP)?
14. Can you find any EDs with an unexpected relation between income (INC_SCORE) and unemployment (UNEMP)?
15. Is the relation between income (INC_SCORE) and being a male resident with a qualification (QUALML) different in the area to the northwest of the city centre and than in the southeast of Bristol?
16. You are a state agent with a client interested in an area where she can start a Gourmet Shop adjacent to a road. Can you recommend 3 locations?

A.2.3.2 Tasks administered: Day 2

1. Can you identify an ED with an *unusually* high proportion of working wives (WKWIFE)?
2. Where is the ED with the highest score for income (INC_SCORE)?
3. Apart from where there are no residents aged 75+ (OLDPSN), are the EDs with the lowest proportions for this variable close to each other?
4. Are the two EDs with the largest proportion of children (CHILDN) close to secondary schools?
5. Identify areas with the highest proportion of students (STUDNT).
6. Close to the river, which are the EDs where most households live in property bought by their tenants (BOUGHT)?
7. Are there any apparent clusters of EDs with low proportions of large houses (BIGACC)?
8. Where are the areas with a low proportion of children (CHILDN) located?
9. Students tend to live in less richer areas. Is there an ED with a high income score (INC_SCORE) which also has an *unusually* high proportion of students (STUDNT)?

10. Typically, EDs with a high proportion of working wives (WKWIFE) also have low proportions of children (CHILDN). Can you locate an ED where this is not the case?
11. In Bristol, EDs where households live in large accommodations (BIGACC) also have large percentages of children (CHILDN) and persons aged 15-24 (YOUTHS). Is this true?
12. Is there a relation between the location of young people between 15-24 years of age (YOUTHS) and secondary schools or colleges?
13. Can you find any EDs with a high proportion of young children (CHILDN) and working wives (WKWIFE) that also has a high level of income (INC_SCORE)?
14. Where do EDs with a high proportion of children (CHILDN) or large percentage of older residents (OLDPSN) tend to be located?
15. Find where EDs with a high proportion of residents aged 75+ (OLDPSN) tend to be located. How do these areas relate with the income score (INC_SCORE) level of adjacent areas?
16. You have been hired by Bristol City Council to recommend one location for building a high school for deprived young people between 15-18 years of age (YOUTHS). Where would you build it?

A.2.4 Questionnaire

Users answered the following questions by ticking one of seven boxes indicating their level of agreement with the statement (i.e. strongly agree, agree, slightly agree, don't know, slightly disagree, disagree or strongly disagree) by checking on the appropriate box (not shown).

1. Using the software would help me find more information from the exploration of spatial data.
2. I found the system unnecessarily complex.
3. Using the software would increase my productivity in investigating spatial data.
4. I found the software easy to use.
5. Using the software would enable me to improve my exploration of spatial data.
6. Using the software would help me investigate spatial data more quickly.
7. It was easy to remember how to use the software.
8. I found the software useful for investigating spatial data.
9. I had to learn a lot of things before I could get going with the system.
10. Learning to operate the software was easy for me.
11. My interaction with the software was clear and understandable.

12. Using the software would make it easier to investigate spatial data.
13. It was easy to become skillful using the software.
14. I found it easy to get the software to do what I wanted it to do.

A.2.5 Interview Schedule

Similar to the interview during the first experiment, participants were told that the interview was a discussion of the environment for spatial data exploration they had just used. The purpose of the interview was to find about their thoughts and impressions about using the environment with the aim of investigating spatial data at the initial stages of data exploration. They were asked for their permission to tape the interview and were advised that the discussion and their results would be kept anonymous. The topics discussed are enumerated below. The interview however was semi-structured meaning that it did not have to abide to any particular order. In addition, comments made by the users which were of particular interest were pursued in more depth. Hence, all the interviews differed slightly as they deviated from these guidelines wherever it was considered relevant or necessary.

Experience with the software

1. What did you think about the environment?
2. Was there anything about the environment or any of its parts that you found particularly useful?
3. Was there anything less useful?

Comparison

1. How would you compare exploring a spatial dataset you know very little about in the environment you just used, and performing the same tasks you did over the past 2 sessions using only a GIS?
2. In what way was the environment better than a GIS by itself for exploring an unknown spatial dataset?
3. In what ways was it worse than using only a GIS?
4. Did the environment enabled you to find the information required from the tasks?
    a. Could you find that information quickly?
    b. Could you have found the same information using a GIS?
        i. If not, Why?
        ii. If so, would that experience be any different?
5. Would it had taken you the same amount of time?

6. Would it had taken you more time?

Data visualisations.

1. Did you understand the visualisations? The scatterplot, the parallel coordinate plot (PCP), and the histogram?
2. Did they help you to understand the dataset?
    a. If not, why?
    b. If so, what did each one help you to understand?
        i. The scatterplot?
        ii. The PCP?
        iii. The Histogram?

Finding information.

1. How did you most often find information?
2. Can you remember an occasion when you felt you could not find the information requested?
    a. Did you not know where to find it?
    b. How to find it?

Suggestions

1. Do you have any suggestions for improving the environment?

A.3 Experiment 3

As detailed in Chapter 3, the aim of this experiment was to investigate the cognitive visual operations in detail, as well as the different user approaches to solving the tasks. The different communications to participants and materials the where required throughout the experiment's life cycle are detailed here for reference.

A.3.1 Invitation to participate

Dear Sir/Madam,

Thank you once again for your participation in the software testing exercise of last summer. The evaluation was a success and we would like to invite you to participate in a follow-up test.

Your participation would be extremely valuable as you have now some experience operating the software environment designed for the visual exploration of spatial data.

If you decide to participate, you will be invited to come to **one** session. It will take place in the usability lab of the University College London Interaction Centre (UCLIC) from **May 19 - 30** at the time which you might find most convenient.

