EVALUATION OF VISUALISATIONS OF GEOGRAPHICALLY WEIGHTED REGRESSION, WITH PERCEPTUAL SCALABILITY

Tommy Burke

A Thesis Submitted for the Degree of PhD
at the
University of St Andrews

2016

Evaluation of Visualisations of Geographically Weighted Regression, with Perceptual Scalability

Tommy Burke

This thesis is submitted in partial fulfilment for the degree of PhD at the University of St Andrews

Date of Submission

17th December 2015

1. Candidate's declarations:

I, Tommy Burke, hereby certify that this thesis, which is approximately 55,000 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for a higher degree.

I was admitted as a research student in [September, 2009] and as a candidate for the degree of Doctoral Degree in [March, 2014]; the higher study for which this is a record was carried out in the University of St Andrews between [2012] and [2014].

(*If you received assistance in writing from anyone other than your supervisor/s*):
I, Tommy Burke, received assistance in the writing of this thesis in respect of [language, grammar, spelling or syntax], which was provided by

Nichola Salmon
Bernadette Burke
John Oliver Burke
Thomas Burke

Date ……..... signature of candidate ……...............…

2. Supervisor's declaration:

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of ……… in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date ……........... signature of supervisor …...........……

3. Permission for publication: (*to be signed by both candidate and supervisor*)

In submitting this thesis to the University of St Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that my thesis will be electronically accessible for personal or research use unless exempt by award of an embargo as requested below, and that the library has the right to migrate my thesis into new electronic forms as required to ensure continued access to the thesis. I have obtained any third-party copyright permissions that may be required in order to allow such access and migration, or have requested the appropriate embargo below.

The following is an agreed request by candidate and supervisor regarding the publication of this thesis:

PRINTED COPY
a) No embargo on print copy
b) Embargo on all or part of print copy for a period of 2 years (maximum five) on the following ground(s):
    - Publication would be commercially damaging to the researcher, or to the supervisor, or the University
    - Publication would preclude future publication
    - Publication would be in breach of laws or ethics
c) Permanent or longer term embargo on all or part of print copy for a period of … years (the request will be referred to the Pro-Provost and permission will be granted only in exceptional circumstances).

Supporting statement for printed embargo request:

ELECTRONIC COPY

a) No embargo on electronic copy

b) Embargo on all or part of electronic copy for a period of 2 years (maximum five) on the following ground(s):
   - Publication would be commercially damaging to the researcher, or to the supervisor, or the University
   - Publication would preclude future publication
   - Publication would be in breach of law or ethics

c) Permanent or longer term embargo on all or part of electronic copy for a period of … years (the request will be referred to the Pro-Provost and permission will be granted only in exceptional circumstances).

Supporting statement for electronic embargo request:

Date …........… signature of candidate …..........… signature of supervisor …......……

Please note initial embargos can be requested for a maximum of five years. An embargo on a thesis submitted to the Faculty of Science or Medicine is rarely granted for more than two years in the first instance, without good justification. The Library will not lift an embargo before confirming with the student and supervisor that they do not intend to request a continuation. In the absence of an agreed response from both student and supervisor, the Head of School will be consulted. Please note that the total period of an embargo, including any continuation, is not expected to exceed ten years.
Where part of a thesis is to be embargoed, please specify the part and the reason.

Abstract

Given the large volume of data that is regularly accumulated, the need to properly manage, efficiently display and correctly interpret, becomes more important. Complex analysis of data is best performed using statistical models and in particular those with a geographical element are best analysed using Spatial Statistical Methods, including local regression. Spatial Statistical Methods are employed in a wide range of disciplines to analyse and interpret data where it is necessary to detect significant spatial patterns or relationships. The topic of the research presented in this thesis is an exploration of the most effective methods of visualising results.

A human being is capable of processing a vast amount of data as long as it is effectively displayed. However, the perceptual load will at some point exceed the cognitive processing ability and therefore the ability to comprehend data. Although increases in data scale did increase the cognitive load and reduce processing, prior knowledge of geographical information systems did not result in an overall processing advantage.

The empirical work in the thesis is divided into two parts. The first part aims to gain insight into visualisations which would be effective for interpretation and analysis of Geographically Weighted Regression (GWR), a popular Spatial Statistical Method. Three different visualisation techniques; two dimensional, three dimensional and interactive, are evaluated through an experiment comprising two data set sizes. Interactive visualisations perform best overall, despite the apparent lack of researcher familiarity.

The increase in data volume can present additional complexity for researchers. Although the evaluation of the first experiment augments understanding of effective visualisation display, the scale at which data can be adequately presented within these visualisations is unclear. Therefore, the second empirical investigation seeks to provide insight into data scalability, and human cognitive limitations associated with data comprehension.

The general discussion concludes that there is a need to better inform researchers of the potential of interactive visualisations. People do need to be properly trained to use these systems, but the limits of human perceptual processing also need to be considered in order to permit more efficient and insightful analysis.

Acknowledgements

The completion of this thesis would not have been possible without the appreciated contribution of the following:

Firstly, gratitude is expressed to my supervisor Dr Urška Demšar and co-supervisor Professor Stewart Fotheringham for their guidance throughout. I would also like to thank the extended staff and my colleagues at the Centre for Geocomputation (CGI), particularly Sila, Wenbai, Maryam and Giota. Appreciation is also extended to the wider Department of Geography and Geosciences at the University of St. Andrews.

The early work for this thesis was undertaken within the National Centre for Geocomputation (NCG) at the National University of Ireland, Maynooth and so I wish to express my appreciation to all who assisted me during my time at NCG. I would like to thank my colleague and friend Peter Foley who gave his time to assist with the design of parts of the first experiment. I would also like to express my gratitude to Dr Gavin McArdle for his time and effort during the development of my research, and to my colleague and friend Dr Conor Cahalane.

During my research I was fortunate to collaborate with experienced researchers in the geo-visualisation field. The knowledge, advice and helpfulness shared by Professor Anthony Robinson (GeoVISTA Centre, Department of Geography, Pennsylvania State University) and Dr Arzu *Çöltekin*, (GIScience Center, Department of Geography, University of Zurich) were integral in the design and completion of the two key experiments which formed the basis of my thesis. Thanks also to Dr Adrian Kavanagh (Department of Geography, Maynooth University) who kindly provided electoral data used in each experiment.

I also wish to acknowledge and thank all the research participants who kindly gave their time to be a part of my experiments. Your contribution is very much appreciated. The funding received from Science Foundation Ireland (SFI) as part of the GeoVis Project and MOVE COST Action funding for a Short Term Scientific Mission is also gratefully acknowledged.

A special note of appreciation is owed to Professor Verity Brown (School of Psychology and Neuroscience and Provost of St. Leonard's College). Thank you for all your guidance and encouragement in the final weeks of my PhD.

Finally, I thank my family for their continued support.

Table of Contents

Tables

Figures

Introduction

Research Context

One million or ten times as many: that is as good an estimate as any of the number of ways information can be presented visually. It is well-known that 'a picture is worth a thousand words' and Tufte (2007) wrote "only a picture can carry such a volume of data in such a small space". If one were to attempt to describe an image using a thousand words, the resulting interpretation would still lack the clarity of the original image. It is indubitable that displaying information in graphical form is better. Extend this thought to the fact that the rate at which humans produce and collect information increases exponentially every day. Finding new ways to manage the information load must be regarded as the highest priority.

The methods used to harness this data are also numerous, and the techniques we use to display data are innovative. However, it is not simply a case of gathering relevant information and using a graphical representation to reveal patterns and to discover new insights: the nature of both the technology and the human beings that use it must be considered together. It is important to understand there is not necessarily a 'best way' to disseminate information, but it is certain that some ways are more effective; others less effective and some are utterly inadequate.

Graphical displays of information have an additional advantage over textual and other forms: they not only present what is already known, they also offer opportunities for new insight. This newly derived insight depends on understanding how to effectively display data in a manner that is useful to those who wish to analyse it and use it. Data incorporating geographical elements add complexity to computational analysis; however, they also allow for different kinds of analysis, because patterns in data can literally be seen by the user.

Kang et al. (2011) highlights the importance of measuring visualisations for their capacity or ability to 'make sense' to end users but also acknowledges the difficulty of this task. This research thesis is concerned with the effective display of geographical data. Existing research in this field includes a range of display options for visualising spatial data so this particular piece of research seeks to draw upon this range of visualisation types to determine the most effective types of visualisations available. The use of newly emerging 3D visualisations is given particular attention. In order to analyse the effectiveness of these visualisations, the spatial statistical method of geographically weighted regression (GWR) is the key focus.

The end user of visual data must also be considered. Human's ability to process information must be taken into account when producing geovisualisations. There is little need for an advanced geovisualisation technique that displays highly complex data if human cognitive limits prevent the visualisation from being used effectively. This research aims to provide deeper insight into the human interactions with a range of geovisualisations.

Research Question, Aim and Objectives

The over-arching research question for this research thesis is to assess and evaluate a range of data visualisation techniques to ascertain the optimal way of geographical data presentation. This originates from a desire to test graphical visualisations over tabular and to compare those to participant task performances using advanced graphical representations. Slocum et al. (2002) suggest part of the difficulty in understanding and improving visualisations exist because human cognitive processes are not easily classifiable. Humans are not as homogenous as visual analytics tool developers think.

This research question contributes to the field of investigation which is a mid-point between presentation of knowns and revelations of unknowns. In other words, the realm between data gathering and data interpretation. The characteristics of the visual display of information affect ease of processing. However, there will also be a limit to the amount of information that can be processed – at some point, data volume in an image will overload humans' limited perceptual processing mechanisms and at this point data will exceed comprehension.

To facilitate the over-arching research question within this thesis, two key aims and related objectives were established. These are critical in attempting to ascertain the optimal approach when presenting geographical data to end-users. Structurally, the empirical work in the thesis is divided into two parts (Experiment One and Experiment Two).

Experiment One seeks to gain insight into visualisations which would be effective for interpretation and analysis of Geographically Weighted Regression (GWR), a popular spatial statistical method. Three different visualisation techniques; two dimensional, three dimensional, and interactive, are evaluated through an experiment comprising two data set sizes. The aim and objectives of Experiment One are outlined below:

Aim 1: To assess the effectiveness of three visualisation types for analysis and interpretation of GWR outputs.

- Objective 1a: To assess the effectiveness of 2D visualisations for display and interpretation of GWR outputs.
- Objective 1b: To assess the effectiveness of 3D visualisations for display and interpretation of GWR outputs.
- Objective 1c: To assess the effectiveness of interactive visualisations for display and interpretation of GWR outputs.
- Objective 1d: To decide upon the most effective of all three visualisations for interpretation and analysis for GWR outputs, offering guidance for producers of GWR outputs.

According to Edsall (2003), geovisualisation grew out of a need to represent and interact with complex data. Edsall also states there is a general agreement among statisticians that visualisations are capable of providing insight into datasets (referring to the research of Tukey, 1977, Hurley and Buja, 1990 and Wegman 2000). The usefulness and thus importance of visualisation methods to display information also extends to spatial data – which include Geographically Weighted Spatial Statistical Methods.

Geographically Weighted Spatial Statistical Methods are employed in a wide range of disciplines to analyse and interpret data where they are used to detect significant patterns or relationships across space. One such method is Geographically Weighted Regression (GWR) which is used to examine processes that vary over space and time (Fotheringham et al., 2001). There is little variation in the types of visualisations which are used to analyse the results of GWR. 2D univariate maps are one of the most common visualisation methods. Dykes et al. (2005) stress the importance of developing knowledge of whether "geovisualisation techniques, tools and solutions" actually work. Despite this, it is not known why other visualisation methods are not employed for GWR.

Two other visualisation methods are selected for comparison in this research, these are: three dimensional visualisations and interactive visualisations. For the past decade, the need for testing usability within geovisualisations is increasing with different types of interactions emerging (Muntz et al., 2000). 3D data visualisation is a novel and relatively unused way of conveying spatial data. Can it be more or less effective than traditional 2D approaches and might it be an emerging aspect within the field of geovisual analytics? This research aims to explore these core questions through actual user experimentation and feedback. Keim et al. (2004) indicate that interactive visualisations can be extremely effective in elucidating insights to the data concerned. Through a comprehensive evaluation, Experiment One aims to offer a deeper insight into the effectiveness of GWR visualisations. It also aims to augment the decision making process of GWR output users so that the most suitable visualisation technique can be utilised to aid user interpretation and understanding of the data.

Referring to early usability work on visualisations, the ease of use, coupled with the need to reduce time to complete an analysis is a measure of visual quality (Bertin, 1983). Through the first aim in this research a contribution is made to knowledge on usability effectiveness by evaluation of visualisations which display outputs of an important geographical spatial statistical method (GWR).

Building on the exploration within Experiment One, the literature acknowledges the increase in data volume can present additional complexity for researchers. Although the evaluation of the first experiment can augment understanding of effective visualisation display, the scale at which data can be adequately presented within these visualisations is unclear. Therefore, the second empirical investigation seeks to provide insight into data scalability, and human cognitive limitations associated with data comprehension. The aim and objectives of Experiment Two are outlined below:

Aim 2: To assess impact of data scale on user interpretation of 2D visualisations, thereby investigating the hypothesised presence of perceptual scalability.

- Objective 2a: To assess the impact of visualisation scale and spatial unit scale on user interpretation of 2D visualisations. Utilising standard metrics to examine potential effects of perceptual scalability.
- Objective 2b: To evaluate the impact of expertise levels on interpretation of 2D visualisations using standard metrics, examining potential effects on perceptual scalability.

- Objective 2c: To evaluate the impact of the order in which data distributions are encountered on interpretation of 2D visualisations through standard metric measurement, examining potential effects of perceptual scalability.
- Objective 2d: To analyse eye movement data to ascertain if there is evidence of perceptual scalability, with a particular emphasis on:
    i. The visualisation scale and spatial unit scale of 2D visualisations, investigating if there is a change in a user's cognitive load.
    ii. The expertise levels on interpretation of 2D visualisations.
    iii. The variation of data distribution on a 3D visualisation, determining if any change in a user's cognitive load occurs according to variations and in doing so, gaining insight into the potential effects of perceptual scalability.

Visual scalability is the ability of visualisations to effectively display large amounts of data and human perception is a facet which affects scalability (Eick and Karr, 2002). This concept is assessed in Objective 2a. It stems from research carried out by Burke and Demsar (2010) which indicated there may be an issue with perceptual scalability when humans are asked to complete visual tasks in differently sized geographic datasets.

According to Hoffman (2000) and Ware (2008), humans use 40% of the brain to provide visual output. This research aims to contribute to the field of human cognition by assessing the change in performance of users when faced with different levels of visual complexity. Through this objective greater insight can be offered into what Hacklay and Tobón (2003) refer to as the 'diversity of human behaviour' when dealing with graphic visualisations. This objective will also enhance the ability of developers of visualisations to produce more cognitively efficient visualisations which cater to the non-homogenous nature of end users. Group differences exist on the basis of parameters such as expertise and sensory differences (Slocum et al. 2001) and so the impacts of these categorisations are also explored in this thesis through Objectives 2b and 2c.

Through Objective 2d, this research contributes significantly to an emerging area in the literature – eye movement analysis. Eye movement analysis can provide valuable insight into psychological and cognitive function in a number of real-world tasks (in this case for visualisations and scalability) including visual reading and exploration of computer displays

(Goldberg at al., 2001). Through eye movement analysis, the changes in participant performance can be assessed beyond standard evaluation metrics.

The above research aims and objectives form the framework for the remainder of this thesis and set out to explore and answer the over-arching research question. Aim 1 and the related objectives are explored in Chapters 3 and 5 while Aim 2 and the related objectives are explored in Chapters 4 and 6, as per below.

Thesis Structure

To set the over-arching research aim into context, Chapter 1 has presented an overview of literature to provide deeper insight into the key components within this research thesis. It necessarily covers the core concepts of spatial statistics and methods and their use in past and current geographical research. Geographically Weighted Regression (GWR) is given special consideration because, as outlined in the research objectives above, it forms an important part of Experiment One in this research. GWR is a widely used spatial statistical method and this research thesis explores how effectively the potential of GWR output is realised by the methods of visualisation that are currently used across the literature. This research acknowledges that human consumption of the visualisations is not as a passive receiver of the image, but rather as active explorers of images. Cartography and visualisation aspects are presented in the second half of Chapter 1 as it is important to understand their relevance in this research thesis.

The literature surrounding the development of Eye Movement science and how it is a valuable tool for researchers is detailed in Chapter 2. Eye movement tracking can be used to assess graphical data exploration, allowing the researcher to learn about and understand the key features of the graphical displays from where the observer takes information. Such human interpretive aspects are outlined in this chapter are closely associated with Experiment Two of this thesis. As suggested in some existing literature, this shows that the display and comprehension of data is not entirely dependent on computer based issues.

Chapter 3 details the methodological approach used to answer the research objectives set within Experiment One. The development of a user-experiment test is outlined here with considerations for ethics, datasets, equipment and participant sourcing. There are many types of visualisations published in the peer-reviewed literature where GWR is used. The purpose of Experiment One was to compare three different types of visualisation 'head-to-head' to attempt to ascertain the most effective one for user interpretation of geographical data. The three visualisation types selected were ArcMap, a 2D semi-interactive visualisation; a 3D visualisation called ArcScene; and a previously unused or currently very rarely used interactive visualisation method called ProVis. This research is contributing to the wider field and discipline by attempting to evaluate the effectiveness of each visualisation type with particular interest emerging in the previously untested area of interactive visualisations.

Chapter 4 discusses the methodological processes of Experiment Two. Again, participant selection, equipment and ethical protocols are considered, with participants requiring knowledge of cartography or GIS. The equipment, including the Tobii eye tracker, is detailed to show what participants would use. Specific task design was an important facet of this experiment, with the task being simplified to minimise difficulty in comprehension. Certain aspects of the experiment design process will have already been covered in Methods 1 (Chapter 3). Given the number of eye fixations would be considered a determining factor in the results chapter, it was necessary to avoid as many distractions as possible which would result in unnecessary eye movement – which in turn would result in noisy data. The procedure followed by participants was straight forward with a detailed briefing being provided before the experiment began. Once the experiments were complete, the data would be analysed using eye movement analysis software, and calculations could be performed on exported data.

Chapter 5 presents the results of Experiment One and outline how 2D, 3D and an Interactive Visualisation System performed. The purpose of the experiment was to discover which of the three visualisation types utilised would be most effective for data interpretation and analysis of GWR output. Task performance is a key measurement, and participant's task times were recorded using specialist software to gauge the differences in task completion times. The correctness of each task was recorded to measure changes in participant responses when faced with tasks of varying difficulty, using the different visualisations. Participants were divided into two main groups based on knowledge of software based visualisation systems and GWR expertise. Comparisons were made between each group to ascertain differences in performances according to knowledge or expertise. Mouse movement and clicks were recorded, providing an indication of the level of participant visualisation interactivity. These movements also help to highlight confusion or difficulty experienced by participants when encountering tasks or visualisations. Finally, participant perception was gauged using post experiment surveys based on their opinion of how fast they completed a task, how easy the task was, and how confident they feel their answer is. These results were compared between visualisations to discover the perceived best performing visualisation system. The results of comprehensibility testing indicated that the largely untested interactive systems have a greater potential for analysis, even while their appearance in published literature is rare.

Chapter 6 displays and describes the results of Experiment Two. Participant knowledge groups show that experience GIS does not necessarily equate to better performance and there is evidence of a scalability effect. Survey responses to gauge perceived participant performance provide an insight into levels of difficulty associated with the visualisations. Correctness ratios indicate differences in difficulty in another way, through the ability of participants to correctly identify a correct cluster within the visualisations presented. These suggest indications of a scalability effect with the geographical data visualisations. The time taken to provide an answer is important to note because it again indicates the possibility of a scalability effect, where participants required more time to answer tasks relating to more complex visualisations. Fixation counts are presented to add further evidence of a scalability effect, while also providing visual examples as a point of reference for the reader.

Chapter 7 concludes this thesis with a discussion of the contribution this research makes to established knowledge in the fields of cartography, spatial statistical methods and geovisualisation. To do this, the results and implications from Experiment One are discussed and recommendations on the type of visualisations for effective interpretation and analysis of GWR are made. Suggestions are applicable to the visualisation of other spatial statistical methods outputs also. Experiment Two augments the first through the investigation of data complexity and resulting issues with user comprehension of data. It contributes to the fields listed above in addition to the cognitive sciences, based on the evidence of perceptual scalability. The general discussion concludes that there is a need to better inform researchers of the potential of interactive visualisations. People do need to be properly trained to use these systems, but the limits of human perceptual processing also need to be considered in order to permit more efficient and insightful analysis.

1: Spatial Statistical Methods, Cartography and Visualisations: A Review of the Literature

1.1 Spatial Statistics

Fisher (1937) was one of the first to recognise the implications of spatial dependence, namely when one attribute in space is dependent upon another. While discussing the shapes of blocks and plots in agricultural experiments he commented,

> *"after selecting an area we usually have no guidance beyond the widely verified fact that patches in close proximity are commonly more alike, as judged by the yield of crops, than those which are further apart."*
>
> (Fisher, 1937: 73-74)

Despite the age of this publication and statement, the basic tenet of spatial dependence has not changed much from Fisher's 1937 characterisation. In the current literature on statistics and econometrics, spatial dependence is further defined as the coincidence of value similarity with location similarity. Tobler (1970: 237) was a key advocate of this definition, understanding and interpretation.

Basic concepts of statistical models are stochastic in nature, meaning that uncertainty is always present. Random variables can represent this uncertainty and the chance that a random variable will have different possible values is owing to its probability distribution. In some cases, the random variable will have a finite number of values making it a discrete random variable. In other cases it can be a value within a continuous range, thereby categorising it as a continuous variable. It is possible to have more than one random variable. In this case, you must consider their joint probability distribution which tells you the likelihood of variables taking a particular value.

It is also common practice in the literature to establish similarity between two variables. This is a measure of their covariance. The covariance of two random variables divided by their standard deviation provides the correlation (relationship between variables). Relationship as a term appears frequently within spatial statistical literature and indeed it is fundamental to

Fotheringham et al. (2001) in their publication on the spatial statistical approach of Geographically Weighted Regression (GWR).

Spatial Statistics can be viewed in two ways: global and local. Fotheringham et al. (2001) suggest there are many global forms of spatial analysis but few local forms of spatial analysis. A local level model differs from a global model in a number of ways and so it is important to understand the differences between the two when researching and applying spatial statistic methods. Global statistics are single valued meaning a singular value is returned for an entire set of data values. An example would be the mean value or the measure of spatial autocorrelation for the entire dataset. A practical example would be asking “what is the temperature of cities in the UK?” Here, the global model would return a single (mean) temperature value for all the cities in the UK. However, a local model of spatial statistics would be more comprehensive and would return a value for each city in the study area.

Global models tend to highlight similarities in a dataset by smoothing noise. Local models highlight final exceptions or anomalies across multiple regions of space. With regard to GIS visualisation, local statistical techniques provide an enhanced platform for data analysis as they return more values, thereby allowing the GIS system the ability to display patterns visually. Since local statistical models provide values for each data point or location (geo-referenced or not) in a dataset, we can say that local statistics are inherently spatial in nature (Fotheringham et al., 2001). Table 1.1 offers a summary of key differences between global and local statistics.

Table 1.1 Distinguishing between Global and Local Statistics (Fotheringham et al., 2001)

Characteristics of Global Statistics	**Characteristics of Local Statistics**
• Summarises data for whole region	Local disaggregation of global statistics
• Single-valued statistic	Multi-valued statistics
• Non-mapable	Mapable
• GIS-unfriendly	GIS-friendly
• A-spatial/Spatially limited	Spatial
• Emphasise similarities across space	

• Searches for regularities or 'laws' • Example: Classic Regression	Emphasise differences across space Searches for exceptions or local 'hot spots' Example: Geographically Weighted Regression.

It may be argued that local statistics are best used to find anomalies in data, but since spatial statistics are intrinsically location-based, it is best to display them in a map-based visualisation so that discoverable patterns can be observed and analysed within the data. It can be suggested that similarities too can be found in a dataset using local statistics.

1.2 Geography and Statistical Methods

Geographical studies require formulated research to provide quantifiable and evidence-based insights into phenomenon and Rogerson (2006) advocates the utilisation of statistical methods to contribute to geographical understanding of phenomena. A general framework can be used to outline particular approaches to geographic problems (e.g. Figure 1.1).

Figure 1.1 Framework for the study of Geographic phenomena

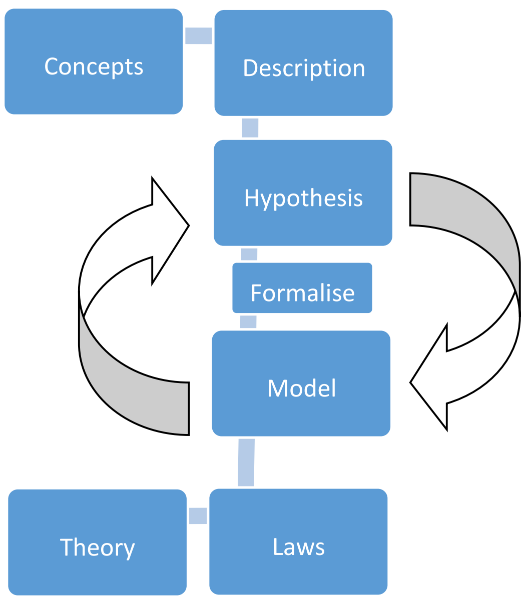


Figure 1.1 shows a general framework that depicts the processes involved in the study of a geographic problem. Initially, a core concept requiring exploration needs to be operationalised into an understandable description. From this, research hypotheses can be derived whereby statistical models can be implemented to test them. Initial results are used to refine the model or formalise it. Validating the model leads to a proof, which could be a law and in turn, laws generate new theory. By implementing this framework in the study of geographic phenomena, valued insights and understandings can be attained.

Giving a geographically relevant example: in the case of a population data set you could hypothesise that there are more males than females in a region where there are a greater number of all-boy schools compared with all-girl schools. This hypothesis is tested by collecting local population data, running local statistical models (this is a good example of where a global model specifically would not be helpful), and thus generating a model which can be used in further work.

A hypothesis is used to test a claim or assertion. A model, deriving from the hypothesis permits the study of the relationship between variables in a dataset by predicting the order against which the real data can be compared. The model can additionally suggest the nature of a relationship between variables. Support for the hypothesis is gathered from the modelling process. The framework allows you to make assertions when carrying out analysis on a dataset which can generally be applied to another similar dataset. If the model is supported by the data, it becomes useful for generalising between studies.

Rogerson (2006) advocates two types of statistical approaches to spatial analysis in geography - exploratory and confirmatory. Confirmatory methods are typically used to verify a hypothesis. Visualisation methods are exploratory in nature and in general it is necessary to adopt exploratory spatial analysis of GWR outputs. The approach taken in this thesis includes both approaches. Along with these statistical approaches, local regression is used. This approach allows for spatial analysis which is essential in studies that are geographic in nature. In the section on Geographically Weighted Regression, descriptive statistics are one of the resulting outputs of this type of local regression. Descriptive statistics are confirmatory in nature because they provide a form of definitive conclusion on the meaning of certain values or scores of that dataset when local regression is performed on it. However, most analysis will be exploratory in nature since visualisations are best used to display spatial data and inferential statistical analysis is used in these situations. Referring back to the example of a question one could ask on a given population dataset, we can call this an inference. Simple descriptive statistical outputs from a local regression model include; the median, the mode, the interquartile range, and squared deviation.

As mentioned earlier, data containing a spatial element is often complex in nature and global models lack the spatial detail for the geographic-based research. For this reason, local models are favoured, particularly when social processes are taken into consideration. Voter behaviour is an example of a non-stationary social process (Agnew, 1996). Political geography research is unanimous in agreement that voter behaviour is location-influenced. Given voter behaviour is therefore a spatial phenomenon, it can be referred to as 'spatial non-stationarity'.

When measuring spatial non-stationarity, the value at any given point in space will depend on the location of the measurement (Fotheringham et al., 2001). Fotheringham et al. (2001) provide examples where it is important to note that relationships vary over space, including sample variation. Local models may also be calibrated for subsets data, which means the parameter estimates, or outputs of the model, will not be the same over space. Fotheringham et al. (2001) also note that relationships can be intrinsically different across space. This indicates the characteristics of the local point will modify the relationships. Considering we are focussing on social processes, the differences across space could be related to population preferences. Voter behaviour is thus a prime example of population preferences requiring it to be analysed by statistical models that can process local data. Kavanagh amongst others demonstrated this point in their work on Irish voter behaviour (Kavanagh et al., 2004; Kavanagh, 2006). The statistical approach used to analyse the data in their research was Geographically Weighted Regression (GWR).

1.3 Correlation

GWR examines the relationship between attributes in a dataset, it is therefore important to understand how these relationships are calculated. The concepts of correlation and spatial autocorrelation must be discussed. Typically, the world comprises of orderliness or patterns, not randomness (Griffith, 2009). Tobler's First Law of Geography "everything is related to everything else, but near things are more related than distant things" (Tobler, 1979: 287) summarises this concept. However, this idea can be expanded upon with a statement: "but not necessarily through the same mechanisms" (Griffith, 2009: 308).

The interaction between attributes in space is a combination of distance and adjacency. This interaction is measured through inverse distance weighting (see Figure 1.2)

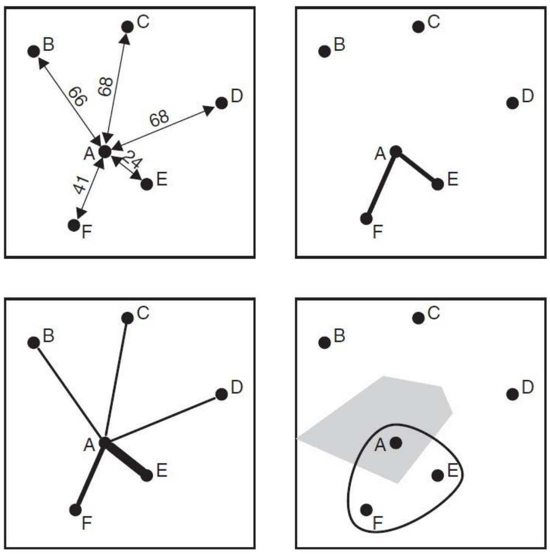
Figure 1.2 Inverse Distance Weighting Formula

$$W_{ij} \propto \frac{1}{d^k}$$

W_{ij} is the interaction between the weight and two entities i and j, at a distance (d) apart with k representing the decline in weight as the entities become further apart in space.

The concept of 'neighbourhood' is inherent in spatial autocorrelation, again referring to Tobler's First Law. A neighbourhood is an entity in space that contains a set of attributes, the number of which depends on the size of the neighbourhood. Figure 1.3 illustrates the neighbourhood concept whilst simultaneously demonstrating the concept of Inverse Distance Weighting (IDW).

Figure 1.3 Neighbourhood Concept (O'Sullivan and Unwin, 2010: 48)



The same schematic is present in each of the diagrams in Figure 1.3. The top left indicates the distance of all points B to F with respect to point A. The top right indicates the two closest neighbours and the bottom left highlights the relative importance that should be assigned to each point according to distance from A. The bottom right diagram shows the selected neighbourhood area of A. It is possible to display relationships in the form of matrices, which become important when understanding the Moran coefficient and Geary's C later. A basic matrix is a table of numbers organised in columns and rows. Distance information can be summarised in a matrix called a "distance matrix".

An adjacency matrix demonstrates the outcome for point A where a rule has been set. The rule states that objects must be less than 50 metres apart. The result shows that two objects are within this distance and the total for columns or rows is two.

An asymmetrical matrix is the result of a different rule using point A. This rule states that each object is adjacent to the nearest three objects. The results for each row and column are different. Using the example in Figure 1.3, the three nearest neighbours for point B may include point E, but the three nearest neighbours to point E may not include point B.

Having discussed how number matrices show relationships between points, the key point of this exercise can be constructed. Figure 1.4 shows a weighted matrix, or an interaction "W", using the Inverse Distance Weighting principle (Figure 1.2).

Figure 1.4 Example of inverse distance weighted matrix (O'Sullivan and Unwin, 2010: 49)

$$\mathbf{W} = \begin{bmatrix} \infty & 0.0152 & 0.0147 & 0.0147 & 0.0417 & 0.0244 \\ 0.0152 & \infty & 0.0196 & 0.0091 & 0.0101 & 0.0099 \\ 0.0147 & 0.0196 & \infty & 0.0149 & 0.0110 & 0.0086 \\ 0.0147 & 0.0091 & 0.0149 & \infty & 0.0167 & 0.0093 \\ 0.0417 & 0.0101 & 0.0110 & 0.0167 & \infty & 0.0222 \\ 0.0244 & 0.0099 & 0.0086 & 0.0093 & 0.0222 & \infty \end{bmatrix} \quad \begin{array}{c} \text{row totals:} \\ 0.1106 \\ 0.0639 \\ 0.0688 \\ 0.0646 \\ 0.1016 \\ 0.0744 \end{array} \tag{2.8}$$

Commonly, the rows in Figure 1.4 are divided by the total of each row so they amount to one. The columns following this calculation then show the level of interaction or the relationship one object has upon another. Predictably, Point A has the greatest influence because of its central location, while Point D is least influential (see Figure 1.5). This matrix principle is applied within the spatial autocorrelation matrix called a spatial structure matrix where spatial weights are assigned to record the spatial relationship between a data value and every other value in the matrix. This is discussed in more detail later.

Figure 1.5 Example of Inverse Distance Weighted matrix (O'Sullivan and Unwin, 2010: 49)

$$
\begin{array}{rc}
\mathbf{W} = & \begin{bmatrix} \infty & 0.1370 & 0.1329 & 0.1329 & 0.3767 & 0.2205 \\ 0.2373 & \infty & 0.3071 & 0.1424 & 0.1582 & 0.1551 \\ 0.2136 & 0.2848 & \infty & 0.2168 & 0.1596 & 0.1252 \\ 0.2275 & 0.1406 & 0.2309 & \infty & 0.2578 & 0.1432 \\ 0.4099 & 0.0994 & 0.1081 & 0.1640 & \infty & 0.2186 \\ 0.3279 & 0.1331 & 0.1159 & 0.1245 & 0.2987 & \infty \end{bmatrix} \\
\begin{array}{c}\text{column}\\ \text{totals:}\end{array} & \begin{array}{cccccc} 1.4161 & 0.7949 & 0.8949 & 0.7805 & 1.2510 & 0.8626 \end{array}
\end{array}
\qquad (2.9)
$$

Spatial properties can also be described by Thiessen or Vornoipolygons. These are proximity polygons in which space is dissected into areas according to the distance between each point and its neighbour. The work undertaken in Experiment Two of this research thesis contains Electoral Division polygons, so it is important to understand how they specify spatial properties.

Two concepts can be taken from proximity polygons. The first is the association of a polygon with the neighbourhood. The second is the association of a pair of polygons within the neighbourhood (O'Sullivan and Unwin, 2010). It is a popular method because of its ability to calculate almost perfectly equilateral triangular polygons.

Having discussed the underlying principles and methods of spatial analysis, it is clear that spatial statistical analysis provides insight into spatial processes. While it is possible to perform spatial analysis manually, the automation of the process is favourable. Spatial processes are categorised as either stochastic or deterministic. Stochastic processes are random in nature, while deterministic are more certain or defined.

1.3.1 The Concept of Spatial Autocorrelation

Spatial autocorrelation can be defined as the degree to which spatial features tend to be clustered together in space and conversely, the degree to which spatial features are dispersed. Measuring spatial autocorrelation requires dataset values with geometric, geographic or topological properties.

Spatial autocorrelation can be used as a diagnostic tool for model-based inferences to confirm a set of valid assumptions. Spatial autocorrelation detects non-linear relationships or in some cases model misspecification. It can help to detect the missing piece(s) of an equation, acting as a ‘surrogate’ for unaccounted variation. For example, where map patterns appear to be spatially autocorrelated and the predictor variables of a model align with the map patterns.

The term correlation can be used to describe ‘redundant information’ (Griffith, 2009). If *x* and *y* are perfectly correlated then knowing *x* will mean you know *y*. Additional items of data can provide less new information than already present in the current data sample (O’Sullivan and Unwin, 2010). Correlation can also signify similarities between at least two variables, i.e. there is a significant relationship between them. The degree of similarity decreases as the correlation coefficient approaches zero. By adding a geographical element to the investigation of the relationship between two or more variables, the term correlation is expanded to include the term ‘spatial’. Non-random distribution of attributes in space has consequences for statistical analysis. Bias may exist towards prevalent values in the sample data. If spatial autocorrelation was not commonplace then there would be little interest in geography and geographical analysis of phenomena.

An excellent example of spatial autocorrelation in action is taken from house prices. Attributes in space are intrinsically linked through their proximity to one another. The construction of an expensive house in close proximity to inexpensive houses will result in reduced house value for the more expensive house, while it increases the value of less expensive houses around it. In terms of spatial variation, there are two kinds; first-order and second-order. First order describes observations across study regions; second order describes the effects of interactions between observations. When analysing crime data, a first-order variation would be the perceived higher incidents in crime where population is most dense. A second-order variation would be the location of crime “hotspots” around bars or clubs. Generally, it is good to model both when analysing spatial data.

In instances where the geographic landscape affects the spread of diseases, they can spread according to their correlation with their topographical surroundings or neighbouring attributes. The process in which a disease spreads can be determined by spatial mechanisms

(Griffith, 2009). Spatial autocorrelation can be interpreted as an 'outcome of areal unit demarcation' (Griffith, 2009: 311).

This is related to the modifiable areal unit problem (MAUP). The example of a chess board can be used here. The pattern represented on this board exhibits negative autocorrelation, there is no discernible spatial relationship between the black and white squares on the board. However, by changing the partitioning of squares to clusters of four and averaging black and white squares so that there is a constant grey colour across the board, the spatial area of the chess board becomes positively correlated.

Statistical relationships may change at different levels of aggregation (O'Sullivan and Unwin, 2010). Another example of this can be found in politics, where 'gerrymandering' (the altering of electoral boundaries) results in pre-specified and often favourable outcomes for the gerrymandering party. Openshaw and Taylor (1979) demonstrated the possibility to aggregate data to produce a +1.0 or -1.0 spatial autocorrelation result. It has been suggested that analysts have sought to ignore MAUP due to a lack of understanding in order to carry out analysis (Openshaw, 1983). This desire to ignore MAUP is likely due to the need for simple explanations, but an attempt to address should be made.

The problem of spatial autocorrelation is understood but not solved. Techniques to measure spatial autocorrelation help to account for its effects. With this understanding of the basis of spatial autocorrelation, the methods used to measure it can now be discussed further.

Spatial Autocorrelation Estimation:

The concept of spatial autocorrelation has been discussed but the methods used to measure it are equally important to understand. Two of the most commonly utilised models (or quantitative indices) used to measure spatial autocorrelation are the Moran coefficient (MC) and the Geary ratio (GR).I evaluated the presence of spatial autocorrelation in the tested maps using a commonly employed index – Moran's I.

Measuring Spatial Autocorrelation**:**

Once again, the basic premise is that spatial data at near locations will be more similar than at distant locations. Building on this explanation, spatial data will have characteristics distances at which it is correlated with itself, or auto correlated (O'Sullivan and Unwin, 2010).

To capture the spatial relationship between all pairs of locations, we use spatial weights or a spatial structure matrix. The first row of this match is the relationship between the first location or data value and every other value in the data set. Based on this explanation we can say the first value of the second row is the relationship between the first value and the second value.

Once a matrix has been created a set of attributes can be assigned to each data value. An example being adjacency where W_{ij} receives a score of 'one' if two locations are right next to each other and conversely receives 'zero' if they are entirely apart. When constructing the weights matrix, it is important to consider the relationship a value has with itself and not just with another value at another location. An adjacency rule can be established to include only values that share edges, or meet at a corner. The contiguity of entities such as polygons can be considered as an attribute, or you can measure the distance between the centroid of one polygon and another.

The relationship between a location and itself is something that should not be considered. To avoid it, symmetry of weights is required, i.e. $W_{ij} = W_{ij}$. A symmetry ruling can be used to overcome the 'k nearest neighbours' issues, where the area A may have three different nearest neighbours to another area, i.e. $W_{BA} \neq W_{BA}$.

A key point to consider is the variety of spatial weight matrices that are possible when measuring the autocorrelation of values to W. An ideal spatial matrix is represented by weights which are meaningful to the processes under consideration. This is not an easy aspect to overcome, particularly when dealing with more fluid processes, e.g. social processes, or where processes are not well understood.

Once variables relevant to the key variable are selected the spatial matrix can be constructed and the method of spatial auto correlation measurement can be decided upon. A popular method of choice is Moran's I. It is most useful for interval or numerical/ratio data. The formidable equation for Moran's I is shown in Figure 1.6.

Figure 1.6 Moran's I Equation

$$I = \left[\frac{n}{\sum_{i=1}^{n}(y_i - \bar{y})^2}\right] \times \left[\frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(y_i - \bar{y})(y_j - \bar{y})}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}}\right]$$

According to O'Sullivan and Unwin (2010), the most important section is the second fraction. A breakdown of the elements of this formula segment are as follows:

$i$and j are areal units, or zones.

y the data value in each area.

$\bar{y}$ is the overall mean.

W_{ij} relates to the spatial weights matrix.

$\Sigma\Sigma w_{ij}$ is the total number of spatial weights.

$\frac{n}{\sum_{1=1}^{n}(y_i - \bar{y})^2}$ is the division by dataset variance, meaning I will be represented by a high number because of high variability in y.

Co-variance of two data values in two areas can be computed by calculating the product of each zone. If both data values y of areas i and j occupy the same side of the mean they are positive, otherwise they are negative.

The co-variance of each y_i and $y_j$are multiplied by w_{ij}which comes from the spatial weights matrix, i.e. it is an element of the spatial weights matrix. In an adjacency matrix a value of '1' is returned where area i and area j are adjacent, and a value of zero when they are not.

The rest of the formula normalises I based on the number of areas under consideration, the range of y values and number of adjacencies.

Data value pairs will be on the same side of the mean if they are positively correlated for *I*. If data values of an area occupy both sides of the mean then we can say that *I* is negatively correlated. A score for *I* of +0.3 or more or -0.3 or less generally indicates the presence of strong autocorrelation.

Local Statistics: Spatial Autocorrelation with a location:

The spatial autocorrelation and Moran's I descriptions thus far have not incorporated a locally varying aspect. Local statistics contain descriptive statistics with spatial data that varies from one location to another (O'Sullivan and Unwin, 2010). Referring back to Figure 1.7, we change the spatial weights matrix to a row matrix with a location element attached such that the matrix becomes:

Figure 1.7 Row Matrix

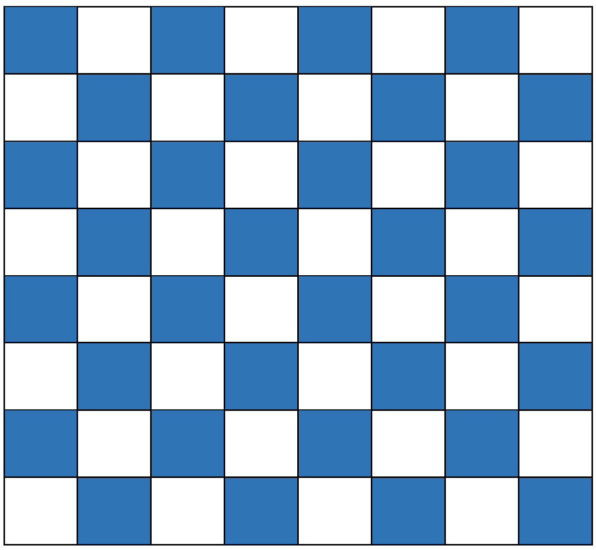
$$W_i = [W_{i1} W_{i2} \dots W_{in}]$$

i denotes each location in a local neighbourhood.

Spatial Patterns: Clustered, Dispersed and Random:

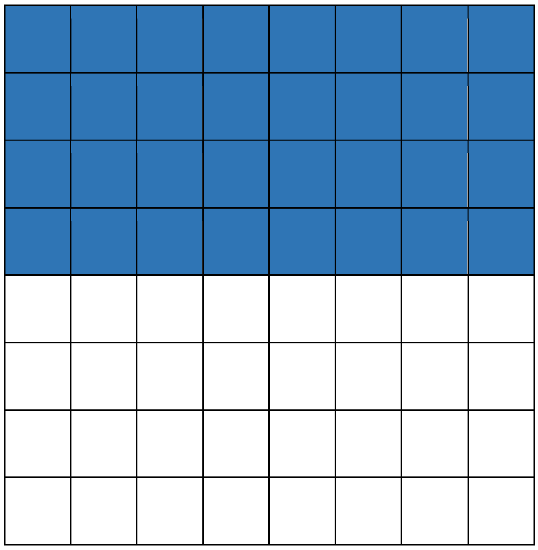
As stated previously, the outcome of spatial autocorrelation is divided into three categories. Clustered, Dispersed and Random (Koppel et al., 2011). Figures 1.8 1.9 and 1.10 visually demonstrate the difference between each of the spatial patterns. Two different attributes are coded in white and blue, 32 of each in an 8x8 or 64 celled grid.

Figure 1.8 Example of a Dispersed Spatial Pattern



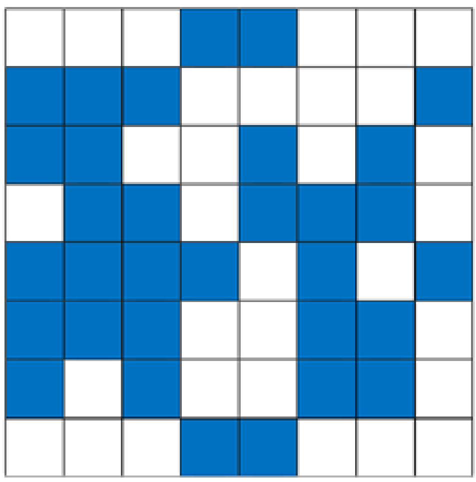
The dispersed pattern has a fully negative correlation of -1.0. This indicates there is no discernible relationship between the values in the dataset based on their proximately to one another.