You will be given a short demonstration of how the software environment works and will have enough time to remember its functionality. Then you will be asked to solve a few tasks or problem scenarios. The session is expected to take about one hour.

Please bear in mind that one person will be conducting the tests so the earlier you reply, the easiest it will be to accommodate the time that suits you best.

If you would like to participate please email Carolina Tobón (c.tobon@ucl.ac.uk). Otherwise, if you have any queries or would like to receive further information about the testing, please send an email to the same address.

Many thanks for your help and interest and we hope to see you again soon.

A.3.2 Data Description: US Census 2000

Pop 2000	Total population in 2000
White	Total White population
Black	Total Black population
Hispanic	Total Hispanic population
p_White	White / pop2000
p_Black	Black / pop2000
p_Hispanic	Hispanic/ pop2000
age_under5	Male and female under 5 years of age
age_5_17	Male and female under 5-17 years of age
age_18_21	Male and female under 18-21 years of age
age_22_29	Male and female under 22-29 years of age

age_30_39	Male and female under 30-39 years of age
age_40_49	Male and female under 40-49 years of age
age_50_64	Male and female under 50-64 years of age
age_65_up	Male and female under 65_up years of age
med_age	Median age both sexes
households	Number of households
ave_hh_sz	Average household size
families	Number of families
ave_fam_sz	Average family size
hse_units	Number of house units
vacant	Vacant housing units
owner_occ	Owner occupied housing units
renter_occ	Renter occupied housing units
Car	People commuting to work on private transport (own car or car pool)
Publictran	People commuting to work using public transport (bus, train, metro)
aggFmlyInc	Aggregate family income sums the incomes of all members 15 years old and over related to the householder
medianFmlyInc	Median family income based on the distribution for the total number of families including those with no income. The median divides the income distribution into 2 equal parts: One half of the cases falling below the median income and one half falling above.
aggHhdInc	Aggregate household income sums the income of the householder and all other individuals 15 years old and over in the household, whether they are related to the householder or not.
medianHhdInc	Median household income based on the distribution for the total number of families including those with no income.
aggSalaInc	Aggregate salary income which includes total money earnings received for work performed as an employee

A.3.3 The Tasks

Part 1

1. Is there a relation between the use of public transport and the location of colleges and universities? Please write down your answer.

2. Is there a relation between the spatial distribution of tracts with a higher percentage of Hispanic population and the distribution of tracts where the majority of its inhabitants are White? If any, please write down what the relation is.

3. Is there a relation between owner occupancy (Owner_Occ) and ethnic origin and does it show a pattern in space? If any, please write down what the relation is.

4. Can you describe the older population in LA by ethnic origin and where they reside? Please write down your answer.

Part 2

1. Where are the tracts with the most users of public transport? Please describe their general location or give tract IDs.

2. Is there a difference in the household size of the 3 ethnic groups? Please write down your answer.

3. Can you distinguish areas in LA where a high proportion of the residents are Hispanic from areas where they are Black? If you can, please write down where they are.

4. Are there tracts of very high aggregate salary income (AggSalaInc)? Please write down their IDs.

A.3.4 Interview

- Can you tell me what you can remember about your thoughts while solving the last task?

- Can you tell me what you can remember about your thoughts while solving task xx?

- Can you think of an occasion when you found an interesting piece of information? Can you describe what you saw or found that brought on that experience?

- Can you think of an occasion when it was difficult or you could not find some information? Can you describe what brought on that experience?

- Can you think of an occasion when you little or no use of the map? Can you describe what brought on that experience?

- Can you think of an occasion when you mostly used the map to solve a task?
- Can you think of an occasion when you felt frustrated? Can you describe what brought on that experience?
- Can you think of an occasion when you made a mistake that had an effect on your answer to the task? Can you describe what brought on that experience?

Appendix B: Code Examples

The following is a selection of routines used in the close coupling of ArcView 8.1 and DecisionSite 7 using component object model (COM) technology for creating a spatial data exploration environment. The coupling is a one-to-one link between 1 ArcMap and 1 DecisionSite application.

For a movie of how it works see http://www.casa.ucl.ac.uk/carolina/sfGIS.htm The code was last updated on 5 June 2002.

```
'*****************************************************************************
'  Function name: UIButtonControl1_Click
'  Description:  Launches DecisionSite from ArcMap as a client
'                Customises visualisations in DecisionSite
'  Important:    Code must be associated to a ButtonControl created in ArcMap
'*****************************************************************************

Private Sub UIButtonControl1_Click()
   'Flag for checking if DecisionSite has been launched as client
   SFstate = 1
   sfFileName = "C:\Documents and Settings\Carolina\My
Documents\UsabilityTests\UsabilityTest April_2003\censusDemogByTract2.txt"
   Dim myFile As String
   myFile = Dir(sfFileName)

   'Column names
   v1 = "p_Hisp"
   v2 = "p_White"
   v3 = "p_Black"
   v4 = "AggFmlyInc"

   'Check if file is there
   If (Len(myFile) = 0) Then
      MsgBox "File does not exist in the specified directory.", _
         vbExclamation, "Error"
      Exit Sub
   End If

   Set theApp = CreateObject("Spotfire.Application")
   theApp.Visible = True
   'Set theVis = theApp.Open("D:\DONOTDELETE\Papers\NewBristol\y1.txt")
   Set theVis = theApp.Open(sfFileName)

   'Remove the default charts: Scatterplot, web win
   Dim s As Spotfire.ScatterPlot
   Set s = theVis.Views("Scatter Plot")
   s.Close
   Spotfire.ControlWindows("Web Window").Visible = False
```

```
    'Add some visualizations

    'A histogram
    Dim theBC As Spotfire.IBarChart
    Set theBC = theVis.Views.Add(sfvtBarChart)
    theBC.Axes.item(sfXAxis).Column = v1
    theBC.ColoringScheme = sfColoringSchemeFixed
    theBC.ColorAll = RGB(255, 0, 0)

    'Details on demand and query devices
    Dim theDetails As Spotfire.DetailsText
    Set theDetails = theVis.Views.Add(sfvtDetailsText)
    Set theCP = theVis.Views.Add(sfvtControlPanel)

    'Sets the enabled state of the ButtonControl once DecisionSite is launched
    UIButtonControl1_Enabled
    'populate combo box with colours
    UIComboBoxControl1.RemoveAll
    PopulateComboBox

    'Prepares ArcView to listen to events
    SetEvents
End Sub


'*************************************************************************
'  Function name: selectRecsInAGISfromSFrecs
'  Description:  Selects features in ArcMap from marked records in DecisionSite
'*************************************************************************

Function selectRecsInAGISfromSFrecs()

    'QIs
    Set pMxDoc = ThisDocument
    Set pMap = pMxDoc.FocusMap
    Set pActiveView = pMap

    Dim pFeature As IFeature
    Dim pFeatureCursor As IFeatureCursor
    Dim pFeatureClass As IFeatureClass
    Dim pFeatureSelection As IFeatureSelection

    'Identify the layer to explore in ArcView
    exploreLayer

    Set pFeatureClass = pFeatureLayer.FeatureClass
    Dim pRow As IRow
    Dim pTable As ITable

    'Need this to get the records in an ArcView table
    Dim pQueryFilter As IQueryFilter
    Set pQueryFilter = New QueryFilter
```

```
'Get the table
Set pTable = pFeatureClass
Set pRow = pTable.GetRow(0)

'Select features in ArcView

Dim pSelectionSet As ISelectionSet
Set pFeatureSelection = pFeatureLayer

Set pFeatureSelection.SelectionColor = pSelectionColor
Set pSelectionSet = pFeatureSelection.SelectionSet

'Loop through the selected features per layer
'Can use Nothing keyword if you don't want to draw them,
'otherwise, the spatial reference might not match the Map's
pSelectionSet.Search Nothing, False, pFeatureCursor
Set pFeature = pFeatureCursor.NextFeature

'Build a comma-delimited list from array theSFrecs
'to use in whereClause for feature selection in ArcView
Dim sWhereClause As String
Dim sList As String
featureID = "ID"
sList = ""

Dim h As Integer
For h = 0 To UBound(theSFrecs)
   If sList = "" Then
     sList = theSFrecs(h)
   Else
     sList = sList & "," & theSFrecs(h)
   End If
Next h

'Use all the selected/marked/highlighted records in DecisionSite for
'feature selection in ArcView
Dim i As Integer
For i = 0 To theNumRecs - 1
   If sList <> "" Then
      sWhereClause = featureID & " IN (" & sList & ")"

   Else
      sWhereClause = ""
   End If
Next i

'Highlight the features in ArcView

'Invalidate only the selection cache
'Flag the original selection
pActiveView.PartialRefresh esriViewGeoSelection, Nothing, Nothing
pQueryFilter.WhereClause = sWhereClause
'Perform the selection
```

```
    pFeatureSelection.SelectFeatures pQueryFilter, esriSelectionResultNew, False
    'Flag the new selection
    pActiveView.PartialRefresh esriViewGeoSelection, Nothing, Nothing

End Function

'*******************************************************************
'  Function name: doRefreshInAGIS
'  Description:     De-highlights selected features in ArcMap, if any.
'*******************************************************************

Function doRefreshInAGIS()
    'QIs
    Set pMxDoc = ThisDocument
    Set pMap = pMxDoc.FocusMap
    Set pActiveView = pMap

    Dim pFeature As IFeature
    Dim pFeatureCursor As IFeatureCursor
    Dim pFeatureClass As IFeatureClass
    Dim pFeatureSelection As IFeatureSelection

    'Identify the layer to explore in ArcView
    exploreLayer

    Set pFeatureClass = pFeatureLayer.FeatureClass
    Dim pRow As IRow
    Dim pTable As ITable

    'Need this to get the records in a table in ArcView
    Dim pQueryFilter As IQueryFilter
    Set pQueryFilter = New QueryFilter

    'get the table
    Set pTable = pFeatureClass
    Set pRow = pTable.GetRow(0)

    'De-Highlight selected features

    Dim pSelectionSet As ISelectionSet
    Set pFeatureSelection = pFeatureLayer
    Set pSelectionSet = pFeatureSelection.SelectionSet

    'Loop through the selected features per layer
    'Can use Nothing keyword if you don't want to draw them,
    'otherwise, the spatial reference might not match the Map's
    pSelectionSet.Search Nothing, False, pFeatureCursor
    Set pFeature = pFeatureCursor.NextFeature

    Dim nada As String