Figure 1.9 Example of a Clustered Spatial Pattern



A clustered spatial pattern yields an entirely opposite score to the dispersed spatial pattern. The grid in Figure 1.9 can exhibit a spatial autocorrelation of +1.0 or close to it. It indicates the presence of a strong positive relationship. Shekhar et al. (2011) define spatial clustering as a process of grouping a set of spatial objects into clusters so that objects within a cluster have high similarity in comparison to one another, but are dissimilar to objects in other clusters. Clustered patterns are distributed dependently in space.

Figure 1.10 Example of a Random Spatial Pattern



For the Random Spatial Pattern in Figure 1.10, a value of 0.0 is returned. This indicates that the spatial patterns exhibits no detectable positive or negative relationship between the two sets of values based on their spatial distribution in space and proximity to each other. Random patterns are equally likely to occur anywhere, they are independently distributed in space and do not interact with each other. (Shekhar et al., 2011). With this understanding of correlation and spatial autocorrelation of spatial distributions, it is now appropriate to discuss GWR more specifically as a measure of the relationship between attributes with a spatial element.

1.4 Geographically Weighted Regression (GWR)

Regression modelling is a reputed quantitative method used in geography (and other disciplines) to analyse data which measures the relationship between dependent variables and a set of independent variables (Bingham and Fry, 2010). A spatial regression technique, GWR incorporates a single dependent variable and one or more independent variables which are location dependent. GWR can therefore be used to detect spatial non-stationarity (Fotheringham et al., 2002).

A global regression model equation can be represented as follows;

$$y=\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots \beta_n x_n+\epsilon$$

The GWR equation is as follows.

$$y_i=\beta_0(i)+\beta_1(i)x_{1i}+\beta_2(i)x_{2i}+\cdots \beta_n(i)x_{ni}+\epsilon_i$$

The global regression model is location independent unlike GWR. It assumes that relationships between attributes do not vary over space. Regression analysis allows you to explore relationships between attributes. The global regression model described above is a popular method utilised to investigate the relationship between geographic variables. Space does not play a role in this model, because location is not taken into account. It is assumed that the relationship of variables will not vary over space. However, it is possible that relationships between geographical variables can vary over space, in other words, spatial non-stationarity may exist.

Ordinary Least Squares is a linear regression model and a good example of a global regression model. It is a method used to make quantitative calculations. A single predictive formula produces a single output which indicates the relationship between at least two geographic variables over an entire data-space. Non-linear regression allows for the partitioning of the data into segments or sections. These subdivisions allow us to gain a better sense of the relationship between geographic variables. Geographically weighted regression goes one step further by assigning a weight to each value in the dataset. Each value is then compared to another to assess the relationship between them.

A more detailed example of the difference between a global regression model and local regression is as follows;

Imagine we need to calculate the temperate of the Island of Ireland. Ireland is split into 32 counties and let us assume there is an instrument in each of the counties to monitor temperature. This means there are 32 data values in the dataset, each representing the temperature of a county.

Using the global model we are returned a single figure for the entire Island. It is an average calculation of the temperature across the country. This may be enough in some cases, but there is no way to tell how varied the temperature is from East to West or from North to South. In the case of Ireland, the South Eastern part of the country is known as the Sunny South East. The average temperature in that region of the country is higher than the rest. The global model would not reveal this information. However, a local model such as GWR is capable of revealing such information. A local regression model will highlight the variation in temperature between each county. It incorporates a spatial element. As long as the attribute values have a spatial element attached, i.e. the data value can be attributed to space, then you will be able to perform GWR on a dataset because of the spatial weight it assigns.

Further to this, GWR incorporate a spatial weight so that the relationship between two or more variables can be examined. All the variables in the GWR model are the same as in global regression with the exception that they are now dependent on geographic location, indicated by (i) in the GWR model outlined above. This addition is a significant contribution to the understanding of geographical phenomena. Location as a core concept in the discipline of geography is now incorporated into the standard global regression model and it is expected this advance in statistics can be a significant aid in understanding geography.

GWR is based on Tobler's First Law of geography (1970: 237); "everything is related to everything else but near things are more related than distant things", which implies the existence of a distance-effect. The GWR model works on the nearest neighbour principle and calibrates a local regression model at each data point based on surrounding data points that are nearest to it in geographic (not attribute) space. Input points are also weighted based on their geographical distance from the target data point model, with each data value measured individually to detect the influence it has in relation to the independent variable based on attributes, or values of its neighbour (Fotheringham et al., 2001).

GWR output can be used to analyse the local relationship of a set of independent variables to a dependent variable through a set of local parameter estimates and local statistical measures (e.g. local R^2). Typical outputs are as follows (Fotheringham et al., 2001):

- A set of location dependent parameter estimates $(\beta_1(i),\dots,\beta_n(i))$.
- A set of T-values and S-values indicating the spatial distribution of reliability for each parameter estimate (i).
- Local R^2 value that indicates the degree of explanation for the entire GWR model at each location, i.e. for all parameter estimates together.

GWR outputs can be visualised, to allow further analysis through visual exploration of the spatial data (Andrienko and Andrienko, 2005) which involves tasks such as identification of spatial proximities, verifying spatial density and obtaining a perspective of a target measurement at a location or neighbouring location (Koua et al. 2006; Ogao and Kraak, 2002; Wehrend and Lewis, 1990). In terms of visualising and visually exploring GWR outputs, the relevant exploration tasks include: identify areas with high/low values in parameter estimate surfaces; identify areas with high/low values in t-value surfaces and link with appropriate values in respective parameter estimates; identify areas of stationarity where all parameter estimates have the same value; identify relationships between parameter estimates and values in the local R^2 surface; identify relationships between two parameter estimates and identify relationships between several parameter estimates (Fotheringham et al., 2001, Demšar et al., 2008).

Using the GWR model, the influence of independent variables on the dependent variable can be tested at every location in the dataset. Each location will have a data point, or set of data points, that are weighted for a regression according to their distance from the point the model is currently measuring, let us say this is x_1. The points that are closer are weighted more heavily than those further away. The weighted regression yields a local parameter estimate for the relationship between the point x_1 and another point we label y_1. For example, you can hypothesise that a house in a city centre with a garage attached would be more expensive because they are harder to find than the equivalent property in a city suburb where the same property is more likely to exist. Such a pattern is hidden in global regression, but GWR will discover it.

In the model above, y is the dependent variable, which is influenced by a number of independent explanatory variables $x_1 \dots x_n$. The measurement of a relationship between the dependent variable and each independent variable is provided through the $_1,\dots$, βn coefficients, and ϵ is a degree of uncertainty.

First, the limitation of the global model is reiterated. As mentioned previously global regression models are capable of detecting the influence of explanatory variables on a dependent variable, they are location independent and assume the relationships are stationary, i.e. they do not vary over space. The same response from the model is returned for every area of the study region or dataset because of this location independence as depicted using the 'temperature of England' as an example. As a result the method is unable to reveal any spatial variation that might occur in the inter-variable relationships in spatial data. However, a geographical data set is likely to contain spatial variation and one solution is to use GWR instead of the global technique.

In addition to its use by Kavanagh et al. (2004), GWR has been used in a wide range of other internationally significant studies and across many disciplines beyond geography. Table 1.2 presents the results of an analysis of GWR use in the wider literature. The search terms "'GWR' and 'Geographically Weighted Regression'" returned 245 articles at maximum in the Science Direct database. The articles that were not accessible or which did not use GWR were rejected from the collation. Relevant literature is listed in the table below and categorised according to the general field of research and the nature of the presentation of the GWR results. The particular type of research or field of study that uses GWR is much less important for present purposes, other than to illustrate just how widely the technique is used. The emergent importance of GWR as a key method in spatial statistics is evident. The literature referenced in Table 1.2, without exception, used a combination of graphs, choropleth maps (i.e. maps displaying values of a single data attribute) and tables to visualise data. The processing power of the human brain far exceeds that of a computer when it comes to pattern recognition: this is not true when it comes to processing tables of numbers. Human comprehension of complex data is facilitated when the data is represented in graphical form. Graphs or choropleth maps may be effective (or perhaps the only option) for presenting results in journal articles. However, other forms of visualisation of GWR data contribute to further analysis of the outputs. The first part of the thesis evaluates this.

Table 1.2 Review of GWR Visualisation Literature

Field/Discipline	Author
Health Geography	Cogdon (2003);
	Gebreab and Diez-Roux (2012);
	Gilbert and Chakraborty (2011);
	Holt and Lo (2008);
	Hsueh et al. (2012);
	Huang and Leung (2002);
	Mandal et al. (2009);
	Nakaya (2001);
	Tu et al. (2012);
	Yang and Matthews (2012);
	Yoo (2012);
	Zhang et al. (2012);
Political Geography	Cahill and Mulligan (2007);
	Darmofal (2010);
	Graif and Sampson (2010);
	Wheeler and Waller (2009);
Urban Geography	Szymanowski and Kryza (2011);
	Brunsdon et al. (2002);
	Cardozo et al. (2012);
	Cho et al. (2007, 2009);
	Crespo and Regamey (2012);

	Gao and Li (2011);
	Kupfer and Farris (2007);
	Luo and Wei (2009);
	Osborne et al. (2007);
	Páez et al. (2008);
	Pearsall and Christman (2012);
	Svenning et al. (2009);
	Terribile and Diniz-Filho (2009);
	Tu (2011);
Ecology	Aguilar and Farnworth (2012);
	Calvo and Escolar (2003);
	Jaimes et al., (2010);
	Lopéz-Carr et al. (2012);
	Ma et al. (2012);
	Wimberly et al. (2008);
	Zhang et al. (2004);
Biological Science	Blanco – Moreno et al. (2008);
	Wimberly et al. (2008);
Environmental Geography	Fotheringham et al. (1998);
	Fotheringham et al. (2001);
	Gilbert and Chakraborty, (2011);
	Harris and Brunsdon (2010);
	Hawkins (2003);
	Nelson (2001);
	Páez et al. (2002);
	Propastin (2009);
	Robinson et al. (2011);
	Tu and Xia (2008);
	Tu and Xia (2008);
	Wentz (2007);
	Zhang et al (2008);
Economic Geography	Li et al. (2007);
	Lu, et al. (2011);

	Mittal et al. (2004);
	Ogneva-Himmelberger et al. (2009);
	Sage and Goldberger (2012);
	Shearmur et al. (2007);
	Wei et al. (2009);
Statistical Model Comparisons	Brunsdon et al. (1998);
	Farber and Páez (2007);
	Fotheringham et al. (1997);
	Gamerman, D., et al. (2003);
	Griffith, D (2008);
	Guo et al. (2008);
	Laffan (1999);
	Leung et al. (2000);
	Tutmez et al. (2012);
	Wang et al. (2008);
	Wheeler (2009);
	Wheeler and Calder (2007);
	Zhang et al. (2005);
Land Use	Páez (2006);
Rural Geography	Leyk et al. (2012);
Migration Geography	Nakaya (2000);
Rural-Urban	Cardozo et al. (2012);
	Crespo and Grêt-Regamey (2012)
	Su et al. (2001);

Table 1.3 shows the type of visualisations found in published peer-reviewed journal articles further emphasising the broad range of fields and disciplines GWR is used in. From an analysis of each of these texts, it is clear there is a preferred set of visualisations used to present GWR related research and advanced visualisation methods such as interactive visualisations are not

utilised. One argument against the lack of an interactive visualisation presence relates to the difficulty in presenting GWR outputs on paper. For example, you cannot demonstrate dynamically-linked views on paper like you can in a live session or presentation. It could be suggested that the effectiveness of interactive visualisations is lost when presenting work for a paper as a result. However, it would still be possible to create set piece images of particular data highlights using an interactive system. In other words a researcher could take a screen capture of a particularly important outlier in their data while it is highlighted in a dynamically linked multi-windowed interactive visualisation.

Table 1.3 GWR Visualisation Types, number of papers

Discipline	Tables	Graphs	Maps	3D
Health Geography	12	3	15	-
Political Geography	1	-	1	-
Crime Geography	3	2	3	-
Urban Geography	18	9	18	-
Ecology	13	9	15	-
Biological Science	2	1	3	-
Economic Geography	10	1	7	-
Environmental Geography	14	8	18	-
Stat Model Comp	14	14	10	3
Marine Geography			1	-
Rural-Urban Geography	1	1	1	-
Land-Use Geography	2		2	-
Migration	1	1	1	-

Table 1.4 shows examples of a table of GWR outputs. This type of display was found in a high percent of the GWR articles included in Table 1.3, which was taken from a GWR workshop.

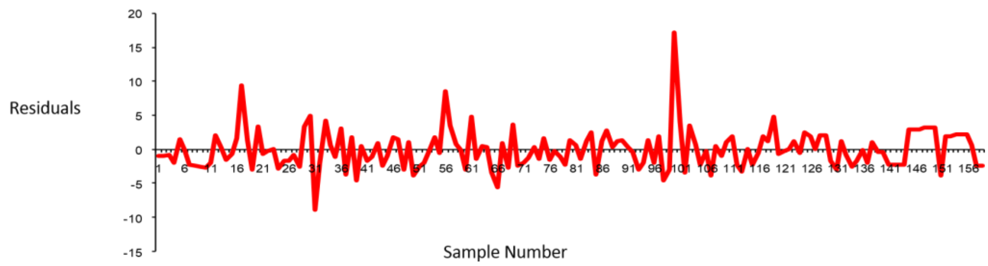
Typically these will be composed of a summary of statistics that include; the Median, Min, Max, Lower Quartile and Upper Quartile for each GWR parameter estimate. Each table will vary according to data used by the researcher. The use of tables as a visualisation method was not considered for the GWR experiment in this research thesis because graphical visualisations were the central focus of the evaluation and so were not deemed to be as significant to understand participant use of them.

Table 1.4 GWR summary statistics

Label	Minimum	Lwr Quartile	Median	Upr Quartile	Maximum
Intrcept	12.620990	13.75425	15.82323	16.31224	16.4894
TotPop	0.000014	0.000018	0.000022	0.000025	0.000028
PctRural	-0.06022	-0.05178	-0.03934	-0.03165	-0.0258
PctEld	-0.25551	-0.20309	-0.1642	-0.12939	-0.0584
PctFB	0.504876	0.82519	1.432738	2.00349	2.417666
PctPov	-0.20451	-0.16479	-0.11004	-0.05626	-0.00424
PctBlack	-0.03619	-0.01358	0.006294	0.031046	0.076566

Graphs were also not considered for evaluation because again the focus was on graphical visualisations that would allow for a better comparison of user understanding. Graphs were used by researchers to highlight data outputs particularly for statistical model comparisons with GWR and Figure 1.11 which was produced using GWR workshop data shows an example of one possible output of Residual values that provide an indication of the ‘goodness of fit’ of the GWR model.

Figure 1.11 Graph of Residual Values of GWR outputs



The most frequent method of visualising GWR outputs are choropleth thematic maps (see Table 1.3). As shown in Table 1.3, 3D surfaces were used on very rare occasions to display GWR outputs. Given the limited use of 3D visualisations in the existing literature, these are one type of potential visualisation that is used as part of the evaluation for the first experiment in this thesis.

3D visualisations can display more than one parameter estimate simultaneously. It allows for the visual comparison of two parameter estimates where one is represented by the colour and the other by the height. However, one major problem exists for 3D visualisations and it has been highlighted by experiment participants during the post experiment interviews of this thesis. It became apparent that 3D visualisations suffer from 'occlusion' which is the obstruction of view based on the angle of perspective.

In general, the level of complexity of GWR outputs means it can be difficult to visualise, (Matthews and Yang, 2012) particularly for publication, even though practical issues such as colour limitation are improving. Negative parameter estimates values can be returned so a grey scale designation from a negative number could be an example of a challenge. Given a large volume of parameter estimate data can be produced, it is advisable that it be mapped to be better understood. Choosing a visualisation type to display the data presents a challenge. For example, advanced GWR parameter estimate comparison would benefit from the use of bivariate and multivariate visualisations. Common cartographic design principle issues such as the placement of legend information can be an issue when faced with particularly large volumes of data. One viable approach to effective display is the mapping of parameter estimate values and their respective t-values simultaneously on a single map. Non-map based visualisations may not be able to effectively display large volumes of data, one example is a scatterplot graph which is highlighted later in this chapter. The most effective methods to visualise and interpret GWR results warrants investigation

There are other useful data representations that have been used to display GWR data. Common examples include a box plot and starplot, both categorised under the graph category. The boxplot is capable of displaying the distribution of data results to aid in the discovery of outliers, while the starplot can display more than one data output on the same plot that can be compared to another starplot that is designed in the same way. The reason they are not evaluated in this research is the same as outlined for tables and graphs earlier. The emphasis was largely on cartographic visualisation of location-based data. It is worth noting the use of other data representation methods but given the extensive use of these in the existing literature, it was deemed more valuable to research the use of 3D visualisations of GWR in this thesis.

1.5 Cartography and Visualisations

1.5.1 Modern Cartography and Visualisation

A major component of this thesis is concerned with the way in which spatial data is visualised. In the sections above, information was offered on the local regression model (GWR); the outputs of GWR are visually represented in the experiments. In this section aspects of cartography and visualisations are discussed. When one couples geographical data with visualisations, geovisualisations are created. This type of visualisation is designed to be more effective in displaying spatial data, and the usefulness of these through GIS-based and tailored geovisualisation tools are discussed in Methodology Chapter 4 in more detail. Before visualisations there were thematic maps, thematic maps are still a commonly produced item in cartography and geography in general, therefore it is important to understand them.

1.5.2 Thematic Maps

Thematic mapping software was used to create most digital maps to begin with. Digital maps are still a major visualisation method. One popular digital map production software is ESRI's ArcGIS Suite for example. Edsall (2003) highlights a general agreement amongst statisticians that visualisations are capable of providing insight into datasets, referring us to the research of (Tukey 1977; Hurley and Buja, 1990; and Wegman, 1990). Furthermore, maps and graphic devices have become key components in the exploration of spatial data, particularly in geovisualisation. Research has been carried out (DiBiase, 1990; MacEachren et al., 1992; and MacEachren and Kraak, 1997) which further supports this point. GWR falls into this category of complexity and stands to benefit from geographic visualisation.

They can be used in three ways: to provide specific information about particular locations, to provide general information about spatial patterns and to compare patterns on two or more maps. A common mistake made by map-makers is to place too much emphasis on the display of inordinate data. Maps typically limit data to a fixed placement, but interactive graphics allow the comparison of distributions in a less fixed manner, i.e. they are more dynamic. However, none of the aforementioned visualisations would be effective without good visualisation design.

Before discussing visual design, it is worth noting the common graphical methods used to represent geographic data. At a more basic level there are thematic maps, these are followed by multivariate maps. Modern methods of data representation include interactive visualisations. As mentioned, thematic maps are used to emphasise the spatial pattern of one or more geographic attributes (Slocum et al., 2009). This is why choropleth maps are often used to display GWR outputs. Thematic maps are primarily used in three ways; for pattern comparison, to provide general information about spatial patterns or to provide specific information about particular locations.

A Choropleth map can be defined as; "a thematic map in which areas are coloured, shaded, dotted, or hatched to create darker or lighter areas in proportion to the density of distribution of the theme subject" (Geography Dictionary, 2014). A common type of thematic map is a choropleth map where data (or enumeration units) are displayed in various colour shades denoting the data value of that area. Choropleth maps sometimes contain symbols which are proportional in their size to the data value they represent and these types of maps can be produced in 2D or 3D form. Figure 1.3 in section 1.3 above is an example of a 2D choropleth map. Data presented on choropleth thematic maps is highly readable and these maps were utilised to display data in Experiment One and Experiment Two of this research thesis. Among the key limitations of choropleth maps are that they do not portray variation which may occur within data units and the unit areas can be arbitrary (Slocum et al., 2009). Other thematic map types described below are not used in either experiment of this research but it is worthwhile to discuss their uses in the wider context of understanding the range of map visualisation types available for researchers to utilise.

Proportional Symbol Maps are used to show quantitative differences. In the case of proportionate symbols, their size would vary according to the data value they represent, i.e. the symbols are scaled in proportion to the magnitude of data they represent. For example, a country's population could be represented by graduated symbols, or graduated circles (i.e. the larger the circle, the greater the population in that area). Symbols can also be suggestive of the data they represent. (Slocum et al., 2009). Once again, using the example of a country's population, you could use an illustration of a person instead of a standard shape. Proportional symbol maps are more valuable than Choropleth maps for the display of raw totals (Slocum et al., 2009).

Dot distributions or dot density symbols are another useful method to augment the display of information on a thematic map, where appropriate. Dot density maps are used to display the distribution of data (Slocum et al., 2009). For example, if you wanted to present the number of cases of a disease in a designated epidemic area, you could use the dot density method. The number of dots placed inside each spatial unit of a map would be determined by the number of recorded infections of the disease. In some instances it is useful to set the number of dots that would appear according to a multiple of a data value. For example, one dot could represent five cases of infection, equally the dot density could be set so that one dot represents 100 cases of infection. A single dot could represent the location of an earthquake or a city.

Isoline or Isopleth Maps essentially use continuous lines to display information such as temperatures or elevation levels. Points of equal value are connected. A contour map is an example of an isoline map. These are specialised maps suitable to select ahead of Choropleth maps when the data is known to be part of a continuous phenomenon (Slocum et al., 2009).

Cartogram thematic maps are less commonly used but are unique in themselves. The main advantage of using a cartogram is that large enumeration units normally hidden on conventional map projections can be emphasised (Slocum et al., 2009). A distorted effect is created where the size of an attribute is portrayed proportional to its value within the dataset. Generally, cartographers will avoid distorting the spatial relationships on a thematic map. Distance or geographic areas are two common types of cartograms where distance or geographic area is displayed proportionally within one of two types of cartograms, contiguous or non-contiguous. Contiguous cartograms attempt to maintain contiguity of shapes while non-contiguous cartograms retain the shape of enumeration units. Non-contiguous cartograms can make shapes difficult to identify depending on the enumeration values.

In some cases, map makers need to display multiple attributes at once. For example, the percentage of owner occupied housing with respect to population and income levels. Displaying these three variables requires the production of a multivariate map while bi-variate maps only display two results. A bivariate map is the cartographic display of two attributes (Slocum et al. 2009). Bivariate mapping can be accomplished using Choropleth maps by overlaying them with additional information using another graphic such as a histogram. It can be more commonly achieved by overlaying proportional/graduated symbols or dot density methods.

Multivariate visualisations comprise more than two different variables. One method is to display single attributes on separate maps within the same window. A famous example of a multivariate symbol is the Chernoff face. Chernoff faces have distinct facial features associated with individual attributes. Starplots/Snowflakes (also known as polygonal glyphs or ray-glyphs) can be overlaid on Choropleth maps to display several additional attributes on a map. Lines representing single attributes extend from a small interior circle proportional to the values they represent (Slocum et al., 2009).

1.5.3 3D visualisations

Just over a decade ago, 3D visualisations were an emerging technology within geographic applications (Pullar and Tidey, 2001) and GIS systems such as ArcGIS (one of the most renowned programmes used to visualise geographic data in 3D). ArcGIS is a convenient system for data management. Geographic 3D visualisations are closely aligned with 3D visualisations produced by computer aided drafting systems (CAD) systems which allows an image to be viewed dynamically (Robins et al., 2005). When performing analysis on 3D visualisations, it is not unusual for the analyst to want to assess a 2D variant either simultaneously or intermittently, as observed in Experiment One in this thesis. In geographic terms, 3D visualisations can also be known as 3D surfaces because the 3D surface shown can be manipulated to show a set of data values, it does not necessarily relate to the physical topology of a geographic area.

3D visualisations are most useful in certain situations, for example, modern medicine, petroleum exploration, gas exploration (Salom et al., 2009) or to display groundwater flow patterns (Robins et al., 2005). Geoscientists would need to produce a 3D model which supports the decisions of geologists to drill in particular locations for oil or gas. 3D visualisations present a more dynamic view of geographic data but have one major drawback mentioned earlier, they suffer from occlusion (Salom et al., (2009) and Tsigas (2007)). Generally, 2D visualisations have two degrees of freedom, they are the x and y axis. 3D visualisations have 6 degrees of freedom; the x, y and z axes. In terms of interaction 3D visualisations are superior (Kok and van Liere, 2007). However, additional viewing angles do not necessarily result in an enhanced ability of non-expert users to analyse data (Cline, 2000). In short, it has been difficult to assess the overall performance levels of 2D visualisations compared to 3D visualisations because of the variety of devices, interactions, techniques and participant expertise (Zudilova-Seinstra et al., 2010). For example, in Experiment One of this research 3D visualisations are assessed by participants with a certain level of expertise for a spatial statistical model. Traditional 3D visualisations revolve around a windows mouse-menu-based interaction paradigm (Kara et al., 2007).

1.5.4 Interactive Visualisations

A majority of our brain's activity deals with processing and analysing visual images. Visualisations are described as a new field of research which examines our innate potential to effectively process visual representations in knowledge intense tasks (Bukhard and Meier, 2005). It is more appropriate to view the use of visualisations as one of the most effective methods of analysing the increasing level of complex data. Previously described as a rapidly advancing field of study (Card et al., 1999; Chen, 1999; Spence, 2000; Ware, 2000), the visualisation of information or information visualisation is an increasingly important aspect associated with geographical analysis. Information visualisation can amplify cognition (Card et al., 1999). The thematic maps and 3D visualisations described above are forms of information visualisation and they can be used to discover patters (e.g. trend clusters, gaps or outliers) concerning individual items or groups of items with the overall goal to derive new insights (Bukhard and Meier, 2005). They have three main characteristics; they are interactive to some extent through manipulation of a user interface where operations such as data selection or filtering can be applied, they are dynamic meaning they can be rendered in real-time and they help with embedding details and context through dynamic zooming. By this logic, more interactive visualisations should allow for improved analysis of geographic data.

Extending this concept, we could say that one of the most current and little used form of visualisation, "interactive visualisation" represents one of the most, if not the most, powerful tool available for analysis of geographic data. Given the emerging use of this type of visualisation in the literature, this research thesis aims to incorporate interactive visualisations into the exploration of user understanding. Interactive visualisations can be purely 2D visualisations, purely 3D visualisations or a combination of 2D and 3D information visualisations. These visualisations can consist of one or more visualisation windows which displays the dataset. Experiment One of this research contains a typical example of a purely 2D interactive visualisation consisting of several interactive windows displaying a dataset using different visualisation methods. 3D examples include cutting edge work on Space-time cubes (Demšar and Virrantaus (2011) and McArdle and Demsar (2011)) to analyse aspects of data including trajectories and densities.

1.6 Visualisation design

“Cartography is about representation” MacEachren (2004). The functionality of a map needs to be taken into account when designing visualisations. Artistic design does not necessarily mean good design and the release of the 1972 subway map of New York (See Figure 1.12) was met with mixed reactions. The design was innovative because it re-imagined the subway system but passengers and visitors could not easily grasp their location by looking at the map. Standardisation could be considered as important to visualisation design, because designer interpretation will vary. Standardisation helps to increase functionality because it focuses on the creation of practises that are most effective for visualisation design. However, cartographers disagree with this idea preferring to form an objective view based on the visualisation they were designing Robinson (1952) and Robinson (1973). The audience a particular visualisation is aimed at would generally dictate how the end product looks.

Figure 1.12 Massimo Vignelli's 1972 New York subway map

MacEachren (2004) produced a simple diagram depicting the process a cartographer undertakes in producing a map. Figure 1.13 shows this diagram and you can see that the recipient is the final link in the design-chain. In other-words, it appears that end users have no input into design. Cartographers saw visualisations as communication tools to be analysed by recipients (Kolacny, 1969). Modern cartographers are now adopting what was previously referred to as a representational approach (MacEachren, 2004) or a user-centered design approach when it comes to visualisation design. Research has been carried out to highlight the necessity of user input and modern visualisation design will be discussed in the next paragraph. Diversity of the user perspective is an inherent difficulty when attempting to communicate through visualisation. Broad interpretations of map detail can occur. Howard (1980) discusses an objective approach to incorporate user related issues including; cultural, psychological, and required communication processes for symbols. Adding to this, perception is a factor in visualisation design as it is linked to user interpretations, or how a user perceives a visualisation. Consider perception as a representation of what we see with our eyes and cognition as the objects and relationships within that scene. You will see that map objects such as symbols are not the only concern cartographers should focus on when designing their representations (MacEachren, 2004).

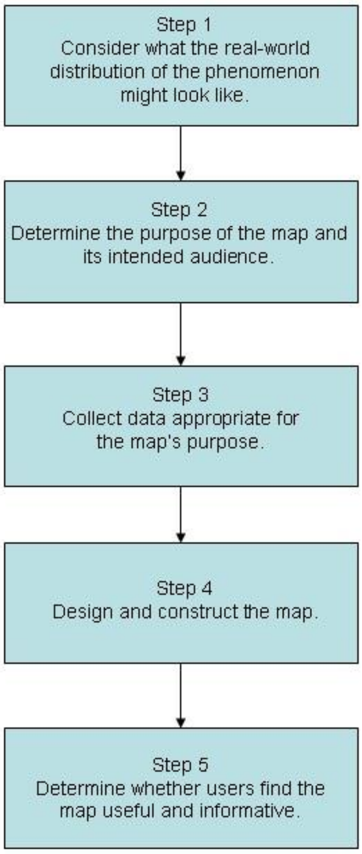
As already mentioned user interpretation of representations will vary between each user. This complicates the ability of the communication model in Figure 1.13 to function effectively. If a user will create assumptions on what a visualisation depicts, then user derived meaning may never correlate with the designers. To aid with map comprehension, Pinker's framework (1990) on graphical comprehension can be considered as a useful tool.

Figure 1.13 Schematic depiction of cartography as a process of information communication (MacEachren, 2004)

Modern Visualisation Design needs to incorporate users in a more meaningful way. This is because a significant proportion of visualisation design problems are related to users. A similar question that was asked about the usefulness of geovisualisation can be repeated. "What is a visualisation good for?" Haklay and Tobon (2003) discussed how we understand the diversity in human behaviour when dealing with graphic visualisations. Research carried out by Wilson et al. (2010) on personalising visualisations shows that we need to account for user differences. This thesis addresses these issues and are discussed in the Experiment Design chapters.

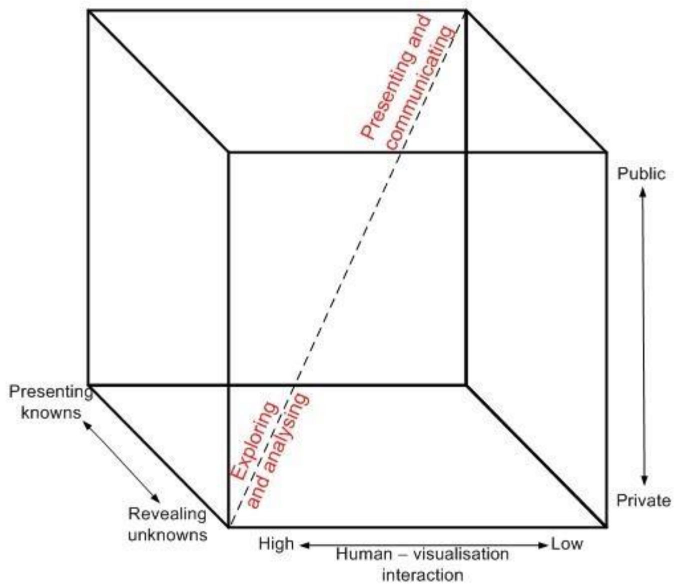
Thematic maps are also used for the acquisition of mapped information. A map conveys information and map creation should follow the model process shown in Figure 1.14. This model was applied during this research, particularly during the experiment design stage to ensure data was appropriately displayed for the end users. Thematic maps should also be considered as a significant piece of interactive visualisation. They are a staple of visualisations to which other types of visualisations are linked.

Figure 1.14 Basic Steps for communicating map information to others (Slocum et al., 2009)



You can split visualisation into two different definitions. The first is geographical visualisation. 'Geographical visualisation [can be defined] as the use of concrete visual representation, whether on paper or through computer displays or other media, to make spatial contexts and problems visible, so as to engage the most powerful human information-processing abilities, those associated with vision' (MacEachren, 1992). The second is based around MacEachren's (1994) Visualisation Cube, displayed in Figure 1.15. MacEachern's argument considered geographic visualisation as a private activity and communication a public activity. The private activity would be highly interactive, while the public would be the opposite, where knowns are presented. More recently the phrase geographic visualisation has been replaced by geovisualisation (MacEachren et al. 1999), and this has become the standard term used. Cyber-cartography was a proposed term by Taylor (1997), and it incorporates some of the elements of geovisualisation such as 'highly interactive user engagements' and 'new research partnerships'. In modern day terms we could describe these types of visualisation as online interactive visualisations. These types of visualisations are an increasingly important driving force behind geographical visualisation advancement.

Figure 1.15 MacEachren's Visualisation Cube (1994)



The importance of visualisations in the advancement of our understanding and analysis of complex data is evident, but it is also crucial to have an adequate set of tools to carry out analysis on data. The next section will discuss the type of visualisations that can be used to successfully carry out the first experiment of this thesis.

1.7 Geovisualisation Systems: design and functionality

Geographic visualisations are often referred to as geovisualisations. The first question to ask would be; "what is geovisualisation effective for?" According to Edsall (2003) the research initiative known as geovisualisation grew out of the need to represent and interact with complex data. Interactive visual representations and their tools encourage creative and intuitive exploration for structures and patterns which are normally difficult to detect through non-visual techniques (Edsall, 2003). GWR data would benefit greatly from this type of exploration. Insights about particular characteristics are gained with the choice and style option. The tools and displays combine with users leading to knowledge construction about complex data via the representations (Rogers, 1999).

Ogao and Kraak (2002) share other authors' opinions that geographic visualisation techniques have linked with increasingly enhanced and interactive dynamic tools. They now complement quantitative approaches to data, particularly in cartographic visualisations. Cartographic animations have been used in mapping geospatial process, and are important for observing geospatial process, not just the end state (Ogao and Kraak, 2002). This is evident from the many authors who have worked with these animations (Kraak and Klomp, 1995; Acevedo and Masuoka, 1997). In the past, a lack in technology and user knowledge mired the incorporation of user needs in design processes (Ogao and Kraak, 2002). Now, early work can be carried out to create a list of user requirements and visualisation goals which could augment visualisation design.

A major part of this research thesis is concerned with the way in which spatial data is visualised. The local regression model GWR has already been discussed along with the outputs of GWR being visually represented in the experiments. In this section, key aspects of cartography and visualisations were discussed, highlight how coupling geographical data with visualisations can create geovisualisation. This type of visualisation is designed to better display spatial data, and the usefulness of these through GIS based and more tailored geovisualisation tools are discussed.

According to Edsall (2003) the research initiative known as geovisualisation grew out of the need to represent and interact with complex data. Interactive visual representations and their tools encourage creative and intuitive exploration for structures and patterns which are normally difficult to detect through non-visual techniques (Edsall, 2003). Insights about particular characteristics are gained with the choice and style option. The tools and displays combine with users leading to knowledge construction about complex data via the representations (Rogers, 1999).

1.8 Usability in GIScience

Thematic maps have evolved into geovisualisations which have been discussed earlier, and these can be used to emphasise spatial patterns, usually of one or more geographic attributes. They can be used in three ways; to provide specific information about particular locations, to provide general information about spatial patterns, and to compare patterns on two or more maps.

A mistake that can be made by map makers is that they place too much emphasis on the display of unnecessary data. Maps typically limited map data to a fixed placement, but interactive graphics allow us to compare distributions. Geovisualisations are also used for information acquisition and memory for mapped information. These maps can be useful for analysis up to a point, but as MacEachren (1995) mentioned that traditional map studies are no longer necessary with the development of interactive visualisations. You can say that many of these interactive visualisations are linked to geovisualisation in the modern day because of the methods or tools available to display data.

In the past, a lack in technology and user knowledge mired the incorporation of user needs in design processes (Ogao and Kraak, 2002). Early work was carried out to create a list of user requirements and visualisation goals which could augment visualisation design. Lucieer and Kraak (2006) consider the development of dynamic visualisation cartographic techniques, a result of improved computer graphics technology. This could be possible through GIS based packages. However, Lucieer and Kraak (2004) state that most GIS packages do not offer tools to "model, visualise and manage uncertainty in classifications". Users now want information map quality, not just the map representations of data. Lucieer and Kraak (2004) designed a tool for interaction with a fuzzy classification algorithm. With this in mind, usefulness and quality aspects are important to consider and became apparent in the analysis of the results of the experiment carried out as part of this research thesis.

A starting point on how to design visualisation operations is illustrated by Shneiderman (1996) and earlier by Goldstein et al. (1994). One of the main problems in designing visualisation systems lies with the levels to which one can explore data. It is difficult to design a visualisation system capable of meeting every user's specific needs, both for broad and narrowly defined data exploration. At the time of publishing Ogao and Kraak (2002) noted a struggle to solve the complexities involved with temporal queries.

The usability of GIS products has improved significantly in recent years (Hacklay and Tobón, 2003). However, GIS products still require considerable technical knowledge to operate them and efforts to increase the ease of use are important. Human Computer Interaction (HCI) work has been carried out with GIS applications since the 1990's (Nyerges at al. (1995a) and Davies and Medyckyj-Scott (1996)). In this early stage of GIS development it appears the same cannot be said of Public Participation GIS (PPGIS), which has not received the attention it requires. As our ability to disseminate GIS based information grows through the use of the internet it can be argued that Public Participation of GIS visualisations are ever increasing. Mapped visualisations displayed in more understandable internet flash or similar based forms have allowed researchers to reach a wider audience with their work. A prime example is the work carried out by the National Centre for Geocomputation with the Irish famine database (Kelly and Fotheringham, 2011).

Ogao and Kraak (2002) share other author's opinion that geographic visualisation techniques have linked with increasingly enhanced and interactive dynamic tools. They now complement quantitative approaches to data, particularly in cartographic visualisations. Cartographic animations have been used in mapping geospatial process, and are important for observing geospatial process, not just the end state Ogao and Kraak (2002). This is evident from the many authors who have worked with these animations (Kraak and Klomp (1995) and Acevedo and Masuoka (1997)).

Geospatial animation functionality has been developed with the increase of computer power, both in speed and graphical elements. A fundamental component of animation presentation is simplicity. The objective is the avoidance of any user distractions such as complex animations. It is important to understand space-time structures and processes within animations. Changes can be continuous, cyclic or discrete, they can also be fast, slow or abrupt (Ogao and Kraak, 2002). An emphasis must be placed on the user's ability to notice changes in their animation. Appropriate spatial and temporal scales are crucial. As you can see work to bring geovisualisations to the fore is ongoing and it emphasises the importance placed on geovisualisations for analysis of increasingly large and complex datasets.

Testing the effectiveness of geovisualisations is a crucial component within this thesis. With the discussion above in mind, it is imperative the design of the experiment is carefully considered and justified in the context of best practice in map and visualisation design.

1.9 Evaluation Design

There are several possible evaluation approaches to be considered. Usability measures such as task performance and task nature (e.g. closed or exploratory) are important factors. Wehrend and Lewis (1990) produced a classification list of operations which is designed to "distinguish problems for which the user's goal in viewing the representation differs" (Wehrend and Lewis, 1990). The effectiveness to which a researcher can extract their required information from a visualisation can be regarded as a criterion, or a usability measurement. This section presents some of the evaluation approaches.

There are common challenges associated with experiments involving information visualisation tools. However, the location of an appropriate testing site is a problem associated with visualisation tools. It is argued that they need to be tested in a real setting, and not solely in a laboratory. Even if this is achieved the tester has to persuade the user to take the leap of faith, and use the tool. The fact a tool may not be specifically designed for the needs of a user is one major obstacle (Plaisant, 2004). Plaisant argues researchers should invest in resources to tailor their tools to specific user needs. This relates to the first experiment. Participants operate tools tailored for the experiment, and the dataset is not chosen by the participants. Also, it can be argued that the experiment laboratory would be appropriate if it replicated a natural working environment.

The previous sections have detailed the significance of Human Cognition in relation to a visualisation interaction. Cognitive aspects are important particularly in the context of information visualisation because they assist in the comprehension of visualisation performance. According to Chen and Yu (2000), "users with a stronger cognitive ability, i.e. high psychometrics, tend to perform better with information visualisation systems than users with weaker cognitive ability in terms of accuracy." However, the work of Gibson (1979) and Schuman (1987) argue that human cognition is guided by their environment and not their head. This indicates that the environment has a role to play in the results of any usability experiments, and as such must be taken into account. Different methods of analysis have been proposed. For example, a person should be observed in a naturalistic setting where they would make perceptual judgements according to Brunswick (1943). Lave and Wegner (1991) have also deemed work in a naturalistic setting appropriate.

When assessing data recorded during an evaluation experiment the technique to measure this data should be appropriate. A correctness of response system was used by Koua et al. (2006), where the correct answer is coded one and the incorrect answer is coded zero. This could be a useful method if coupled with the interaction logs, it would be easier to analyse the experiment results. Time can be used as a measurement of performance, and can reveal some important differences in experiments as Koua et al. (2006) discovered. This method of recording correctness of user tasks could be applied to the experiments carried out in this thesis. However, given the complexity of the tasks in the first experiment this response system should be adapted.

Generally empirical evaluations only include simple tasks; for example, identify and locate tasks. Performance reports on a task by task basis are the preferred method, compared to overall performance reports. This allows for in-depth analysis of performance. Tools can be matched with particular tasks. This is something that could be achieved through evaluations of visualisations in this thesis to an extent. The second experiment involving eye tracking technology would work most effectively with these types of tasks.

Given the nature of the GWR outputs in the first experiment the inclusion of a set of exploratory tasks would represent a more accurate testing platform. The evaluation methods used by Koua et al. (2006) emphasised exploratory tasks and knowledge discovery support. The authors presented an approach for assessing the usability and usefulness of the visual-computational analysis environment. It is important to understand user cognition and how users make interpretations and inferences when conducting analysis. Crucial to this is the choice of an appropriate visualisation metaphor. The evaluation results serve as a guideline for the design of geovisualisation tools which integrate several visualisation types simultaneously. It is evident from the evaluations conducted by Koua et al. (2006) that certain visualisations perform better for each task. This suggests a multiple representation environment would be best for analysis of varying spatial and temporal data. It could also be suggested that any experiment including multi-windowed representations would be augmented by comparative single windowed representations. A more comprehensive set of results could be obtained for analysis.

Gabbard et al. (1999) have produced methods for virtual environment (VE) usability engineering. Their methodology steps includes; user task analysis, expert guidelines based evaluation, formative user-centric evaluation and summative comparative evaluations. The user task analysis is the process of identifying a complete description of tasks, subtasks and methods required to use a system, as well as other resources necessary (Gabbard et al., 1999). The user task analysis represents insights gained through organisational and social workflow, and a general understanding of the needs of the user. This step of user analysis was often overlooked in the 1990s. Developers and evaluators operate on a 'best guess' to interpret how VE applications should be designed. It highlights the need for a more user inclusive approach to visualisation design.

Heuristic evaluations or usability inspections are capable of identifying potential problems by measuring established design guidelines to an applications user interaction design. More than one person should perform these evaluations since it is unlikely that one person will identify all of the problems in an application. Evaluators inspect the application individually and the results are then combined. Gabbard et al. (1999) have found the expert guidelines to be too general, and have designed a set of guidelines specifically for VEs within a framework of usability characteristics.

The expert-guideline-based evaluations are critical to the effectiveness of formal and summative evaluations (Gabbard et al. 1999). They are useful for streamlining the two latter evaluation types. Time is not wasted on identifying obvious usability problems.

According to Andrienko and Andrienko (2005) geographical analysis is considered to be a set of operations or tasks to achieve data exploration aims. The process involves the identification of different element spatial proximities, verifying spatial density, and obtaining a perspective of a targets measurement at a location or neighbouring location (Koua et al., 2006).

Weldon (1996) describes operations in a more specific manner, the tasks include:

- Identification of clusters in data, and relationships between elements.
- Comparison of values at different spatial locations, distinguishing a range of values.
- Relation of the value, shape and position of identified object.
- Analysis of extracted relevant information.

There are several taxonomies for visualisation that have been suggested by authors (Zhou and Feiner, 1998; Ogao and Kraak, 2002). An earlier taxonomy for visualisation operations can be obtained from reading Wehrend and Lewis (1990). It is presented in Table 1.5. Table 1.6 shows a more modern taxonomy that condenses the operation keywords (Ogao & Kraak, 2002). However, this taxonomy can be streamlined further for the purposes of the experiments in this thesis, and Table 1.7 shows an adapted version.

In the first experiment, for Task 1, participants identified a particular Electoral Division. However for Task 2 and Task 3 an additional appropriate operation class ‘relationship’ within each task was added to Table 1.19 below. These operations aided the creation of a series of tasks to represent the types of questions a user of GWR would want to answer when analysing a complex data set.

Table 1.5 Taxonomy of Visualisation Operations

Operation	Operation Description
Identify	to establish characteristics of a piece of data recognise or select by analysis.
Locate	to discover the extract place or position of data of object.
Distinguish	to discern, to identify characteristic differences.
Categorise	to order/split or arrange a group (of data in this case).
Cluster	a group of similar things positioned or occurring closely together.
Distribution	the way in which something is dispersed, diffused or spread.
Rank	a position within a fixed hierarchy.
Compare	to estimate, measure, note similarity or dissimilarity.
Associate	a logical or casual connection between two things.

Correlate	have or bring into a relationship in which one thing affects or depends on another.
Determine	to ascertain, to decide by research or calculation.
Look up value	similar to identify according to Roth and Mattis (1990).

Source: Reproduced from Wehrend and Lewis (1990).

Table 1.6 Four Generic visualisation operations (reproduced from Ogao and Kraak, 2002).

Visualization operator	Visualization sub-operator	Example of definitive characteristics of results.
Identify	Spatial identification	Length, area, irregularity, minimum, maximum range, distance, pattern of distribution
	Temporal identification	Extent: longest, shortest; sequence: first, last; category: nominal, ordinal, interval/ratio; movement: continuous, cyclical, intermittent.
	Thematic identification	Name, symbols (legend)
Locate	Spatial Location	(x,y), (φ, λ), grid locations, (rows, columns), near, within, between, above, below, neighbourhood of.
	Temporal Location	Event, valid time t, observed interval (t1-t2) before, after, together, next.
Associate/ compare	Spatial association/ comparison	Topological relations, spatial collection, covariance, correlation
	Temporal association/ Temporal comparison	Temporal relations, time between objects, orientation (before, after), adjacency (just before, just after), causality (correlation).
*Relationship		Relationships between two variables.