    Dim sWhereClause As String
    nada = "0"

    featureID = "ID"
```

```
   sWhereClause = featureID & "=" & nada

   'Highlight features in ArcView from the selected records

   'Invalidate only the selection cache
   'Flag the original selection
   'pActiveView.PartialRefresh esriViewGeoSelection, Nothing, Nothing
   pQueryFilter.WhereClause = sWhereClause
   'Perform the selection
   pFeatureSelection.SelectFeatures pQueryFilter, esriSelectionResultNew, False
   'Flag the new selection
    pActiveView.ScreenDisplay.Invalidate Nothing, True,
    pActiveView.ScreenCacheID(esriViewGeoSelection, Nothing)

End Function

'*********************************************************************
'  Function name: UIToolControl1_MouseMove
'  Description:    Records map coordinates as the cursor moves
'*********************************************************************

Private Sub UIToolControl1_MouseMove(ByVal button As Long, ByVal Shift As Long,
ByVal x As Long, ByVal y As Long)
   'record x,y to file
   Dim provx As String
   Dim provy As String
   provx = Str(x)
   provy = Str(y)

   theView = "x " & provx & " y " & provy & " select tool on in ArcGIS"
   writeToFile (theView)
End Sub

'*********************************************************************
'  Function name: UIToolControl1_MouseDown
'  Description:  Selection tool in ArcMap that works by drawing a polygon to
'                enclose features of interest or by clicking on such features.
'                Also, detects selected features in ArcMap and calls code to mark
'                them in DecisionSite.
'  Important:    A ToolControl must be created in ArcMap and this
'                code associated to it.
'*********************************************************************

Private Sub UIToolControl1_MouseDown(ByVal button As Long, ByVal Shift As Long,
ByVal x As Long, ByVal y As Long)

 ' QIs
 Set pMxApp = Application
 Set pMxDoc = ThisDocument
 Set pMap = pMxDoc.FocusMap
 Set pActiveView = pMap
```

```
 'Select features in ArcView by drawing polygon

 Dim pEnvelope As IEnvelope
 Dim pRubberEnv As IRubberBand
 Dim pSelEnv As ISelectionEnvironment
 Dim pSelEnvStore As ISelectionEnvironmentStorage
 Dim saveSelect As Boolean
 Set pSelEnv = pMxApp.SelectionEnvironment
 'pSelEnvStore.SaveSelections = True

'RubberEnvelope for capturing geometry enclosed by polygon drawn by user
 Set pRubberEnv = New RubberEnvelope
 'Return a new Envelope from the tracker object using TrackNew
 Set pEnvelope = pRubberEnv.TrackNew(pMxDoc.activeView.ScreenDisplay, Nothing)

 'Check if polygon was drawn by user or user only clicked on a feature
 Dim emp As Boolean
 emp = pEnvelope.IsEmpty

 If emp = False Then
  'If polygon is drawn without pressing Ctrl at the same time, using the tool
  'creates a new selection of features in ArcView
  If Shift = 0 Then
     pSelEnv.CombinationMethod = esriSelectionResultNew
  ElseIf Shift = 2 Then
     'add features if selection is done pressing the cntrl key
     pSelEnv.CombinationMethod = esriSelectionResultAdd
  ElseIf Shift = 1 Then
     'add features if selection is done pressing the cntrl key
     pSelEnv.CombinationMethod = esriSelectionResultSubtract
  End If

  'Perform the selection using a polygon created on mouse down
  pMap.SelectByShape pEnvelope, pSelEnv, False
  'only refresh a specific region on the display
  pActiveView.ScreenDisplay.Invalidate Nothing, True,
pActiveView.ScreenCacheID(esriViewGeoSelection, Nothing)

 'If user clicks on a feature, also select it by using "point" where user clicked
 ElseIf emp = True Then
  'Dim pFeatureLayer As IFeatureLayer
  Dim pFeatureClass As IFeatureClass

  'Identify the layer to explore in ArcView
  exploreLayer

  Dim pRow As IRow
  Dim pTable As ITable
  Set pFeatureClass = pFeatureLayer.FeatureClass
  Set pTable = pFeatureClass
  Set pRow = pTable.GetRow(0)

  'get the current location and ID
  Dim pPoint As esriCore.IPoint
```

```
   Set pPoint = pMxDoc.CurrentLocation
   Dim pSpatialfilter As ISpatialFilter
   Set pSpatialfilter = New SpatialFilter

   With pSpatialfilter
      Set .Geometry = pPoint
      .SpatialRel = esriSpatialRelWithin
   End With

   Dim pFeatureCursor As IFeatureCursor
   Set pFeatureCursor = pFeatureLayer.Search(pSpatialfilter, False)
   Dim pFeature As IFeature
   Set pFeature = pFeatureCursor.NextFeature

   Do While Not pFeature Is Nothing
      pMap.SelectFeature pFeatureLayer, pFeature
      Set pFeature = pFeatureCursor.NextFeature
   Loop

   If Shift = 0 Then
      pSelEnv.CombinationMethod = esriSelectionResultNew
   ElseIf Shift = 2 Then
      'add features if selection is done pressing the cntrl key
      pSelEnv.CombinationMethod = esriSelectionResultAdd
   ElseIf Shift = 1 Then
      'add features if selection is done pressing the cntrl key
      pSelEnv.CombinationMethod = esriSelectionResultSubtract
   End If
   'only refresh a specific region on the display
   pActiveView.ScreenDisplay.Invalidate Nothing, True,
pActiveView.ScreenCacheID(esriViewGeoSelection, Nothing)

  End If

   'record x,y to file
   Dim provx As String
   Dim provy As String
   provx = Str(x)
   provy = Str(y)

   theView = "x " & provx & " y " & provy & " selected recs in ArcGIS - BoxBegins"
   writeToFile (theView)
End Sub

'*****************************************************************************
'  Function name: theVis_MarkedRecordsChangedEx
'  Description:  Detects when user is marking or unmarking records in a visualization
'                Calls a function that selects them in ArcMap.
'*****************************************************************************

Private Sub theVis_MarkedRecordsChangedEx(ByVal pRecords As Spotfire.IRecords)

   Dim item As Variant, i As Integer
   theNumRecs = pRecords.Count
```

```
ReDim theSFrecs(pRecords.Count)
i = 0
featureID = "ID"

Dim theViewBC As IView
Set theViewBC = theApp.ActiveVisualization.Views(sfvtPlots)
Dim activeView As Spotfire.IView
Dim theSP As Spotfire.ScatterPlot
Dim thePC As Spotfire.ProfileChart
Set activeView = theApp.ActiveVisualization.Views(sfvtNone)

If UIComboBoxControl1.ListIndex = 1 Then
  If activeView.ViewType = sfvtScatterPlot Then
    Set theSP = theVis.Views(sfvtScatterPlot)
    theSP.ColorMarked = RGB(255, 0, 0)
  ElseIf activeView.ViewType = sfvtProfileChart Then
    Set thePC = theVis.Views(sfvtProfileChart)
    thePC.ColorMarked = RGB(255, 0, 0)
  End If
ElseIf UIComboBoxControl1.ListIndex = 2 Then
  If activeView.ViewType = sfvtScatterPlot Then
    Set theSP = theVis.Views(sfvtScatterPlot)
    theSP.ColorMarked = RGB(255, 255, 0)
  ElseIf activeView.ViewType = sfvtProfileChart Then
    Set thePC = theVis.Views(sfvtProfileChart)
    thePC.ColorMarked = RGB(255, 255, 0)
  End If
ElseIf UIComboBoxControl1.ListIndex = 0 Then
  If activeView.ViewType = sfvtScatterPlot Then
    Set theSP = theVis.Views(sfvtScatterPlot)
    theSP.ColorMarked = RGB(0, 255, 255)
  ElseIf activeView.ViewType = sfvtProfileChart Then
    Set thePC = theVis.Views(sfvtProfileChart)
    thePC.ColorMarked = RGB(0, 255, 255)
  End If
End If

If theViewBC.ViewType <> sfvtBarChart Then
  'Check if recs are selected in DecisionSite
  If pRecords.Count = 0 Then
    'refresh Arcview to deselect the selected
    doRefreshInAGIS

  Else
    'Add the ID of the selected records to the array theSFrecs
    For Each item In pRecords.Columns(featureID).RecordValues
      If i <= pRecords.Count - 1 Then
        theSFrecs(i) = item
        i = i + 1
      End If
    Next item

  'function call to select relevant records in ArcView
  selectRecsInAGISfromSFrecs
```

```
      End If
    End If

    theView = "MarkedRecsChanged"
    writeToFile (theView)
End Sub


'*********************************************************************
'  Function name: theVis_ActiveRecordsChanged
'  Description:  Detects when user is activating records in a visualization.
'                Calls function to highlight the relevant record in ArcMap.
'*********************************************************************

Private Sub theVis_ActiveRecordsChanged(ByVal pRecords As Spotfire.IRecords)

    Dim item As Variant
    strItems = ""
    featureID = "ID"

   Dim theViewBC As IView
   Set theViewBC = theApp.ActiveVisualization.Views(sfvtPlots)

   If theViewBC.ViewType <> sfvtBarChart Then

      For Each item In pRecords.Columns(featureID).RecordValues
         If strItems <> "" Then strItems = strItems & vbCrLf
            strItems = strItems & item
         Next item

         'Highlight feature in ArcView temporarily
         getGeometryToFlash2
    End If

    theView = "ActiveRecsChanged"
    writeToFile (theView)
End Sub

'*********************************************************************
'  Function name: theVis_HiLiteMarker
'  Description:  Detects when user highlights records in DecisionSite.
'                Calls function to highlight the relevant record in ArcMap.
'*********************************************************************

Private Sub theVis_HiLiteMarker(ByVal pRecord As Spotfire.Record)
    strItems = ""
    featureID = "ID"
    If strItems <> "" Then strItems = strItems & vbCrLf
    strItems = strItems & pRecord.Field(featureID)

   Dim theViewBC As IView
   Set theViewBC = theApp.ActiveVisualization.Views(sfvtPlots)

   If theViewBC.ViewType <> sfvtBarChart Then
```

```
      'Highlight feature in ArcView
      getGeometryToFlash2
   End If

   theView = "HiLiteRecs"
   writeToFile (theView)
End Sub

'*****************************************************************
'  Function name: markRecsInSFfromAGIS
'  Description:  Marks records in DecisionSite that are selected in ArcView
'*****************************************************************

Function markRecsInSFfromAGIS()
   Dim pRecords As Spotfire.IRecords
   Dim recsToMark As Spotfire.IRecords4
   Dim theMarkedRecs As IRecords

   Set theVis = theApp.ActiveVisualization
   Set pRecords = theVis.Records

   Dim iCol2 As Spotfire.IColumn2
   featureID = "ID"
   Set iCol2 = pRecords.Columns(featureID)
   Dim k As Integer

   Set recsToMark = iCol2.FindByValues(theAGISrecs)
   recsToMark.MarkExclusive

End Function

'*********************************************************************
'  Function name: hltRecinSFfromAGIS
'  Description:  Highlights records in DecisionSite that are highlighted in ArcMap
'*********************************************************************

Function hltRecinSFfromAGIS()
  Set theApp = Spotfire.Application
  Set theVis = theApp.ActiveVisualization
  Dim iCols As Spotfire.IColumns2
  Set iCols = theVis.Records.Columns
  Dim iCol As Spotfire.IColumn
  featureID = "ID"
  Set iCol = iCols.item(featureID)

  iCol.Find(strItems, strItems).item(1).HiLite

End Function

'*****************************************************************
'  Function name: doWithMarked
'  Description:  Clears marked records in DecisionSite
'*****************************************************************
```

```
Function doWithMarked()
   theApp.ExecuteCommand ("MarkedRecordsClear")

End Function

'*************************************************************************
'  Function name: FlashGeometry
'  Description:  Highlights (changes colour) of features in ArcMap temporarily
'                as user passes over features with the pointer.
'                Needs the geometry of the features to be highlighted as input.
'*************************************************************************

Public Sub FlashGeometry(pGeometry As esriCore.IGeometry)
  On Error GoTo ErrHand:

  Dim pFeature As IFeature
  Dim pSimpleLineSymbol As ILineSymbol, pSimpleFillSymbol As ISimpleFillSymbol
  Dim pSymbolBorder As isymbolborder
  Dim pSimpleMarkersymbol As ISimpleMarkerSymbol, iInterval As Integer
  Dim pSymbol As ISymbol, pDisplay As IScreenDisplay
  Dim pColor As IColor
  Dim pColor2 As IColor

  'Sets the delay (time the feature is highlighted after mouse has been passed over)
  iInterval = 200

  'If features in ArcView had no colour coding, highlighting would be in red,