Table 1.7 Adapted Taxonomy of Visualisation Operations

Operation	Operation Description
Identify	to establish characteristics of a piece of data recognise or select by analysis.
Locate	to discover the extract place or position of data of object.
*Relationship	to discover the type of interaction between two variables.

While designing tasks, it is important to prevent any overlapping of content to minimise the possibility of a learning effect occurring. The learning effect experienced by participants who recall information from memory as they become familiar with the layout and content of an experiment had to be minimised. This familiarity could alter results (Robinson and Griffin, 2010). The order of the experiment tasks example - from the Latin square technique by adapting its concept of randomisation to ensure GWR tasks, some of which are complex, would not be repeated. The Latin squares technique (Dénes & Keedwell, 1974) is a matrix of letters or numbers used to prevent experiment participants from learning the details of an experiment.

Since this experiment contains GWR related tasks, i.e. bivariate and multivariate tasks it complicates our efforts to use a textbook Latin square technique for the experiment. A standard Latin Squares task randomisation sequence of ABC CAB BCA is ideal for univariate tasks. However, given the number of tasks and the fact that they were also bivariate and multivariate in nature, this was not possible. As you can see, the technique is designed to prevent a task from recurring. In the case of the GWR experiment there are five parameter estimates to account for and their respective T-values were used as an additional complexity to simulate typical GWR task analysis. Tasks could be arranged in a sequence that helps to prevent the learning effect whilst providing the necessary complexity required. In some cases, more than one task contained the same attributes but participants would be asked to look for highest or lowest value relationships so tasks would still remain different. In other words, no two tasks were the exact same.

The next chapter will address in greater depth the human-related aspects that are important to consider for effective visualisation of data. It should be noted that the key goal of this research was not to design a visualisation, but rather to compare and test the utility of visualisations. The first experiment tests these for GWR, and the second experiment tests if human perception is affected by scalability. It was found that participant performances partly reflected familiarity with and ease of use of the visualisation tools, as well as their understanding of GWR and GWR outputs.

1.10 Glossary of Terms

2D visualisations: A two dimensional graphical display of data values.

3D visualisations: A three dimensional graphical display of data values.

Choropleth Map: a map in which data is shaded with varying colour intensities according to the value of data of that unit.

Geographically Weighted Regression: A spatial statistical model used to examine processes which vary over space.

Geovisualisation: Is a simplified version of geographic visualisation, a term simplified by MacEachren et al. (1999). It is the use of visual representation of spatial information on paper or in digital form. It allows humans to engage their processing spatial data. Unknown spatial information can be revealed in a Geovisualisation.

Interactive Visualisation: A set of dynamically linked windows in which data is displayed using different types of visualisations. These visualisations can be manipulated to highlight one or more data values simultaneously across every dynamically linked window.

Spatial non-stationarity: relationships which are not constant across space.

Spatial Statistics: A set of statistics that contain a spatial element or geographic properties. They give explicit consideration to spatial properties, e.g. location, patterns and distance.

Thematic Map: Used to highlight spatial patterns of one or more geographic attributes. They include; Choropleth maps, Dot Distribution/Dot Density maps, Graduated symbol maps and Isoline maps.

Visualisations: Any graphical technique that is used to display data and communicate information of that data.

2: Cognitive Processing of Graphical Data

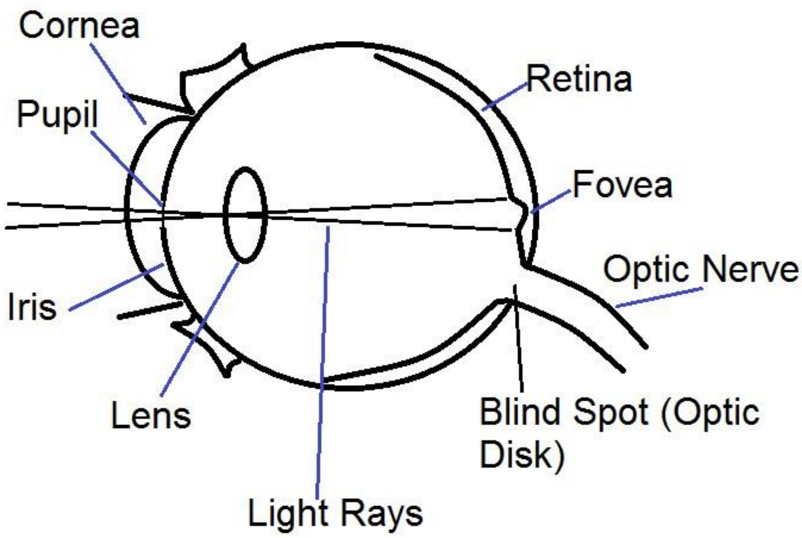
The previous chapter reviewed literature in relation to the first experiment on visualisations within this research thesis. Chapter two is concerned with literature relevant to the second part of the research within this thesis. Here, cognitive processing of graphical data is considered in depth to contextualise the likely output from Experiment Two of this research. André Du Laurens, in the 16th century, referred to the eyes as the 'window into the mind' because the eyes move to objects of interest and salience to the mind. Therefore, analysis of eye movements are a critical part of the second phase of the research carried out in this thesis.

The chapter is structured into seven key sections. Firstly, the structure of the human eye is described given this is the organ of sight and Experiment Two is concerned with how users view and then interpret visualisations. Next, a discussion on eye movement research is provided to help contextualise the research aims of Experiment Two in this thesis. The emerging theme of geovisual analytics is then considered before focusing on human computer interaction. Stemming from this, a discussion on human cognition and perception follows before the theme of visual scalability is introduced. Each of the above sections are critical in framing the research carried out in Experiment Two and will aid analysis of the research output later.

2.1 The Structure of the Eye

Yarbus (1967), a pioneer of eye movement research, provides an account of why it is important to understand how the structure of the human eye (see Figure 2.1) relates to its function and movement. The shape of the eye is maintained by the firm tissue called Sclera and the Cornea is a transparent membrane that protects the inner structure of the eye. The Iris changes the size of the pupil determining the amount of light transmitted into the eye and reflected onto the rods and cones – the receptors – of the retina. Where the axons of the neurons gather to form the optic nerve which projects to the brain, there are no receptors, this part of the retina is therefore a "blind spot". Another area of the retina – the fovea centralis – has thinner retinal layers and greater cone density. This is the region of greatest acuity and therefore eye movements align the eyes so that light from areas of interest are focussed on the fovea. This is called 'foveation'. Although attention can be directed to other parts of the visual field, generally one foveates attended locations. For this reason, eye movements are a proxy measure of covert attention. For some time, eye movements have been researched in an attempt to better understand user interest and interpretation of visual phenomenon. The next section discusses this research in-depth.

Figure 2.1 of detailing parts of the eye



2.2 Origins of Eye Movement Research

Human eyes make dozens of movements – saccades – every second. Analysis of eye movements therefore provides valuable insights into psychological and cognitive function in a number of real-world tasks, including reading and visual exploration of computer displays (Goldberg et al., 2002). Eye movements are generally studied through acquisition and analysis of eye movement trajectories via eye tracking.

A saccadic eye movement involves the synchronised movement of both eyes at a rapid speed from one position to another. Up to 173,000 saccades are made each day (Abrams et al., 1989). The exact duration of a saccade will vary according to the visual input, although the speed of a saccade can be as high as 500 degrees per second (Rayner, 1998).

Any object or scene the eyes pass over during a saccade may not be properly visualised due to saccadic suppression (Bridgeman et al., 1975). This implies that information is derived from the visual scene during fixations. We study eye movements on the assumption that fixations are foveating points of interest.

There are a variety of techniques for measuring eye movements, which vary according to the extent to which they are invasive and in their accuracy. Unfortunately, these factors are inversely correlated. The most basic method of eye movement measurement is direct observation. We are reasonably good at being able to tell, broadly speaking, what another human is looking at. However, although this is the least invasive method, it is also the least accurate. Trajectory and velocity of saccades are largely impossible to track by observation, and this method would only be somewhat useful in recording the final fixation point of the eyes. Our capability to accurately record eye movement has improved considerably. Wade and Tatler (2005) provide a summary of eye tracking technologies and note that remote devices – not head mounted, but rather at a comfortable and not visually interfering distance from the participant – provide sufficient accuracy to record patterns of fixations while minimising interference with vision. This is the type of eye tracker used in the second experiment.

Çöltekin et al. (2009) studied map designs using eye movement analysis as well as usability metrics (e.g., speed and accuracy of comprehension). Çöltekin et al. (2009) state that ‘softcopy’ (digitally-based) maps are becoming the norm with the spread of internet access and mobile devices. The need for usable interface tools for digital cartographic mapping is therefore an important factor to consider for the future of maps. Basic map design principles, and to an extent the usability principles applied in the work of Çöltekin et al. (2009), are important to consider because maps are one of the oldest examples of a visual display of complex information. The evaluation carried out by Çöltekin et al. (2009) used two different interactive maps services, Natlas (2008) and Carto.net (2008), which have different layouts, in order to compare their usability. Çöltekin et al. (2009) state that it is a common assumption that more fixations “indicate a less efficient search strategy, longer fixations indicate difficulty with the display, and plotting scan paths and fixations will allow documenting” (Çöltekin et al., 2009). The authors provide an account of a user’s typical visual search pattern where two processes occur. The first is perceptual, during which the user locates the target area of interest, and the second is cognitive, where the user will processes the information at the target location. The same principle applies for any visual processing of a complex graphic.

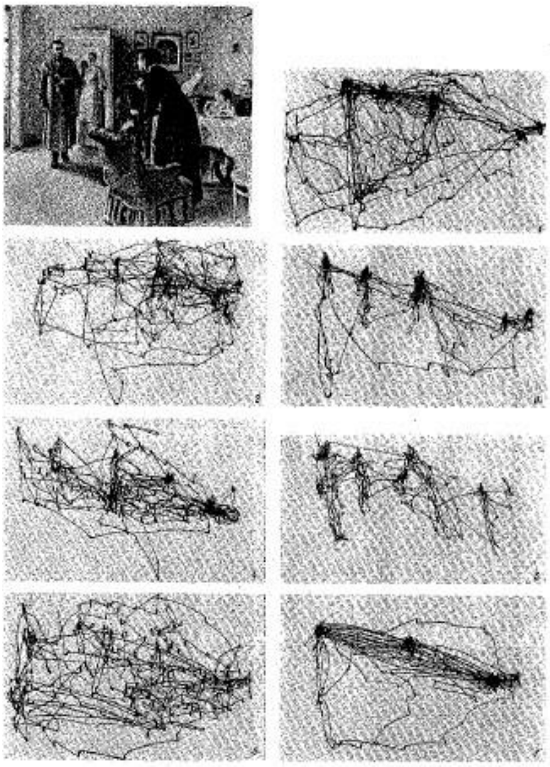
The visual processing of complex objects or a scene with many items of potential interest will not generate the same response from every person. People’s eyes will fixate on certain elements of the objects, and typically complex objects or scenes will contain many elements.

An early study by Yarbus (1967) on the 1884 art work by Ilya Repin called 'An Unexpected Visitor' shows that participants focused more on elements of the painting than they thought; as Yarbus put it, 'useful and essential for perception'. Figure 2.2 shows the original work, while Figure 2.3 shows an example of participants recorded eye movement patterns over the course of a three minute period. Through a series of experiments such as this one Yarbus found participant eye movements were not necessarily determined by the number of details. This is interesting to note because it is linked to more modern research, where the number of elements presented to participants vary greatly. Yarbus also noted that light and dark details did not attract participant attention unless essential information lay within those details. In relation to the experiment conducted for this thesis, the lightest and darkest areas would be the ones containing the most information and therefore participants' attention and foveation would be attracted to these areas of the visualisations.

Figure 2.2 Repin's 'An Unexpected Visitor', 1884

Yarbus (1967) concluded in his ground-breaking work that eye movement is at most only slightly dependent on the contents of a scene being observed by the eyes. However, in his experiment certain details were not highlighted as being particularly important. This experiment was exploratory in nature because there was no task requiring an answer, participants did not have to search for key elements in the scenes presented to them for three minutes at a time.

Figure 2.3 Illustration of participant eye movement, Yarbus 1967



A large component of this research is linked to the perception of a human being.

2.3 Modern Eye Movement Science

Eye movement science is used to augment research on human behaviour. Eye tracking technology is more sophisticated and accessible than ever before resulting in more precise readings during experiments. Approximately 40% of our brain is used to process the visual environment (Ware, 2008). Therefore it would be advantageous to attempt to measure eye movement to evaluate, understand, and improve. Ware (2008) focuses on the idea of visual thinking and, how the human cognitive pattern of thinking and eye movement are linked in 'acts of attention'. Attracting and keeping the attention of a person is better achieved through graphical representations; "graphics can be more precise and revealing than conventional computations" (Tufte, 2007). A valid assumption is that graphical representations are more efficient than non-graphical visualisations (Çöltekin et al., 2010) to display information.

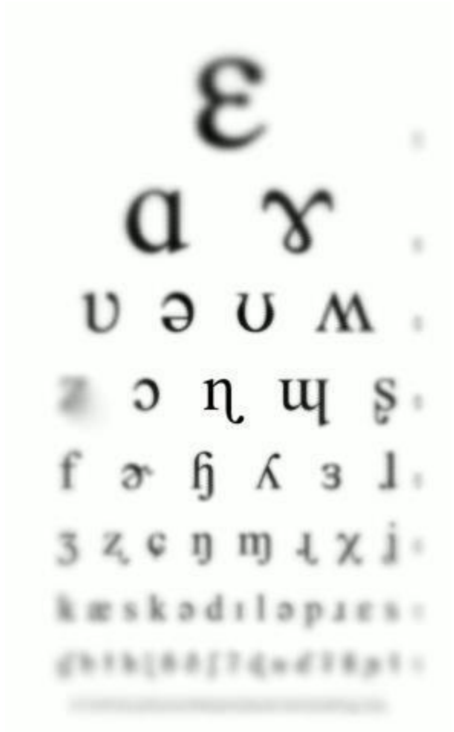
The comprehension of cognitive perception is a key point of eye movement research and is consequently strongly relevant to the research presented in this thesis. There are two types of perceptual process to note: Bottom-Up and Top-Down (Ware, 2008). The bottom-up process consists of information filtering so that meaningful objects can be visualised in the environment. Generally the number of objects that a human can focus on at one time is three. This is our visual working memory and in the design of the eye tracking experiment in this research it was important to carefully construct the visualisations because they contain unimportant objects that participants would potentially focus on. The top-down approach is also present somewhat in the bottom-up approach to a certain extent because the eyes will work to clearly establish what the objects of interest are in the particular environment. However, instead of using a series of filters to better visualise what would be a set of objects present in front of us, the top-down approach is related to instances where we have a specific task to complete (Ware, 2008). When attempting to complete a task we will search for the relevant objects in our environment.

What is important in the study of perception through eye movement are the sequences or potential patterns our eyes perform. Abbott (1995) mentions perception as the first major cognitive sequence topic. Ware's "*Visual Thinking for Design*" (2008) cites an experiment by Hayhoe and Ballard (2005) on the sequencing of eye movement who also cite the work of Yarbus (1967) on the ability of eye movement to offer an insight into cognitive thought patterns. Hayhoe and Ballard (2005) mention the evidence of sequential eye movement patterns, which again can be attributed to a pattern associated with possible task completion. A more recent example of eye move sequential studies is Çöltekin et al. (2010) where eye movement pattern recognition is explored through interactive and dynamic graphical representations. The graphical representations, or visualisations in this research are geographical in nature and the work by Çöltekin et al. (2010) is one of the few studies relating to cartography and geovisualisation. Swienty and Reichenbacher (2008) is another example of investigating the complexity of geovisualisations in relation to cognitive processing and visual scanning. Çöltekin at al. (2010) made a distinction between Bottom Up and Top Down processes that is different from the Swienty and Reichenbacher (2008) study. A top down process relates to theory or hypothetically driven patterns of scanning of a visual scene, while bottom up processes are data driven. These processes can be identified by observing scanning patterns of participant eye movements. More on the approaches used in the eye movement research of this thesis is discussed further in Methodology Chapter 5, but it is important to mention its place among eye movement literature here.

Perception is linked to cognitive aspects of our brain. According to Lai et al. (2013) the visual perception of a person is achieved through all three sources of information; foveal, parafoveal and peripheral vision. Since these aspects help form our perception of the environment it is important to understand their relevance. As mentioned earlier, Landolt (1879) carried out very early experiments on visual acuity in which he suggested our visual ability deteriorates as it gets further away from the central point of focus. The fovea is the area with greatest clarity or visual acuity in the eye, followed by the parafovea and then the peripheral region. Our fixations and saccades are made through the fovea, which is why the clearest images in our environment are always those we focus upon. Work carried out by Duchowski and Çöltekin (2007) show examples of the foveated vision through Gaze-Contingent Displays (GCDs). One example used can be seen in Figure 2.4 where the foveated vision is simulated on a scene from the Movie 'Gladiator' which was cited by Duchowski and Çöltekin (2007) and was originally produced in Geisler and Perry (2002). The original image can be seen in the bottom left, and the visual field on the bottom right. Figure 2.5 effectively provides an example of what your eyes would see when they focus on an element in a scene. Figure 2.5 was created using Giesler and Perry's Space Variant Imaging software from the Center for Perceptual Systems (2014) at the University of Texas. It is a foveated screen capture of a phonetic alphabet often found on the wall of an optometrist, the focus of the eye is simulated on the 'n' at the centre of the image. Although these displays are simulated artificially they still offer a good insight into the visual acuity of our eyes.

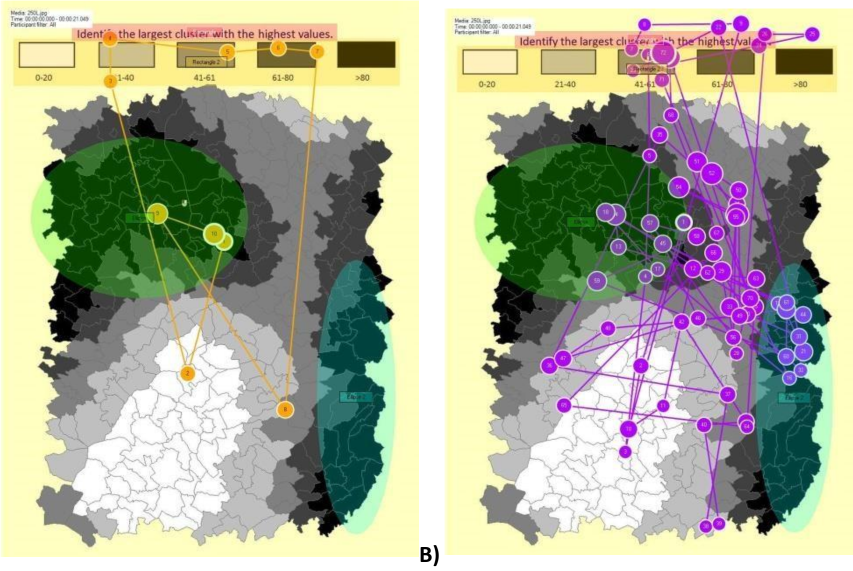
Figure 2.4 adapted by Duchowski and Çöltekin (2007) from Giesler and Perry (2002). The original image is ©DreamWorks SKG and Universal Studios.

Figure 2.5 An example of our area of vision, produced using Visual Field Simulator 1.7 created by Giesler and Perry (2002)



It has already been mentioned that fixations and saccades can occur at areas of interest (AOIs) (Çöltekin et al., 2009). When analysing fixations the researcher will check for two things. The first is the number of fixations a participant made. Figure 2.6 shows two examples with different numbers of fixations completing the same task with the type and scale of the visualisation the same for both. Referring back to the AOIs you can see the circled areas that would typically be areas of primary interest when analysing eye movement.

Figure 2.6 An example of eye fixation count variety A) fewer fixations, B) more fixations



B)

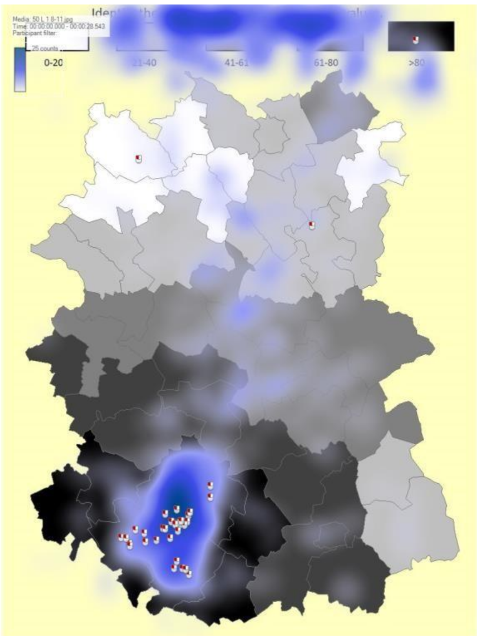
Table 2.1 provides a summary of the measurements that can be obtained from eye movement fixations and shows an example of these fixation measurements. Some have already been discussed as these are the measurements that were primarily focused on during analysis of the second experiment data. The spatial density or coverage of eye movement on a visualisation is another useful method that will now be discussed.

Table 2.1 Summary of eye movement measurements.

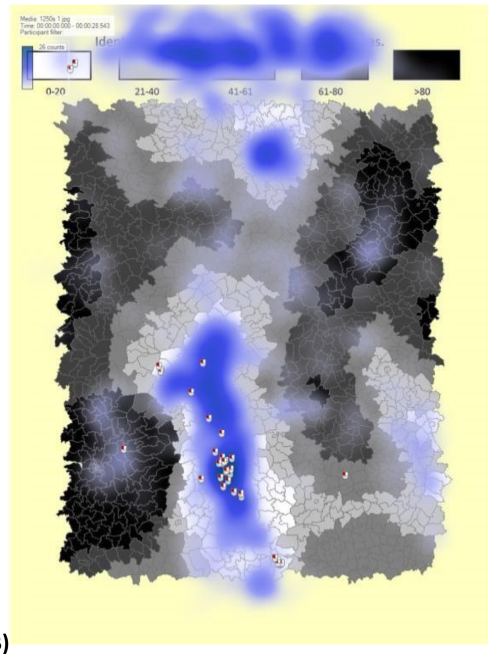
Type Description	Measurement
No. of fixations	Number required by participant to process the information
Location of fixations	Where the spatial focus occurs
Fixation duration	Time taken to interpret a visual representation
Total fixation time	Total time on an area of interest
Scanpath length	Shorter paths indicate more organised information
Scanpath duration	Indication of time taken to process complexity
Spatial density	Coverage of eye attention of the representation

It is possible to display the coverage of eye movement on a visualisation in the form of a spatial density map or 'heat map. Figure 2.7 shows two examples of these heat maps. A heat map is a reflection on the frequency an area on the visualisation is attended by a participant. In other words, the amount of time a participant focuses on a particular area of the map is indicated by the intensity of the heat colour. In the case of the second experiment the heat colour is blue with darker blue signifying a greater amount of attention and lighter blue indicating less attention given. In instances where the map is entirely visible with no heat map colour you can assume that no attention was paid to that area by the participant during the experiment.

Figure 2.7 Heat Map of all participant eye movement for Clustered 50 spatial unit (A) and 250 spatial unit (B) visualisations from Experiment Two



A)



B)

Once eye movement data is collected, it will need to be analysed to derive meaningful insight. In addition to fixations, the saccadic trajectory is also important to include in analysis of eye movements. The meaning of a saccade has already been mentioned above, and generally eye movement trajectory interpretation comes from ideas on 'scanpath' theory. Norton and Stark (1971) were the first to propose scanpath theory: the theory is based on the sequence of fixations, labelled as "scanpaths", which reoccur. Stark (1994) found experimental participants would repeat the same pattern of eye movements when first presented with a visualisation on a screen followed by a blank screen. In another piece of research Ellis and

Smith (1985) had suggested scanpaths are completely random but this was not tested at the time. According to Pieters et al. (1999), scanpath theory “predicts that a subject scans a new stimulus during the first exposure and stores the sequence of fixations in memory as a spatial model, so that a scanpath is established.” If this is true, then a type of learning effect could be observed if participants were faced with the same visualisation with the same data distribution. This aspect is important to note because experiment design best practise seeks to minimise participant learning effects to avoid skewing of the results. Velocity and distance of scan paths can also be useful to analyse, however, for the purposes of this research it is not essential to facilitate analysis of scanpaths.

Related works discussed are associated with a graphical environment and although text based eye movement occurs during the eye tracking experiment it forms only a small part and it is not the primary concentration of this research. Literature based on saccadic eye movements over text has already been mentioned under the description of saccadic eye movement. The notion of eye movement perception, effects of lighting, speed of movement of the environment, potential cognitive patterns and sequences have also been addressed. As Çöltekin et al. (2010) suggests there is a need for work on graphical representations, or visualisations, with eye movement to be carried out. It is important to understand concepts on graphical representation.

Additional concepts to consider when exploring graphical representation include Human computer interaction (HCI), human cognition – within a visual context – and human perception. An extension of the latter includes our ability to perceive data, and how it can be scaled. However, before discussing these, the emergent theme of geovisual analytics is explored given its relevance to this research thesis.

We use more than 40% of our brain to provide visual output (Hoffman, 2000; Ware, 2008). Such a considerable percentage explains why so much time is spent understanding and improving visual representations. Cognitive Fit Theory (CFT) (Vessy, 1991) is one example of research into how technology is suited to a task. CFT is related to this research because we are assessing the fit of visualisation technology to a set of tasks. Dennis and Carte (1998) found maps faster and more efficient for task completion than tabular representations for multi-criteria decisions. The research of this thesis originates from a desire to test graphical visualisations over tabular and to compare those to participant task performances using advanced graphical representations. Part of the difficulty in understanding and improving visualisations exist because human cognitive processes are not easily classifiable. Humans are not as homogenous as visual analytics tool developers think (Slocum et al., 2001). Individual and group differences exist on the basis of parameters such as expertise, culture, sex, age, sensory differences, ethnicity and socioeconomic status.

Slocum et al. (2001) suggest when differences in visual analytic strategies are identified then two approaches are possible to address the question ‘what to do with them’; Integrate the insights from findings in educations (training) of potential user, or, modify the design to meet the needs of the user. In the case of this research we integrate the findings in education of potential users through discussion in this thesis and from that suggestions to improve visualisation designs can be made. Future direction could be to study the task and stimuli-based clustering more in depth and then to modify the design of the interfaces for further comparative testing. It is also important to systematically compare and contrast methods, thresholds and tools that are used in sequence analysis of eve movement recordings to establish benchmarks and guidelines for this kind of empirical work. The extent to which a graphic can be complicated can be referred to the level of detail (LOD). LOD management requires decisions on objects location, geometrical properties, and viewer’s location and visual capabilities (23). We need balance between perceptual (visual fidelity, visual quality, visual clutter) and technical (storage, bandwidth, processing power) fidelity. This is linked to visual scalability and human perception with regards to how scalable visualisations are. Visualisations used in this research are cartographic in nature and predominantly relate to GIS based systems. Compression techniques of LOD already implemented in GIS but don’t consider human visual system (HVS) properties and don’t address limits and strengths of human spatial perception (2). HVS doesn’t process visual information uniformly so a uniform resolution is pointless (27). Referring to Slocum et al. (2001) this is another reason why understanding and implementing improvements for visualisations can be difficult.

We can visualise non-uniform information through Gaze contingent displays (GCDs) which is a type of Variable Resolution Displays (VRDs) used in eye tracking research to identify Points of Interested (POIs). These may be more effective for the display of information. These types of displays are more capable of supporting what is termed perceptually lossless degradation of an image resolution in computer vision literature. In eye movement literature this can be called foveation. As the computer vision field understand it, the brain processes the functional field of view, and not just where the eyes are looking. GCDs are similar to multiple adjacent displays not unlike those found in interactive visualisations such as ProVis used for this research. A step further for GCD displays include image fusion based displays. Medical imaging, remote sensing and geovisual applications use multi-resolution image sampling from hierarchical data structures in GCDs. Fusion image output consists of some input images from the same region of interest but the available input may have different data structures such as raster or vector. They could also be produced by different sensors to give you satellite imagery at different spatial intervals. In short, if fusion image based GCDs are designed adequately then an observer will not notice a transition from one structure and/or resolution to another. Fusion based imaging is related to research carried out on map change detection but it is not studied in this research. The GCD alternatives are worth considering for research now that their cost has been reduced and reliability has been improved. Duchowski (2004) reported some early implementations suffered from technical difficulties and high equipment costs.

2.4 Human Computer Interaction

Surveys, usability evaluations and interviews represent some of the more formative methods of Human-Computer Interaction. Increasingly, these methods are being applied online and can yield similar results (Brush et al., 2004). In the early human computer interaction, many HCI experiment participants were professional users from a corporate environment, this group would have been financially motivated to improve system processes (Lazar et al., 2010: 368). Importantly, systems and products are now more widely tested.

The usability of GIS products has improved significantly in recent years (Hacklay and Tobón, 2003). GIS products still require considerable technical knowledge to operate them. Human Computer Interaction (HCI) work has been carried out with GIS applications since the 1990's (Nyerges at al., 1995a and Davies and Medyckyj-Scott, 1996). The development of more usable online thematic maps, geovisualisation and interactive visualisations is evident.

Eastman (1985a) working on a more conceptual level, provides a clear sketch of the information processing perspective and its potential applicability to user research in cartography. He also deals with the issue of linking a semiotic approach to map symbolisation with an information processing view of map reading. In this presentation of system and process models for map reading, Eastman is careful to point out the fundamental difference between an information theory approach to map communication that measures bits of information transmitted (with the goal being communication of the most bits) and attention to information processing (with the goal being to understand how people actively see and conceptualise about map information. For cartographers to facilitate visual and cognitive information processing, we must understand both the system that does the processing and the processes themselves.

2.5 Evaluation and Eye Tracking for Geovisualisation

Standard metrics which measure the effectiveness of visualisation methods (including geovisualisation methods) are useful for evaluating their viability in terms of displaying information. However, these can be augmented by more modern metrics, such as eye movement. Eye tracking evaluations are capable of offering previously unattainable information on the effectiveness of geovisualisations for displaying results. This is particularly useful given the fact on how little work has been carried out to assess user perceptions of complexity (Schnur et al., 2010).

Evaluations have been conducted to assess the most effective data visualisation methods (Slocum et al., 2001; Robinson et al., 2005) with modern evaluations assessing interactive visualisations (Lobo et al., 2015). It is important to continue to improve on usability as the rate at which we can collect and gather increases. Virrantaus et al. (2009) suggest there is a need to move towards an understanding of the role of the user in the process. This furthers the case for a user centred design approach to map design and geovisualisation evaluations. Eye tracking recordings will augment these evaluations.

Evaluations such as those completed by Lobo et al. (2015) help to inform the design of geovisualisations. Human Computer Interaction researchers focus on; how users interact with maps, the techniques used to visualise content and to navigate in terms of space and scale (Lobo et al., 2015). Modern methodologies for evaluating the effectiveness of interactive visualisations can include metrics taken from eye movements which have been discussed earlier (Çöltekin et al., 2009). These further the ability to create the most effective visualisations. Reliably designing a highly usable interactive visualisation has proven to be difficult as explained by Robinson et al. (2005). In many instances an element of context provided through user input is helpful at every stage of the design.

The usefulness of evaluations containing eye movement records is evident from work carried out by Harrie and Stigmar (2009). Work by Çöltekin et al. (2010) demonstrates evidence of how humans work with interactive geovisualisations, through monitoring of their search pattern techniques. This systematic approach highlights the potential for improving interactive geovisualisations. An evaluation of more than one geovisualisation using eye tracking technology is effective for the design decision making process. Çöltekin et al. (2009) performed a comparative test of two different geovisualisation interactive designs, using eye-movement recording to inform how each design performed.

Advanced forms of geovisualisations such as the space-time cube have also been evaluated using eye-tracking technology. Li et al. (2010) combined more commonly used approaches such as interviews and think-aloud protocols with eye tracking. The authors recognise that current eye-tracking procedures and analysis are still relatively cumbersome and the overwhelming level of data gathered on eye movements can result in over-plotting. This problem can be mitigated by the scale of the geovisualisation being evaluated but despite this issue the benefits of incorporating eye tracking technology warrant its' inclusion where possible.

More advanced geovisualisations are interactive as described above, this means they contain some form of map animation, i.e. the geovisualisation changes as it is manipulated by the user. This change can be categorised into a set of visual variables (Slocum et al., 2009), they are: duration, rate of change, order, display date, frequency and synchronisation. Robinson and Griffin conducted research which examined the effect of leader lines with a geovisualisation. Several of these map animation variables were utilised for this evaluation. 3D geovisualisations including Space-Time Cubes and spatio-temporal visualisation contain many of these visual variables. The point to note here is, eye movement analysis can be an important asset to geovisualisation evaluations and it is not limited to the analysis of less animated geovisualisations. That is not to say that analysis would be considerably more difficult on an animated 3D Space-Time Cube displaying the famous visualisation of, for example, the long march of Napoleon's army to Russia and back compared to a 2D geovisualisation version. Returning to the reason why eye tracking is used in this research, it significantly augments the evaluation study carried out on geovisualisations in Experiment two. Particularly in relation to the four levels of details or “Information Intensity” of geovisualisations. As Slocum at al. (2009) ask, how much detail is appropriate?

2.6 Human Cognition

According to MacEachern (2004: 23) “Human vision and visual cognition remain incompletely understood”. The dominant consensus on vision is that we use it as a system to process information.

> “*Visual cognition encompasses issues of how cognitive processes interact with vision to enable us to interpret the world and our apparent ability to mentally manipulate visual information in the form of images*” (MacEachren, 2004: 33).

Visual cognition work carried out by Marr (1982) describes a framework for vision that converts and transforms simplistic representations into complicated scenes. Figure 2.8 shows an adaptation of this framework. The primal or basic sketch was the construction of basic objects, lines or boundaries. The 2.5 D sketch establishes surfaces according to visual acuity. The 3D section provides detail of that surface and objects can become three dimensional. Pinker (1990) extended upon Marr’s framework for graphical comprehension. MacEachren (2004) states that Pinker’s framework simultaneously addresses comprehension issues and we can cite this work on human cognition as an important part of visual design. Figure 2.9 shows an adaptation of what Pinker’s framework would look like when visually presented, as you can see it is similar to Marr’s. Pinker suggests users create physical dimensions of graphical marks, or a visual description and users map physical dimensions onto mathematical scales. As mentioned earlier user interpretation of a visualisation will vary and Pinker proposes four factors to determine the most likely visual description of a representation. MacEachren (2004: 35) summarises these factors and he said the first was borrowed from the work of Kubovy (1981). Certain visual variables are indispensable, these are; space and time. Both attributes are perceptually dominant. MacEachren mentions that time is important in non-static graphs and since interactive visualisations are becoming the mainstream technique used to visualise data the attribute ‘time’ is increasingly apparent. The attribute space can be translated into ‘spatial location’ particularly in a geographic context and geographic data analysis. Users visualise a scene with the assistance of spatial locations within that scene.

MacEachren goes on to say that our attention is more; 'selective', meaning we will focus on a location better than an object or shape. It also links to Bertin's work on graphics and perception (Bertin, 1967 and Bertin, 1983).

Figure 2.8 Marr's Stages of Vision Framework (Marr, 1982)

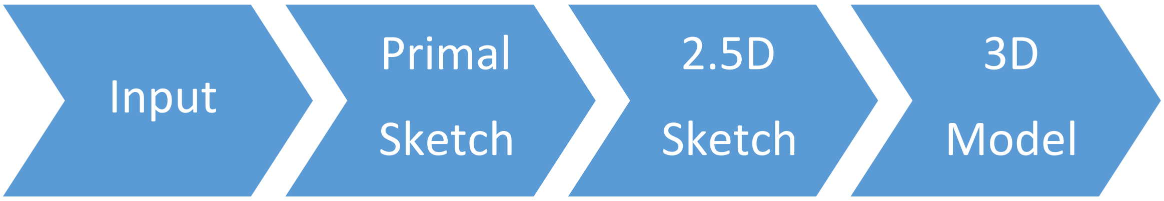
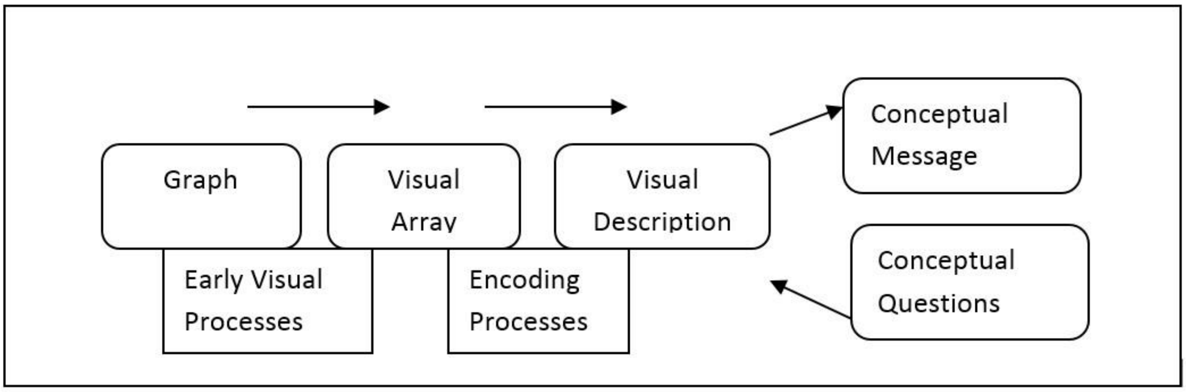


Figure 2.9 Pinker's Stages for Graphical Comprehension Framework (Pinker, 1990) Redrawn for this thesis



Differences and similarities found in fixation sequences may be parallel to cognitive differences and similarities of the viewers (Stark and Ellis, 1981; Brandt and Stark, 1997; West et al., 2006).

2.7 Human Perception

MacEachren (2004) discusses visual magnitude as a factor that affects the visual description of a scene, this would include various types of visualisations. Eastman (1985) and Pinkman (1990) suggest that maps are facilitators rather than communicators whilst Gibson (1979) believed that visual perception is a highly viable alternative to the information-processing perspective of cognition. Eastman (1985) and Pinkman (1990) believe a clear perceptual organisation of the display will be most effective for the display of information. This is linked to visualisation design, in that human perception is affected by the characteristics of a representation. MacEachren (2004) states; 'geographic visualisations require a higher level of interaction between the user and the display'. Users must analyse or search the displayed data and a different information processing model designed by MacEachren and Ganter (1990) is most effective for geographic visualisations. It is also important to consider perceptual limits of visual retention. For example, Ware (2008) states that visual working memory can hold three objects at any one time. In some cases objects with multiple features are treated by the mind as a single object (Luck and Vogel, 1997).

2.8 Visual Scalability

In the modern day it is common for data to be collected in large volumes. Previously issues of data scalability and terms such as 'big data' did not exist because there was no need to consider potential problems now associated with data. Visual scalability is linked to issues with data scalability (geographic scale, grain and extent) because data needs to be displayed effectively if the user is to successfully interpret it. This is particularly important if data is to be disseminated to a wide audience including the general public. It could be suggested that the time a websites presentation has to impress a person could also be applied to a visualisation of data. In order for data to be successfully publicised the general public will have to like what they see and find it usable almost instantly. Researchers considered experts in a field involving 'big-data' will afford more time to understanding this data when it is visualised.

Visual scalability is designed by two things according to Eick and Karr (2002). The first are responses or impacts of a visualisation. The second are factors or the characteristics of visualisations. A simple formula was devised for this definition that could be used to measure the visual scalability of a visualisation;

$$\text{Responses} = F\ (\text{factors, data}).$$

Eick and Karr (2002) critique this formula however by suggesting it isn't feasible due to a lack of measureable response data for visualisations. It is important to note that this research focuses on the performance of visualisation systems to display data in the first experiment and on the change of participant performance in the second. Visual scalability is a key focus for investigation and analysis in Experiment Two of this thesis.

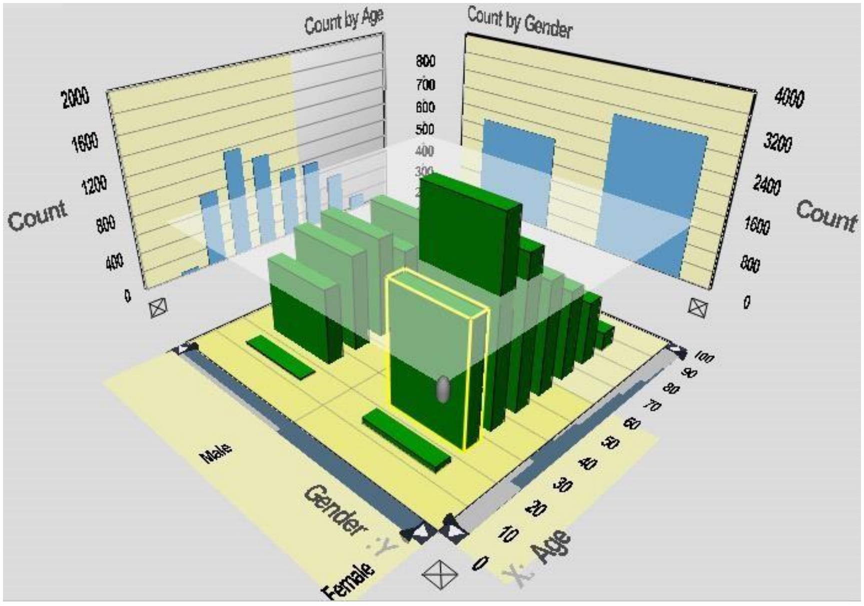
To measure visual scalability in a more practicable manner, Eick and Karr (2002) divide visual scalability into two classes; database metrics and visualisation characteristics. Database metrics measure dataset size in terms of rows and bytes while visualisation characteristics relate to elements and attributes displayed on screen. In this thesis it can be suggested that measuring visualisation characteristics is more relevant than database metrics. From the human perspective several principles must be taken into consideration, they are; human perception, monitor resolution, visual metaphors, interactivity, data structures and algorithms and computational infrastructure (Eick and Karr, 2002).

As previously mentioned, the human eye accounts for approximately 40% of brain activity and is one of the most powerful tools humans have to process the visual environment. This is linked to human perception and it is no surprise to see Eick and Karr include it as a factor to consider in visual scalability. Monitor resolution through physical size and pixel count affects the visual scalability of data, or as Yost et al. (2007) suggest, it affects visual acuity. The techniques used to display data, e.g. attribute charts and maps are visual metaphors that require thought. The scalability of a visualisation is greatly affected by its interactivity and therefore must be taken into account. Algorithms that allow different methods of data representation are also important and the power of a computer to process and display data cannot be underestimated. In the case of the second experiment of this thesis it would not be a problem. However, for the larger 3D dataset in experiment one, the computational power of the experiment host machine reached its limits on several occasions. The effect on overall experiment results would be negligible or minimal but it is a factor that would need to be assigned a greater weight of importance if larger datasets are used in the future.

The general principle remains the same. Visualisation plays an essential role in dealing with large data-sets. Eick and Karr in their 2002 work stated ‘scalability analyses of visualisations are almost entirely absent’ despite a need to scale effectively. Eick and Karr provide examples of bar charts and scatterplots which are two popular data display types. Bar-charts cannot cater for more than two to three dozen data attributes and scatterplots are overwhelmed with too much data.

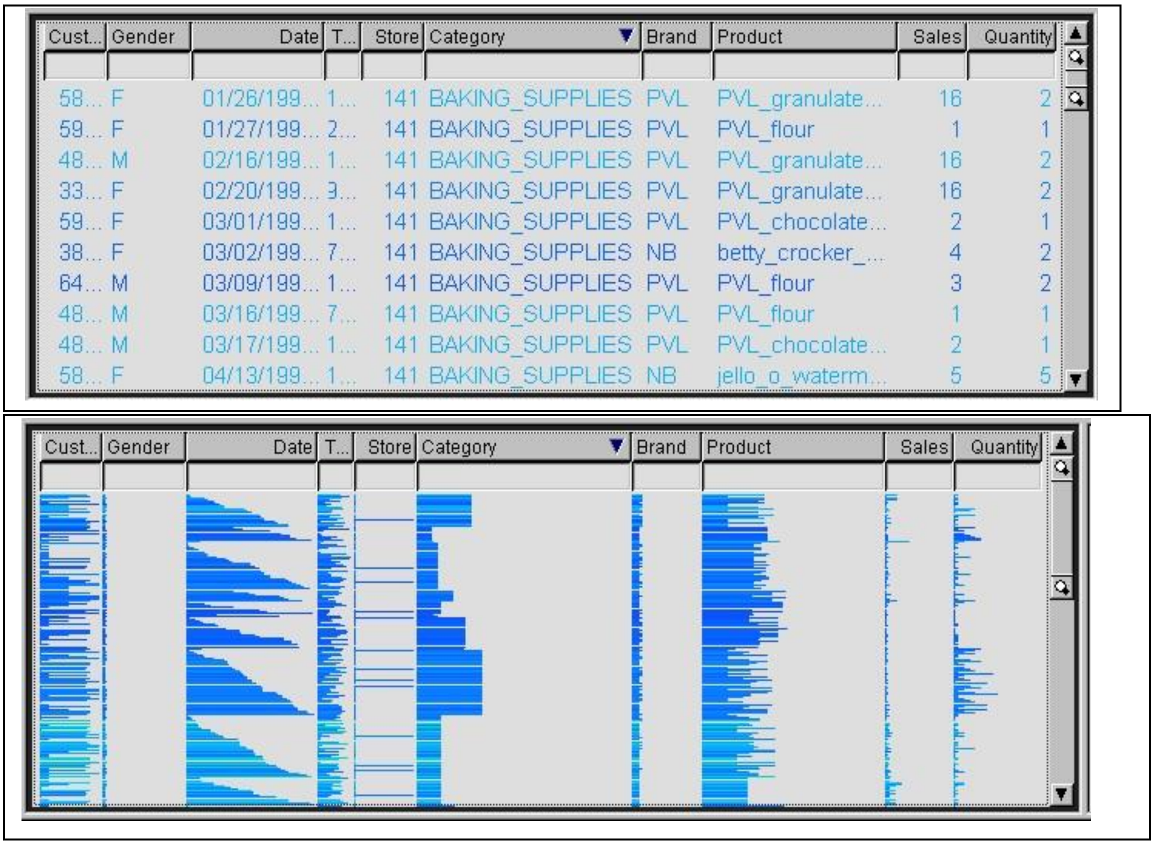
Increasing visual scalability is possible provided a correct set of measures is taken. Interactivity is regarded as the crucial aspect that would allow for improved visual scalability. Eick and Karr suggest that intelligent design of data display would be beneficial. For example, when analysing a bar-chart graph the detail presented on it will be grainy (less detail will be shown) as the user zooms out. Conversely more specific detail will be presented to the user when they zoom in. Computer algorithms could augment the implementation of this idea. They talk about using different representation techniques to improve visual scalability, one example being 3D representations. They offer an alternative 3D technique to help overcome occlusion and Figure 2.10 shows an example of this. It contains 3D scrollbars and a water-level plane across the central 3D model. Attributes displayed on the 3D model are separated into two vertical bar-charts on the back walls of the multi-scape view.