  'as the default colours in the DecisionSite visualisations. These methods however
  'will combine the red with any other colour if feature is colour coded by some attribute
value
  Set pColor = New RgbColor
  Set pColor2 = New RgbColor
  pColor.RGB = RGB(100, 250, 250)
  pColor2.RGB = RGB(250, 0, 0)

  Set pMxDoc = ThisDocument
  Set pMap = pMxDoc.FocusMap
  Set pActiveView = pMap
  Set pDisplay = pActiveView.ScreenDisplay

  pDisplay.StartDrawing 0, esriNoScreenCache

  'Account for various possible feature geometry types that could be highlighted
  Select Case pGeometry.GeometryType
    Case esriGeometryPolyline
      Set pSimpleLineSymbol = New SimpleLineSymbol
      Set pSymbol = pSimpleLineSymbol
      'pSymbol.ROP2 = esriROPNotXOrPen 'erase itself when drawn twice
      pSimpleLineSymbol.Width = 4
      pSimpleLineSymbol.Color = pColor
      pDisplay.SetSymbol pSimpleLineSymbol
      pDisplay.DrawPolyline pGeometry
      Sleep iInterval
      pDisplay.DrawPolyline pGeometry
```

```
    Case esriGeometryPolygon
      Set pSimpleFillSymbol = New SimpleFillSymbol
      Set pSymbol = pSimpleFillSymbol
      pSymbol.ROP2 = esriROPNotXOrPen
      pSimpleFillSymbol.Color = pColor
      pSimpleFillSymbol.Style = esriSFSSolid
      'pSimpleFillSymbol.Style = esriSFSHollow
      pSimpleFillSymbol.Outline.Color = pColor2
      pDisplay.SetSymbol pSimpleFillSymbol
      pDisplay.DrawPolygon pGeometry
      Sleep iInterval
      pDisplay.DrawPolygon pGeometry

    Case esriGeometryPoint
      Set pSimpleMarkersymbol = New SimpleMarkerSymbol
      Set pSymbol = pSimpleMarkersymbol
      'pSymbol.ROP2 = esriROPNotXOrPen
      pSimpleMarkersymbol.Color = pColor
      'pSimpleMarkersymbol.Style = esriSMSCircle
      'pSimpleMarkersymbol.Size = 12
      pDisplay.SetSymbol pSimpleMarkersymbol
      pDisplay.DrawPoint pGeometry
      Sleep iInterval
      pDisplay.DrawPoint pGeometry

    Case Else   ' Other type
  End Select

  pDisplay.FinishDrawing

  Exit Sub

ErrHand:
  MsgBox "FlashGeometry - " & Err.Description

End Sub

'*********************************************************************
'  Function name: UIToolControl2_MouseMove
'  Description:   Tool for highlighting features in ArcMap when pointer passes over.
'                 Highlights corresponding records in DecisionSite simultaneously.
'   Important:    ToolControl must be created in ArcMap and this code
'                 associated to it.
'*********************************************************************

Private Sub UIToolControl2_MouseMove(ByVal button As Long, ByVal Shift As Long,
ByVal x As Long, ByVal y As Long)
  'Flag if tool is being used
  moveState = 1

  'Dim pFeatureLayer As IFeatureLayer
  Dim pFeatureClass As IFeatureClass
  Set pMxDoc = ThisDocument
  Set pMap = pMxDoc.FocusMap
```

```
  Set pActiveView = pMap

  'Identify the layer to explore in ArcView
  exploreLayer

  Dim pRow As IRow
  Dim pTable As ITable
  Set pFeatureClass = pFeatureLayer.FeatureClass
  Set pTable = pFeatureClass
  Set pRow = pTable.GetRow(0)

  'get the current location and ID
  strItems = ""
  Dim pPoint As esriCore.IPoint
  Set pPoint = pMxDoc.CurrentLocation
  Dim pSpatialfilter As ISpatialFilter
  Set pSpatialfilter = New SpatialFilter
  With pSpatialfilter
    Set .Geometry = pPoint
    .SpatialRel = esriSpatialRelWithin
  End With
  Dim pFeatureCursor As IFeatureCursor
  Set pFeatureCursor = pFeatureLayer.Search(pSpatialfilter, False)
  Dim pFeature As IFeature
  Set pFeature = pFeatureCursor.NextFeature
  Do While Not pFeature Is Nothing
    pMap.SelectFeature pFeatureLayer, pFeature
    If strItems <> "" Then strItems = strItems & vbCrLf
    featureID = "ID"
    strItems = strItems & pFeature.Value(pRow.Fields.FindField(featureID))

    hltRecinSFfromAGIS
    getGeometryToFlash2

    Set pFeature = pFeatureCursor.NextFeature
  Loop

  'Flag if no features are being highlighted
  moveState = 0

  'Record x,y to file
  Dim provx As String
  Dim provy As String
  provx = Str(x)
  provy = Str(y)

  theView = "x " & provx & " y " & provy & " hiLite tool on in ArcGIS"
  writeToFile (theView)

End Sub

'*********************************************************************
'  Function name: SetEvents
'  Description:   Makes ArcMap start listening to events
```

```
'*****************************************************************

Public Sub SetEvents()
   Set pMxDoc = ThisDocument
   Set pActiveViewEvents = pMxDoc.FocusMap
End Sub

'*****************************************************************
'  Function name: pActiveViewEvents_SelectionChanged
'  Description:   Detect if feature selection has changed in ArcMap
'*****************************************************************

Private Sub pActiveViewEvents_SelectionChanged()
  'Check if features are being highlighted in ArcMap
  If moveState = 1 Then
      'If they are, don't get the selection
  ElseIf moveState = 0 Then
      'If they are not, then if the selection changed it is because the
      'polygon selection tool is used and corresponding selection needs to be
      'marked in DecisionSite
      getAGISrecs
  End If

End Sub

'*************************************************************************
'  Function name: getAGISrecs
'  Description:   When features are selected in ArcView, pass IDs to DecisionSite
'                 and mark the records there
'*************************************************************************
Function getAGISrecs()
  Set pMap = pMxDoc.FocusMap
  Set pActiveView = pMap

  Dim pFeature As IFeature
  Dim pFeatureCursor As IFeatureCursor
  'Dim pFeatureLayer As IGeoFeatureLayer
  Dim pFeatureClass As IFeatureClass
  Dim pFeatureSelection As IFeatureSelection
  Dim pSelectionSet As ISelectionSet

 'Identify the layer to explore in ArcView
 exploreLayer

 Dim pRow As IRow
 Dim pTable As ITable
 Set pFeatureClass = pFeatureLayer.FeatureClass
 Set pTable = pFeatureClass
 Set pRow = pTable.GetRow(0)

 'Loop through the selected features per layer

 Set pFeatureSelection = pFeatureLayer