Figure 2.10 Example of a multiscape view by Eick and Karr (2000)



A secondary example provided is of scalable text and graphic data windows (see Figure 2.11). Animated movements of the level of detail presented on this type of representation increases visual scalability and it is another useful technique.

Figure 2.11 Data sheet and scalable text view



Cust...	Gender	Date	T...	Store	Category	Brand	Product	Sales	Quantity
58...	F	01/26/199...	1...	141	BAKING_SUPPLIES	PVL	PVL_granulate...	16	2
59...	F	01/27/199...	2...	141	BAKING_SUPPLIES	PVL	PVL_flour	1	1
48...	M	02/16/199...	1...	141	BAKING_SUPPLIES	PVL	PVL_granulate...	16	2
33...	F	02/20/199...	3...	141	BAKING_SUPPLIES	PVL	PVL_granulate...	16	2
59...	F	03/01/199...	1...	141	BAKING_SUPPLIES	PVL	PVL_chocolate...	2	1
38...	F	03/02/199...	7...	141	BAKING_SUPPLIES	NB	betty_crocker_...	4	2
64...	M	03/09/199...	1...	141	BAKING_SUPPLIES	PVL	PVL_flour	3	2
48...	M	03/16/199...	7...	141	BAKING_SUPPLIES	PVL	PVL_flour	1	1
48...	M	03/17/199...	1...	141	BAKING_SUPPLIES	PVL	PVL_chocolate...	2	1
58...	F	04/13/199...	1...	141	BAKING_SUPPLIES	NB	jello_o_waterm...	5	5

Eick and Karr (2002) emphasise the importance of interactivity for larger datasets. Zooming and panning, highlighting selections of data, dynamic window resolutions and intelligent labelling are key aspects to consider when designing a data scalable interactive visualisation. In the second experiment involving eye movement it was important to understand the underlying principles of visual scalability because an element of scalability was assessed. Eye movement data is best recorded when potential distractions or problems are minimised for participants. To explain it in some more detail, the experiment process had to be designed to be as simple as possible so it would be easier to measure, analyse and then extract results from the experiment instead of wondering what participants may have been thinking or looking at during the experiment. It would all be comparable. To have provided more than one 'window', with participants moving between, would have increased the difficulty of obtaining comparable measurements. This problem would increase with more windows, particularly if interactivity and multi-scape or multi-windowed visualisations were considered. As Eick and Karr (2002) stated it is not possible to definitively measure visual scalability but we can try. The addition of eye movement can help us measure one element of it which is perceptual scalability. We are not interested in the performance of the visualisations but how the participant performs with different levels of data complexity displayed in those visualisations. It is a fine line similar to the first experiment where the effectiveness of the visualisation were being tested and not the participant.

The 'footprint' is the number of pixels occupied by a visualisation on screen, i.e. how much space they take up. If a map footprint is large is means it occupies most of the available screen space and consequently contains more pixels. If a map footprint is smaller it will occupy less overall pixels. Theory on map footprint is linked to Visual Scalability where research was carried out to assess how many pixels a display needs to effectively display information. The first sentence of a paper published by Yost and North (2006) was "Larger, higher resolution displays can be used to increase the scalability of information visualizations". Their work concluded that the visualisations used in their experiment were perceptually scalable. They found participant performance did not decrease when the display used to present each visualisation to experiment participants changed in size.

Yost et al. (2007) carried out work which suggests an increase in performance efficiency is observed when larger and higher resolution displays are used to 'scale up' information beyond the limitations of what is known as 'visual acuity'. "Visual acuity is the sharpness of vision, measured by the ability to discern letters or numbers at a given distance according to a fixed standard" (Online Oxford Dictionary, 2014). Visual acuity is linked to visual scalability and perceptual scalability. Visual scalability research attempts to measure visualisation performance at varying pixel rates compared to perceptual scalability which attempts to measure human performance at varying visualised data complexity rates.

2.9 Glossary

Fixations: The action performed by the eye when it focused on visual stimuli. Associated with attention being given by a human to stimuli.
Foevation: A commonly used image and video compression technique. It is an imitation of the eyes fovea or band of vision in computer generated form (Çöltekin, 2006).
Geovisual analytics: Analytical reasoning facilitated by visualisations and the exploration of complex spatial datasets.

Human cognition: The thought process which takes place in the human brain. It includes problem solving, evaluation, reasoning and computation of stimuli.

Human Computer Interaction: The study of how people interact with computers.

Human Perception: The capacity for a human to sense or comprehend stimuli.

Perceptual Scalability: The degree to which humans can view and comprehend scale or complexity, where scale could be a number of data values. The human capacity to perceive a number of objects may be limited and exposure to too much stimuli could alter the capacity to comprehend. Visual acuity imposes a limit on perceptual scalability (Yost et al., 2007).

Saccadic eye movement: The rapid movement of the eye, it is a jerk and jump type of movement as the eye moves from one fixation to another.

Scanpath: An internal mechanism of the eye that produces eye movements where the structure of human visual perception is understood. Scanpaths are repetitive sequences of saccades and fixations by the eyes when they are exposed to visual stimulus.

Spatial perception: The human ability to sense the stimuli; e.g. size, movement, shape or orientation.

Visual acuity: A measure of your central vision where your ability to distinguish details and shapes of objects is clearest.

Visual scalability/spatial resolution: Defined as the Responses that measure the number of insights of a visualisation and the Factors (characteristics) that affect Responses (Eick and Karr, 2002). This is difficult to measure.

3: Methodology – Experiment One

Chapter 2 outlines the usefulness of GWR and how it is used to examine processes that vary over space (Fotheringham et al., 2002). The common types of visualisations used to display GWR have also been discussed. The purpose of Experiment One was to evaluate the effectiveness of graphical representations of GWR output data. The aim was to discover the most appropriate way to facilitate interpretation and analysis of GWR. It is not difficult to display GWR data because there is a wealth of visualisation choice for the researcher with the right knowledge to choose from. However it is not fully understood which visualisation method is most effective for analysing GWR outputs. It can be suggested that current choices in the field are largely down to personal researcher preference and knowledge. The research in Experiment One can serve as a possible guide for future GWR visualisations.

3.1 Participants

Participants (n=13) were recruited from the staff and student body at the Maynooth University based on participants having at least some knowledge of GWR.

MSc. students (n=2) were asked to volunteer having completed a set of GWR, GIS and visualisation modules during their study for a Masters in Geocomputation. This meant they had operational knowledge of the GIS systems used in the experiment, an understanding of GWR principles, GWR output and were familiar with an interactive visualisation system. These features made them suitable candidates to take part in this experiment.

A cohort of doctoral research students (n=9) had knowledge of GWR, GIS and/or interactive visualisations. Some participants were members of University staff (n=2). Similar to the PhD student cohort, members of staff had knowledge of visualisations, GWR, GIS and/or interactive visualisations.

The selected participants answered a background profile questionnaire to gauge the level of participant expertise and these were then divided into two groups; A and B. There were six levels of knowledge of visualisation systems (Group A, 1-4): and secondly GWR (Group B, 1-2). Group A1 had knowledge of the ArcMap (2D) visualisation system; Group A2 had knowledge of both ArcMap and ArcScene (3D) visualisations. Group A3 had knowledge of all three types of visualisations used in the experiment. Group A4 were familiar with ArcMap and an Interactive Visualisation system. Group B1 consisted of participants who said they had a more basic understanding of GWR while Group B2 consisted of those who said they had a good to expert understanding of GWR. Table 3.1 provides a summary of these groupings.

Table 3.1 Knowledge Groups of Participants according to their background information

Group	Subset	Knowledge Bracket	Number of Participants
	1	2D	5
A	2	2D & 3D	5
	3	3D, 3D & Interactive	2
	4	2D & Interactive Vis'	3
B	1	Poor and Basic GWR	9
	2	Good & Expert GWR	6

Participants required a particular knowledge base to enable them to participate successfully in Experiment One. Basic understanding of GWR was necessary so that participants would be able to comprehend the data being displayed. It was also necessary that participants had some working knowledge of GWR-based or visualisation-based data display systems. The experiment script provided a refresher for GWR and the visualisation systems used in the experiment but the refresher was limited to explanations of basic feature functionality. An ethics process also required completion prior to the commencement of the experiment. The next section discusses this process.

3.2 Ethics

Every experiment undertaken in a third level institution in Ireland must be first given permission by that institution’s ethical body. Two main types of ethical committees exist: one processes medically-based experiments and the other facilitates all other types of research. A specific set of forms had to be completed and a list of documents for use in the experiment had to be drawn up. These included a participant consent form and an experiment briefing sheet. The ethics committee also required a complete description of the experiment via a protocol form. Once the correct paperwork was gathered, completed and submitted, the experiment received approval and a letter was sent from the ethics committee as proof of permission (See Appendix 1).

The sections on the protocol form included: Name of researcher, Position and Department, Research Objectives, Methodology, Participants, Persons Under 18, Vulnerable Persons, Risks, Informed Consent, Follow-up, Confidentiality/Anonymity of Data, Ethics in subsequent outputs, Professional Codes of Ethics and included an instruction template to follow when creating the experiment briefing sheet and consent form. Full details of the ethics protocol form can be found in the Appendix. No substantial risks were identified in the experiment and the two most important elements of this section were to ensure participants would be as relaxed as possible to prevent any psychological stress during the experiment and to emphasise that all experiment profiles, content, discussion and output would be held in the strictest confidence and used solely for research purposes.

3.3 Pilot Experiments

Pilot experiments were carried out in order to gain feedback on the original design and adjust the experiment accordingly. They were used to improve experiment design and minimise problems that could be experienced during the course of the actual experiments.

From the pilot experiments carried out as part of this research thesis, it was clear the time taken to complete the initially designed tasks was too long in duration. The greatest concern here related to participant fatigue as stressed by authors such as Steinborn et al. (2009). Experiment participants on average can perform at an optimal and constant level for a set period of time, once that time is exceeded their ability to perform decreases. This aspect could skew results. It was estimated a minimum of 45 minutes and maximum total time of 90 minutes in a single session was acceptable based on the information needed as part of this experiment.

Initially the experiment was designed with five questions per visualisation with a total of 30 questions, but to reduce the time taken we cut the question number down to three per visualisation for part A and part B with a total of 18 questions (content discussed below). In the revised experiment, one univariate, one bivariate and one multivariate task were included. This meant typical GWR analysis could still be simulated which was important. The experiment script was also streamlined and discrepancies in visualisation details were corrected.

3.4 Experiment Set Up

It has been mentioned in Chapter 1 that the experiment environment should be as natural as possible. The experiment was carried out on the premises of the National Centre for Geocomputation (NCG) which had a private meeting room with sufficient space and lighting deemed suitable for the experiment conduction. A secondary suitable location was also used during times when the meeting room was unavailable. The meeting room consisted of a set of large tables that the experiment equipment could be placed upon. The secondary room layout was similar. A 24 inch flat-screen monitor was placed in front of the participant at a distance of approximately 18 inches. A mouse and keyboard along with the monitor were connected to an experiment host laptop where the researcher could follow the participant's experiment discretely. A comfortable chair was positioned directly in front of the monitor. This set up replicated a typical workplace station. An additional personal laptop was connected to the host laptop via a cable that allowed for observation of participant progress; this is elaborated upon in the next paragraph.

3.5 Materials

Data were recorded using several methods as summarised in Table 3.2. A major component was "Morae", an experiment management system designed by Techsmith. It is a reasonably affordable software package (costing approximately €1,200). It consists of three components; the Manager, Recorder and Observer. Morae Manager is used to analyse data gathered by Morae Recorder, which was installed on the experiment computer. Morae Observer allowed for the monitoring of participant progress in real time and logged anything of interest for later analysis.

Table 3.2 Experiment Recording Methods Used

Measurement Taken	Record Method
Task Time	Morae Usability Software
Mouse Movement	Morae Usability Software
Mouse Clicks	Morae Usability Software
Correctness Ratios	Written responses on script
Perception of Task Ease	Morae Usability Software
Perception of Task Speed	Morae Usability Software
Perception of Task Confidence	Morae Usability Software
Participant Activity A	Webcam - Morae Feature
Participant Activity B	Sony HD Camera
Part C Interviews	HTC Phone Voice Recorder

Every recorded experiment could be viewed in Morae Manager. Figure 3.1 shows an example of the 2D visualisation. A 3D visualisation example can be seen in Figure 3.2 and Figure 3.3 displays the Interactive visualisation type. The different visualisation styles are clear and the menu and logging systems are consistent in each. In the bottom right corner of each figure is the view recorded by an attached webcam.

Figure 3.1 Example of Morae Project: 2D Visualisation

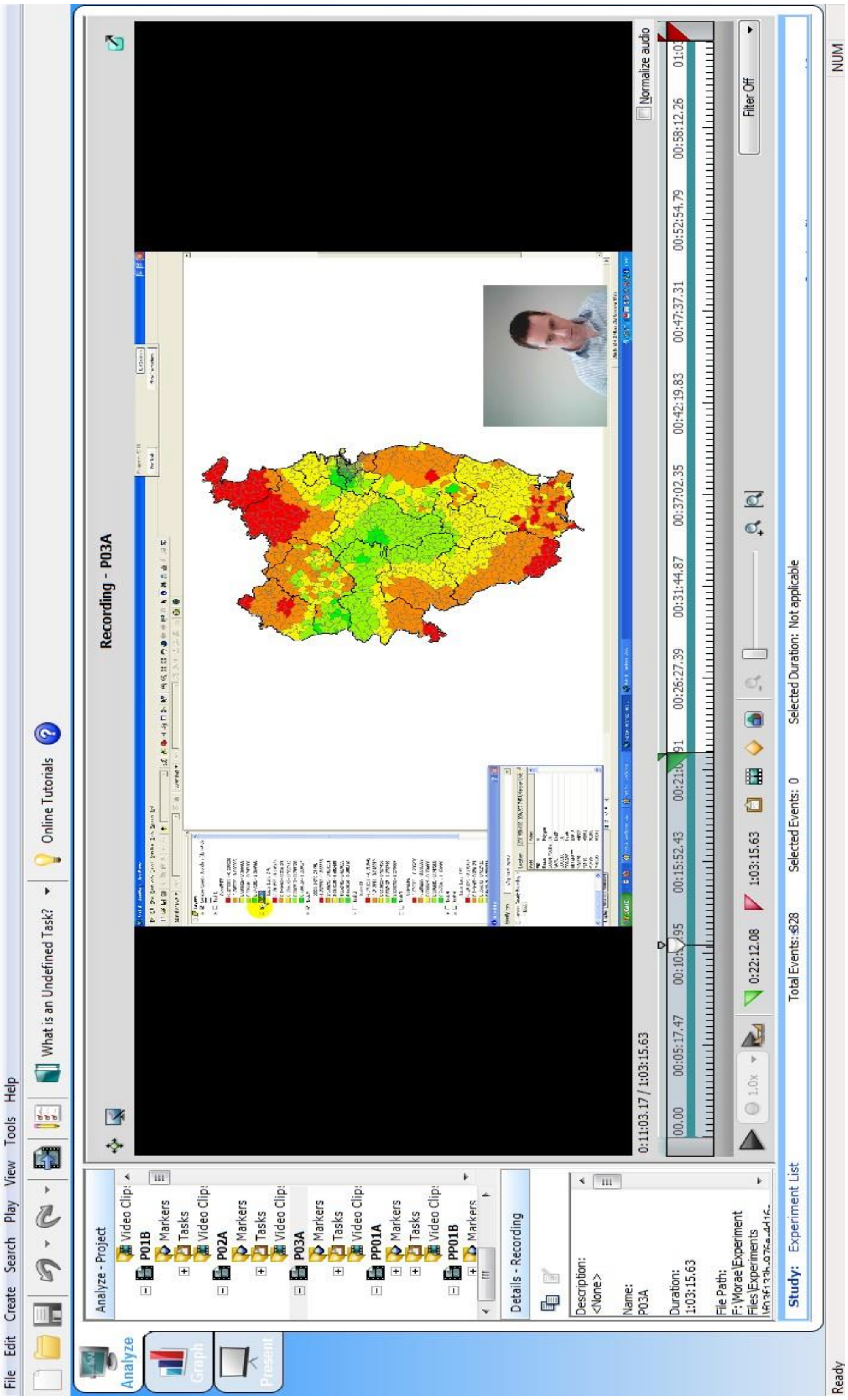


Figure 3.2 Example of Morae Project: 3D Visualisation

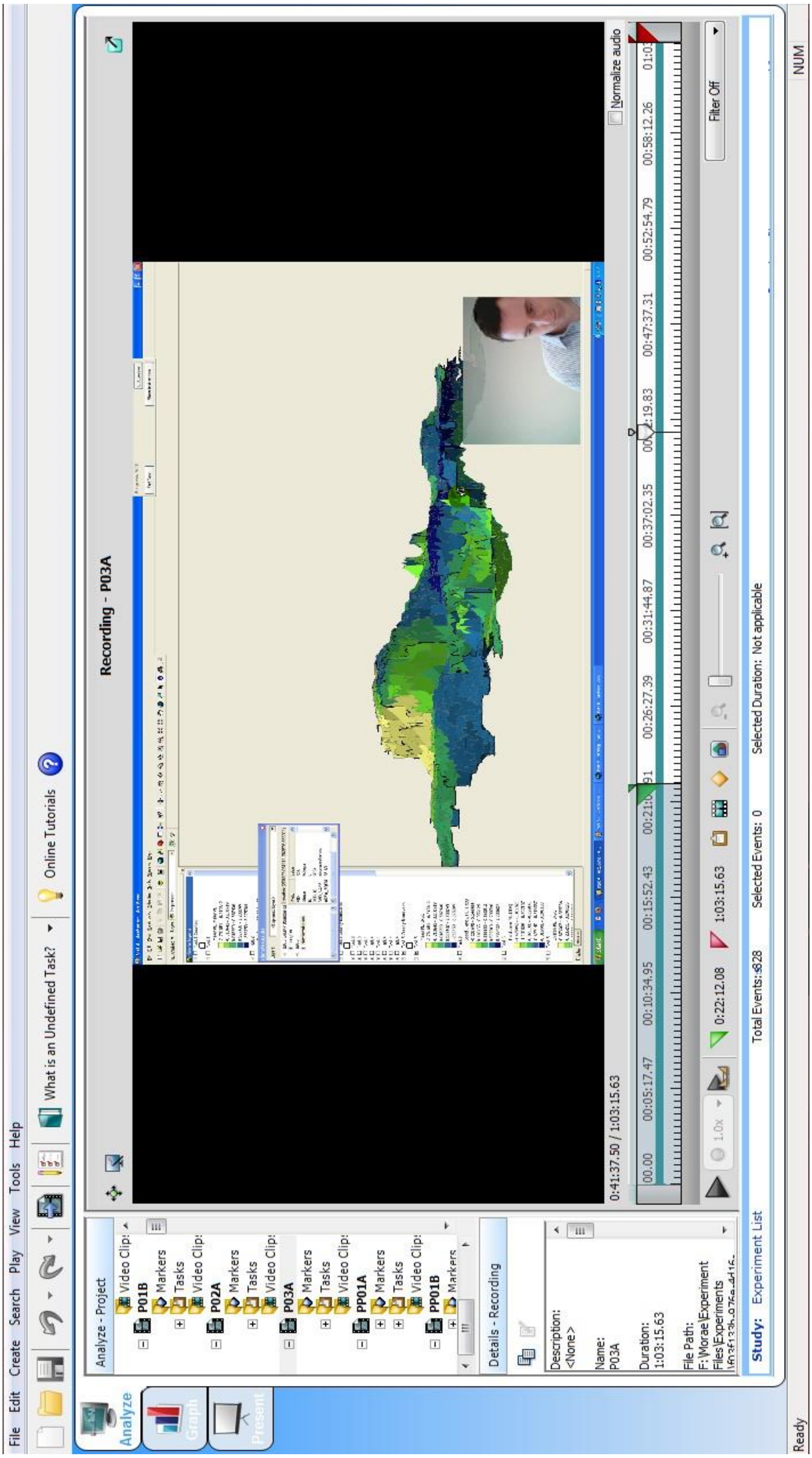
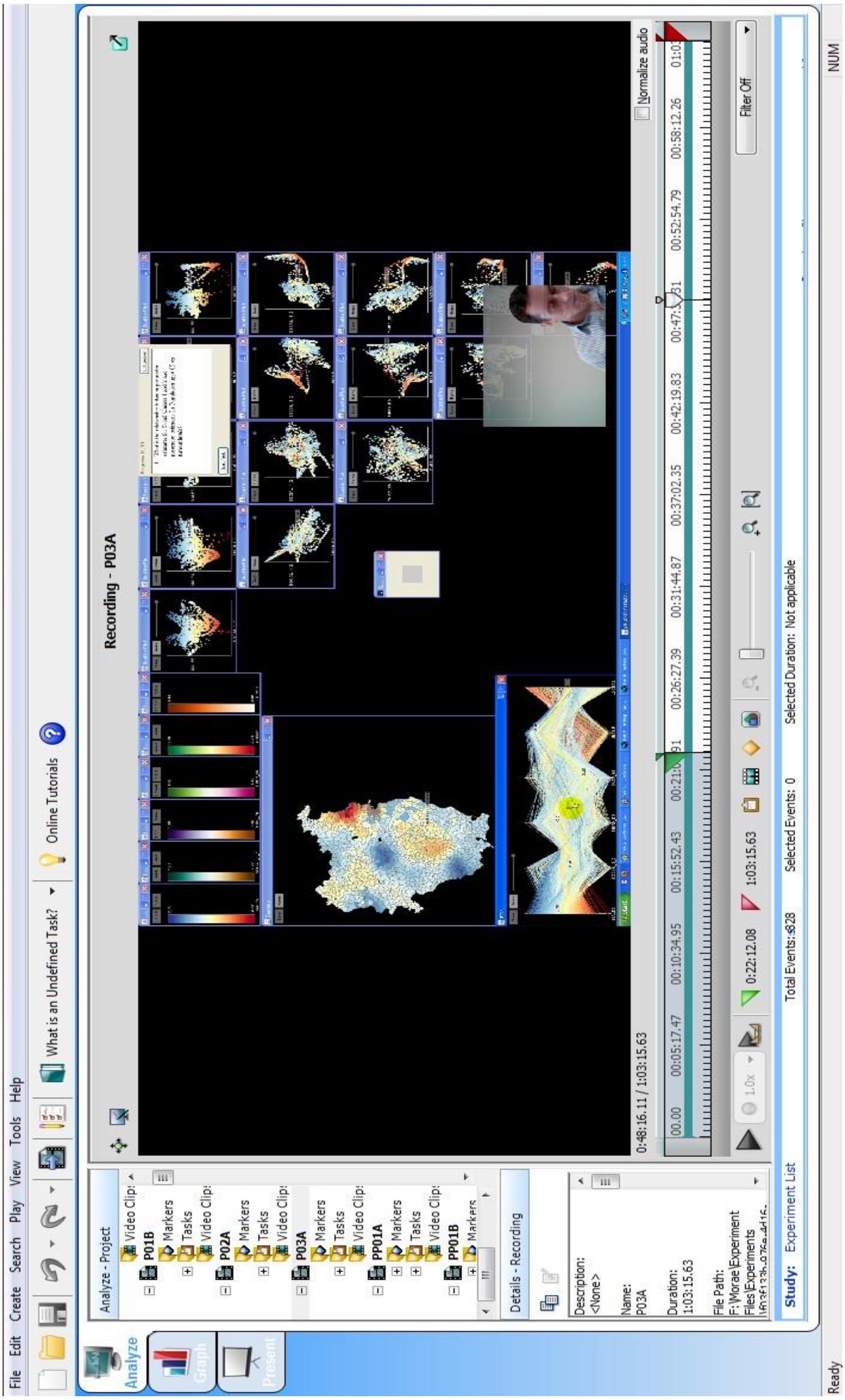
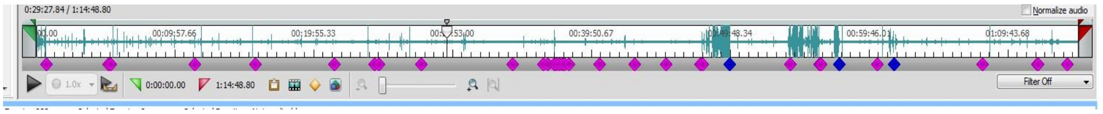


Figure 3.3 Example of Morae Project: Interactive Visualisation



At the bottom of the above figures is a timeline menu indicating the current time in the experiment (see Figure 3.4 for a magnified version of this timeline) along with any audio data picked up by the webcam and a series of logged events created during the experiment to assist in analysis. Different coloured loggings were made with pink being an event of interest for example and blue being a point where the participant required assistance from the experiment host. It is also possible to create clips of experiment sections which could be useful for live demonstrations but for the purposes of this research and thesis this was not necessary.

Figure 3.4 Morae Project Timeline Bar



Morae Recorder was installed on the host laptop in addition to; ArcMap; ArcScene and ProVis. Morae Recorder has the ability to set up a progress panel that can be overlaid on other software, in this case, the visualisation software. The progress panel contained start, end and continue buttons that helped instruct participants and it also measured participant task times. It provided the questions participants were required to answer followed by a short survey relating to the task they just completed. While post-task survey responses were logged on the host laptop, task answers were hand written in space provided on the experiment script.

Morae Observer allowed for visual on-screen monitoring of the experiment by the researcher. A set of logs could be created while monitoring participant activity to reduce data filtering later. It was also helpful to note participant expressions of frustration such as a 'sigh' or where their facial expression may have indicated high levels of concentration or a degree of annoyance. If a participant was experiencing an observable problem without asking for assistance, the experimenter volunteered to provide a solution so that the experiment would continue to run as smoothly as possible. In some instances participants were reminded to use the progress panel so that task time measurements could be recorded properly.
As mentioned previously, a Morae webcam module was used during the experiments to record participant activities. The webcam was a Microsoft Lifecam VX-3000. Clarity and function of the webcam were adequate for the purposes of the experiment. The use of a webcam was deemed useful if there was a need to determine reasons for a lack of on screen activity. The webcam would assist in the smooth running of the experiment as participant behaviour could

be monitored more carefully. In the event a participant exited the webcam view window, a second camera was used to record participants, from a different angle. The second camera (a Sony HDR-SR10) could still record participant behaviour with a built-in storage space of 30GB.

Both cameras were used to record participant activities. Facial expressions that could highlight participant confusion or frustration could be noted if anomalies in recorded Morae Task data occurred. It would also help to explain a sudden lack of onscreen activity caused by participant reading, writing or questioning of the experiment supervisor. The second camera would be able to record this type of activity even if the participant moved out of the webcam view. Any anticipated or unexpected activity by a participant could therefore be accounted for and would help reduce any possible errors in the analysis of the various experiment indicators (correctness ratios, time-taken, mouse clicks etc.).

Finally, post-experiment interviews were carried out with each participant and these were recorded using a digital recording device.

3.6 Stimuli

The experiment consisted of three parts; A, B and C. Participants worked with two separate data sets of GWR outputs in controlled parts A and B, which allowed the measurement of participant performance on each visualisation type, and each task type. Part A and B took approximately one hour each and part C, which consisted of a short, semi-structured interview, lasting no more than 25 minutes. Participants performance would be tested using two different data set sizes, with Part A's data set being approximately one third the size of part B's, as outlined in the following section.

To account for issues of visual scalability, as discussed by Yost et al. in 2007, two different data set sizes were used; one containing spatial units for the entire Republic of Ireland, the second smaller dataset being a subset covering the Leinster province of them. The larger Republic of Ireland dataset contains 3,412 Electoral Divisions (EDs), split over 26 counties. EDs are the smallest spatial units in the Republic of Ireland for 2007 census data. The Leinster Province (one of the four provinces in Ireland) dataset contained over 1,214 EDs split over 12 counties. Data were obtained from the Irish Central Statistics Office (CSO, 2011), Ordnance Survey

Ireland (OSI, 2011) and Kavanagh (2011). These were used as input in a GWR model to explain the voter turnout based on census variables.

Based on previous work by researchers such as Kavanagh and Fotheringham, it was decided a similar real life data set and a similar model would be used in Experiment One of this thesis. Five variables were chosen to form the model which was combined with 2002 voter turnout data. The GWR was then performed on these data using the following GWR procedure. A stepwise regression method, or Principle Components Analysis (PCA), was followed to construct an appropriate model for these input data similar to one used by Kavanagh et al. (2002). A PCA involves model building to include the set of attributes that would provide the most significant explanation or that most affect the independent variable. The variables were entered into (SPSS) and added into the final model according to their level of significance. Based on previous research, seven primary attributes were selected for consideration. Resulting outputs of each variable included an R^2 value.

The variable with the highest R^2 value was then paired with every other selected variable and run though the GWR model again. The resulting outputs which resulted in the highest R^2 value would then be paired with every other selected variable again until the five variables best explaining the model were selected. The dependent variables which were calculated as a percentage of population are as follows;

i. Male population
ii. Social Classes 1 and 2
iii. Third Level Education
iv. Population Aged 5 and Over
v. Population Unemployed.

The parameter estimate surfaces belonging to these dependent variables were then visualised along with their respective T-values and the local R^2 in the 2D, 3D and interactive visualisation types used for analysis in Experiment One. The understandings of these terms are outlined in Chapter 1.

Once the factors had been selected they were combined in a separate dataset so that it would be easier to create the choropleth map visualisations for the experiment content. Once GWR output data was integrated with the voter turnout data, it could then be visualised. ArcMap and ArcScene project menus display maps in a layered order with each layer selectable at any time. For the purposes of Experiment One, maps relating to each task were placed in an order starting with task one at the top. ProVis attribute integration was kindly contributed by a former colleague.

3.7 Experiment Protocol

3.7.1 Participant's Tasks

Task design was based on the Latin square that was adapted for the purposes of this experiment. Participants were required to answer three types of tasks for each visualisation in Part A and Part B. Table 3.3 provides examples of each task. Task 1 is a univariate task which examines only one variable, Task 2 is a bivariate task examining two variables, and task 3 is a multivariate task with three or more variables. The difficulty of each task increased as the participant progressed from one to three for each visualisation type in each part. Although it was important to avoid a learning effect as discussed in Chapters 1 and 2, the task format required a degree of consistency and uniformity (i.e. task one was always a univariate, task two was always a bivariate, and task three was always a multivariate). Since the task questions varied, the concern of a learning effect could still be minimised.

Table 3.3 Visual Exploration Tasks Relevant to GWR Results

Task Type	Task Number	Relevant Exploration Task for GWR
Univariate Task	1	Identify high/low values in different spatial locations for each parameter estimate.
		Identify high/low values in different spatial locations for Local R^2 value.
		Identify high/low values in different spatial locations for T-values.
Bivariate Task	2	Identify spatial relationships between each parameter estimate and T-value.
		Identify spatial relationships between each parameter estimate and Local R^2 value.
Multivariate Task	3	Identify spatial relationships between several parameter estimates.

Table 3.4 provides detail on the task breakdown for each part of the experiment. It includes an initial task set prior to pilot experiment work which is discussed later.

Table 3.4 Task breakdown for the GWR experiment

Part A	Pilot 1	Pilot 2	Pilot 3	Experiment
ArcMap	5	3	3	3
ArcScene	5	3	3	3
ProVis	5	3	3	3
Total	15	9	9	9
Part B				
ArcMap	5	3	3	3
ArcScene	5	3	3	3
ProVis	5	3	3	3
Total	15	9	9	9
Overall Total	30	18	18	18

3.8 Experiment Procedure

The participants were provided with a script detailing the steps they were required to complete as they progressed through the experiment. Table 3.5 outlines this experiment procedure.

Prior to commencement of the experiment, participants were presented with an explanation of the purpose of the experiment, details on the GWR Model along with a brief refresher on GWR itself. Details on the relevant characteristics of the visualisations were also provided and a quick walkthrough of an example task was provided for the interactive visualisation system as it was anticipated this visualisation system would be the least familiar, but potentially the most complicated. The task questions were written on the experiment script and displayed in a progress panel on screen. Participants responded to each task in a space provided within the experiment script. The following is an example of a specific task as given in the script;

> "Identify the relationship between the parameter estimates for the population of Social Classes 1+2 and the parameter estimates for Third Level Education on voter turnout levels?"

A consent form was designed and provided for participants to sign before they began the experiment in accordance with the ethical requirements outlined in Section 3.2 of this thesis. A profiler questionnaire was created to enable the categorisation of experiment participants, also outlined above. Once participants completed part A of the experiment they had a break to prevent any possible onset of fatigue. The experiment script was retained by the researcher and reissued to participants at the beginning of part B.

Table 3.5 Experiment Script Process

Title of Research
Aim
Experiment Schedule
Experiment Length
Experiment Hardware & Software
The Dataset
GWR Refresh
Part A
Interacting with ArcMap
ArcMap Tasks 1-3

Interacting with ArcScene
ArcScene Tasks 1-3
Interacting with Processor (ProVis)
ProVis Tasks 1-3
Part B
ArcMap Tasks 1-3
ArcScene Tasks 1-3
ProVis Tasks 1-3

Experiment recording began when participants clicked on the “Start” button of the progress panel associated with Morae Recorder (discussed above). The panel displayed the task that required completion by the participant, this was the same task written on their experiment script. Each post-task survey question was displayed in the same order as the previous post survey order, so participants would be well prepared to answer it, avoiding unnecessary complication. Table 3.6 displays the three types of questions asked of participants following the completion of each task in the experiment. These questions are adapted from a standard System Usability Survey.

Table 3.6 Post Task Survey Questions

How easy do you think the task was to complete?
How fast do you think the task was to complete?
How correct do you think the answer is?

Once the second part of the experiment had been completed, participants were asked to take part in a short review interview. The interview method and design was constructed using a semi-structured interview approach. Here, a general interview guide was prepared to discuss with each participant but there was flexibility to allow the individual participant raise new topics not in the interview guide. This approach, as opposed to a structured interview, meant there was greater exploration of the issues experienced by the participant within the experiment and points not initially considered by the researcher could be flagged and identified as important in the research analysis.

Participants were shown a set of possible questions they could be asked to instigate discussion on the experiment. Table 3.7 shows this set of questions. In some cases, participants answered more than one question spontaneously before it was asked by the researcher. The interviews were semi-structured in nature because as mentioned above, this allows for a degree of researcher control over content and yet participants are free to expand upon a topic that is not covered in the interview questions. From previous work (Burke, 2009) semi-structured interviews had proved to be successful in eliciting useful information from participants.

An important element of interview design is to remember to avoid asking a leading question because that can influence the interviewee's response. A leading question could be "you enjoyed that didn't you?" It is also good interview practise to avoid overly vague questions when possible. The first question interviewees responded to was the only one of two general questions. The second was the final question asking interviewees if they had any comments to make relating to the experiment that had not already been covered.

Table 3.7 Interview question template

1. In general, what did you think about the visualisations? What did you like or not like about them?
2. What did you like or not like about the 2D Visualisation, ArcMap?
3. How did you find the 2D visualisation "ArcMap" to answer Task 1, the univariate task?
4. How did you find the 2D visualisation "ArcMap" to answer Task 2, the bivariate task?
5. How did you find the 2D visualisation "ArcMap" to answer Task 3, the multivariate task?
6. How did you find the 3D visualisation "ArcMap" to answer Task 1, the univariate task?
7. How did you find the 3D visualisation "ArcMap" to answer Task 2, the bivariate task?
8. How did you find the 3D visualisation "ArcMap" to answer Task 3, the multivariate task?
9. How did you find the interactive visualisation "ArcMap" to answer Task 1, the univariate task?
10. How did you find the interactive visualisation "ArcMap" to answer Task 2, the bivariate task?
11. How did you find the interactive visualisation "ArcMap" to answer Task 3, the multivariate task?
12. If you had to pick one visualisation to use for univariate task, which visualisation would that be?
13. If you had to pick one visualisation to use for bivariate task, which visualisation would that be?
14. If you had to pick one visualisation to use for multivariate task, which visualisation would that be?
15. Do you have anything you would like to mention that was not already discussed?

The interviews were usually carried out in the meeting room in which the experiment had just taken place but in the instances where the secondary room had to be used, the interview had to take place in an adjacent common room. The interview was conducted as courteously as possible to help the interviewee feel comfortable whilst maintaining the semi-structured approach deemed so useful earlier.

3.9 Experiment Measurements and Data Analysis

A correctness of response system adapted from a system used by Koua et al. (2006) was used when analysing participant task responses. Correctness was measured through comparison to correct answers for each task, and a correctness score was assigned to participant responses. Correctness is the percentage of the task a participant answered accurately. Generally empirical evaluations only include simple tasks, for example; identify and locate tasks and although the interaction logs can be effective for analysis in this experiment, the spatial nature of the majority of tasks mean tasks responses must be assessed differently. Participant answers were coded according to the percentage of correctness, with 100% being a fully correct answer and 0% being an entirely incorrect answer. Participants were required to provide the name of one or more Electoral Divisions in their answer. For example if a participant is required to submit the names of four Electoral Divisions in their answer but only submit two, that task correctness score will be 50%. The largest cluster of spatial units on the map (when identifying highest or lowest value clusters) is the correct answer. This method was also applied for Random maps, where it was estimated that participants would find it more difficult to identify clusters (where only a few spatial units could be grouped together).

Mouse data were collected on the number of pixels the participants moved their mouse during each task. Both right and left mouse clicks were recorded as part of the mouse clicks data. Movement and clicks were acquired for the duration of each task. These logged clicks and mouse movements offer an insight into how the data were explored by the participants and also on the degree to which they may have been confused depending on their performance on a task. The participants completed a short post task survey that allowed us to measure the perception of the user on their performance of each task. Information included; the ease in which they completed the task, the speed in which they completed the task, and the confidence they had in their answer to the task. The survey contained three questions scored on a Likert scale of 1 to 5 with 1 being “Strongly Disagree” and 5 being “Strongly Agree”. Using this recorded data on perception it was intended to discover if participants thought their performance changed when they had to contend with a data set with a higher number of variables. This indicated if a scalability effect could be detected and/or observed.

During analysis of Experiment One output, the average of each combined post-task survey results were used to gauge overall participant perception on their performance and satisfaction levels relating to the task they had just completed. This perceptual element is important because it provides an indication on the likelihood that a participant would use a particular visualisation to answer a task. A participant may complete a task quickly with 100% accuracy but if they found it very difficult and felt the time taken to complete the task was long they would be less likely to re-use that visualisation. The post-experiment interviews augment the survey results.

Participant post-task survey responses and task responses were compiled. First the written answers had to be transcribed and then each task was assigned a correctness score depending on the extent to which the answer was correct. There was no straightforward way of applying a correctness score to every task because participants answered differently. The correct answers were compared to participant answers and scores were based on the extent to which participant responses matched the correct answer. For example, where a correct answer required the participant to rank the parameter estimates from the most negative to most positive in relation to voter turnout and they ranked three of the five in the correct order then that answer would receive a correctness score of 60%.

Analysis of the interview data was carried out by transcribing each interview in full and the collation of interview content by identifying common or recurring themes using both quantitative (e.g. counting, frequency) and qualitative approaches (e.g. meaning, interpretations, themes).

The general frequent themes identified in the interview data were divided into visualisation groups, and ranked in order of appearance. Qualitative information obtained through interviews is always subjective in nature, meaning the understanding derived from collected qualitative data is dependent on the researcher. Additional information from qualitative data has always proven to be useful and in the case of this experiment, common themes emerging in the interview transcripts were clear. These are discussed in the results chapter for this experiment.

4: Methodology – Experiment Two

The idea and basis for Experiment Two stemmed from research carried out in Experiment One of this research thesis. Results indicated there may be a change in user performance when more information is displayed on a screen. While it might seem logical to suggest that the performance of a person to complete a task becomes more difficult as the information they are required to comprehend increases, we do not know to what extent this is true. Results of the previous experiment suggest changes in participant performance when the level of data or type of visualisation encountered became more complex. However, without attempting to measure the cognitive behaviour of participants it would have been difficult to understand the results beyond standard measured task times and correctness ratios.

The European Cooperation in Science and Technology (COST) provided the opportunity to apply for a Short Term Scientific Mission (STSM) under the Knowledge Discovery from Moving Objects (MOVE) group. Dr. Arzu Çöltekin of the University of Zürich, Switzerland was the host collaborator on this research strand. This chapter outlines the processes involved with the experiment development, implementation and analysis.

In the first section, the composition of experiment participants is described. In the second section, the ethical procedures required to carry out the experiment are discussed. Section three highlights the improvements derived from the pilot experiments carried out. The fourth section describes materials used in the experiment which include dataset spatial units and eye tracking equipment. Section five details the stimuli involved. Section six describes the process of spatial autocorrelation performed on the resulting visualisations to determine the patterns of each choropleth map. Terminology on 'Clustered' versus 'Random' is discussed and the outputs of the Moran's I spatial autocorrelation calculations are explained in this section too. The protocol for the completion of tasks within the experiment is detailed in section seven. Section eight describes the procedure followed throughout the experiment. Finally, section nine on Data Analysis highlights the recorded metrics obtained from the experiment which can be used for analysis. Findings and analysis from Experiment Two are discussed in Chapter 6: Results 2.

4.1 Participants

A total of 32 participants were sourced to complete Experiment Two. Three of these acted as pilot participants in the first three days of the experiment procedure. Three of the participants were excluded from the experiment due to the low number of recorded eye movement fixations in their experiment. Therefore, the remaining 29 participants formed the final set of participants for the experiment. This experiment utilised choropleth maps. Therefore, canvassing for participants with an understanding of GIS minimised the potential for encountering user difficulties during the experiment. Participants comprised of undergraduate students, postgraduate students, doctoral students and University research institute staff.

A request for participation of undergraduate students was made to those with at least a basic understanding of GIS. Suitable and eligible participants consisted of: Undergraduate students who were taking a GIS module; Postgraduate students who were either research assistants or studied in the host geographic institute; Post-doctoral researchers working in the host institute; and tenured staff of the host institute. Each participant had the required level of understanding of GIS principles and so were deemed eligible for participation in the experiment.

4.2 Ethics

Appropriate ethical procedures, permissions and risk assessments were completed (see Appendix). A written description of the experiment was produced for participants to read before they began the experiment. This ensured participants were fully aware of the experiment aim, what to expect from their involvement and helped ensure they had no reservations about participating in the experiment. A consent form was created for each participant to sign once they had understood the experiment description. Again, this outlined their involvement in the experiment and assured the participant their involvement was voluntary and confidential and that they were free to withdraw from the experiment at any time.

4.3 Pilot Experiments

As outlined above, three participants assisted with the set of pilot experiments. Improvements were made after each test. Initially, the map legend was positioned at the bottom of each visualisation and the task was positioned at the top. The positions of each of these were changed following pilot participant feedback to better suit the user's visualisation of the overall experiment screen. The second pilot test experiment resulted in a change in the visualisation background. Previously the background was coloured white which was the same colour as the lowest valued spatial units on each map. This may have caused user confusion so the colour was changed to a pale yellow to provide a better visual contrast. Questions asked by pilot participants allowed the experiment researcher to clarify participant instructions for the experiment.

4.4 Experiment Set Up

When conducting an experiment involving human participants it is important to create the most natural working environment possible as outlined in a previous Chapter. The experimental laboratory was occupied by the researcher and the participant only. The use of non-invasive eye-tracking technology was important to minimise participant discomfort.

4.5 Materials

The eye tracking machine was attached to the desk and located under the computer screen. The eye tracker was a 'Tobii X120' complete with analysis software on the desktop computer. The eye tracking machine was calibrated for the screen size and extent of tracked eye movements in advance of the experiment. The chair was an adjustable computer chair, with the castors removed to limit erratic participant movement. A small glass weight was used as a hand rest to stop the chair from swivelling. The participant was seated approximately 18 inches from the screen. A webcam was used to record participant facial behaviour and can be seen on top of the desktop.

A pen and paper were used to observe and record participant behaviour and performance during the experiment itself. A laptop was used to back-up the experiment recording after each session had finished and the participant had left the room.

Tobii eye tracking analysis software was installed on the desktop. This software recorded a number of metrics including: time-to-first fixation; number of fixations; time-to-first mouse click; number of mouse clicks; number of fixations in an area of interest (AOI); post-task survey responses; time taken for each task; and participant profiler responses. This meant that almost all of the recordings were taken by the Tobii analysis software and it was expected this would offer somewhat greater ease of data collation following completion of the entire set of experiments.

4.5.1 Map Production

The maps were produced using ESRI's ArcMap software and followed the same procedure as outlined for map production in Experiment One, with the exceptions outlined below. The procedure for the calculation of the spatial autocorrelation for these maps is described later.

4.5.2 Dataset

The dataset used in Experiment One was also used to produce the maps in Experiment Two. This allowed for more accurate comparisons between the results of the two experiments if needed. This was a real-life dataset rather than a generated simulated dataset. This was assumed to be beneficial to the participant as it was more likely to reflect the types of data being used in real life work. Specifically within the dataset, Geographically Weighted Regression (GWR) outputs were selected to represent Clustered and Random patterns in the dataset. The outputs included parameter estimate values of the five attributes used in the first experiment (Social Class 1 and 2; Owner-occupied homes; Over-65 age category; 3rd level education and unemployed persons; S-Values and T-Values. The attributes used to produce the maps were selected based on observable patterns in this data. The default data classification system was used in all instances.

4.5.3 Spatial Units

The ArcMap shapefile used to create the spatial map units in Experiment Two was the same shapefile as used in Experiment One again allowing the possibility of comparisons. During the design stage of Experiment Two a rigid rectangular-shaped map was decided upon as opposed to a freeform or more naturally-occurring shape. This permitted a more uniformed set of maps for participants with the intention of reducing a potential learning effect which can occur with more recognisable map shapes (e.g. shape of a country). The spatial unit polygons used are real world Irish “Electoral Divisions” (EDs) as described in Experiment One.

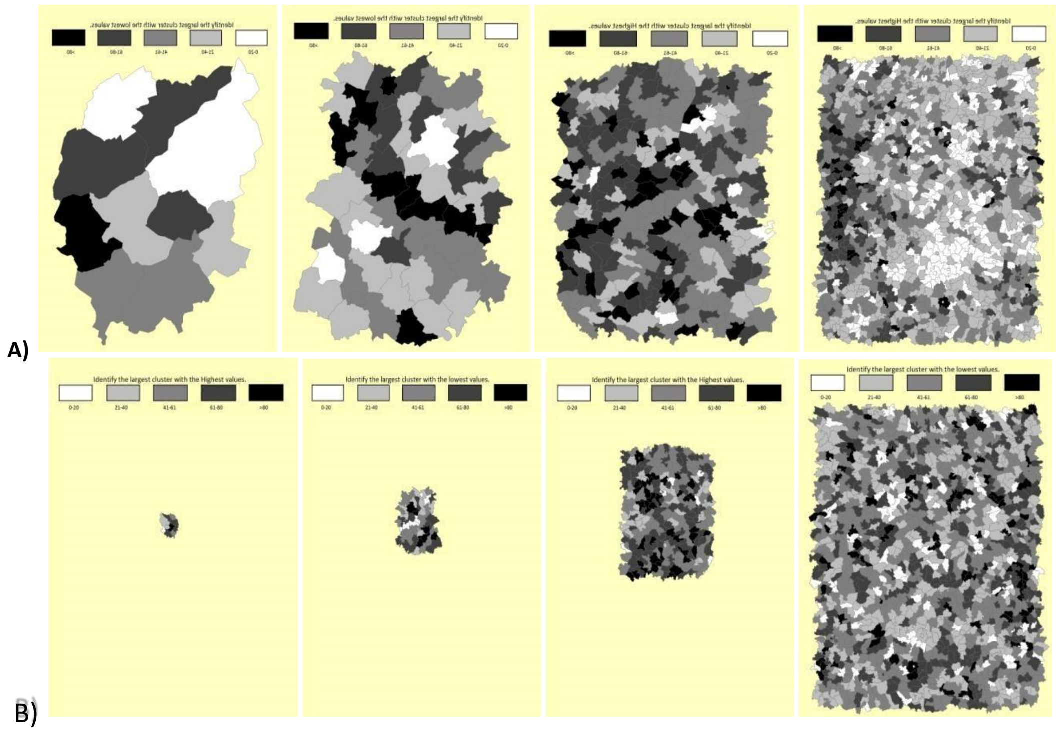
4.6 Stimuli

Since we were testing perceptual scalability in this experiment, a systematic approach was adopted in order to enhance the comparability of the visualisations thereby improving the experimental design. Four different sized maps were produced with incrementally increasing scales by order of five (10; 50; 250 and 1,250). Real life map topography of Electoral Divisions (the smallest area unit) in Ireland were used to replicate a more natural visualisation and experience for participants. At the time, Electoral Divisions were the smallest spatial units used in Irish datasets. The data on the choropleth maps presented to participants was also based on real life data. This was a mix of Geographically Weighted Regression outputs and Census data obtained from the National Census of Ireland Database (CSO.ie). However, the origin of the data was not a key point of information for the experiment participants. The visualisations produced and used in the experiment were designed to be as simple as possible, without distractors or extraneous clutter.

Three visualisation sets were created so the experiment would consist of three different sections; Random, Clustered and Paired. As mentioned above, four data scales were used in the experiment with increments of x5 to allow for more comparable measurements and analysis. Stimuli were created to vary two factors: Random versus Clustered and the size of the spatial units; large/small vs. constant. That is to say, spatial unit size decreases as more units are included, unless the size of the units remain constant in which case the overall map area will change as more units are added. As outlined in the opening paragraphs of this chapter, it was hypothesised that participants would find it more difficult to answer tasks for Random data visualisations because the visualisation appeared to be more complex.

Figure 4.1 shows an example of the eight different types of visualisations used. The same map structure was used for representation of Clustered data. It is important to note the difference between A) and B). In Figure 4.1 A) the size of each spatial unit decreases as the level of data increase. The spatial unit size remains the same throughout Figure 4.1 B.

Figure 4.1 of the "Random" Visualisation maps



4.7 Spatial Autocorrelation

Spatial Autocorrelation calculation of experiment visualisations (the choropleth maps) was used to differentiate between the levels of visual randomness. Visual randomness was calculated so that a comparison of the performance of participants between Clustered and Random could be performed. This comparison will assist in answering the question of the effect Clustered maps have on participants compared to Random maps. Changes in performance are expected to be an indication of the difficulties encountered by participants. Changes in participant cognitive load could not be reliably assessed without knowing the spatial autocorrelation group that each visualisation belonged to. This is a principal reason spatial autocorrelation was calculated.

Moran's I was chosen as the method to calculate spatial autocorrelation. It is used to measure the degree of similarity between data values. In the case of the second experiment, the data values are spatial units called Electoral Divisions (EDs). Moran's I is calculated on a global level so an overall assessment of spatial autocorrelation for each visualisation is achieved.

It is estimated that participant interaction with "Random" choropleth maps will require a greater cognitive load than "Clustered" choropleth maps.

4.7.1 Moran's I

Based on the relevant literature (Slocum et al., 2013; Moore and Drecki, 2013; O'Sullivan and Unwin, 2010) and further drawing on Tobler's First Law of Geography (1970) one can gain an understanding of methods used to understand and assess spatial autocorrelation. The simple outcomes are; positive, negative or none/zero autocorrelation. The description of spatial autocorrelation is of key importance in spatial analysis as indicated by Slocum et al. (2013: 35).

Geary's C and Moran's I are two common methods used to calculate autocorrelation of data. Moran's I can be used to study stochastic phenomena of data that is distributed in space. It is heavily linked to the standard correlation coefficient. Geary's C is based on a comparison of contrasting map values. Geary's C establishes whether there is a correlation between neighbouring values in a dataset. Geary's C is more sensitive to local autocorrelation compared to Moran's I. For Experiment Two, Moran's I was chosen as the method to calculate spatial autocorrelation so that an overall picture of spatial autocorrelation can be taken for each visualisation used within the experiment.

4.7.2 Spatial Autocorrelation calculation

As with the creation of the maps, ESRI's ArcMap was utilised to calculate the spatial autocorrelation of each visualisation. ArcMap's toolbox contains a feature to calculate spatial autocorrelation using Moran's I. Some of the changeable options were as follows;

Table 4.1 ArcGIS Toolbox Moran's I Spatial Autocorrelation selected options

Input Feature Class	Left to the default working directory.
Input Field	Set according to the active parameter.
Conceptualisation of Spatial Relationships	Inverse Distance
Distance Method	Euclidean
Standardisation	Standardisation

The input class and input field choices are self-explanatory. Inverse Distance was chosen because it would best measure clusters or patterns. The further one spatial unit is from another, the smaller the impact it has. The closer two spatial units are in space to each other, the more likely they are to interact and affect each other.

Euclidean Distance was selected as it calculates the distances between the central points of each spatial unit to its neighbouring spatial units. The limitations of Euclidean distance are not relevant here, e.g. slopes or physical landscape variations. Manhattan distance is restricted to straight line connections, comparable to right angles only, e.g. North to South or East to West. This makes it less appropriate to use because the dataset is less grid-like and is very much a real life dataset. Standardisation was tested but the difference in the Moran's I outcome was negligible.

4.7.3 Spatial Autocorrelation Outputs

Essentially, Moran's I can indicate whether a spatial pattern is Clustered or Random. There are three main outputs from Moran's I: Index, Z-Value and P-Value. Table 4.2 below demonstrates the categorisation of each of the main outputs in terms of confidence levels.

Once the spatial autocorrelation calculations are complete, the first step is to assess whether the null hypothesis is true or false. In this experiment, the "null hypothesis" states;

"There is no spatial clustering of the values associated in the study area."

Positive Index scores indicate clustering, while negative index scores indicate. Scores closer to zero indicate a Random distribution. The p-value is probability. The probability that observed spatial patterns are Random processes. A small p-value indicates observed spatial patterns are unlikely to be a result of Random process. The Z-values are standard deviations. P-values and Z-values are suggestions on whether you can reject the null hypothesis or not. Generally, the null hypothesis will be that there are no spatial processes present in data.

High positive or negative Z-values combined with small P-values are normal. Small p-values with very high or low Z-values indicate that a spatial pattern may not reflect the null hypothesis random pattern. In this instance, a subjective decision is made. A typical confidence level table is a follows:

Table 4.2 Moran’s I Confidence Measurement

Z-score	P-value	Confidence
< -2.58 or +2.58	< 0.01	99%
< -1.96 or > +1.96	<0.05	95%
<-1.65 or > +1.65	< 0.10	90%

The above description demonstrates the necessary completed steps so that analysis can be carried out on participant expertise, fixations, time-to-task completion and the spatial distribution of experiment visualisations.

4.7.4 Experiment Terminology

The difference between Random and Clustered as presented to participants for the experiment. The standard technical definition of ‘Random’ and ‘Clustered’ is different to their use in Experiment Two. In general statistical literature, the term “Clustered” can relate to spatial autocorrelation where a spatial dataset can be assessed for the extent to which a perceived pattern has occurred by chance. The term “Random” can be applied to patterns which have occurred more by chance rather than having a quantifiable statistical spatial connection.

While the above definition of each term is understood, the displayed terms of “Clustered” and Random” in the tasks presented to participants in Experiment Two are used to facilitate understanding by participants according to their perceptual perspective. The terms “Random” and “Clustered” are used to describe the perceived or observable spatial pattern of visualised data on each map output. The distributions were not initially created using a mathematical formula for Clustered or Random map generation as may be assumed in the statistical definition of the terms. However, for analysis purposes, spatial autocorrelation was performed on the visualisations in order to categorise them correctly.

4.7.5 Analysis Terminology

Referring back to Chapter 3, we know what qualifies a spatial distribution as “Clustered”, “Random”. Through this literature we can say that:

- An ideal “Clustered” map is described as a map with an Index Score of “1.0”and P-Score of “less than 0.01”.

- An ideal “Random” map is described as a map with an Index Score of “zero”, more and a P-Score of “less than 0.01”.

4.7.6 Experiment Visualisations Spatial Autocorrelation Outputs

Table 4.3 shows the outputs of the spatial autocorrelation calculations for the experiment map visualisations. A description of why each map has been categorised as “Clustered” or “Random”” is provided below. Studying each primary output, it has been established that P-values (assist in the detection of the reliability of the spatial autocorrelation result) of zero indicate that the resulting Index Scores (level of clustering) are not likely to be a result of a random process. A choice was made for purposes of labelling and analysis that maps would be classified as “Random” is their Index score was less than 0.4. Even though there is statistically significant clustering on certain maps, e.g. “1250 L r” is labelled Random because the Index scores are relatively low.

Table 4.3 of Visualisation Spatial Autocorrelation

Map	Index	P-Value	Indicated Pattern
10s	0.493	0.001553	Clustered
10s 1	0.492541	0.001407	Clustered
10L	0.492541	0.001407	Clustered
10L 1	0.516546	0.000842	Clustered
10s r	0.018652	0.400002	Random
10L r	0.028002	0.414583	Random
10s r1	-0.39483	0.132513	Random
10L r1	-0.27857	0.386992	Random
50s	0.831886	0	Clustered
50s 1	0.858707	0	Clustered
50L	0.809499	0	Clustered
50L 1	0.833335	0.000213	Clustered
50s r	0.312483	0.117649	Random
50s r1	0.121791	0.117649	Random
50L r	0.059247	0.380477	Random
50L r1	0.059247	0.380477	Random
250L	0.935263	0	Clustered
250L 1	0.922299	0	Clustered
250s	0.931089	0	Clustered
250s 1	0.914997	0.413945	Clustered
250L r	0.027359	0.001859	Random
250L r1	0.116653	0.001859	Random

250s r	0.128894	0.000571	Random
250s r1	0.229048	0	Random
1250	0.948047	0	Clustered
1250L	0.960195	0	Clustered
1250s	0.937812	0	Clustered
1250s1	0.95829	0	Clustered
1250L r	0.363214	0	Random
1250L r1	0.056256	0.000835	Random
1250s r	0.055902	0.000899	Random
1250s r1	0.055902	0.000899	Random

A threshold was established for the categorisation of each visualisation. Most of the Clustered maps were identified as having an Index score that was closer to 1, with a Z-score greater than 2.58 and a p-value of zero or close to zero. In most instances, this description fits the maps categorised as Clustered. Some of the maps were straightforwardly classified as being Random. These maps had a much lower Index score, with a Z-score that is still greater than 2.58, but much closer to that figure. While the Random p-value was still close to zero, it was much greater in relative terms to the majority of the Clustered P-values. Only two maps recorded a negative spatial autocorrelation. Relative to the Clustered maps, the P-value was much greater than zero.

It is difficult to be entirely confident with the spatial autocorrelation outputs for the ‘10’ spatial unit maps category given the minimum recommended number for spatial autocorrelation calculation is 30. It could be suggested that the comparatively lower Index scores for the ‘10’ spatial unit maps demonstrate the difficultly associated with discerning patterns which are not likely the result of a random process in so few data values. Similar to ‘50L r’ and ‘50 L r1’ the Z values for ‘10s r’ and ‘10L r’ are below the 2.58 threshold so there is a degree of uncertainty surrounding their Index score to determine if Clustered or Random patterns are observed. The high P-values indicate that the observed patterns are likely to be the result of a random process. The lack of a presence of a Clustered pattern means that ‘10s r’ and ‘10L r’ can be categorised as being “Random”.

For the ‘50L r’ and ‘50L r1’ maps, the Index score is quite low. It is much closer to zero than 1.0. This indicates that the presence of a Clustered pattern is unlikely. The p-values for these maps are over the 0.1 threshold which indicates that the observed patterns are likely to be the result of a random process. They can still be accepted into the “Random” process category.

For the ‘250’ set of Random maps, the distinction is less certain because the Index scores are still quite high. But with a much lower Index score than their Clustered ‘250’ spatial unit counterparts, they have been classified as being Clustered. The Z-Score for all but one of the ‘250’ maps is above 2.58 indicating that the observed patterns are unlikely to have occurred by chance. This is least certain for the ‘250L r’ because the Z-score is below 2.58 indicating there is a degree of uncertainty about its place as a Random map. The Index score is much lower than the other ‘250’ spatial unit Z-scores and it indicates that the presence of Clustered distribution is not very likely. The p-value is also higher than 0.1 which indicates that the observed pattern for the ‘250L r’ maps may have occurred by chance. This may indicate the presence of a random distribution, so the map can be categorised as being “Random” here for the purposes of analysis.

The classification for some of the visualisations was less straight forward. For example, in the case of ‘1250L r’, the Index score was considerably less than the Index score of the other Clustered ‘1250’ scale maps. However, the Z-value is much higher than the Random ‘1250’ so we can be more certain that the map is Clustered.

4.8 Experiment Protocol

4.8.1 Participant’s task

In experiment segments lasting approximately 10 minutes, participants were asked to complete tasks that a typical geographer would find very simple, but which required them to inspect the maps. Clear instructions were given.

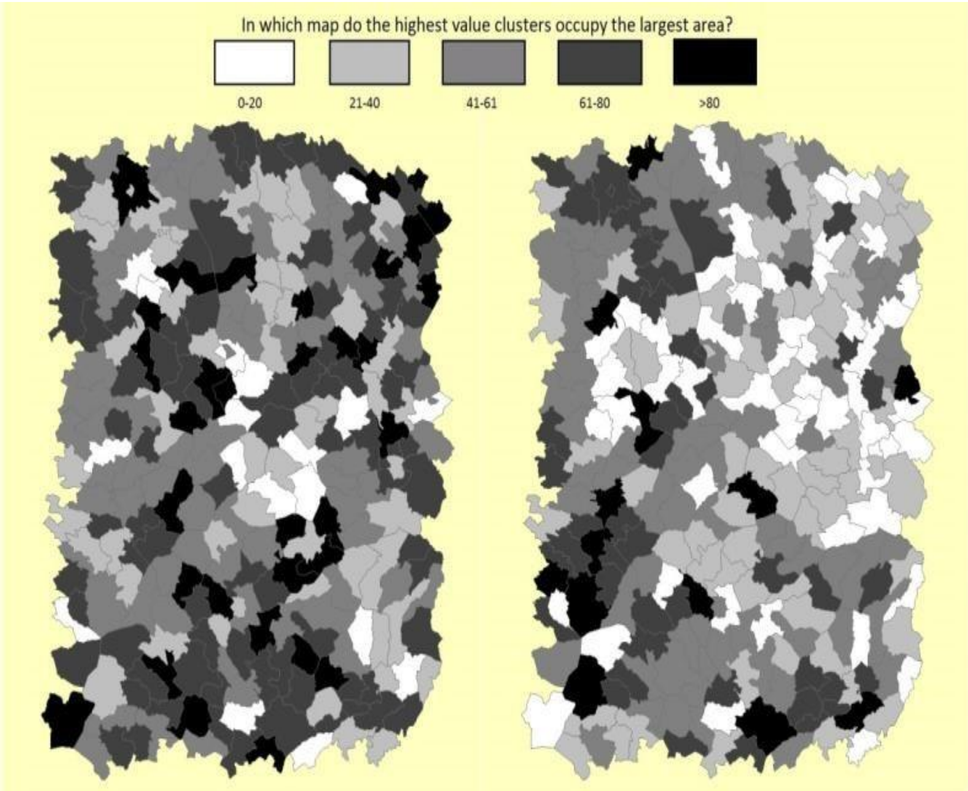
An example of a single choropleth visualisation task is as follows: Identify the large cluster with the highest values.

An example of paired choropleth visualisation task is as follows:

Identify the map containing the largest cluster of the highest values.

Tasks to complete were considered straight-forward and participants were required to identify the highest or lowest valued clusters of Electoral Divisions on the choropleth maps presented to them. In order to see if perceptual scalability was apparent when participants had to compare two maps of the same scale and information, a set of paired visualisations was included in the experiment (see Figure 4.2 for an example).

Figure 4.2 Paired choropleth maps for Random distribution



These were choropleth maps of the same scale. Participants were asked to complete similar types of tasks as before, except now they had to choose a map instead of a cluster. Following guidance in the literature, the order in which participants were shown the visualisations in each spatial distribution group was randomised using the Tobii Studio experiment system to prevent a possible learning effect. Table 4.4 provides additional details on this visualisation, data and task design.

Table 4.4 Experiment visualisation, data and task details

	Display Type	Spatial Unit Size	Spatial Unit Distribution	No. of Display Sizes	No. of Tasks per Display	Total tasks for Part
Part 1 A	Remains the same	Progressively smaller	Spatially auto correlated	4	2	8
Part 1 B	Remains the same	Progressively smaller	Random	4	2	8
Part 2 A	Progressively larger	Remains the same	Spatially auto correlated	4	2	8
Part 2 B	Progressively larger	Remains the same	Random	4	2	8
Part C	Remains the Same	Remain the same	Spatially auto correlated and Random	4	1	8

It was important to maintain visualisation uniformity in terms of shape so that results could be more readily compared. An example where this uniformity of shape was successfully implemented and can be seen in either of the '1250' scale Random visualisations shown in Figure 4.1. Although it was difficult to make a selection of EDs that occupied a rectangular or uniformed shape for the smallest scale ('10' spatial units), the selections made were relatively successful.

In order to record measureable results the methods of scaling would have to be considered carefully. Initially a maximum of 2,500 EDs were planned for. However, the scaling was not consistent because the smallest number of spatial units would have been 50. It would increase to 250 and then to 500 followed by 1000 and then either 2000 followed by 2500.

While results could be compared between visualisations with the same scale it would have been difficult to measure perceptual scalability across visualisations with different scales. An additional practicality regarding the dataset had to be considered when using the largest number of spatial units. A real life data set based on the shape of Ireland was used. This meant that a uniform square shape would not have been possible if the number of spatial units exceeded more than 2000. If the shape/rectangular of any visualisation looked unique and was reused to display different data or the same data in a different colour then a learning effect could occur.

4.9 Experiment Procedure

Participants were briefed prior to starting the experiment and thereafter proceeded to complete the tasks in a self-paced manner. The experiment was expected to last between 5 and 10 minutes dependent on the pace of the individual participant. Each participant performed the tasks in segmented blocks of Random data distribution and Clustered data distribution, with the order of blocks counter-balanced.

Upon arrival, participants were instructed to take a seat at a desk upon which they would be presented with the experimental brief (see Appendix). The brief was explained by the researcher and participants were told that the visualisations were being evaluated for comprehensibility: this served to relax participants. When any participant questions had been answered a consent form was shown and explained to them. Again, once they were happy with the explanations and understood what would be required of them in order to complete the experiment, the consent form was signed.

The participant was seated in front of the experiment desktop computer and the eye tracking machine, which was calibrated for each individual. Depending on the dominant hand a participant would use, the mouse and the stabilising globe were arranged so that the participant used their preferred hand to respond with the mouse and their non-dominant had was placed on the globe to stabilise movements. Participants were instructed to look at the screen at all times (this was to prevent the infrared light from the tracker losing the target (i.e. the participant's eyes). The height of the chair was adjusted to compensate for differences in participant height so that the eye tracking machine would successfully locate and track the eye movement.

A profiler questionnaire was completed, with several questions using a Likert scale of 1 to 5. Table 4.5 details the profiler questions asked as part of the experiment. The participant was allowed to ask any number of questions during the course of the experiment. They were informed that a single mouse click would be sufficient when answering tasks but that their final mouse click would be taken as their answer. The tasks described above were written in straightforward, non-technical English language so that any participants with less experience of the English language would not find it difficult to comprehend what was asked of them.

Table 4.5 Experiment Profiler Questions

How many hours of sleep did you get last night? <4, 4-6, 6-8, >8
Please rate your training on Cartography, 1-5
Please rate your training on GIS software, state verbally which type(s), 1-5
Please state your level of experience in the English language, 1-5

The same post-task survey from Experiment One was attached to each visualisation in the experiment. The survey questions asked participants to rate their responses using a Likert scale of 1 to 5 (see Table 4.6).

Table 4.6 Post Task Survey Questions

How easy was that task to complete?
How fast do you think you completed that task?
How confident are you that you answered the task correctly?

In accordance with the randomised system of visualisation display discussed above, the Clustered visualisation set would be shown first to one half of total participants and the randomised visualisation set would be shown first to the other half of total participants. The paired map visualisations would be shown once the Clustered and randomised visualisations segments were completed. All three map distribution types were shown to participants in a single sitting because the time taken to complete the experiment was not expected to result in participant fatigue. The participant was instructed to notify the experiment host when they had completed the first part of the experiment. The host would then start the second part immediately. This was repeated at the end of the second part for the final, third segment.

Once the experiment was completed the participants were asked if they had any additional thoughts or feedback to contribute. Every participant was offered a nominal reward (chocolate) and the Undergraduates were also rewarded with a 5 Swiss Franc university café voucher in acknowledgment of their participation. Participants were not aware of this nominal gift prior to their participation in the experiment.

4.10 Experiment Measurements and Data Analysis

The cognitive element of Experiment Two was less complex because there were no functioning map features. There were just mouse clicking, reading the question and finding the target area they believe to be the correct answer. This experiment process is one that needs to be taken into consideration when analysing results.

Each recorded measurement was extracted for every participant for further analysis. Participants were divided into perceived brackets of expertise using participant profiler data. Post-task responses were also categorised for each visualisation. The number of eye fixations and times to first mouse click were exported and filtered to remove irrelevant data. Eye fixations provide an indication of difficulty experienced by a participant in answering a task. It is expected that the number of fixations will be greater for non-clustered distribution maps over comparable maps scales. The time to first mouse click will suggest the length of time required by a participant to discern where the correct answer is located. This data was added to the divided group and a comparison of these groups was made to assess differences in performance.

Areas of interest (AOIs) were created for each map within Tobii Studio. The AOIs are selected areas containing more relevant information than the rest of the map. Rectangle and ellipse shapes were drawn around the cluster relating to the task for each visualisation. Ellipse AOIs were adjusted to better mimic the shape of the cluster where necessary. The number of ellipses varied according to the number of relevant clusters. Participants answered a question through a mouse click - they would click on the cluster they believed to be the correct answer. In section three of the experiment, merely clicking the map was sufficient, but in most cases participants selected the cluster on one of the maps.

The duration of fixations within the Areas of Interest (AOIs) provide an indication of required attention. If a participant has spent more time visiting the AOIs for each task then it can suggest difficulty in deciding on the correct answer. This is also true for mouse click recordings.

The “time to first mouse click” is typically used to measure speed of participant decision-making. Again, this was recorded within the AOI to determine if a participant answered correctly or not. In the case of this experiment, this speed would help determine the efficiency of a visualisation to present information.

Correctness levels were determined based on whether a participant clicked in the visualisation cluster with the largest grouping of spatial units, as required by the task question. The simple tasks described above varied in levels of difficulty based on the visualisation type presented to the participant. The short survey answered by participants following the completion of each set of visualisations was also exported to Microsoft Excel for collation and analysis. Suggestions on perceived task difficulty, speed to complete and confidence in answers can be derived from this data.

The Tobii eye movement analysis software allowed for the collation of almost all experiment data as mentioned above, with the exception being post-experiment participant feedback and written observations made by the host during the sessions.

Using the above methods of data collation, participant data were also divided according to their perceived level of GIS expertise and the order in which they viewed each map during the experiment. This is described in more detail below and would allow for further comparison of results. From the set of experiment participants, two GIS expertise groups were created according to the answers provided by participants on their perceived level of GIS expertise possessed. The two groups were labelled GIS Non-Experts and GIS-Experts. Non-Experts rated themselves between 2 and 3 on a Likert Scale. Experts rated themselves between 4 and 5 on that same Likert scale. A group was also created to differentiate participants according to the order in which they encountered each distribution of visualisations within the experiment. One group was presented with visualisations from the “Clustered” category first, the other group was presented with visualisations from the “Random” category first.

The metrics (number of fixations, correctness, time to complete task, time to mouse click) described above were compared within the groups of expertise and within the groups which encountered either “Clustered” or “Random” visualisations first. Results for visualisations in which the spatial unit remains the same versus visualisation in which the spatial unit size becomes progressively smaller. The next chapter will discuss the results derived from data collected as part of Experiment Two.

5: Results and Analysis – Experiment One

The purpose of Experiment One was to investigate the performance levels of people with knowledge of GWR with respect to their interactions with visualisations. A trio of visualisations were selected for testing and these acted as a medium to assess the popular visualisation techniques of 2D, 3D and interactive visualisation types. The performance of all three visualisations are evaluated and discussed in this chapter. To overview, visualisations performed equally in some instances. However, results did indicate that interactive visualisations are better suited for more comprehensive or advanced analysis of GWR outputs. The chapter is structured with seven core sections outlined below to aid analysis, interpretation and understanding of Experiment One's results.

First, the composition of the experiment participants is outlined according to knowledge levels (section 5.1). Next, the experiment correctness ratios of participants are analysed in section 5.2. Section 5.3 highlights the speed of task completion for each visualisation type and each visualisation scale tested during the experiment. Data obtained on mouse movement and mouse clicks which offers an insight into the difficulties encountered by participants is investigated in section 5.4. In section 5.5, participant perception of their own performance on task completion is considered through analysis of post-task surveys within the experiment itself. Then, qualitative results of the semi-structured interviews and participant's own observations compared to their actual performance results is discussed in section 5.6 before a discussion piece in section 5.7 to conclude the chapter.

5.1 Categorisation of Participants According to Expertise Levels

Table 5.1 outlines the divisions created for analysis purposes according to participant expertise. As highlighted later in this chapter, these divisions are split into two sections, one for Part A of the experiment and the other for Part B.

The explanation of each division is as follows; G1 contains participants with knowledge of the 2D visualisation system used (ArcMap). Participants with knowledge of the 2D and 3D visualisation systems, ArcMap and ArcScene, are found in G2. G3 comprises participants with knowledge of all three visualisation systems (ArcMap, ArcScene and ProVis). Participants with knowledge of ArcMap (2D) and ProVis (Interactive) are located in G4.

G5 and G6 are related to participant GWR expertise. G5 contains participants with a perceived "poor" or "basic" knowledge of GWR while participants with "good" or "expert" knowledge of GWR are in G6.

Table 5.1 Divisions of participants according to their knowledge base on the experiment visualisations and GWR

	Knowledge bracket
G1	2D
G2	2D & 3D
G3	3D, 3D & Interactive
G4	2D & Interactive Vis'
G5	Poor and Basic GWR
G6	Good & Expert GWR

5.2 Correctness Ratios Analysis

Correctness ratios varied across the knowledge groups. Tables 5.2, 5.3 and 5.4 show the average correctness percentages of the knowledge groups for each visualisation on every task. In general, participants attained the highest correctness ratios in Task 1 of the 2D visualisation for Part A. Task 3 proved the most difficult to answer when using the 2D and 3D visualisations. The interactive visualisation performed best for the multivariate task. This section gives details on each of the participant groups and visualisation types.

Table 5.2 Average Correctness Ratios for 2D Visualisations

	Part A			Part B		
A	2D T1	2D T2	2D T3	2D T1	2D T2	2D T3
G1	90	85	57.5	56	52	56
G2	92	85	52	64	44	45
G3	100	100	65.00	60	55	37.5
G4	86.67	100	46.67	53.33	41	48.33
G5	92.5	81.25	51.25	58.89	50.56	47.72
G6	90	90	53.33	60	51.66	50.83

Table 5.3 Average Correctness Ratios for 3D Visualisations

	Part A			Part B		
A	3D T1	3D T2	3D T3	3D T1	3D T2	3D T3
G1	68.75	75	42.5	66	52	39
G2	51.25	100	20	76	54	33
G3	80	75	40	55	50	65
G4	86.67	66.67	28.33	60	30	30
G5	75.63	87.5	23.13	66.67	43.75	28.33
G6	65.83	83.33	45.83	71.67	51.67	36.67

Table 5.4 Average Correctness Ratios for Interactive Visualisations

	Part A			Part B		
	Int T1	Int T2	Int T3	Int T1	Int T2	Int T3
G1	60	52.5	62.5	72	84	72
G2	88	53	72	72	84	96
G3	80	37.5	75	100	65	80
G4	61.67	41.67	60	70	75	80
G5	87.5	48.75	61.25	77.78	78.33	85.56
G6	60.83	76.67	75	71.67	78.33	76.67

5.2.1 Differentiation between Knowledge Groups

Summarised results for both parts are shown per group in Tables 5.2, 5.3 and 5.4. Those with knowledge of ArcMap only (Group 1) contained some of the worst performers when using the interactive visualisation for Part A and Part B. Those with ArcMap and ArcScene knowledge (Group 2) recorded the lowest average scores for Part A Task 3 of the 3D visualisation despite having knowledge of this visualisation type. This group was also the worst performer on Task 1 of the 3D visualisation. This is unusual because this group also recorded the highest average correctness ratios for Part A, Task 1 in addition to Part B, Task 1 and 2 of the 3D visualisations.

Group 6 recorded better correctness ratios than Group 5 for most tasks in Part A and Part B with approximately 17% of Group 5 ratios being higher. Group 6 performed better than Group 5 in Part A and Part B of the 2D visualisation. The results show those with a good understanding of GWR recording having higher correctness ratios for every task on Part B of the 3D visualisation. Correctness ratios for Part A and Part B, the interactive visualisation indicate that participants with a more basic understanding of GWR recorded a slightly better score than participants with a good understanding of GWR.

5.2.2 Differentiation between Visualisation Types

2D maps performed best for Task 1 on Part A data set. Both Tasks 1 and 2 in Part A received correctness scores of over 80%. This suggests that this type of visualisation is well suited to simple GWR tasks. The 2D map correctness ratio for Part A's Task 3 was less than the interactive visualisation, while Part B's percentage difference with the interactive visualisation was over 15%. The 2D visualisations are therefore less appropriate for multivariate GWR exploration and these are significant differences since GWR analysis is likely to be complex.

The 3D visualisation failed to record the highest correctness percentage for any of the tasks in Part A or B with the exception of Task 2 of Part A (Group 2). In particular, Task 3 had the lowest correctness ratios across all groups in Part A and again in Part B with respect to ratios for Task 1 and 2 for all visualisations. This can be attributed to the ArcScene layout which incorporates a 3D visualisation rotation feature. Technically the participants did not have to utilise this rotation feature for Tasks 1 and 2 and could observe the maps from a top-down perspective. Task 3 however required the use of interactive rotation and even knowledge of ArcScene (groups G2 and G3) didn't seem to help with rotation, although these groups performed better than most on Tasks 1 and 2 in Part A. This indicates a difficulty in using a 3D visualisation to complete multivariate tasks. 3D surfaces with rotation therefore seem least suitable for a majority of GWR tasks.

In addition, a potential geographic scalability effect can be observed for 3D surfaces between Parts A and B in that participant's correctness ratios were in general lower in Part B, with each task average for all visualisations being between 55 and 68% compared to a Part A average of almost 70%. This includes an average of approximately 80% for Task 1 of Part A and an approximate average of 75% for Task 2 for all visualisations. Correctness ratios for Part A Task 3 were lower than Task 1 and 2 with averages between 50 and 56% for Part A and Part B.

On average for all tasks, participants performed better on Part B than Part A. When the average correctness ratios of all knowledge groups are compared, it is evident the interactive visualisation performed best overall for Part B. This suggests that participants perform better with this visualisation type once they gain operational experience when they complete Part A.

It is more difficult to determine if there is a scalability effect present between Part A and Part B of this visualisation type. For instance, the univariate correctness ratios are slightly lower in Part B, but the multivariate ratios are slightly higher in Part B. Since the correctness ratio for Task 3 was better for the interactive visualisation in Part B compared to the 2D and 3D visualisation, it offers an indication that interactive visualisations could be best for more complex GWR analysis.

5.3 Speed of Task Completion

Tables 5.5, 5.6 and 5.7 provide details on knowledge groups for every task of both parts for the three visualisations. The average time-to-task completion and standard deviation for each group and task is recorded. In the following, results per group are outlined below the set of tables.

Table 5.5 Average Task Times for 2D Visualisations (in minutes)

	Part A						Part B					
	T1 Time	T1 SD	T2 Time	T2 SD	T3 Time	T3 SD	T1 Time	T1 SD	T2 Time	T2 SD	T3 Time	T3 SD
G1	5.51	1.45	8.64	5.61	6.16	1.96	7.39	4.89	4.4	2.01	4.43	1.73
G2	5.37	4.08	3.52	1.54	6.99	5.71	4.9	5.2	4.4	1.67	4.48	2.02
G3	4.38	1.97	3.12	1.97	2.68	1.26	4.1	2.33	5.36	2.22	2.48	0.91
G4	5.66	1.62	8.59	8.6	6.24	3.88	5.85	5.32	5.47	1.83	4.13	0.89
G5	5.61	3.34	3.79	1.38	6.54	5.02	3.67	3.26	4.6	1.39	4.1	1.93
G6	5.75	1.19	8.02	5.51	7.44	3.06	7.53	5.72	4.03	2.52	3.35	2.23

Table 5.6 Average Task Times for 3D Visualisations (in minutes)

	Part A						Part B					
	T1 Time	T1 SD	T2 Time	T2 SD	T3 Time	T3 SD	T1 Time	T1 SD	T2 Time	T2 SD	T3 Time	T3 SD
G1	5.66	1.85	5.26	3.38	16.37	5.62	4.74	2.51	11.2	8.74	9.29	5.46
G2	3.9	2.72	2.34	1.38	9.34	3.39	4.04	3.25	7.2	3.13	8.31	1.81
G3	4.67	0.98	5.4	3.97	3.75	2.37	4.64	0.03	7.15	2.55	8.89	0.25
G4	5.15	1.52	3.23	0.38	12.28	11.3	3.08	0.24	6.48	3.43	12.18	5.89
G5	3.57	1.39	3.59	1.65	8.27	4.39	4	2.49	6.8	3.02	9.51	3.82
G6	5.43	1.79	5.02	3.41	12.97	8.46	5.21	2.1	6.09	3.65	10.23	4.97

Table 5.7 Average Task Times for Interactive Visualisations (in minutes)

	Part A						Part B					
	T1 Time	T1 SD	T2 Time	T2 SD	T3 Time	T3 SD	T1 Time	T1 SD	T2 Time	T2 SD	T3 Time	T3 SD
G1	4.31	3.37	3.49	2.28	5.19	3.52	2.65	1.57	5.55	3.1	5.37	4.65
G2	4.38	1.97	3.12	1.97	2.68	1.26	4.1	2.33	5.36	2.22	2.48	0.91
G3	2.6	0.11	3.46	2.18	3.03	3.27	3.9	2.23	4.11	1.07	2.85	1.13
G4	5.41	1.93	0.49	0.42	5.68	1.99	1.66	0.92	6.8	0.71	2.43	0.47
G5	4.21	2.02	3.19	1.57	3.72	2.35	3.63	2.02	5.47	1.69	2.71	1.17
G6	4.05	3.07	3.39	2.24	5.01	3.44	3.31	1.79	5.95	2.78	4.48	3

5.3.1 Group 1 – Knowledge of ArcMap only

Participants with knowledge of ArcMap took the longest time on average to complete tasks using the 3D visualisation. This was particularly apparent for the Part A multivariate task and on the bivariate and multivariate tasks in Part B. This group also managed to complete the Part A 2D tasks faster than Part B's. This is likely a result of users gaining a sense of familiarity with the visualisation. The ProVis task times were quite consistent over both parts; this suggests these users do not experience any greater difficulty when faced with a more complex dataset despite having no prior experience in the operation of an interactive visualisation. On average the ArcMap knowledge group completed tasks quicker using the interactive visualisation compared to the 2D visualisation.

5.3.2 Group 2 – Knowledge of ArcMap and ArcScene

For participants with knowledge of both ArcMap and ArcScene (Group 2) it was difficult to discern any noticeable difference in task times for the 2D visualisation between Parts A and B. There is an increase in the time taken to complete 3D visualisation tasks between Task 1 and Task 3, this increase can also be distinguished when comparing Part A and Part B. Part B tasks took longer to complete. With the exception of the Part B bivariate task for ProVis every task was as quick if not quicker than the 2D visualisations bivariate task. Tasks were completed faster using ProVis in comparison to ArcScene except for the univariate and bivariate task in Part A where participants could employ the same completion strategy already used with the 2D visualisation univariate and bivariate tasks.

5.3.3 Group 3 – Knowledge of all three visualisation types

Two participants signified they had experience with all three visualisation types but this was not necessarily indicated by the time taken to complete tasks. For example, when using the 2D visualisation these users took as long if not longer to complete tasks than the other groups. They completed the 3D tasks in a similar time to other groups, they were not significantly faster. The same can be said about the Interactive visualisation task times. For some tasks they were faster on average for the multivariate tasks than the other groups.

5.3.4 Group 4 – Knowledge of ArcMap and the Interactive Visualisation

The group with knowledge of ArcMap and Interactive visualisations (Group 4) struggled to complete the multivariate tasks using the 3D visualisations. In particular, this knowledge group was the slowest to complete Task 3 for Part B of the 3D visualisation. This could be due to a lack of knowledge with this visualisation type. The average task completion time for Part A Task 2 of the interactive visualisation was less than half a minute, making Group 4 the fastest group to complete this task. The group also completed the interactive visualisation Part B tasks one and three faster than any other group, meaning Group 4 completed half of the interactive visualisation tasks faster than any other group.

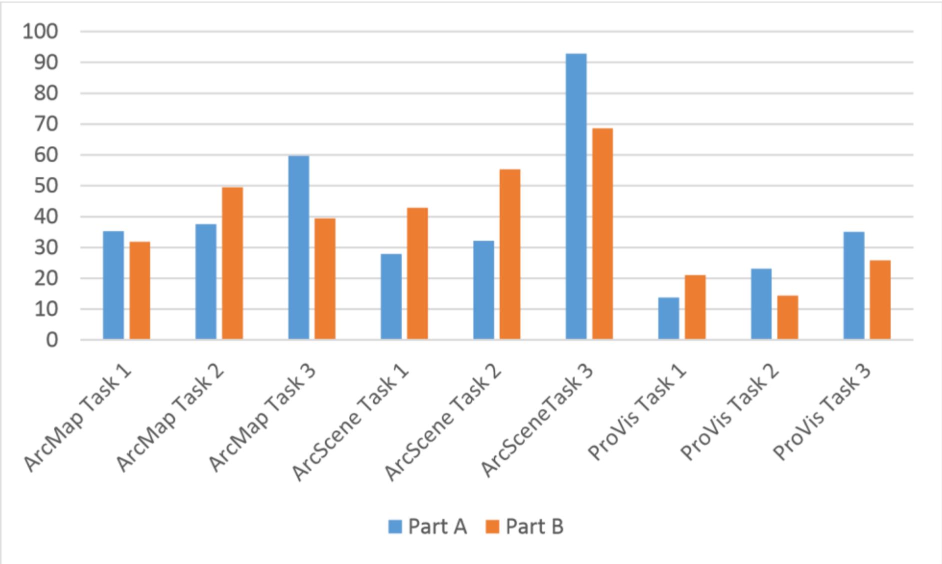
5.3.5 Group 5 and 6 – GWR knowledge

Groups 5 and 6 can be compared because they are a combination of participants' knowledge of GWR, with Group 5 being the basic knowledge group, and Group 6 being the more proficient knowledge group. The most noticeable difference between these groups is that Group 5 completed most tasks with each visualisation type faster than Group 6 (refer back to Tables 5.5, 5.6 and 5.7). The exceptions are 2D Part B Tasks 2 and 3, 3D Part B Task 2, Interactive Visualisation Part A Task 1 and Part B Task 1. This means that only 16.67% of the group with superior knowledge completed tasks quicker than the group with a more basic knowledge. Standard deviations for Group 6 are slightly more varied compared to standard deviations for Group 5 which may suggest that there is an outlying participant, but since Group 6 deviations are not consistently more varied it is difficult to discern if this is the case.

5.4 Analysis of Mouse Movement and Mouse Clicks

Figure 5.1 illustrates the number of clicks made by the individual on each task. Immediately noticeable is a relatively low number of left mouse clicks when using ProVis. This is because participants did not need to click on any item to see information, except when switching between which parameter estimate was viewed or when they performed the select feature task. Otherwise the relevant information was displayed by holding the mouse over the graphic elements.

Figure 5.1 Average number of Mouse Clicks of each application by for Parts A and B



There is no clear indication when analysing the entire group together of any definitive difference between Part A and Part B for univariate tasks.

A pattern that stands out is a comparatively large number of clicks when performing the bivariate and univariate tasks (Tasks 2 and 3) using 2D Maps. Participants were required to analyse up to 5 maps for the multivariate task when using the 2D visualisation and this is indicative of the number of mouse clicks executed by participants. It can also indicate a degree of confusion or extra cognitive effort since users are required to perform navigation between the five maps.

It is interesting to observe the number of mouse clicks for 3D visualisations decrease between

Part A and Part B on the univariate and multivariate tasks. It is possible they were more familiar with the visualisation system by the time they were answering tasks for Part B which meant they did not need to click as frequently.

With the number of mouse clicks for interactive visualisations being considerably lower on average compared to the 2D and 3D visualisation, it can be suggested that the interactive visualisation system (ProVis) represents a more efficient visualisation technique for analysis of GWR results.

Figure 5.2 Average number of Pixels of Mouse Movement of each application for Part A and Part B

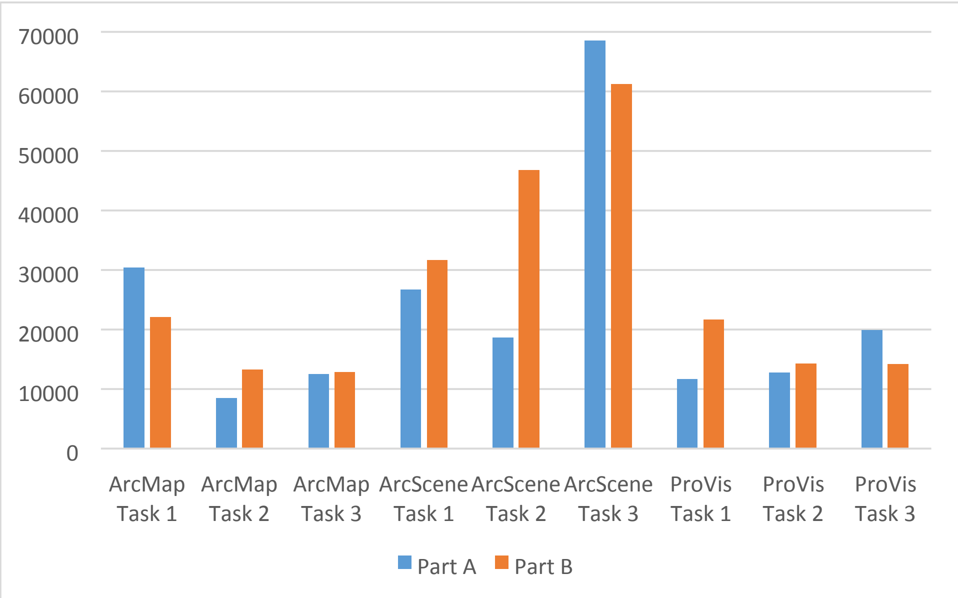


Figure 5.2 shows the average level of mouse movement on each task for each visualisation. The obvious outlier in this figure is the multivariate task (Task 3) for the 3D visualisation. The levels of mouse movement are considerably higher here because of the interaction with the visualisation. It indicates the significantly greater effort required by participants to complete the most ArcScene task. ArcScene mouse movement levels are higher for every task compared to ArcMap and ProVis.