 Set pSelectionSet = pFeatureSelection.SelectionSet
```

```
 'Can use Nothing keyword if you don't want to draw them,
 'otherwise, the spatial reference might not match the Map's
 pSelectionSet.Search Nothing, False, pFeatureCursor
 Set pFeature = pFeatureCursor.NextFeature

 'loop through selected records in ArcView table, get their ID value and add to array
theAGISrecs
 ReDim theAGISrecs(pSelectionSet.Count)
 Dim k As Integer
 k = 0
 featureID = "ID"
 theNumRecs2 = pSelectionSet.Count
 Do Until pFeature Is Nothing
  If k <= theNumRecs2 - 1 Then
       theAGISrecs(k) = pFeature.Value(pRow.Fields.FindField(featureID))
       k = k + 1
  End If


  Set pFeature = pFeatureCursor.NextFeature

 Loop


 'Check if recs have been marked in DecisionSite

 'If DecisionSite client application is on
 If SFstate = 1 Then
     'Mark records in DecisionSite
     markRecsInSFfromAGIS
 'If there is no DecisionSite client application
 ElseIf SFstate = 0 Then
     MsgBox "Please start your data exploration session before continuing by clicking on the
bull's eye button in the tool bar to start DecisionSite."
     'De-select features in ArcView if any
     doRefreshInAGIS

 End If

End Function

'*****************************************************************************
'  Function name: exploreLayer
'  Description:  Identify the layer to explore from any other loaded layers in ArcMap
'*****************************************************************************

Function exploreLayer()

  'QIs
  Set pMxDoc = ThisDocument
  Set pMap = pMxDoc.FocusMap
  Set pActiveView = pMap
  spatialData = "censusDemogByTract1"          'Change to be the relevant ArcView layer
```

```
  'Dim pFeatureLayer As IFeatureLayer

  'Identify the layer to explore from any other loaded layers in ArcView
  Dim k As Integer
  For k = 0 To pMap.LayerCount
    If (pMap.Layer(k).Name = spatialData) Then
      Set pFeatureLayer = pMap.Layer(k)
      pMap.Layer(k).Cached = True
    Exit For
    End If
  Next k
End Function

Private Sub theCP_FocusChanged(ByVal qd As Spotfire.IQueryDevice)
  theView = "QD focusChanged"
  writeToFile (theView)
End Sub

'*********************************************************************
'  Function name: theCP_selectedRecordsChanged
'  Description:   Event to identify filtered records in SF
'*********************************************************************

Private Sub theCP_selectedRecordsChanged(ByVal pRecords As Spotfire.Records)
  Dim pRecs2 As IRecords3
  Dim pRecsCompl2 As IRecords3
  Dim theNumRecsDiff As Integer
  Dim theNumRecsCompl2 As Integer
  Dim item As Variant, i As Integer
  Dim h As Integer
  theFiltList = ""

  If theNumRecsFilt = 0 Then
    'Find the records that where filtered out
    Set pRecs2 = pRecords
    Set pRecsCompl = pRecs2.Complement
    theNumRecsFilt = pRecsCompl.Count
    ReDim theFiltRecs(pRecsCompl.Count)
    i = 0
    featureID = "ID"

    'Add the ID of the selected records to the array theSFrecs
    For Each item In pRecsCompl.Columns(featureID).RecordValues
      If i <= pRecsCompl.Count - 1 Then
        theFiltRecs(i) = item
        i = i + 1
      End If
    Next item

  ElseIf pRecsCompl.Count <> 0 Then

    'Record previous filtered recs
    Set pRecsCompl2 = pRecsCompl
    theNumRecsCompl2 = pRecsCompl2.Count
```

```
    'Find actual filtered recs
    Set pRecs2 = pRecords
    Set pRecsCompl = pRecs2.Complement
    theNumRecsFilt = pRecsCompl.Count

    If theNumRecsFilt > theNumRecsCompl2 Then
        'If new set is larger than old one then
        'only filter the difference
        Set pRecsDiff = pRecsCompl.Difference(pRecsCompl2)
        ReDim theFiltRecs(pRecsDiff.Count)
        i = 0
        featureID = "ID"
        For Each item In pRecsDiff.Columns(featureID).RecordValues
            If i <= pRecsDiff.Count - 1 Then
                theFiltRecs(i) = item
                i = i + 1
            End If
        Next item

    ElseIf theNumRecsFilt < theNumRecsCompl2 Then
        'If new set is smaller than old one then
        'refresh and select again

        Set pRecs2 = pRecords
        Set pRecsCompl = pRecs2.Complement
        theNumRecsFilt = pRecsCompl.Count
        ReDim theFiltRecs(pRecsCompl.Count)
        i = 0
        featureID = "ID"

        For Each item In pRecsCompl.Columns(featureID).RecordValues
            If i <= pRecsCompl.Count - 1 Then
                theFiltRecs(i) = item
                i = i + 1
            End If
        Next item
        geogRefresh
        'flag1 = 1
    End If
End If

'Build a comma-delimited list from array theSFrecs
'to use in whereClause for feature selection in ArcView
For h = 0 To UBound(theFiltRecs)
    If theFiltList = "" Then
        theFiltList = theFiltRecs(h)
    Else
        theFiltList = theFiltList & "," & theFiltRecs(h)
    End If
Next h

'Show filtering in ArcView
grayOutRecsInAGISfromSFsql
'write to file
```

```
   theView = " moved QueryDevice"
   writeToFile (theView)

End Sub

Function geogRefresh()
   'QIs
    Set pMxDoc = ThisDocument
    Set pMap = pMxDoc.FocusMap
    Set pActiveView = pMap
    pActiveView.ScreenDisplay.Invalidate Nothing, True,
pActiveView.ScreenCacheID(esriViewGraphics, Nothing)
    'pActiveView.PartialRefresh esriViewGeoselection, Nothing, Nothing
End Function


'*********************************************************************
'  Function name: writeToFile
'  Description:  writes details of user interaction with system to a log
'*********************************************************************

Function writeToFile(whatView As String)

Dim nFile As Integer
On Error GoTo err_handler
nFile = FreeFile          'get the next free file number

'open the file to append to
Open "C:\Documents and Settings\Carolina\My Documents\UsabilityTests\UsabilityTest
April_2003\log1.txt" For Append As #nFile

'write to the file
Print #nFile, Format$(Now, "mm-dd-yy hh:mm:ss") & " - " & whatView

'close the file
Close #nFile

Exit Function

err_handler:
MsgBox Err.Number & " " & Err.Description
Err.Clear
Exit Function

End Function
```