Despite the interactive nature of ProVis, the level of mouse movement is comparable with the 2D mouse movement. Mouse movement levels for ArcMap and ProVis are similar for Tasks 2 and 3. ProVis required less movement for Task 1 than ArcMap.

5.5 Participant Task Perception

Quantitative survey scores provide an insight into the perception participants have of the tasks completed during the experiment. These survey scores are based on the participant's opinion of how they felt they performed on every task. The scores attained on each task are provided in Tables 5.8, 5.9 and 5.10. Part B average scores indicate participants were more satisfied with their performance on univariate and bivariate 2D visualisation exercises. This could be a result of increased familiarity with this visualisation type since all 2D tasks and the first two tasks of the 3D visualisations could be answered using a top down visualisation view. Participants were less confident with their performances on 3D and on Part B than their Part A counterparts. In the literature, 3D visualisations are known to be poor performers for analysis but in the case of GWR analysis and interpretation of the survey results here, it suggests 3D visualisations are just as useful.

When analysing the knowledge groups for surveys, they indicate a trend of increased satisfaction with performance when progressing from Part A to the more complex Part B. This seems counterintuitive because Part B's dataset is much larger in size than Part A's. It could suggest participants are more familiar with the visualisations the second time around. It is only when you compare correctness ratios and task times that the issue of perceptual scalability becomes visible. In any case it is interesting to compare measures of user perception to actual task time and task correctness, it is clear a gap exists between these.

Table 5.8 Average Survey Scores for 2D Visualisations, Parts A and B

	Part A			Part B		
	T1	T2	T3	T1	T2	T3
Av Ease	3.07	2.86	2.57	3.71	3	3.79
Av Speed	2.86	2.93	2.86	3.43	3.14	3.79
Av Confidence	3.07	2.86	2.64	3.5	2.93	3.14

Table 5.9 Average Survey Scores for 3D Visualisations, Parts A and B

	Part A			Part B		
	T1	T2	T3	T1	T2	T3
Av Ease	3.1	3.57	1.8	3.5	2.214	2.2
Av Speed	3.2	3.5	1.7	3.5	2.214	2.2
Av Confidence	3.3	3.29	2.4	3.1	2.286	2.1

Table 5.10 Average Survey Scores for Interactive Visualisations, Parts A and B

	Part A			Part B		
	T1	T2	T3	T1	T2	T3
Av Ease	3.6	3.29	2.3	4	3.286	2.7
Av Speed	3.4	3.5	2.9	3.9	3.214	3.2
Av Confidence	3.6	3.36	2.5	3.6	2.571	2.6

5.6 Insights from Post-Experiment Semi-structured Interviews

Short semi-structured interviews were carried out at the end of the experiment with each participant. Interview data was used to investigate and to try understand more qualitatively participant's impression about the visualisations tested. For greater validation, results were compared to the quantitative measurements of correctness and time to assess if and how users' perception corresponds to their performance.

Table 5.11 presents a set of coded categories which indicate the general themes most commonly observed during the post experiment interviews. While a positive theme was evident in the top five topics of each visualisation type there were clearly some outlying results. The interactive visualisation system received the most positive set of remarks, while the 3D visualisation followed the trend of the survey analysis by featuring as the least preferred choice for completing tasks according to the points raised during the post-experiment participant interviews.

The 2D visualisation was noted as being preferred for Task 1, or univariate visualisations. The 2D visualisation is also an adequate choice to answer a task that involves just one parameter estimate and is reflected in the number of times that 2D visualisation received a positive comment from participants when they spoke about Task 1. This makes sense when participant familiarity is taken into account, particularly when it is the simplest task. It is important to remember however that GWR output analysis often involves more than a single univariate parameter estimate but this served as a good baseline of feedback.

The trend of the 3D visualisation largely continues through the participant interviews. There was a recognisable difficulty in using the 3D visualisation to complete the tasks. 3D was mentioned 18 times in a general negative fashion, meaning there was some problem encountered with it. More specifically it was possible to note that the 3D visualisation was difficult to navigate in terms of visualisation rotation by eleven out of thirteen participants, and height was seen as an issue in the 3D visualisation by eleven participants also. As the tasks became more complex particularly with the multivariate task it could be suggested that visualisation height issues coupled with difficulties with rotation ultimately lead to occlusion problems and thus the comments received from participants post-experiment. There were some positive comments in general related to the 3D visualisation however, and these are most likely associated with the simpler tasks where it is technically possible to avoid the use of the rotation capabilities. It effectively places the 3D visualisation on a similar line with the 2D visualisation for Task 1 and Task 2, and the majority of participants were noted as having avoided the use of full 3D functionality in order to answer the first two tasks. This would explain the number of general positive comments number, and also the fact that 3D was mentioned as being preferred to answer Task 2.

The Interactive visualisation comments were largely positive as depicted by the top five remarks made when this visualisation type was discussed during the interviews. A general positive comment was made over a dozen times by participants relating to interactive visualisation which is double the number of general positive comments made for the two other visualisations. The interactive visualisation was mentioned as the preferred or easiest to use visualisation six times each. While both the 2D and 3D visualisations were mentioned as being preferred, their preference depended on the type of task they had to complete. It is interesting to note this 'ease of use' theme given it is the visualisation cited by most participants as being the least familiar to them.

Table 5.11 Frequency of Selected Interview Topic Content

2D (times mentioned)		3D (times mentioned)		ProVis (times mentioned)	
2D Map difficulty	7	3D problem	18	Int positive	13
2D preferred for task 1	6	3D rotation/navigation problem	11	Int preferred	6
2D positive	6	3D height problem	11	Int easiest to use	6
2D positive task 1	5	3D positive	7	Int PCP positive	5
2D problem	4	3D preferred for task 2	4	Int preferred for task 2	4

Participant satisfaction in using 2D visualisations to answer a univariate question is visible in Table 5.11 adding to the positive measured participant performance levels obtained through methods already discussed. Participant preference for the bivariate task was split between the 3D and Interactive visualisations. Overall though, the Interactive visualisation was cited as being the easier to use in many instances by participants and was generally preferred to answering tasks with the aspect of familiarity the apparent obstacle to placing the Interactive visualisation as the clear preferred choice. The integration of this visualisation type also led users to say it was the fastest to use compared to the other two visualisation techniques in order to complete tasks. Despite these positive comments on speed to complete a task, 3D and Interactive visualisations were mentioned in relatively equal measure when participants were asked which visualisation type they would most like to use to answer multivariate tasks.

In terms of negative comments, there were some for each of the visualisation types. The 2D visualisations were not without problems which participants highlighted when using the choropleth maps to answer more complex bivariate and multivariate tasks.

3D visualisations were the most problematic overall according to the participants. Participants found these visualisations difficult to understand or confusing, yet using this visualisation type was less of an issue for the univariate and bivariate tasks compared to the multivariate task that required use of the 3D rotation feature which caused the most problems for participants according to interview feedback. This occurred once participants were required to navigate the 3D surface through the use of the rotation feature on the multivariate task.

The interview content can be linked to actual performances on tasks, not just perceptual. 3D visualisations did not perform as well as the 2D and interactive visualisations on task times, and correctness ratios. The lack of positive commentary using the 3D visualisation to answer multivariate tasks is also clear and this can be related to the survey perception scores recorded for Task 3 in both Parts A and B. They are among the lowest averages between all participants.

Although the top 5 mentioned topics relating to Interactive visualisations were positive, some participants did encounter problems. These problems occurred mainly due to a lack of familiarity with the system and the PCP was highlighted in particular because it is not such a well-known visualisation technique compared to other methods used in the experiment, for example, a choropleth map. Problems with more complex tasks of a bivariate and multivariate nature were linked to this theme of familiarity. In fact, the aspect of familiarity was a common theme through all interviews.

5.7 Discussion of Results

Referring back to Chapter 1, Hacklay and Tobón (2004) asked "What is a visualisation good for?". This is what this experiment helps to answer for analysis of GWR outputs. To this end, the usefulness of all three visualisation techniques need to be considered and compared.

The interactive visualisations performed well overall. This is indicated by the metrics discussed in this chapter. Correctness levels are generally higher for the interactive visualisation compared to the 2D and 3D visualisations. Participants with a greater perceived knowledge of GWR (G6) recorded higher correctness scores which indicates that a level of GWR knowledge will improve proficiency in successfully completing tasks. This is not to say that a higher level of GWR knowledge results in faster completion times because those with less GWR knowledge (G5) were faster to complete tasks. Where approximately 17% of Group 6 recorded higher levels of correctness, approximately 17% of the same group completed tasks faster than Group 5. Perhaps there is a medium between speed of completion and levels of correctness.

The post task results which record participant perception offer an insight into the likelihood that they would want to use one particular visualisation type over another, this is because they rate their own ease of task completion, speed of task completion and confidence in their task answer. The 3D visualisation can be ruled out of this perception comparison because perception scores for the 3D visualisation were considerably lower than the 2D and Interactive Visualisation, with the exception of the univariate task which was more on par with the other two visualisation types and did not require any 3D rotation or typical 3D interaction. Participants were slightly happier with their performance using the 2D visualisation compared to the Interactive visualisation according to their perception scores, despite exerting a greater effort to complete the task in terms of mouse movement and mouse clicks. Overall task completion times were higher for 2D visualisations compared to Interactive visualisations.

One emerging trend from these results is that 3D visualisations performance is poorer than 2D and Interactive visualisations. Despite this trend, some participants speak favourably of the 3D visualisation. In some cases the 3D visualisation is preferred to the Interactive Visualisation. This could be due to the visual appeal of the 3D method, it also encourages interactivity for the most complex experiment task type (Task 3). Ultimately, participant preference is likely to decide whether one visualisation type is used over another without a more comprehensive comparison this experiment provides. This is reflected in Table 5.11 where the frequency that the 3D visualisations is spoken about positively is similar to the 2D visualisation. However, the frequency of negative 3D visualisation is three times higher than the frequency of positive 3D visualisation comments.

It is important to remember that the specific visualisation systems used in this experiment are not being assessed. They are merely chosen mediums which facilitate the evaluation of the visualisation techniques. As mentioned in Chapter 1, Plaisant (2004) said visualisation tools may not be specifically designed with the needs of the user in mind. This is why it is important to complete this assessment, to better understand which visualisation technique is best for interpretation and analysis of GWR outputs. Based on the results, there are clear indications that the interactive visualisation performs better overall and that it presents the least number of problems according to participant feedback.

6: Results and Analysis: Experiment Two

The Second Experiment began on the 4th November 2012 and finished on the 18th November 2012. Experiment Two was designed and carried out following results of the research and output derived from Experiment One, as discussed earlier in this research thesis. Results indicated that a visual scalability effect may be present. Visualisations play an essential role in dealing with large data sets. Despite this, visual scalability analyses are almost entirely absent according to Eick and Karr (2002) and this result aims to contribute to this gap in knowledge. This visual scalability effect could be dependent on a participant's level of expertise, the map scale encountered, or the order in which a set of spatial distributions are presented to a participant. These aspects will now be discussed.

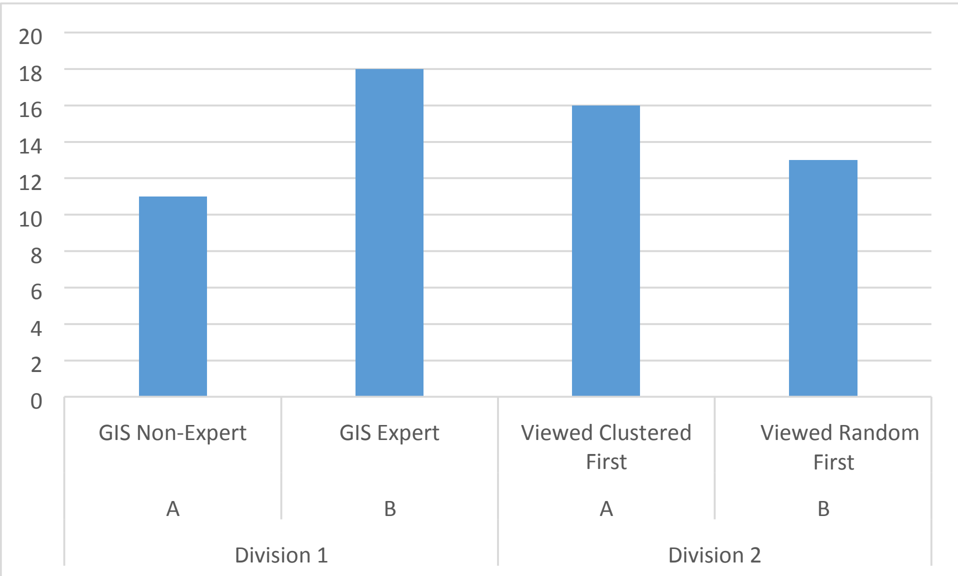
Following from Chapter 5, this chapter details the results of Experiment Two. The chapter is split into seven key sections. The first describes various divisions for the purposes of data analysis. Section two analyses the effect of the various spatial unit scales used in the visualisations in the experiment. Possible visual scalability effects are discussed here using correctness ratios, time to answer and eye fixations as indicators. Section three focuses on the categorisation of participants according to GIS expertise. Similar to section two, the discussion here is shaped around key indicators such as correctness ratios, spatial unit scales, eye fixations, time to answer and user perception of tasks. Section four explores the data in relation to type of visualisation first encountered in the experiment. Again for comparability of analyses, indicators of correctness ratios, time to answer, eye fixations and user perception form the framework for the discussion. Section five outlines the results of the experiment with two paired maps and section six a visual comparison of fixation patterns. Finally, section seven concludes the chapter with an overall discussion of the findings and context in relation to the experiment aim.

6.1 Data divisions for data analysis

Data were analysed using three different divisions of tasks for participants. The first division (section 6.2) was defined by task properties and was by the number of spatial units in the display. The other two divisions were defined by participant properties and were based on either the participants' expertise in GIS (section 6.3) or on the order of viewing in which they saw the tasks (section 6.4). In this section we describe these three divisions.

Figure 6.1 shows the division of participants into two main divisions: Division 1 and Division 2. Division 1 and 2 are separated into two participant groups which are marked A and B. Division 1 is separated according to the participant's perceived level of expertise in Geographic Information Science (GIS), group 1A being GIS Non-Experts and group 1B being GIS Experts. There are a total of 11 participants in the "GIS Non-Expert" and 18 in the "GIS Expert" groups. Division 2 is divided according to the order in which Participants encountered 'Clustered' (Group 2A) and 'Random' (Group 2B) visualisations. There are 16 participants in the Clustered group (these saw a Clustered visualisation first) and 13 in the Random group (these saw a Random visualisation first).

Figure 6.1 Participant Divisions for Experiment Two



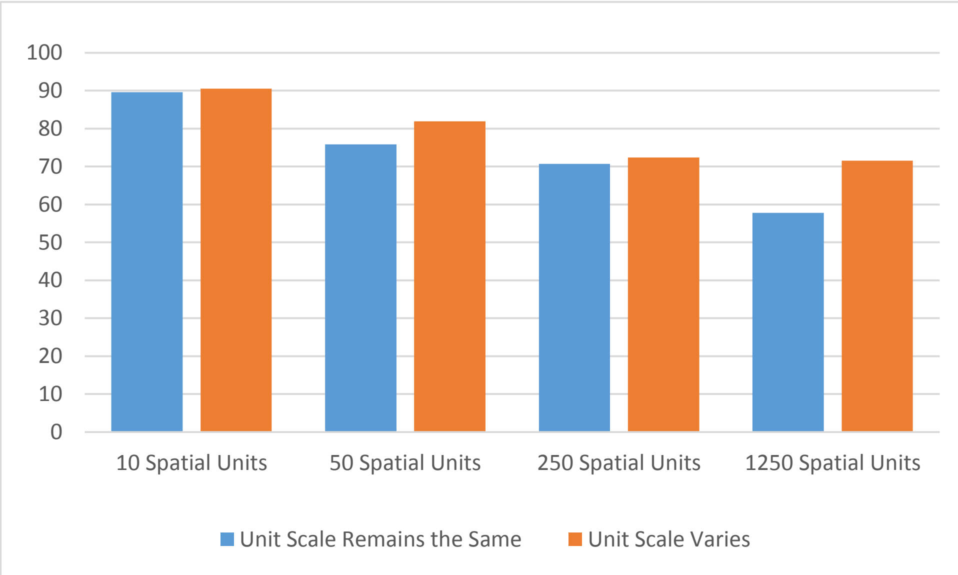
6.2 Spatial Unit Scale

In this section we analyse the effect of the various spatial unit scales used in the visualisations of the experiment for all participants as one group. Possible visual scalability effects are discussed here using correctness ratios, time to answer and eye fixations as indicators of complexity in each set of spatial unit maps.

6.2.1 Correctness Ratios across each spatial unit scale

Comparing results within each spatial unit type scores (i.e. between spatial units that vary in their size and between spatial units that do not vary in their size) highlights several aspects to correctness ratios (see Figure 6.2). Participant correctness scores clearly decrease from the 10 spatial unit display to the 1250 spatial unit display, regardless if the unit scale varies or remains the same. Maps in which the spatial unit size varies tend to have a higher correctness score than maps in which the spatial unit size remains static or unchanged.

Figure 6.2 Spatial Unit Scale: Correctness Ratios (in %)

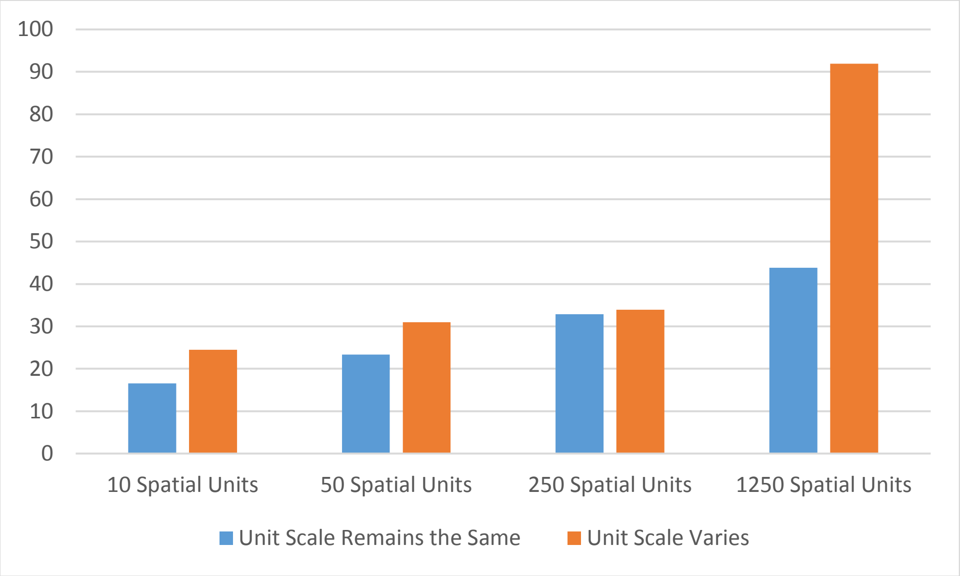


6.2.2 Eye Fixations on each spatial unit scale

Figure 6.3 shows a bar chart of the average number of fixations per spatial unit scale. The average number of fixations is higher for maps in which the spatial unit size varies compared to the maps in which the spatial unit size remains the same. This is true for each spatial unit scale and indicates that a greater cognitive effort is likely required to answer tasks in which the spatial unit size varies. The average number of fixations is greater for almost all of the visualisations in which the size of spatial units vary. The difference in the average number of fixations between displays where unit scale varies and those where it remains the same is smallest for the 250 spatial units and largest at the 1250 spatial unit scale. This suggests that more complex Clustered visualisations present a more complex challenge of pattern recognition. A similar finding was outlined in the literature review sections of this thesis

(Abbott, 1995; Hayhoe and Ballard 2005; Tufte, 2007; Swienty and Reichenbacher 2008; Çöltekin et al., 2009; Çöltekin et al. (2010).

Figure 6.3 Spatial Unit Scale: Number of Fixations

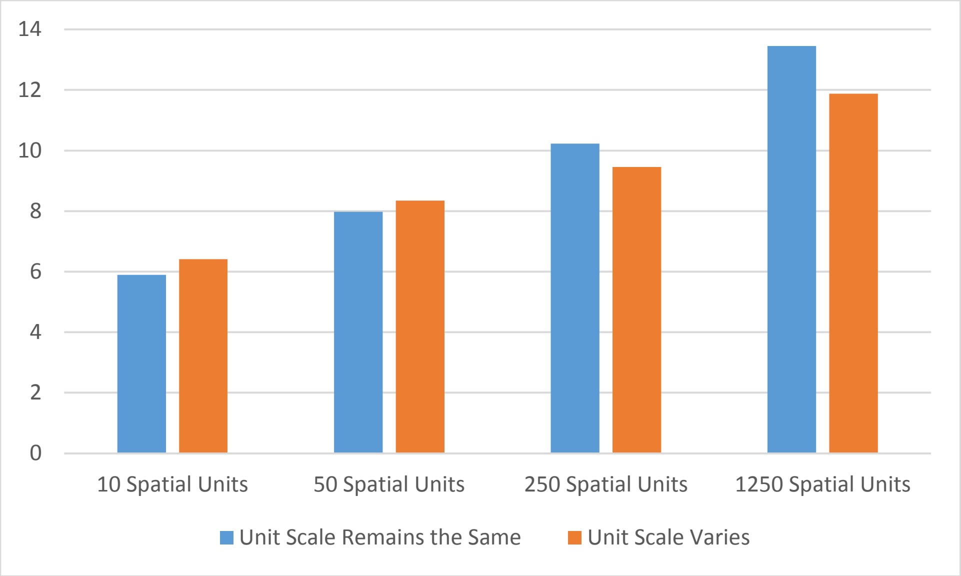


6.2.3 Time to Answer within each Spatial Unit Scale

For visualisations in which the spatial unit size remains the same, there is a clear increment in the time required by participants to achieve task completion (Figure 6.4). The mean values for time to task completion for varying spatial unit size maps shows there is an increase through each spatial unit scale.

The average time taken to complete a task is greater on non-varying spatial unit sized maps for 10 and 50 unit scales. This changes once the scale reaches the 250 scale (third level of complexity). For the 250 and 1250 spatial unit scales, tasks were completed faster for the varying unit size maps. It is interesting to note that it took participants a greater amount of time to complete tasks on the 10 and 50 spatial unit scale maps in which the spatial unit size is designated as “Large”.

Figure 6.4 Spatial Unit Scale: Time to Answer a Task (1 second = 1.0)

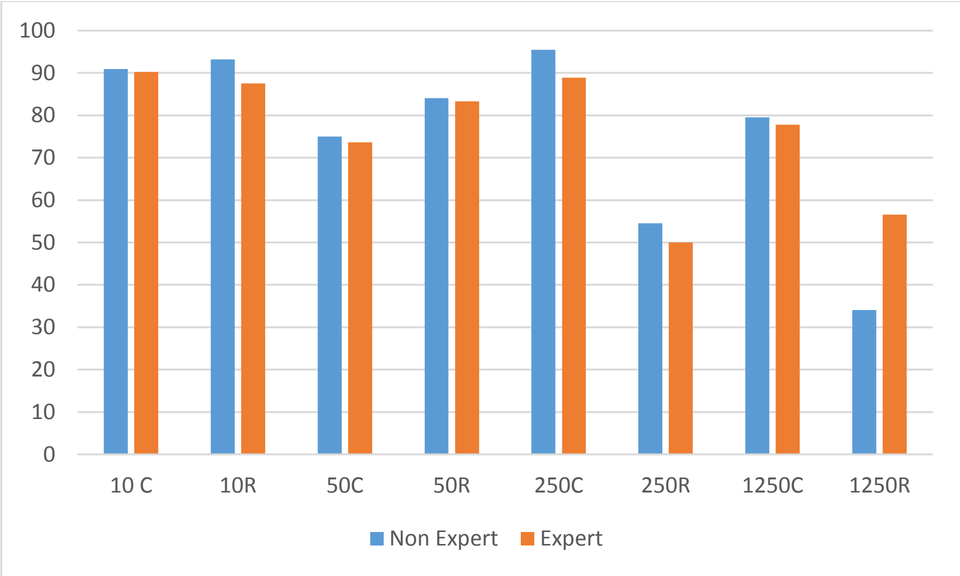


6.3 Division of Expertise

6.3.1 Division of Expertise: Correctness Ratios

Figure 6.5 shows the average correctness ratios for participants grouped according to their GIS expertise, grouped by the type of maps (Random vs. Clustered) and number of spatial units (10 to 1250). Non Experts and Experts have similar average correctness values, with the exception of the Random 1250 Spatial Unit Map. The average Correctness ratios for Experts were slightly lower overall compared to Non-Experts. This finding suggests that participant level of expertise does not result in a difference in performance of completing a task correctly, however, a more complex (based on map size) or randomly distributed data may result in increased levels of difficulty.

Figure 6.5: Division of Expertise: Average Correctness Ratios (in %)



We were further interested in the distribution of correctness values in order to investigate the spread of intra-participant differences. To do this, we created box plots for correctness values, per map type and size, where participants were grouped according to their expertise (figures 6.6. and 6.7).

Correctness ratios for Non-Experts in Figure 6.6 highlight the spread of Non-Experts' correctness values. There is a general compactness to the results with correctness levels occupying an approximate 20% band or less for five of the eight boxplots below. The exceptions being Clustered and Random 50 spatial unit maps and 1250 spatial unit maps which occupy an approximate 40% spread of correctness. The wider boxplot for 1250 units suggests a higher level of difficulty, associated with larger amounts of data and increased complexity of data distribution.

Figure 6.6 Division of Expertise: Correctness Ratios for Non-Experts (in %)

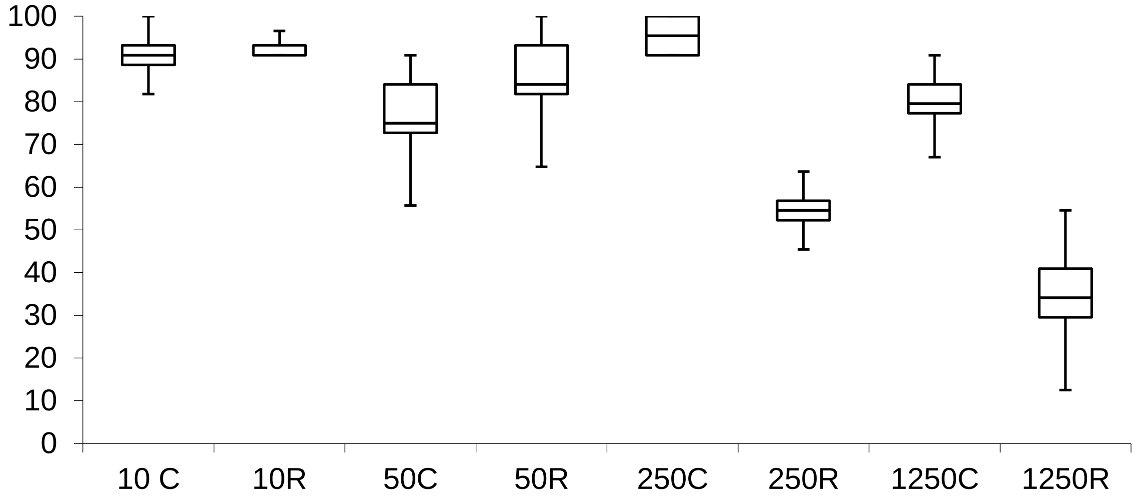


Figure 6.7 shows the same boxplots for GIS experts. Interestingly, the spread of values here is much larger than for non-experts, in particular for maps of the following three types: Clustered 50, Clustered 1250 and Random 1250. The wider spread of values could indicate the variety in levels of knowledge in the Expert division, compared to the Non-Expert division.

Figure 6.7 Division of Expertise: Correctness Ratios for Experts (in %)

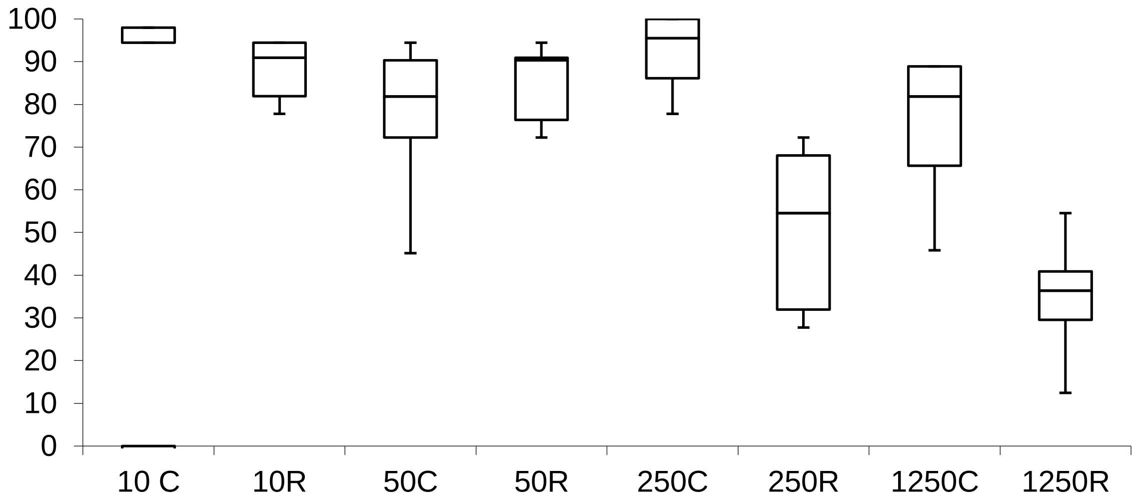
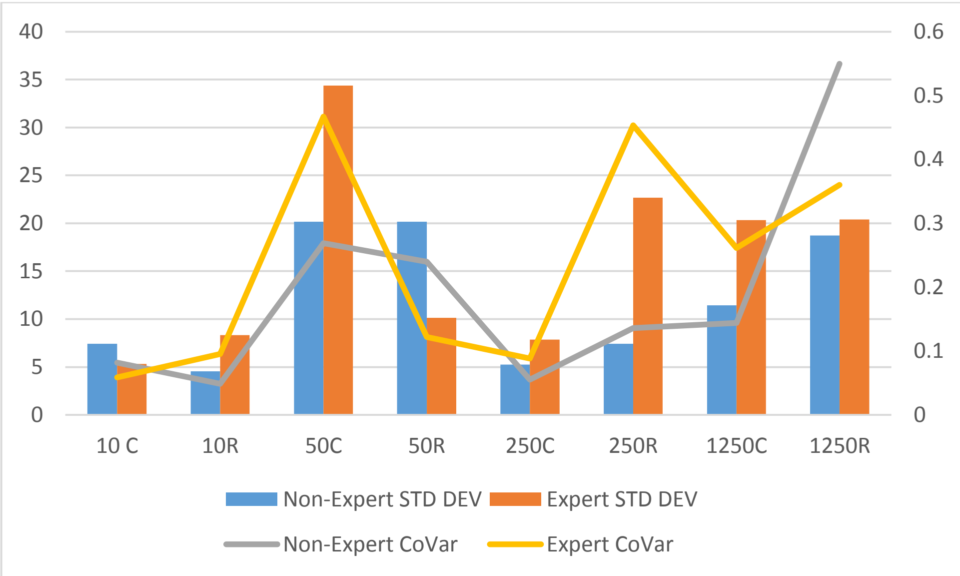


Figure 6.8 shows the Standard Deviations and Coefficient of Variation for Spatial Unit Scale according to division of expertise. Standard Deviations for Non Experts and Experts indicate there are considerable variations in participant correctness ratios. Experts have a higher level of consistent variation for more complex Clustered and Random maps (250 and 1250 spatial unit maps).

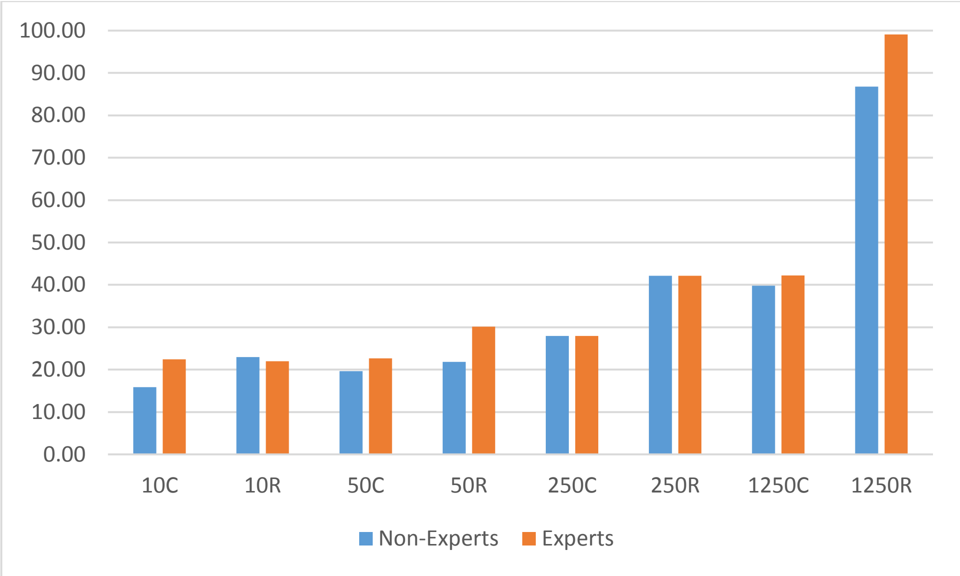
Figure 6.8 Division of Expertise: Standard Deviation and Coefficient of Variation for Correctness Ratios (in %) – Left axis for Standard Deviation, Right axis for Coefficient of Variation



6.3.2 Division of Expertise: Number of Fixations

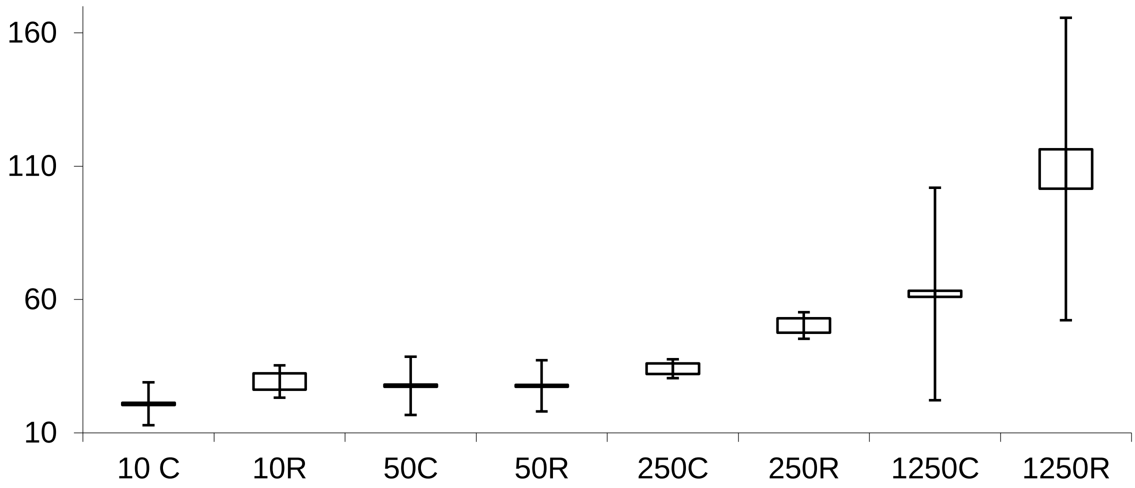
The second metric to assess the complexity of tasks was the number of fixations. A larger number of recorded eye fixations can indicate a higher cognitive load and can be considered as a proxy for complexity of the task. Figure 6.9 shows this effect clearly, i.e. there is an increase in the average number of fixations and therefore cognitive load on participants when faced with more complex data. 10 spatial unit maps (both Clustered and Random) and 50 spatial unit Clustered maps have a similar average number of fixations, which could be described as the first level of difficulty. The second level begins at the Random 50 spatial unit maps and the third is reserved for the Random 1250 spatial unit maps.

Figure 6.9: Number of Fixations for Non-Experts and Experts



To investigate distributions of the fixation numbers within each participant group, we again created relevant boxplots (figures 6.10 and 6.11). Figure 6.10 shows the distribution of eye fixations for Non-Experts on every spatial unit scale. The figure has been split into two graphs in order to visualise the two very different scales in one figure. Again, there is an observable increase in the number of fixations vs. the increase in the number of spatial units and there are some variations in the number of fixations recorded for participants. Fixation counts are spread for 50 and 1250 spatial unit maps in particular, while fixation counts for 250 spatial unit maps are more similar between participants. There are more fixations for Random maps (except the 50 spatial unit map) than for Clustered maps which can indicate difficulty with search patterns or displayed data (Çöltekin et al., 2009).

Figure 6.10 Division of Expertise: Boxplot of the Number of Fixations for Non-Experts



For experts (see Figure 6.11), the spread of data is very large for the Random 1250 maps, which is due to a pair of individuals with substantially more or substantially less fixations on Random 1250 maps (the two whiskers in the plot). There is an observable increase in the number of fixations from 10 to 1250 maps, with some Clustered maps (e.g. 50 or 250) requiring less fixations or less cognitive effort to complete tasks.

Figure 6.11 Division of Expertise: Boxplot of the Number of Fixations for Experts

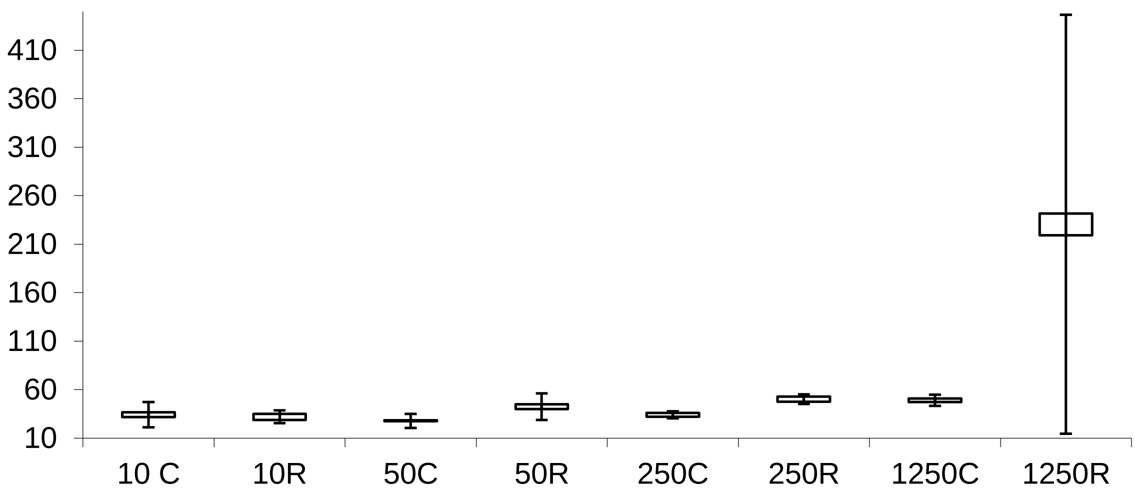
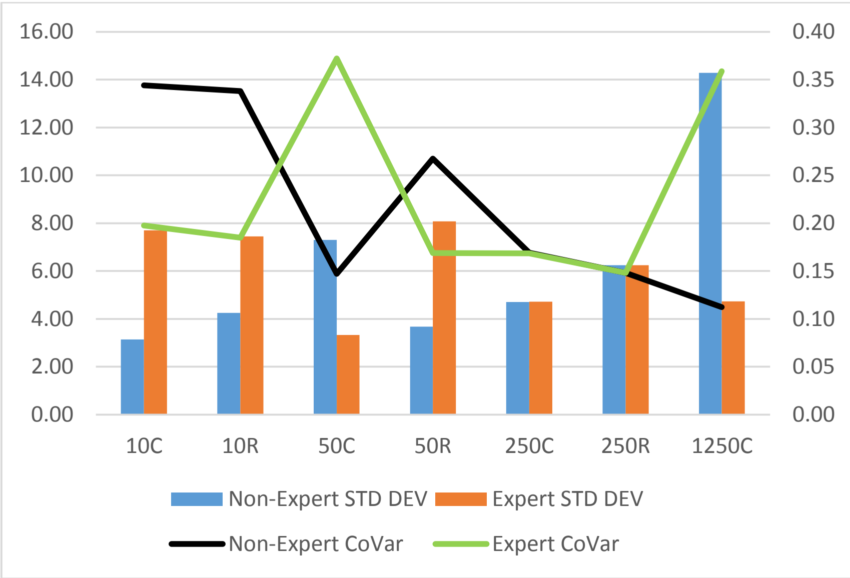


Figure 6.12 Division of Expertise: Standard Deviation and Coefficient of variation for Number of Fixations – Left axis for Standard Deviation, Right axis for Coefficient of Variation



In Figure 6.12 the Random 1250 maps were removed due to the outlying nature of the values. The Standard deviation for Random 1250 maps were 68.58 and 82.85 for Non-Experts and Experts respectively, while the coefficient of variation is 0.84 and 0.79 respectively for Non-Experts and Experts. The coefficient of variation and standard deviation values follow a similar trend, where values are higher for standard deviations, values for coefficient of variation are higher. Expert values for standard deviations and coefficient of variation are higher in general than Non-Experts for less complex 10 and 50 maps, except for the clustered 50 spatial unit maps. This greater spread on less complex maps could indicate some participants found tasks straightforward to complete, while others required slightly more time. Standard Deviations are unusually similar for Non-Experts and Experts on 250 spatial unit maps, with a slightly greater spread for Random 250 maps. Even with the 1250 Random maps removed, the Non-Expert Clustered 1250 maps have a great spread of values according to standard deviations and coefficient of variation. This suggests a possibility of participants finding complex data difficult to interact with, particularly if they are less experienced. The standard deviation and coefficient of variation values of Non-experts on 1250 maps are at least double the values of 10, 50 and 250 spatial unit maps. For Experts, standard deviation and coefficient of variation values on Random 1250 maps are also double the values of 10, 50 and 250 spatial unit maps.

6.3.3 Division of Expertise: Time to Answer

The third metric to explore was the time it took each participant to find the answer to the task and click on respective areas on the map. Average times to answer are shown in Figure 6.13, broken down by participants' expertise. There is a clear general increase in time taken to answer a task for both divisions of expertise from 10 to 1250 spatial units. Non-Experts did not perform significantly worse than Experts and outperformed Experts in some instances (e.g. 250 and 1250 Clustered). Similar to Correctness ratios, it can be said that participant level of expertise does not necessarily result in better performance. Experts take more time to answer the task, possibly because they feel they have an ability to analyse data in more depth.

Figure 6.13 Division of Expertise: Time to Answer a Task (1 second = 1.0)

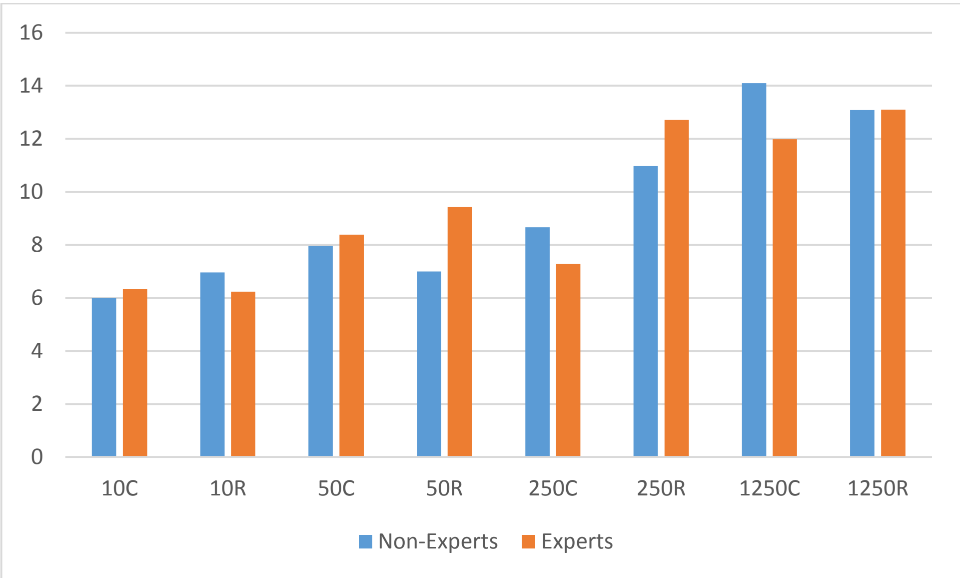


Figure 6.14 shows a general increase in the time required to complete a task from 10 to 1250 spatial maps, similar to other Figures in this subsection on Time to Answer a Task. The greatest spread of time occurs on the 1250C map. Upon studying the boxplot on the Number of Fixations (Figures 6.10 and 6.11) it appears the level of difficulty for participants is similar between Random 250 and Clustered 1250 maps. However, in Figure 6.14 there is an observable increase in the spread of participant times to complete a task.

Figure 6.14 Division of Expertise: Box plot of Time to Answer a Task (1 second = 1.0)

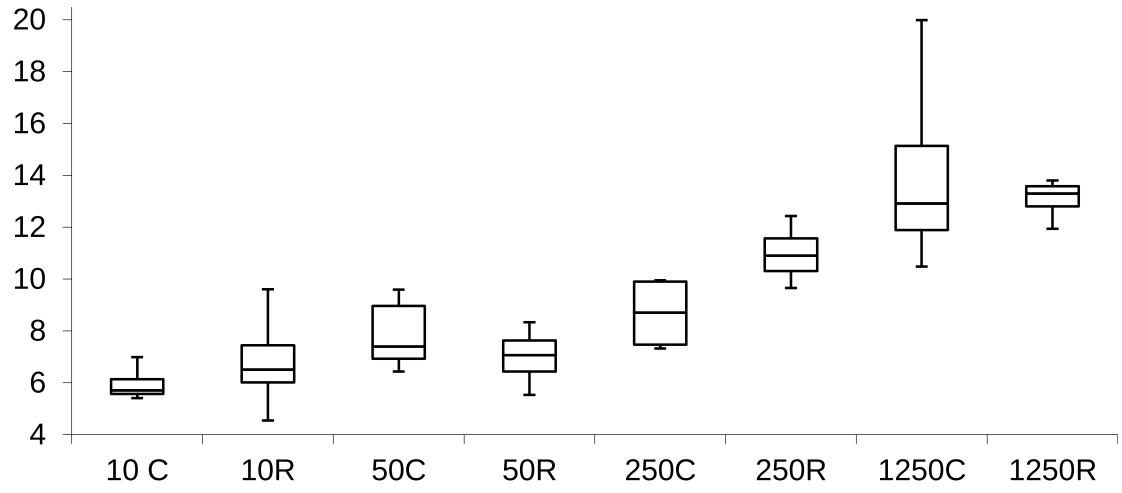


Figure 6.15 shows that there is a general trend towards an increase in time to completion as participants completed tasks from 10 spatial unit maps to 1250 spatial unit maps. The spread of time to complete a task falls within a tight band for several maps. The greatest spread occurs on the Clustered 1250 maps which indicates there could be some difficulty involved in completing more complex maps. This interpretation is also observable for Random 250 maps where the band spread is greatest. Similar to Random 250 and Clustered 1250 maps, Random 1250 maps also recorded longer times on average compared to other map scales.

Figure 6.15 Division of Expertise: Box plot of Time to Answer a Task (1 second = 1.0)

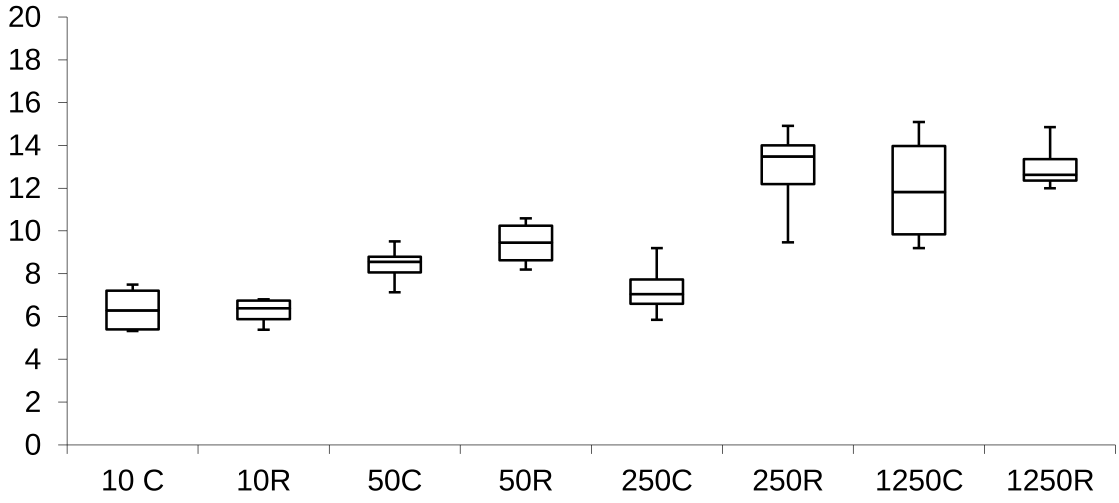
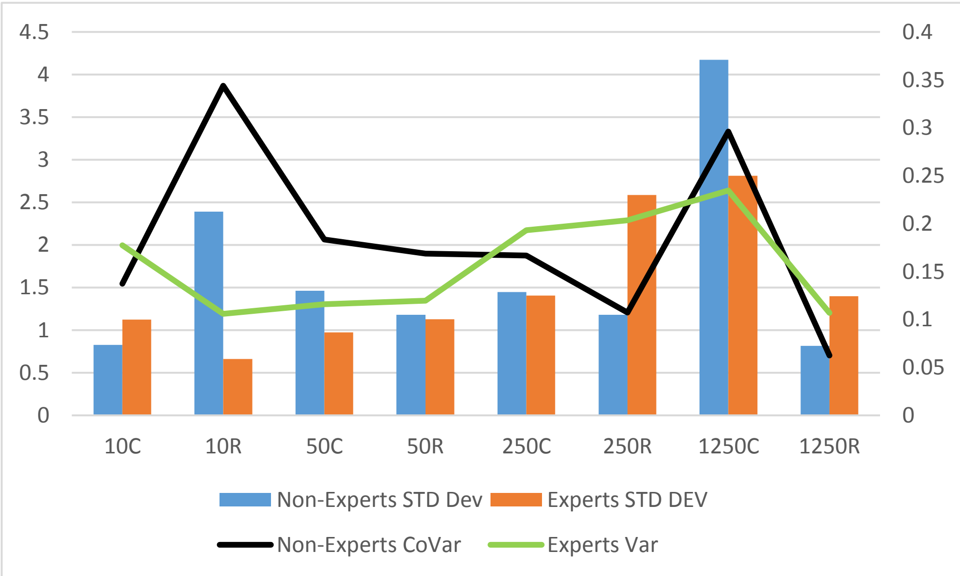


Figure 6.16 highlights the Coefficient of Variation (lines) and Standard Deviation (bars) levels. The Non-Expert Coefficient of Variation is higher for Random 10 spatial unit maps than the coefficient for experts. The coefficient of Variation levels are also lowest for Random 1250 maps for Non-Experts and one of the lowest for Experts. The spread of relative standard deviation predominately lies between 0.1 and 0.3. The standard deviation spread is higher overall for 250 and 1250 maps. There are less than 2 seconds of standard deviation for 12 of the 16 standard deviation values indicating a close spread of values. An exception being the outlier for Clustered 1250 maps, which matches another standard deviation outlier for Random 1250 maps in Figure 6.13 (Number of Fixations). The higher standard deviation value for the 1250C map suggests there are some participant results that do not fit into the general trend.

Figure 6.16 Division of Expertise: Standard Deviation and Coefficient of Variation for Time to Answer – Left axis for Standard Deviation Right axis for Coefficient of Variation



6.3.4. User perception of tasks

This section presents the results of the post-task survey, where the participants were asked to rank the level of ease, the perceived speed of solving the task and their confidence in the correct solution. This was done on a scale from 1 to 5 with 1 being the worst and 5 being the best (e.g. for task ease, 1 is very difficult and 5 is very easy). Figure 6.17 shows the average results broken down by participants' expertise. Non-Experts were less confident about their answers and felt the questions were a little more difficult than Experts (see Figure 6.16). This indicates a level of expertise will determine the type of perception participants have when completing tasks. The indications on the speed of task completion are that Experts feel they take longer, which is consistent with Average Time to Task Completion metrics already discussed in this section.

Figure 6.17 Division of Expertise: Participant Task Perception (rating out of 5)

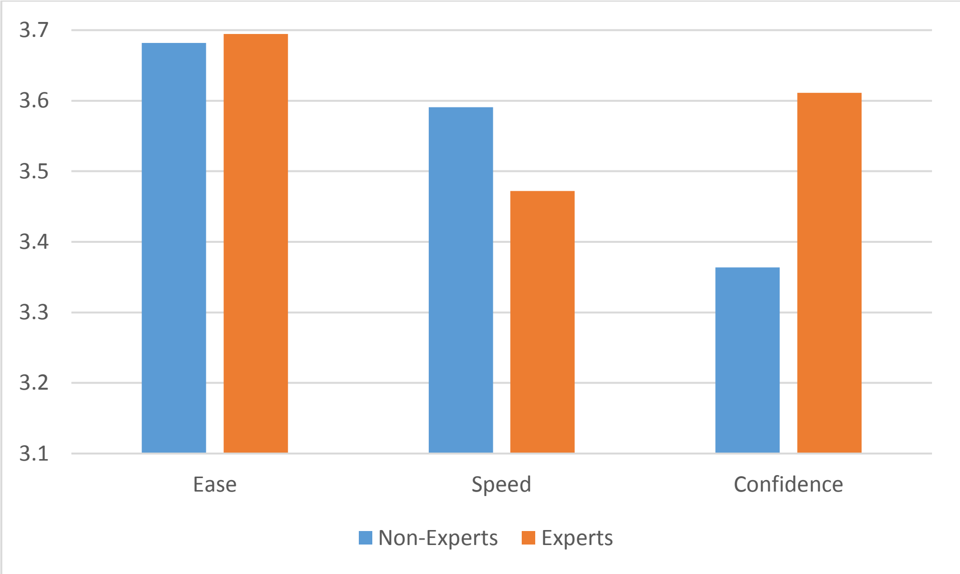
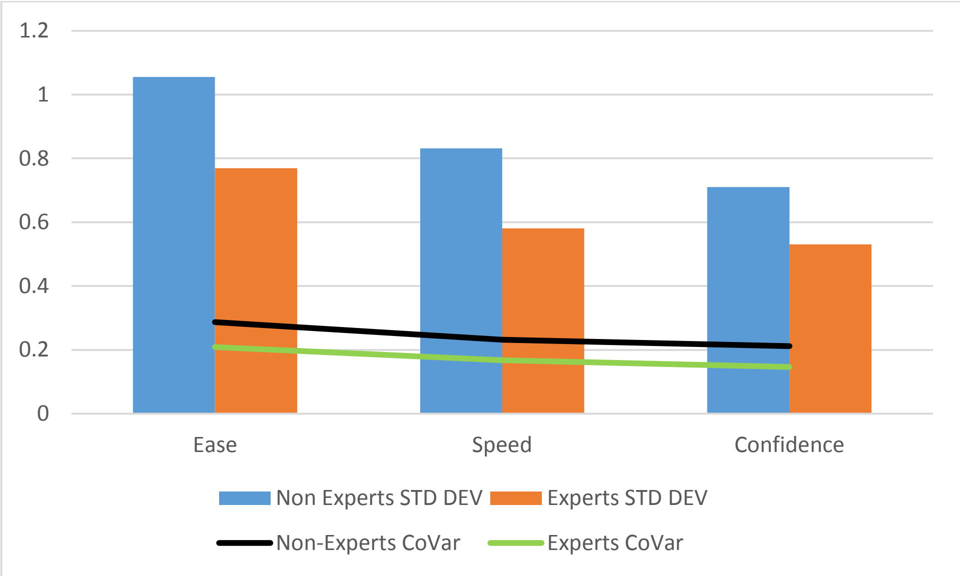


Figure 6.18 shows the standard deviation and coefficient of variation for the survey. The coefficient of variation levels are similar with minor changes for each question. There is a greater spread in the perception of ease of completion for both divisions of expertise according to the Standard Deviation for Non-Experts and Experts. Expert Standard Deviations are lower compared to Non-Expert. This could indicate participants with less GIS experience will experience a greater degree of variation on the Ease of completion. There is also a greater spread of Standard Deviation values for Speed and Confidence levels of Non-Experts, the coefficient of variation is also consistently higher for Non-Experts. Again this indicates the potential for greater variation of perception among Non-Experts compared to Experts.

Figure 6.18 Division of Expertise: Participant Perception: Standard Deviation and Coefficient of Variation for Non-Experts and Experts



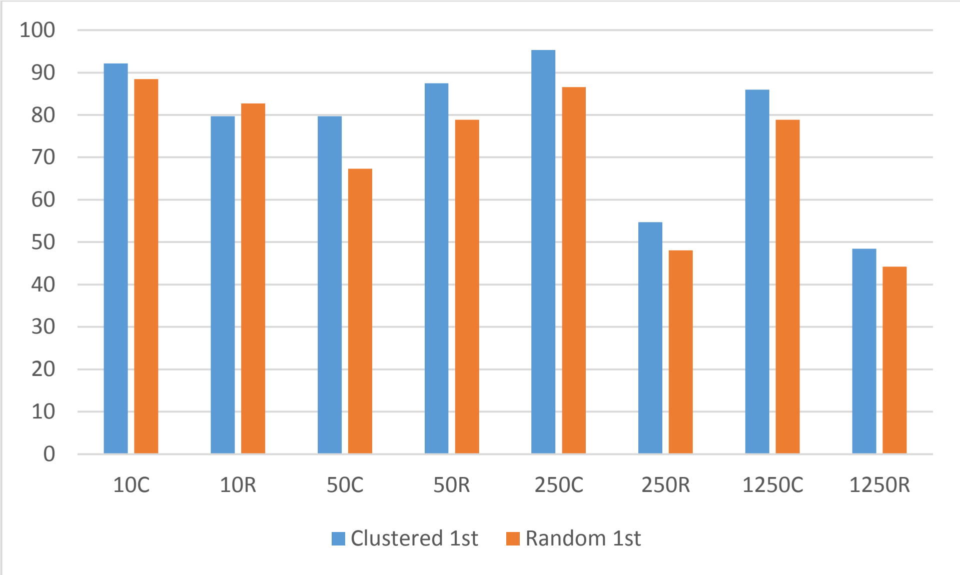
6.4 Order Encountered

In the next step we assessed the effect of the order in which participants were shown certain map distributions, in order to see if there was any learning effect present in the experiment. This learning effect is related to participants' increased familiarity with displayed maps as the experiment progresses. To counter this, we defined the schedule of the experiment so that participants were either shown a set of Random maps first, followed by the Clustered maps or vice versa. We assessed the performance of each group of participants (i.e. those that were shown Random maps first and those that were shown Clustered maps first) using the same set of metrics as before: correctness ratios, time spent on task and number of fixations.

6.4.1 Order Encountered: Correctness Ratios

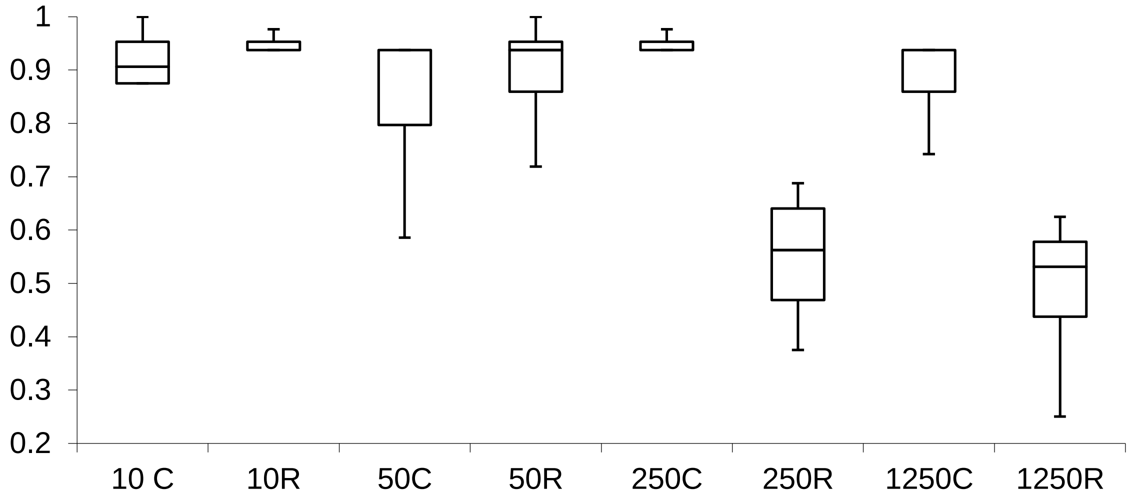
Figure 6.19 shows that the average correctness ratios for participants who saw Random maps first is lower for almost all map distributions and map scales, with the exception of the Random 10 spatial unit maps. This suggests a possible learning effect on participants who encounter the Random set of maps first, versus second. For participants who encountered Clustered maps first, their correctness ratios were higher on Random maps (again, except for spatial unit maps 10). The difference is minor overall, approximately 5-10%.

Figure 6.19 Order Encountered: Average Correctness Ratios (in %)



The spread of data in Figure 6.20 shows there are some outlying participants in terms of correctness of tasks. The results indicate that there is potential for participants to perform poorly when completing tasks, particularly for more complex maps where the two lowest average distributions can be observed (Random 250 and 1250 maps). This suggests there is a greater level of difficulty associated with more complex maps and with more complex data distributions.

Figure 6.20 Order Encountered Clustered 1st Boxplot (1 = 100%)



Participants who encountered Random maps first exhibit a greater distribution of performance levels than participants who encountered Clustered maps first according to Figure 6.21. Levels of correctness are shown to be lowest for the Random 250 and 1250 maps which correlates with results in Figure 6.22. The spread of these two maps is greater for participants who encountered Random maps first which suggests the potential to perform poorly is higher for more Random (or more complex) maps, particularly if Random maps are shown first.

Figure 6.21 Order Encountered Random 1st Boxplot (1 = 100%)

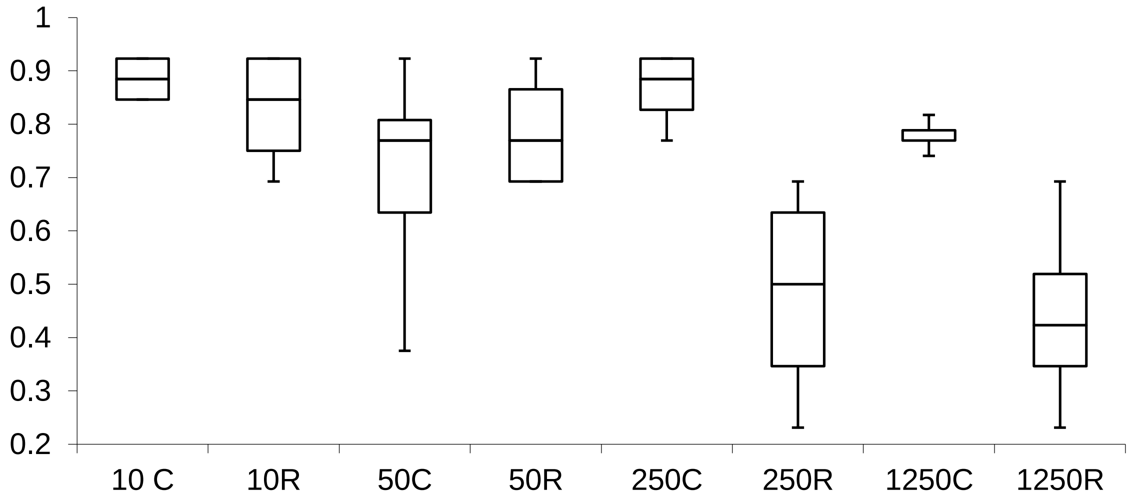
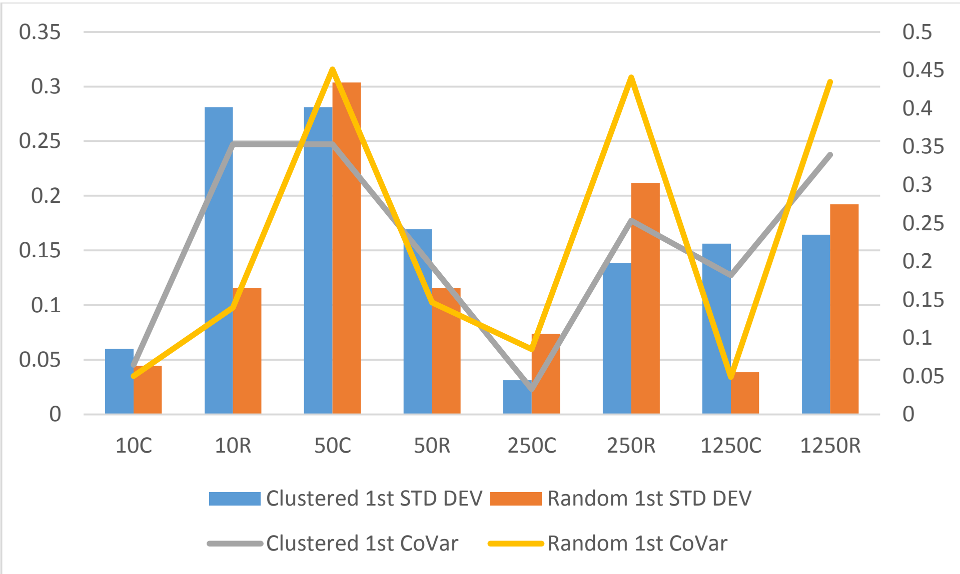


Figure 6.22 shows a considerable variation in the Standard Deviation and Coefficient of Variation values compared to most other charts of a similar kind in this Chapter. The Clustered 50 map exhibits the highest variation in participant correctness for Standard Deviation and Coefficient of Variation. For Random 10, 250 and 1250 maps, both statistics vary more than for their Clustered counterparts. This indicates a greater level of difficulty occurs for some participants for Random maps, values of correctness vary to a greater extent.

Figure 6.22 Order Encountered: Average Correctness Ratios: Standard Deviation and Coefficient of Variation: Left axis for Standard Deviations and Right axis for Coefficient of Variation



6.4.2 Order Encountered: Number of Fixations

Figure 6.23 shows the average number of fixations for the two groups of participants. Fixation counts for participants who were shown Random maps first are higher for most of the map distributions and map scales. In some instances there are considerable differences, with the 1250 maps showing the greatest difference. This indicates that participants who encountered Random maps first experienced a greater cognitive effort in order to analyse a map and provide an answer.

Figure 6.23 Order Encountered: Average Number of Fixations

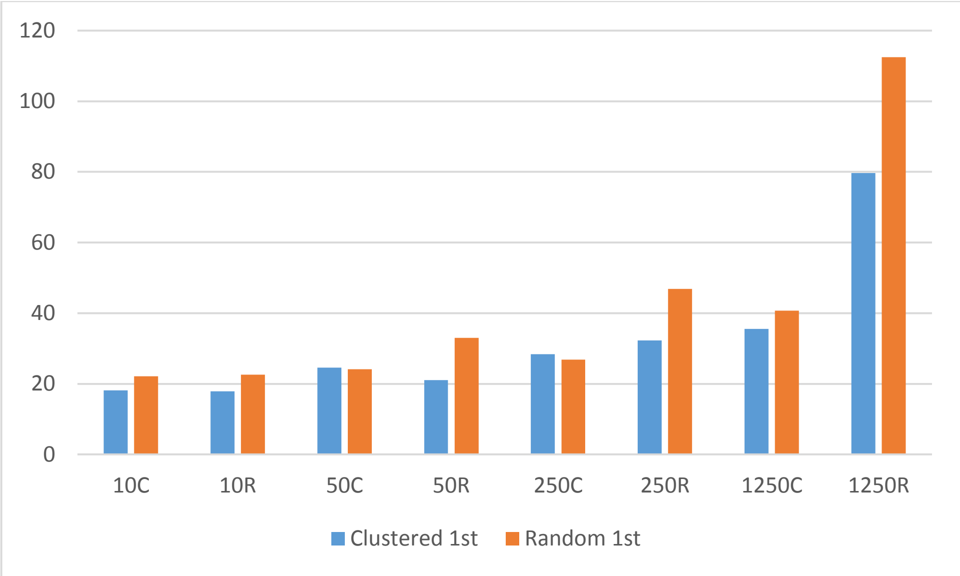
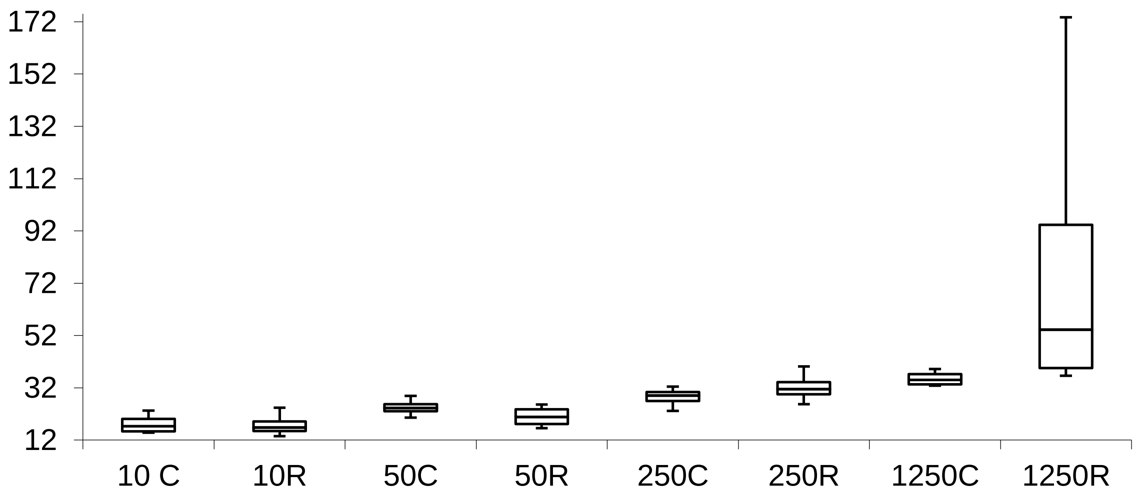


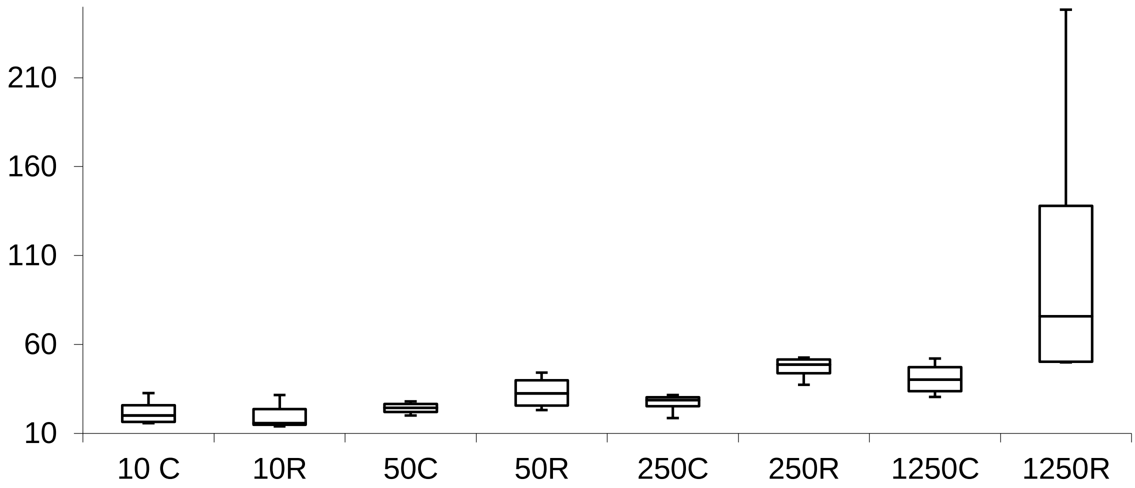
Figure 6.24 shows the spread of fixation counts for participants who encountered Clustered Maps first. Overall, there is a compactness to this participant grouping for every map type, with the exception of Random 1250 spatial unit maps which indicates participants may have encountered a higher level of difficulty when completing tasks for the Random 1250 maps.

Figure 6.24 Order Encountered: Clustered First: Number of Fixations



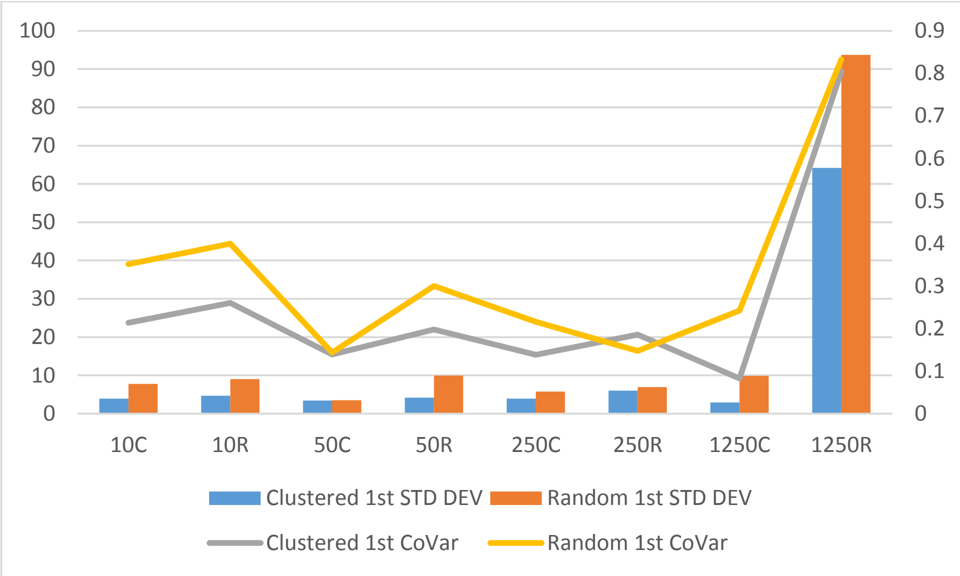
The results of the Random 1st boxplot (Figure 6.25) show a compactness to the spread of participant performance levels, thought the spread is greater than it is for participants who encountered Clustered maps first. Random 1250 spatial unit maps shows a similar distribution of performance with some outlying results. The spread of participant performance is far greater for this map which again indicates participant difficulty levels may increase when required to complete tasks using more complex maps.

Figure 6.25 Order Encountered: Random First: Number of Fixations



Random 1250 maps show significantly greater Standard Deviation and Coefficient of Variation compared to every other map (Figure 6.26). This suggests that more complex maps require a significantly greater cognitive effort to analyse, this is an effect of perceptual scalability which is exacerbated if a participant encounters a Random map first. There is still a large variation according to the Standard Deviations and Coefficient of Variation for other map types, particularly for Random maps (except Random 250). The variation for participants who encountered Random maps first is also greater, suggesting a difficulty in analysing Random maps first.

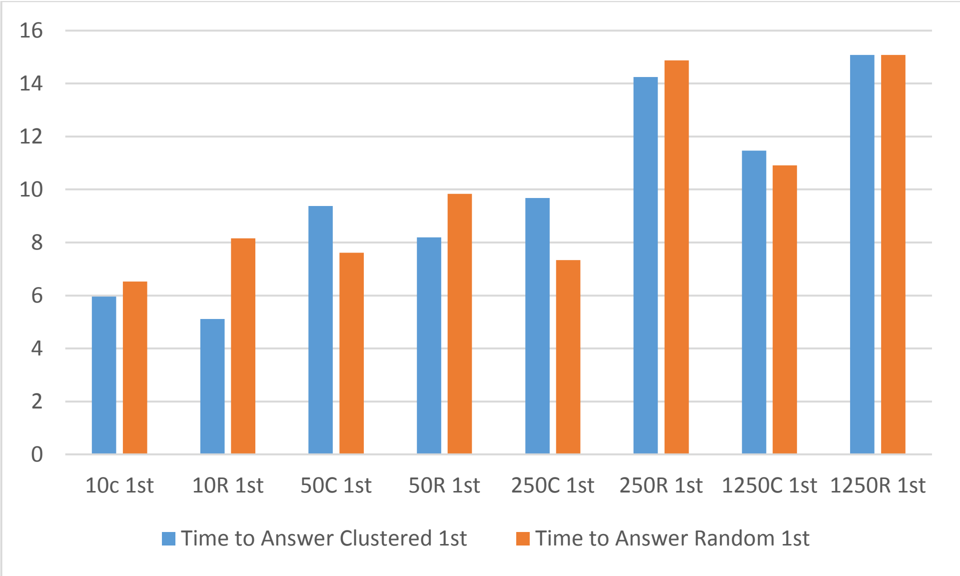
Figure 6.26 Order Encountered: Number of Fixations: Standard Deviation and Coefficient of Variation– Left axis for Standard Deviation, Right axis for Coefficient of Variation



6.4.3 Order Encountered: Time to Answer

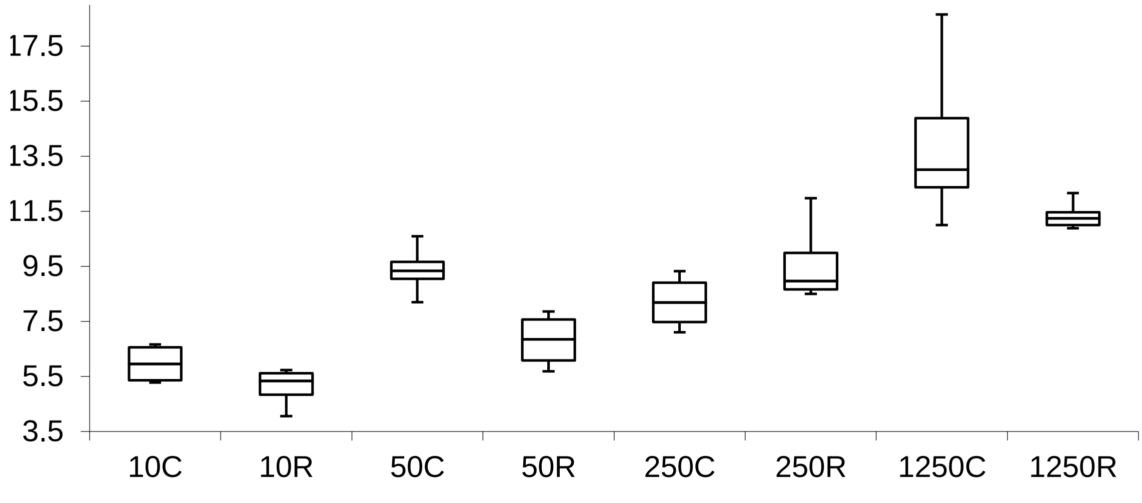
Figure 6.27 shows the effect encountering Random maps first has on participant performance in terms of average time spent on task. Times to complete a task were higher for participants who encountered Random maps first, the exceptions being: 50C and 250C 1250C maps. This suggests that participants who encountered Random maps first were faster to complete tasks on Clustered maps highlight a potential learning effect throughout the experiment.

Figure 6.27 Order Encountered: Time to Answer (1 second = 1.0)



There is a general trend towards a longer time to completion according to the distribution in Figure 6.28 for more complex spatial unit maps. The spread is greatest for Clustered 1250 maps which indicates level of complexity is detrimental to consistent participant performance.

Figure 6.28 Order Encountered: Clustered First: Time to Answer (1 second = 1.0)



For participants who encountered Random maps first (Figure 6.29), the range of performance varied more significantly than for participants who encountered Clustered Maps first. In general, a greater time was taken by participants who encountered Random maps first according to Figure 6.30. Similar to results in Figure 6.28, the time require to complete a task is higher for more complex maps (the Clustered 1250 map and Random 250 and 1250 maps. There is an observable scalability effect, with the potential for participants to perform at a lower level when faced with more complex data.

Figure 6.29 Order Encountered: Random First: Time to Answer (1 second = 1.0)

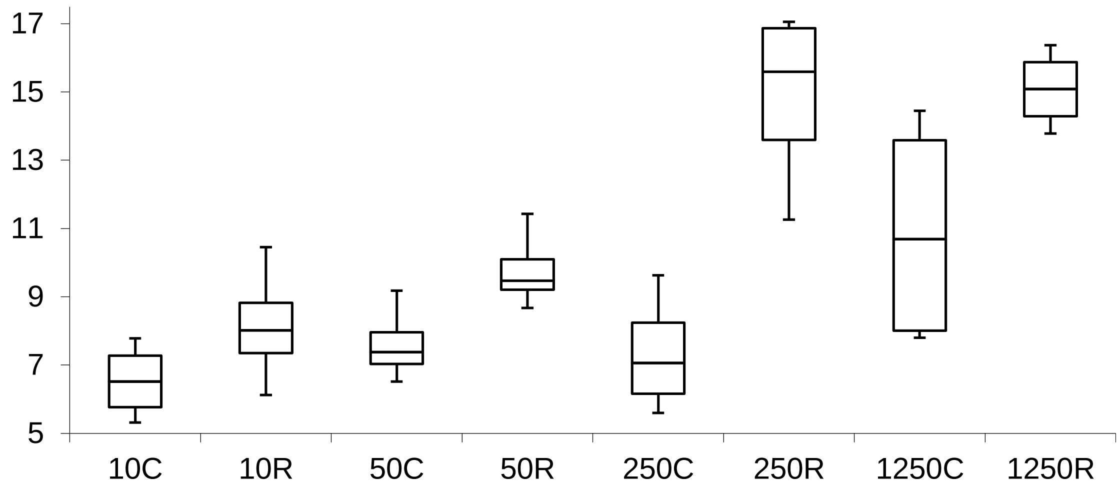
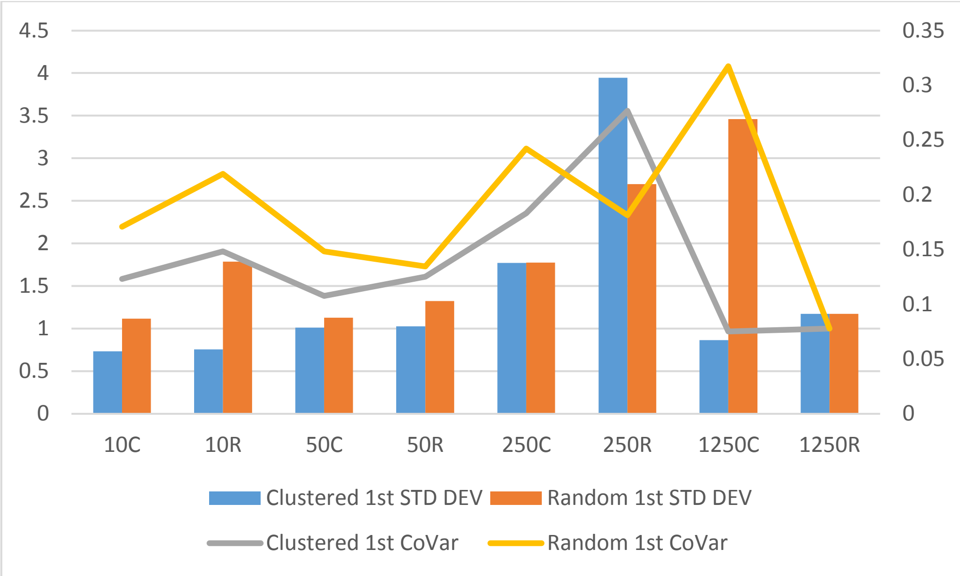


Figure 6.30 shows that variation is greatest for two of the four more complex maps (250 and 1250) for both Standard Deviation and Coefficient of Variation. This suggests there is a greater spread of results and therefore potentially greater difficulty associated with more complex map types. The variation is greatest for Random 250 maps which could indicate a difficulty associated with this type of data distribution.

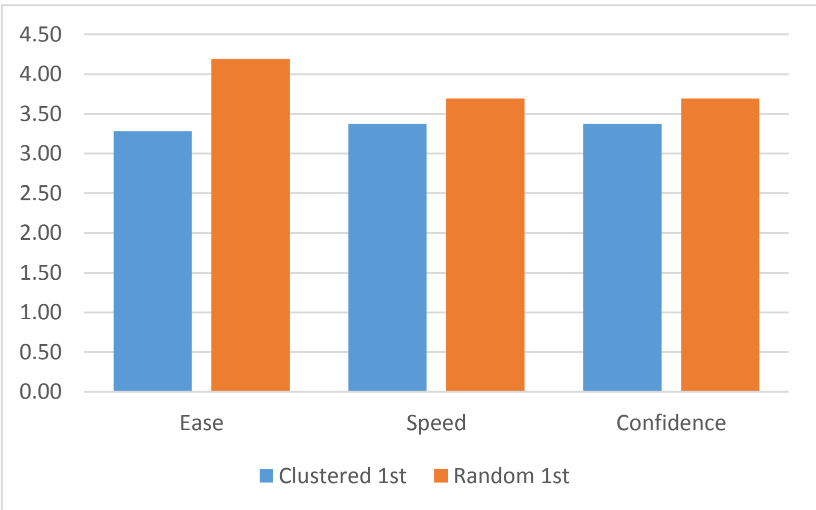
Figure 6.30 Order Encountered: Time to Answer (1 second = 1.0): Standard Deviation and Coefficient of Variation – Left axis for Standard Deviation, Right axis for Coefficient of Variation



6.4.4 User perception of tasks

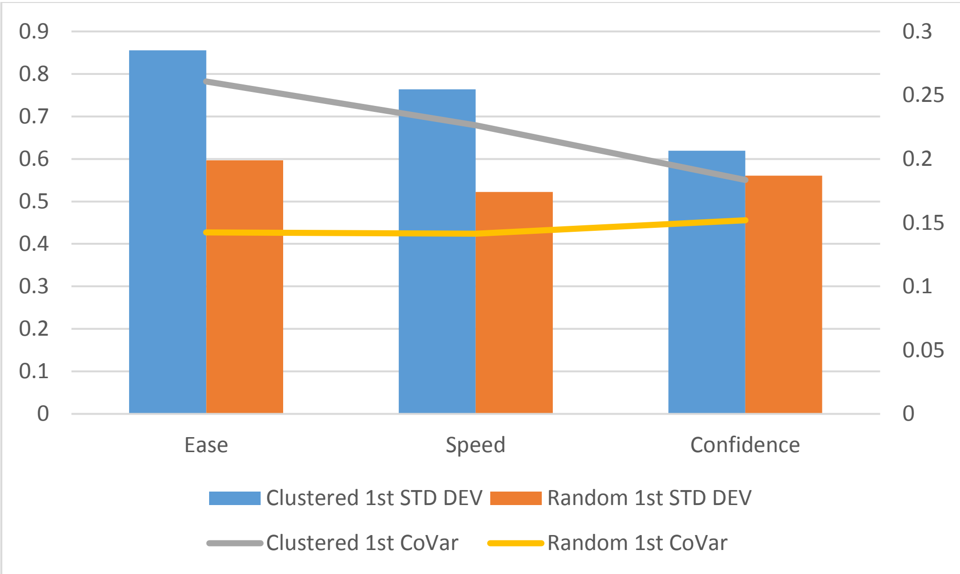
As the results of Figure 6.31 suggest, participants who encountered Random maps first are more confident in completing tasks, with an "Ease of Completion score being greater than 4.0. They also produced results to show they felt their speed of task completion is faster than participants who encountered Clustered maps first. The ease at which they completed tasks again demonstrates the confidence of participants that encountered Random maps first.

Figure 6.31 Order Encountered: Participant Task Perception (rating out of 5)



The Coefficient of Variation shows there is greater variation in participants who encountered Clustered maps first for Ease of Completion, than Confidence (Figure 6.32). This indicates that participants who encountered Clustered maps first felt they were able to complete tasks quickly, but were less confident in their answer. The Standard Deviations are greater for Ease of completion for participants who encountered Clustered maps first, with lower Standard Deviation values for this group's levels of confidence in their answers. Participants who encountered Random maps first show less variation between Standard Deviation Values and Coefficient of Variation values. This indicates participants in this group were in agreement with performance levels on a more consistent basis.

Figure 6.32 Order Encountered: Participant Perception: Standard Deviation and Coefficient of Variation: Left axis for Standard Deviations and Right axis for Coefficient of Variation



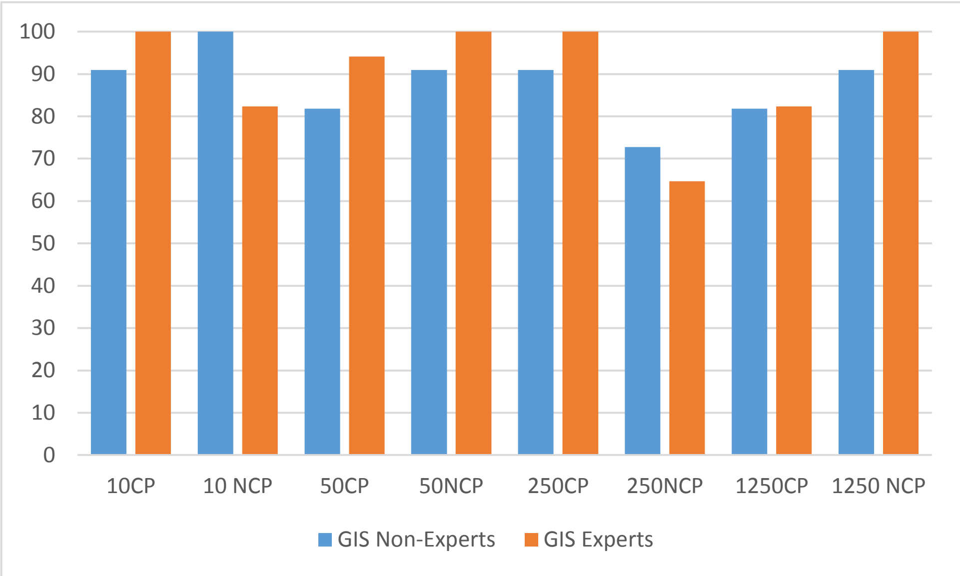
6.5 Paired Maps

6.5.1 Paired Maps: Division of Expertise: Correctness Ratios, Number of Fixations and Time to Answer.

Paired maps required the completion of comparative distribution tasks. Two maps of the same scale, spatial variation and data distribution were presented side by side.

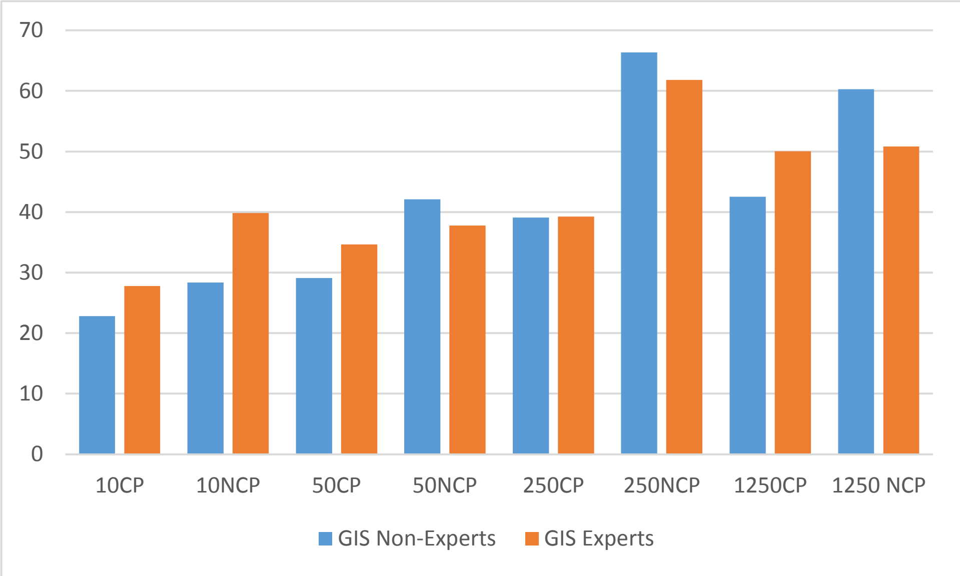
Contrary to previous Non-Experts versus Experts results, Paired Maps show that Experts perform better on average compared to Non-Experts in terms of average Correctness Ratios. With some, with two exceptions, Experts attained a higher overall correctness ratio than Non-Experts (Figure 6.33). This suggests that higher levels of GIS knowledge is advantageous when analysing more than one visualisation simultaneously.