Appendix C: Publications

C.1 Refereed Papers

Haklay, M. and C. Tobón (2003) "Usability Engineering and PPGIS: Towards a User-Centred Design Approach," *International Journal of Geographical Information Science*, 17(6), 577-92.

Longley, P. and Tobón, C. (Submitted). The Scale of Urban Deprivation: Spatial Dependence and Heterogeneity in Patterns of Hardship in Urban Areas.

Tobón, C. (Submitted). "Geovisualization Task Characterisation".

C.2 Refereed Book Chapters

Tobón, C. (2003) "Evaluating Geographic Visualization Tools and Methods: An Approach and Experiment Based Upon User Tasks", in J. Dykes, A. M. MacEachren and M. J. Kraak (ed.), *Exploring Geovisualization,* London: Elsevier, in press.

Fuhrmann, S., P. Ahonen-Rainio, R. M. Edsall, S. I. Fabrikant, E. L. Koua, C. Tobón, C. Ware and S. Wilson (2003) "Making Useful and Useable Geovisualization: Design and Evaluation Issues", in J. Dykes, A. M. Maceachern and M. J. Kraak (ed.), *Exploring Geovisualization,* London: Elsevier, in press.

Tobón, C. and M. Haklay (2003) "Human-Computer Interaction and Usability Evaluation in GIScience", in P. A. Longley and M. Batty (ed.), *Advanced Spatial Analysis: The CASA book of GIS,* Redlands: ESRI Press.

Lloyd, D., M. Haklay, M. Thurstain-Goodwin and C. Tobón (2003) "Visualizing Spatial Structure in Urban Data", in P. A. Longley and M. Batty (ed.), *Advanced Spatial Analysis: The CASA book of GIS,* Redlands: ESRI Press, 267-86.

C.3 Refereed Conference Proceedings

Tobón, C. and Longley, P. (2003). Patterns of Spatial dependence and heterogeneity and the scale of urban deprivation. GeoComputation 2003. Southampton, 8-10 September.

Tobón, C. (2003). Evaluating Geographic Visualization Tools and Learning about User Tasks. CogSci2003, Boston, USA. 27 July - 3 Aug.

Longley, P. and Tobón, C. (2003). Spatial dependence and heterogeneity in patterns of urban deprivation. ERSA 2003 Congress. Jyvaskyla, Finland. 27-30 August.

Haklay, M. and Tobón, C. (2002). Usability Engineering and PPGIS: Towards a User-Centred Design Approach. PPGIS (URISA) Conference. July.

Tobón, C. (2002). Usability Testing for Improving Interactive Geovisualization Environments. GIS Research UK, University of Sheffield. April.

C.4 Conference Papers

Haklay, M. and Tobón, C. (2002). Usability and e-Government Web sites: the Good, the Bad and the Ugly. AGI Conference. September.

Tobón, C. (2001). Using interactive Visual Exploration of Spatial Data to Inform Spatial Statistical Analysis. AGI Conference. September.

Tobón, C. (2001). Interactive Displays for the Visual Exploration of Point Data, GIS Research UK, University of Glamorgan. April.

C.5 Working Papers

Tobón, C. (2002). Usability Testing for Improving Interactive GeoVisualization Techniques. CASA Working Paper Series 45.

Tobón, C. (2001). Visual and Interactive Exploration of Point Data. CASA Working Paper Series 31.