Figure 6.33 Paired Maps: Expertise: Correctness Ratios for Paired Maps



In some instances, Experts produced less fixations for two of the four more complex map sets (250 and 1250 spatial unit maps), which a third showed similar fixation averages for both Divisions of Expertise (Figure 6.34). This indicates a higher level of expertise could result in requiring less cognitive effort to complete tasks. This trend is contradictory to the fixations for less complex maps, where Experts required more fixations to complete three of the four tasks (for 10 and 50 spatial unit maps).

Figure 6.34 Paired Maps: Expertise: Number of Fixations for Paired Maps



According to Figure 6.35, time taken to complete tasks is higher on average for Non-Experts compared to Experts. The simultaneous comparison of maps shows what single map analysis above do not. Experts were capable of completing Paired Map tasks quicker for every complex spatial unit map, with the exception of the 10 spatial unit maps. Experts also produced a similar time to answer for the Clustered 50 spatial unit map. The trend is still apparent though for five of the eight map sets. This additional level of complexity – the comparison of maps simultaneously shows that level of expertise could be an advantage.

Figure 6.35 Paired Maps: Expertise: Time to Answer a Task (1 second = 1.0)

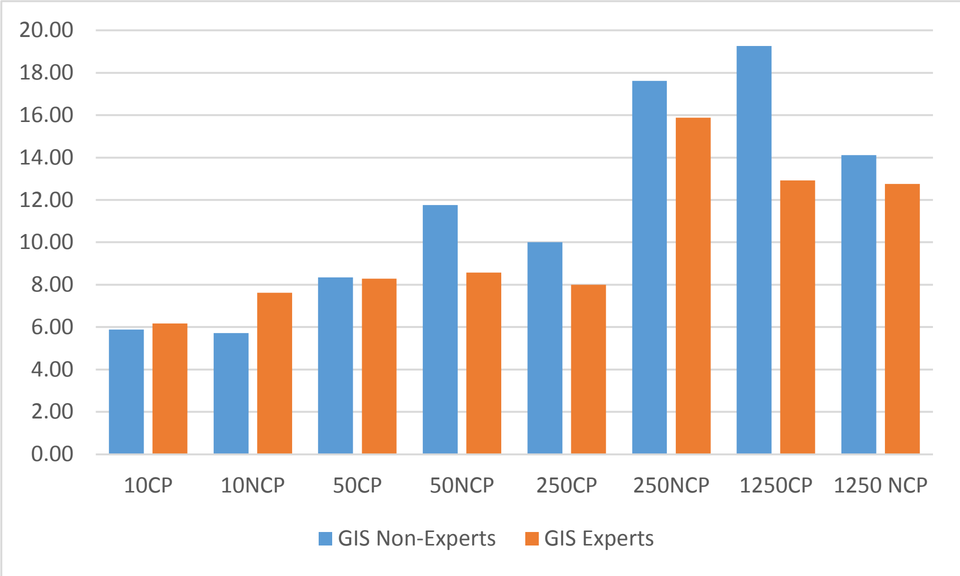
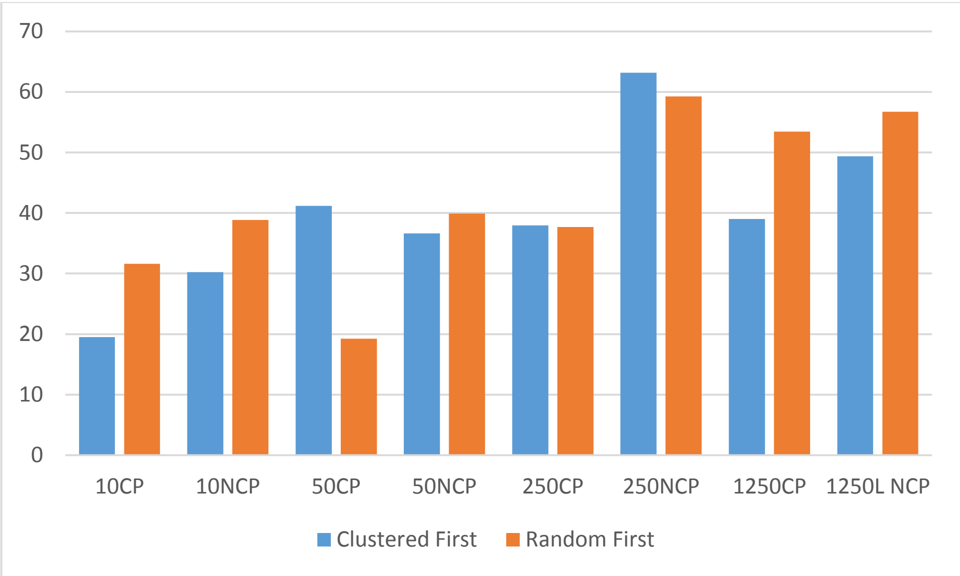


Figure 6.36 shows the correctness ratios of Paired Maps for order encountered. Participants were required to select one of the two maps as the correct answer. Those who encountered Random maps first attained a higher percentage of correctness on average compared to participants who encountered Clustered maps first. This indicates that participants who encounter the arguably more complex Random maps first are better equipped to correctly complete tasks where two maps are compared simultaneously, side by side.

Figure 6.36 Paired Maps: Order Encountered: Correctness Ratios

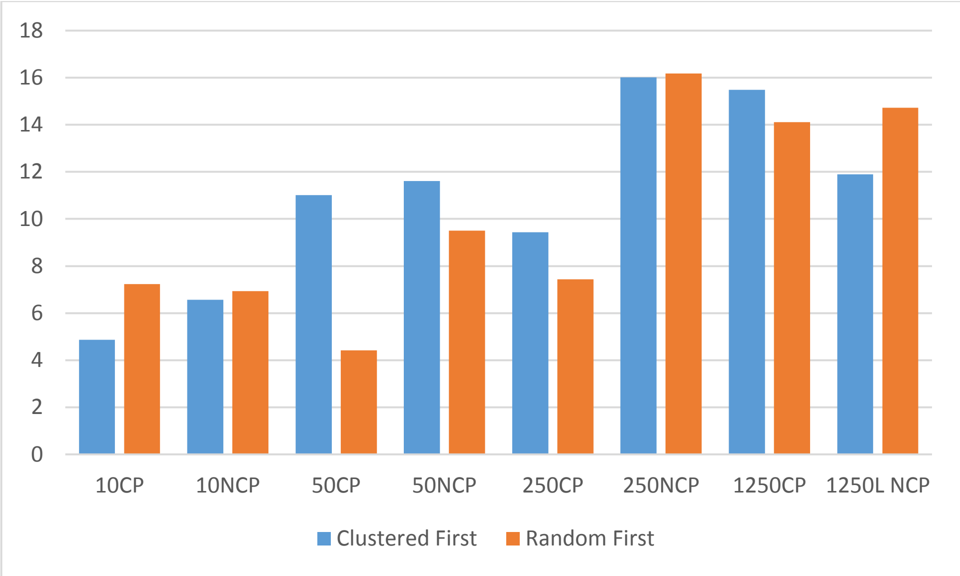
According to Figure 6.37, there is no discernible pattern of fixations for those who encountered Clustered or Random maps first. There is a general increase from 10 spatial units to 1250 spatial units as observed in Figures 6.3, 6.9 and 6.19. For more complex 1250 spatial unit maps, participants who encountered Random maps first produced more fixations in order to complete the tasks, resulting in a higher cognitive load.

Figure 6.37 Paired Maps: Order Encountered: Number of Fixations



Understandably, the time taken to complete Paired Map tasks generally increase in tandem with the number of fixations observed in Figure 6.37. Figure 6.38 shows an observable pattern between those who encountered Clustered or Random maps first. In general, those who encountered Random Maps first were faster to complete tasks. This evidence related to a lower average number of fixations for participants who encountered Random Maps first, as shown in Figure 6.37). One exception being the 50 spatial unit Clustered map, where participants who encountered Random maps first required much less fixations on average to complete the task.

Figure 6.38 Paired Maps: Order Encountered: Time to Answer a Task (1 second = 1.0)



6.6 Visual Comparison of Fixations

As an example of eye-tracking trajectory analysis, in this section we visually compare example scanpaths (sequences of fixations and saccades) of one participant on different map types. Further trajectory analysis is beyond the scope of this thesis, but similar differences in scanpaths as outlined for this particular individual are generally observable among all participants. Figures 6.39 and 6.40 show the scanpaths on Clustered and Random maps respectively. The areas corresponding to correct answers (i.e. where the participants were supposed to click if they answered the task question correctly), are marked as Areas of Interest (AOIs, Çöltekin et al. 2009) and shown with circles and oval shapes, overlaid on each of the maps.

There is an observable search pattern of groups of fixations on Clustered maps (figure 6.40), while the search pattern on Random maps which is more scattered (Figure 6.41). This difference is a consequence of the complexity of the visual distribution of data values and places an additional cognitive load on participants when they are completing tasks (Çöltekin et al., 2009 and Çöltekin et al., 2010).

In this case we can also observe an effect of the size of the maps in terms of the number of spatial units. Less complex maps, i.e. 10 spatial unit maps, do not show a large difference in search patterns, while 50 spatial unit maps only show minor differences as additional clusters of fixations appear. The search patterns on 250 and 1250 spatial unit maps show the difference between the grouped fixations search and a scattered search.

Figure 6.39 Example of the difference in fixation counts on Clustered maps

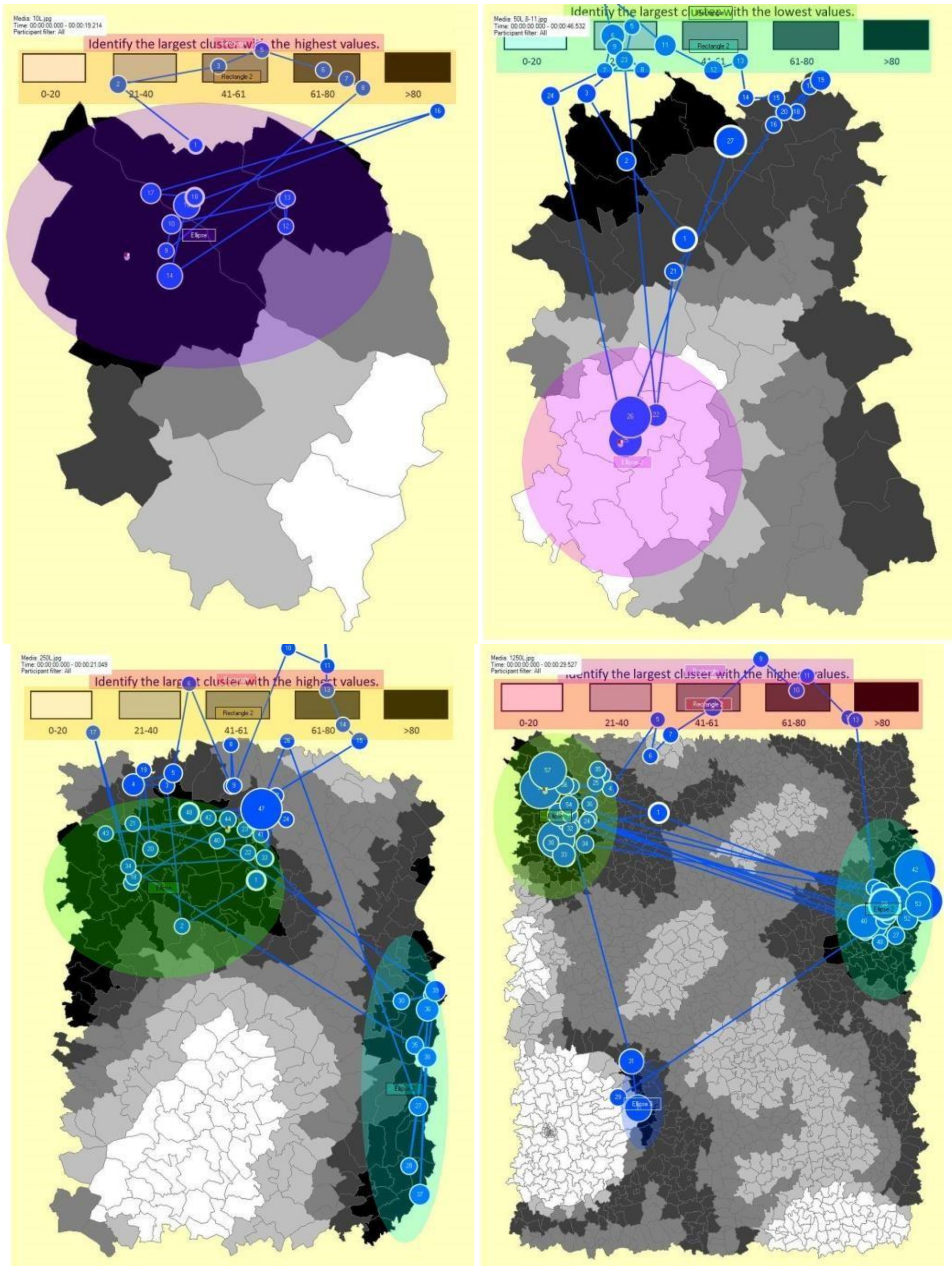
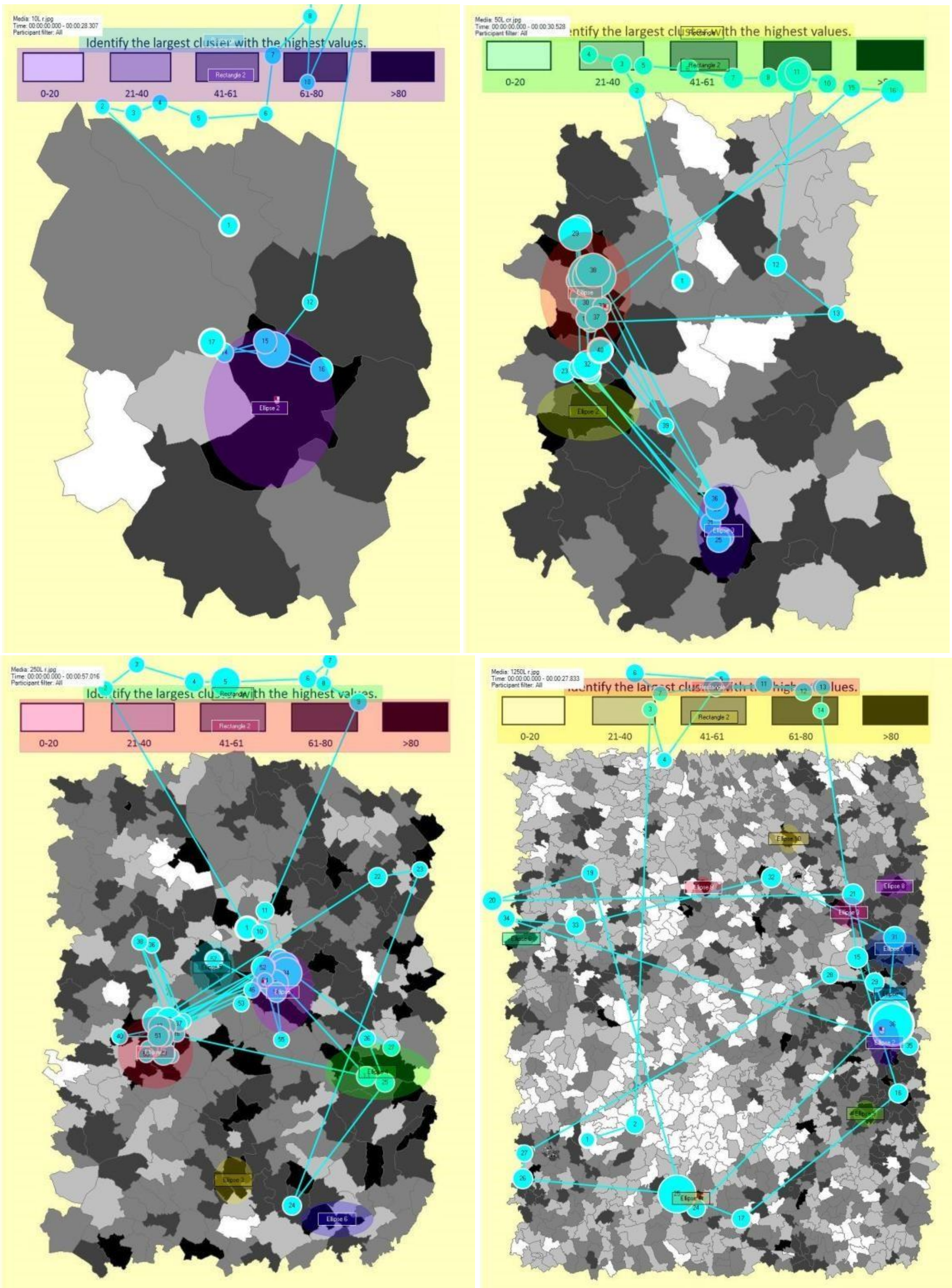


Figure 6.40 Example of the difference in fixation counts on Random maps



6.7 Discussion

As Yost et al. stated in 2007 (see Chapter 2). Visual scalability issues can be resolved by presenting data on a larger set of displays, but this isn't practical and it essentially limits the volume of displayable data according to screen size. The results in this chapter displays some evidence for the change in performance of individuals who interact with spatial data at different levels of visual complexity. We assessed this using three standard eye-tracking analysis metrics: correctness, time spent on task and number of eye fixations. We performed the analysis by grouping the data in three different manners: spatial unit scale, levels of GIS expertise of the participants, and the order in which participants were shown two types of displays. The results indicate an influence of the visual scalability on the task performance and fills a gap in literature as mentioned in Chapter 2 (Eick and Karr, 2002). In this section we briefly summarise and discuss some of the findings.

One finding is that a greater effort is required by individuals to complete tasks on Random maps compared to Clustered maps. This is indicated by the time to answer metrics, which increases for Clustered maps. As the data becomes more complex, in the form of additional spatial units, this therefore slows down the performance. Our finding echoes the results of related studies, such as work by Abbot (1995), Hayhoe and Ballard (2005), Swienty and Reichenbacher (2008) and Çöltekin et al. (2009), which have all investigated eye movement with respect to visual complexity. In particular, Çöltekin et al. (2010) suggest there may be an observable increase in time spent on task in eye movement experiments when a difficulty in perception occurs.

Further, we found that the level of participant expertise does not necessarily reflect what would normally be expected, i.e. a higher level of performance. To an extent this is shown through all three metrics, the correctness ratios, time to task completion and the number of fixations. Correctness levels are marginally in favour of the GIS Experts group which could indicate a higher level of performance on more complex maps (in this case, the 1250 spatial unit maps). A greater cognitive effort is expended by GIS Experts than GIS Non-Experts as evident by the greater number of fixations for each visualisation scale.

Interestingly, the results from the post-task survey on participant perception, that is on the ease, speed and confidence of task completion by GIS Non-Experts and Experts for Clustered and Random maps do not make it possible to differentiate between groups. GIS Non-Experts were slightly less confident in their answers but this was not reflected in their overall correctness levels and might indicate perception of ability to complete tasks. In fact, GIS Non-Experts felt they completed tasks faster than GIS Experts which largely coincides with the time taken to complete tasks by both groups. Non-Experts were faster overall.

In terms of the order of maps shown to the participants, the participants who encountered Random visualisations first recorded a greater average number of fixations compared to participants who encountered Clustered visualisations first. There was no observable consistency in the average number of fixations to suggest that participants who encountered the Random visualisations first would require a greater cognitive effort over the course of the entire experiment. The only consistency was the increase in the overall average number of fixations from the 10 spatial unit scale through to the 1250 spatial unit scale. Comparing with the survey on user perception for this section, participants who encountered Random visualisations first felt their tasks were easier to complete despite believing it took them longer to complete and they were more confident in their task responses. There is a noticeable difference in correctness levels however. Participants who encountered Clustered visualisations first recorded a higher level of correctness for every task on every scale with the exception of the Random 10 spatial unit scale map. I refer to my earlier discussion which eludes to potential familiarity with the experiment format and increased skill in completing experiment tasks. Also, most of the correctness scores for the order of map encounters (Clustered First or Random First) were lower for 250 and 1250 spatial unit scales compared to 10 and 50 spatial unit scales. This suggests that Random visualisations are a more difficult introductory spatial distribution which appears to affect participant performance. Less discernible patterns require a greater cognitive effort.

The Spatial Unit Scale analysis produced some insightful results. The average level of correctness was greater for all of the “Large” unit scale visualisations, i.e. visualisations with a greater map footprint (with the exception of the 1250 scale). This suggests the presence of a visual scalability effect with respect to the map footprint. To strengthen the case for a visual scalability effect, a greater average number of fixations were recorded for every visualisation in which the spatial unit scale varies. Although it is not entirely possible to compare the 1250 scales because their map footprint does not differ significantly, it can be said that; a larger map footprint results in a greater cognitive load, which in turn yields a greater correctness score for tasks that may or may not take more time to complete than tasks on maps which occupy a smaller map footprint where the spatial unit size remains constant. The spatial unit scale of an interactive visualisation would typically vary in size, depending on the extent to which a user elects to examine the visualisation. It would change to occupy the greatest amount of space where appropriate, i.e. the map footprint would be as large as is appropriate. This related to Experiment One where the interactive visualisation performed best where the visual scalability effect was first observed.

As highlighted in Chapter 2 under the Geovisual Analytics section, more than 40% of our brain power is used to provide visual output (Hoffman, 2000; Ware, 2008). This explains why so much time is spent understanding and improving on visual representations, not least geographical representations as geographical data becomes more accessible and increases in volume. According to MacEachren (2004) “Human vision and visual cognition remain incompletely understood” (see Chapter 2). The combination of metrics in this experiment which include eye fixations and perception of participants contribute to understanding the effect of data complexity on visual cognition.

7: Discussion and Research Conclusions

In this chapter, the research is summarised and contextualised within the wider field of research. First a summary of the thesis is provided before the research question and aims are addressed along with an in-depth discussion on the overall findings and implications of this research thesis.

Chapter 1 introduced the research area and provided an insight into the aims and objectives of the research.

The effective display of data is an important aspect of this research. As discussed earlier, the existing research and literature in this field do include a variety of visualisation types for spatial data representation. This research sought to draw upon this range of visualisation types to ascertain the optimal or most effective types of visualisations available, with a particular emphasis on the use of newly emerging and novel 3D visualisations. The assessment of these visual analytics is specifically focused on the spatial statistical method of geographically weighted regression (GWR).

Furthermore, the human element is equally important to understand. The human ability to process information has to be taken into account to truly understand the outputs and to produce the most effective outputs. There is little need for an advanced geovisualisation technique that displays highly complex data if human cognitive limits prevent the visualisation from being used effectively. This research aimed to provide deeper insight into the human interactions with such geovisualisations so that future research can be guided on most effective types of data display techniques.

To reiterate from Chapter 1, based on current knowledge, the over-arching aim of this research thesis was to assess and evaluate a range of data visualisation techniques to ascertain the optimal way of geographical data presentation with a particular emphasis on GWR and perceptual scalability. To attempt to answer this aim, two empirical experiments were constructed and carried out, each with a specific set of objectives designed to contribute to deeper understanding of the overall thesis aim and research question. Experiment One asked what visualisation type is best for interpretation and analysis of GWR results, focusing on three particular types (2D, 3D and interactive). Experiment Two explored at what point does visual scalability result in a change in human performance? Below a discussion of the key research findings from these experiments are offered.

Chapter 2 explored the relevant literature in the field. This research was interdisciplinary in nature, with aspects of geography, computer science and psychology. Given its inherently encompassing nature, geography as a discipline has a tendency to relate to almost every other discipline and GWR is incorporated into a lot of research (not exclusively geographical) that has a related spatial element. GWR uses geographical weighting to run the spatial statistical model (Fotheringham et al., 2002). Visualisation techniques to effectively display data and the interaction with these visualisations are embedded within computer science. These techniques are developed to include geographical aspects to produce a geovisualisation. Methods such as GWR adopts a geographically local approach to model spatial non-stationarity and Crespo's work outlines spatial non-stationarity or spatial heterogeneity as a way to understand the concept of spatially varying relationships (2009).

Chapter 3 and 4 outlined the methods employed to design, implement and analyse the two key experiments. Each experiment was researched and designed in line with best practice in the field. Each experiment was innovative and attempted to explore and answer a unique set of questions to better understand the field of geovisualisation, particularly that of GWR data.

Chapters 5 and 6 presented the key findings of the research and explored their significance and contribution to the field. Specific research aims and objectives are addressed in each of these chapters.

7.1 Discussion of Key Research Findings

To attempt to answer the research question, first a discussion on the key experiment findings is required to contextualise the overall conclusions before the findings and implications can be positioned in the wider literature in this field of research.

The main research question within this thesis was to evaluate visualisations of Geographically Weighted Regression and assess 2D visualisations for perceptual scalability. There are two core aims for the experiment, each with several objectives. These are now discussed.

Aim 1: To assess the effectiveness of three visualisation types for analysis and interpretation of GWR outputs.

The motivation of Aim 1 was to assess which visualisation technique works best for interpretation and analysis of GWR output. This aim was explored and evaluated in the first experiment. Three different visualisation types (2D, 3D, Interactive) were utilised as benchmarks for visualisations in general.

- Objective 1a: To assess the effectiveness of 2D visualisations for display and interpretation of GWR outputs.

The 2D visualisation is regarded as the most commonly used visualisation type for GWR outputs according to research carried out in Chapter 2. Specifically, thematic maps are a popular example of these 2D visualisations. Choropleth or thematic maps are primarily used in three ways; for pattern comparison, to provide general information about spatial patterns or to provide specific information about particular locations. These thematic maps can be created through digital means. Statisticians generally agree that visualisations are capable of providing insight into datasets (Edsall, 2003). Thematic maps have become key components in spatial data exploration and research showing the value of geovisualisations has been carried out (Hurley and Buja, 1990; Wegman, 1990 and Tukey, 1977).

Despite popular use of thematic map 2D visualisation with GWR outputs, their effectiveness for interpretation and analysis is not fully understood. Given their popularity of use with GWR outputs (see Table 1.3 in Chapter 1), Objective 1a seeks to provide insight into the effectives of these 2D visualisations and uses metrics such as correctness ratios, mouse clicks and eye fixations to interpret and draw conclusions. Having assessed the results of Experiment One, we can say that 2D visualisations are effective for interpretation and analysis on a basic level. 2D visualisations are less appropriate for multivariate GWR exploration and these are significant differences since GWR analysis is likely to be complex.

Expanding on what a "more basic level" refers to, 2D maps are effective completion of non-complex tasks like univariate tasks (tasks which involve one parameter estimate which is displayed on a thematic map visualisation). 2D visualisations are also relative effective for bi-variate tasks, where two different thematic map visualisations are compared. GWR outputs are complex and geovisualisations are key to deriving meaning from complex data (DiBiase, 1990; MacEachren et al., 1992; and MacEachren and Kraak, 1997). A significant question therefore is, are 2D visualisations effectivness for complex analysis of GWR outputs, i.e. multivariate analysis in the form of task three in Experiment One, containing three different parameter estimates. The multivariate display which is discussed below (the interactive visualisation) outperformed 2D visualisations on correctness ratios which suggests more modern interactive visualisation techniques such as this could be more effective for interpretation and analysis of GWR results. The difference in correctness levels between 2D visualisations and interactive visualisations is more than 15%. Assessed on their own, with comparison, it can be argued that 2D visualisations could perform adequately, particularly for more basic analysis, but not for more complex analysis.

A significantly higher number of mouse clicks are required for more complex task completion with 2D visualisations, indicating they are not perfect for use in these instances. For simpler tasks, this visualisation type one again appears adequate. The higher number of mouse clicks are an indication of an increased cognitive effort on the part of the user to complete more complex tasks. For example, navigation between five different 2D maps is not ideal.

As mentioned, thematic maps are used to emphasise the spatial pattern of one or more geographic attributes (Slocum et al., 2009). This is why choropleth maps are often used to display GWR outputs. Based on findings in Experiment One, there is no reason to discourage the use of 2D visualisations for less complex tasks. The inclusion of 2D visualisations in interactive visualisation systems is to be encouraged. As Plaisant (2004) describes, no one (visualisation) tool is specifically designed to cater for the needs of a user. 2D visualisations can be useful for GWR output interpretation and analysis to a certain extent.

This analysis is designed to provide initial guidance to GWR outputs users and to demonstrate 2D visualisation effective for interpretation and analysis of GWR results. A more comprehensive method of the popularity of GWR output analysis would be through direct communication with GWR output users. A database containing analysis methods utilised by GWR outputs users would offer further insight into GWR output visualisation methods.

- Objective 1b: To assess the effectiveness of 3D visualisations for analysis and interpretation of GWR outputs.

3D visualisations are a more recent method of displaying data compared to 2D visualisations. This is an emerging technology within geographic applications (Pullar and Tidey, 2001) commonly displayed in the popular geographic software tool ArcGIS Suite. They are arguably more effective for displaying at least two different sets of data values. As such, it is important to compare the effectiveness of 3D visualisations to other common mediums of display for 2D and Interactive visualisations – of which the latter is discussed later in this chapter.

According to Slocum et al. (2009), 3D visualisations are particularly useful in the display of certain kinds of data, e.g.: petroleum exploration, gas exploration or modern medicine. However, they have one major drawback, they suffer from occlusion (Slocum et al., 2009; Tsigas, 2007). Assessing the degree to which this is true for GWR data is therefore important because it will affect the overall effectiveness of this visualisation type for interpretation and analysis of GWR outputs. The additional viewing angles presented in 3D visualisations compared to 2D visualisations (Kok and Liere, 2007) do not necessarily result in an augmented ability of users to analyse information (Cline, 2000). Again, using the same metrics, this was explored further.

3D visualisations were unable to record the highest correctness percentage scores for any task in Part A or Part B of Experiment One. The exception being the Task 2 in Part two (Group 2). That is five out of six tasks. A scalability effect was also observed for 3D surfaces between Part A and Part B, with correctness ratios being lower for Part B than Part A. Furthermore, 3D visualisations recorded some of the lowest correctness ratios across all expertise groups for Task 3. This is partly due to a 3D rotation feature incorporated into the ArcGIS 3D visualisation system which is necessary to use when completing the multivariate task (Task 3).

As Zudilova-Seinstra et al. (2010) state, it's difficult to assess 3D visualisation performance compared to 2D visualisations because of the variety of devices, interactions, techniques and participant expertise. This is true to a certain extent based on the results of Experiment One, but there are indications which show that there is no advantage to using simple 3D visualisations for simple or complex GWR tasks than compared to 2D visualisations. Tasks 1 and 2 for 3D visualisations could be completed using the same methods applied to Task 1 and 2 for 2D visualisations.

Having knowledge of 3D visualisation techniques didn't seem to help the problem encountered with occlusion, exacerbated by the 3D rotation feature. Essentially, the rotation feature and occlusion issue is linked because the 3D rotation feature was used by participants to manoeuver the visualisation into a desired position for greater perspective to interpret and analyse the visualisation and data. In summary, 3D surfaces with rotation seem less suitable for a majority of GWR tasks. As Slocum et al. (2009) and Tsigas (2007) suggest, 3D visualisations are hampered for the problem of occlusion and Experiment One results indicate this issue is also prevalent for GWR results. As previously mentioned, advanced research using cutting edge 3D visualisation techniques has been carried out. 3D examples include cutting edge work on space-time cubes (Demsar and Virrantaus, 2011; McArdle and Demsar, 2011) to analyse aspects of data including trajectories and densities. Perhaps it is in this domain that 3D visualisations are in a better position, where interpretation and analysis of data containing a temporal aspect is required.

- Objective 1c: To assess the effectiveness of interactive visualisations for analysis and interpretation of GWR outputs.

We already know the majority of brain activity is related to the processing and analysing of visual images. Visualisations allow us to explore our innate potential to process visual representations in knowledge intense tasks (Bukhard and Meier, 2005). The emergence of interactive visualisations render traditional map studies unnecessary (MacEachren, 1995). Geographic visualisation techniques are linked with increasingly interactive dynamic tools (Ogao and Kraak, 2002) and it is possible to display geovisualisations within these interactive dynamic systems. These interactive visual representations are designed to encourage intuitive and creative data exploration and in identifying difficult to find patterns (Edsall, 2003). Interactive visualisation methods could be useful for interpretation and analysis of GWR outputs, since interactive visualisation methods are becoming a popular medium of data display, it was important to include them as part of the evaluation in Experiment One.

Evaluating interactive visualisations it is evident that participants performed better on Part B than Part A. It is apparent that the interactive visualisation performed best overall for Part B, the more complex of the two datasets used. This indicates interactive visualisations could be better suited to interpretation and analysis of more complex data. It is less limited. For simpler tasks, such as the bivariate task, it was as quick if not quicker to complete using interactive visualisations compared to traditional 2D visualisations, despite 2D visualisations performing well for less complex tasks. Group 4 contained participant with knowledge of all visualisation types, these participants completed Part B tasks using the interactive visualisation faster than other groups. This suggests a knowledge of visualisation types or systems will lead to faster analysis and interpretation, while maintaining similar levels of correctness. Participants with knowledge of ArcMap only completed tasks faster using the interactive visualisation system, compared to 2D and 3D visualisation tasks. This highlights the potential usefulness of interactive visualisations for users with no prior knowledge.

Mouse click counts were considerably lower for the interactive visualisation system, particularly for multivariate tasks. Interactive visualisations demonstrate a reduction in effort required by users to complete tasks, suggesting it could be an efficient visualisation technique for analysis of GWR results. It is difficult to determine if there is a scalability effect present using interactive visualisation when analysing the performance of this visualisation technique for Part A and Part B. For instance, Task 1 correctness ratios are lower in Part B, while multivariate correctness ratios are higher in Part B. Again this indicates interactive visualisation techniques could be best suited to complex GWR output interpretation and analysis.

Interactive visualisation results indicate the potency of this particular visualisation type. Overall, participant performing with this visualisation type is at least on par with 2D and 3D visualisations for the completion of simple tasks. For multivariate tasks, the interactive visualisation is either on par or outperforms 2D and 3D visualisations. There are instances where interactive visualisation task completion times were faster and required less effort to complete. User perception feedback also highlights a greater number of positive thoughts or comments related to interactive visualisations than 2D or 3D visualisations.

- Objective 1d: To decide upon the most effective of all three visualisations for interpretation and analysis for GWR outputs, offering guidance for producers of GWR outputs.

Collating content from Objective 1a, 1b and 1c, we can surmise the performance of each visualisation type to draw conclusions on Aim 1. The metrics used to evaluate the effectiveness of the visualisations in Chapter 5 (Results of Experiment One) provide evidence of the interactive visualisation's superior performance. Interactive visualisations are less used and perform best overall, despite the apparent lack of researcher familiarity. 2D visualisation performance levels are largely on par with interactive visualisations for simpler tasks and 3D visualisations are also on par to an extent with 2D visualisations for simpler tasks. This means the defining task is the multivariate task. The 2D visualisation became more cumbersome when participants were required to complete Task 3. It is clear that 3D visualisation performance was poorest for Task 3, while the interactive visualisation performance was higher for Task 3. If users can overcome the lack of familiarity with interactive visualisation then this would be the display method of choice, for smaller or larger datasets, since performance levels for Part A and Part B on the interactive visualisation were not significantly different. Furthermore, the Knowledge group containing participants with knowledge of ArcMap only – participants who have no prior knowledge of interactive visualisations – were faster to complete tasks using the interactive visualisation system. Again, this suggests that a lack of familiarity does not necessarily equate to reduced performance when interpreting and analysing complex geographic data. Figure 7.1 provides a simple visual breakdown of suggested visualisation use based on the results of Experiment One.

To summarise findings from Aim 1, overall correctness levels are lower for the 2D and 3D visualisation types in part B than in Part A which indicates participants had a more difficult time in correctly completing tasks. The interactive visualisation type actually shows an increase in correctness levels which could be attributed to an increased familiarity with the interactive visualisation system, resulting in higher levels of performance. It was hypothesised that there would be a scalability effect, whereby participant performance would change when faced with a larger dataset. The change in dataset size had a negative effect on performance levels for all of the visualisation types, although it is less noticeable for the interactive visualisation system. Speed of completion does not necessarily indicate the presence of a scalability effect because the time to task completion is so varied but the number of mouse clicks and level of mouse movement do indicate the presence of a scalability effect, particularly the level of mouse movement which is higher for the majority of Part B tasks compared to Part A tasks. Post task survey scores do not indicate an increased level of difficulty or any positive or negative change in terms of scalability but this could be offset somewhat due to increased levels of familiarity with the visualisation systems. Returning to a description of the 3D visualisation during interviews, the discussions with participant indicate that Part B multivariate task was toughest to complete, particularly using the 3D visualisation type.

To conclude the discussion on Aim 1, the range of metrics used to evaluate the three visualisation techniques were beneficial in assessing the effectiveness of each in terms of the ease of use, interpretation and analysis. The relatively unused technique of interactive visualisations within GWR research has positive indicators of its effectiveness and therefore, this research suggests further use of this technique in emerging studies to ensure users are presented with the optimal geovisualisation for the spatial data concerned. Stemming from Aim 1, a need to study potential impacts of scalability of data became apparent and this formed the basis for Aim 2 and therefore Experiment Two.

Aim 2: To assess the impact of data scale on user interpretation of 2D visualisations, thereby investigating the hypothesised presence of perceptual scalability.

The motivation of Aim 2 was to investigate the presence (if any) of a perceptual scalability effect with visualisation of spatial data. This aim was addressed through the research, design and user testing of Experiment Two. The performance of participants completing tasks in Experiment Two was measured using standard evaluation metrics in addition to eye movement monitoring, a novel approach to assess user interaction with the subject. It is useful to briefly reiterate the definition of Random and Clustered at this point. Clustered refers to a higher level of spatial autocorrelation in a map as opposed to a spatial pattern that has occurred by chance. The term Random relates to spatial patterns which have occurred more by chance than as a result of a quantifiable statistical spatial distribution.

In general, there is a noticeable difference in performance between Clustered and Random map distributions. This is prevalent on more complex data scales (e.g. 250 and 1250 spatial unit scales). Results in Experiment Two suggests an increased cognitive load is required by participants to complete tasks on more complex maps. Evidence of this is observable in Figures 6.3, 6.10 and 6.19 of Chapter 6.

The pattern of eye fixations is also affected by the spatial distribution of data. Random maps generally required a higher number of fixations compared to Clustered maps, thus requiring a greater cognitive effort.

- Objective 2a: To assess the impact of visualisation scale and spatial unit scale on user interpretation of 2D visualisations, utilising standard metrics to examine potential effects of perceptual scalability.

Research shows data complexity requires management of a display resolution or screen size. Yost (2006) carried out research on the size of a screen displaying visualisations and how they are scaled. This research offers further insight into the effect of visualisations in terms of scale and size. Experiment Two built on Yost's work. The screen size remained the same however with the focus being on the change in performance of participants when presented with different data value (or spatial unit) scales – among other aspects discussed in objective 2b.

Correctness score for participants decrease from 10 spatial unit maps to 1250 spatial unit maps, regardless of whether the spatial unit scale varies or not. These correctness results do indicate the presence of a scalability effect, a concept which was of particular interest in this research thesis. Though correctness levels decrease from 10 to 1250 spatial unit maps, correctness scores are higher in general for maps in which the spatial unit scale remains the same (see Figure 6.4). Time taken to complete tasks on these map types increases with each spatial unit scale. Participants were slower to complete tasks on maps where the spatial unit size remained the same. By utilising available space for the map footprint, participant performance will be less affected by the change in the number of spatial units being displayed on screen. Significantly, this objective suggests perceptual scalability does exist and is a factor in geovisualisation (of 2D GWR data in this case).

- Objective 2b: To evaluate the impact of expertise levels on interpretation of 2D visualisations using standard metrics, examining potential effects on perceptual scalability.

Assessing the changes in cognitive load can be difficult to understand, "Human vision and visual cognition remain incompletely understood" (MacEachern, 2004: 23). Objective 2b provides evidence of a scalability effect for 2D visualisations and the extent to which cognitive load of users with different levels of knowledge changes, contributing to our understanding cognition. Firstly, it is important to note that knowledge of GIS did not result in an overall processing advantage. Here, cognitive load is measured by the change in metrics recorded during Experiment Two. Increases in data scale did result in a change in the cognitive load of participants, this is most noticeable on the largest data scale of 1250 spatial units. However, average correctness values were similar between Experts and Non-Experts which indicates there may be no significant advantage to having more or less knowledge of 2D visualisations or varying dataset scales.

In a number of cases the 250 spatial unit maps demonstrated an increased cognitive load. The differences in performance by participants according to their level of GIS expertise are not consistently in favour of less or more experience. In some instances, GIS Experts are capable of completing more tasks correctly, but they take longer to complete these same tasks. In other instances, GIS Non-Experts record a higher correctness level and complete tasks faster than GIS Experts. This is not what would normally be expected as experts typically use their greater level of knowledge to work more efficiently and produce more accurate results. Non-Experts were less confident in their answers, presumably because they believed their comparative lack of expertise would yield poor performance. Interestingly, Experts believed they took longer to complete tasks on average, conversely to Non-Experts, perhaps Experts felt a need to study each visualisation in more detail before providing an answer – which could also indicate why they felt more confident in their task completion performance.

- Objective 2c: To evaluate the impact of the order in which data distributions are encountered on interpretation of 2D visualisations through standard metric measurement, examining potential effects of perceptual scalability.

Robinson and Griffin (2015) introduced a rotation feature into their visualisation to avoid a learning effect; participants would not be presented with the same kind of map visualisation to the extent a learning effect would occur. This is similar to the idea adopted for this experiment. A control was introduced to Experiment Two to assess the impact of presenting different participants with Clustered or Random distribution maps first. As briefly mentioned in the outline of this section, encountering Random maps had a generally negative effect on performance levels and resulted in a greater cognitive load as tasks took longer to complete. However, according to participant perception the effect of encountering Random maps first was negligible. Opinions on the ease of task completion, speed of task completion and confidence in their answer were higher for those who encountered Random maps first.

- Objective 2d: To analyse eye movement data to ascertain if there is evidence of perceptual scalability, with a particular emphasis on:
    i. Spatial unit scale of 2D visualisations, investigating if there is a change in a user's cognitive load.
    ii. The expertise levels on interpretation of 2D visualisations.
    iii. The variation of data distribution on a 2D visualisation, determining if any change in a user's cognitive load occurs according to variations and in doing so, gaining insight into the potential effects of perceptual scalability.

Eye movement science is an emerging area across disciplines such as psychology, business, education and geography, providing valuable insight into psychological and cognitive function into a number of real world tasks Goldberg et al. (2002). This particular strand of the research is novel in the field of geovisual analytics and findings made are insightful and beneficial to better understanding the data and results of the experiment. The invaluable contribution of eye movement recordings is because the number of fixations for each spatial unit scale clearly indicate the presence of a scalability effect – a concept of interest throughout this thesis. They also highlight the extent to which perceptual scalability is an anomaly proportionately related to the level of data complexity. The number of eye fixations increase when progressing from 10 spatial units to 50 spatial units, the same occurs when progressing from 50 spatial units to 250 spatial units. Once again, when moving from 250 spatial units to 1250 spatial units the number of eye fixations increases. This indicates an increased cognitive load on the individual (see Figures 6.3, 6.10 and 6.19 of Chapter 6).

i.

There is evidence to show the presence of a scalability effect when the spatial unit scale varies, compared to when spatial unit scale remains the same. A greater cognitive effort is required to complete tasks on maps where the spatial unit scale varies. As highlighted in

Chapter 6, the difference in the average number of fixations for displays in which the spatial unit scale varies and displays in which it remains the same is greatest for 1250 spatial unit maps. This suggests the effort required to recognise patterns is higher on more complex maps. This finding coincides with existing literature, where more complex challenges of pattern recognition are presented by more complex clustered visualisations (Abbott, 1995; Hayhoe and Ballard 2005; Tufte, 2007; Swienty and Reichenbacher 2008; Çöltekin et al., 2009; Çöltekin et al.,2010). It's possible that findings on cognitive effort required to complete tasks on more complex Clustered data distribution visualisations are transferrable to Random maps as described in this thesis and presented through 2D visualisations.

ii.

Eye movement metrics are useful as proven by Çöltekin et al. (2009). By assessing the number of fixations for different level of expertise we can assess how the users' cognitive load is affected by changes in data complexity. Greater numbers of eye fixations may indicate a higher cognitive load and can be considered as a representation for task complexity. The number of fixations for Non-Expert and Experts show no significant difference which suggests higher levels of knowledge result in a better performance or decreased cognitive load when completing tasks. Figure 6.10 highlights the degree to which spatial unit scale affects user cognitive load for both expertise levels and there is a clear change in participant cognitive behaviour between the 10 spatial unit scale and the 1250 spatial unit scale. As the dataset becomes more complex, the effort required by either expertise group to complete a task increases. For five of the eight maps presented, Experts demonstrated a higher increase in the number of fixations required to complete a task. This increased level of effort relates to the time taken by Experts to complete tasks, as discussed earlier in this section. It's also worth noting that the higher number of fixations for Experts did not equate to a higher average of correctness.

iii.

There is an increase in the average number of fixations and therefore cognitive load on participants when faced with more complex data. 10 spatial unit maps (both Clustered and

Random) and 50 spatial unit Clustered maps have a similar average number of fixations, which could be described as the first level of difficulty. The second level begins at the Random 50 spatial unit maps and the third is reserved for the Random 1250 spatial unit maps.

Again, there is an observable increase in the number of fixations vs. the increase in the number of spatial units and there are some variations in the number of fixations recorded for participants. Fixation counts are spread for 50 and 1250 spatial unit maps in particular, while fixation counts for 250 spatial unit maps are more similar between participants. There are more fixations for Random maps (except the 50 spatial unit map) than for Clustered maps which can indicate difficulty with search patterns or displayed data (Çöltekin et al., 2009).

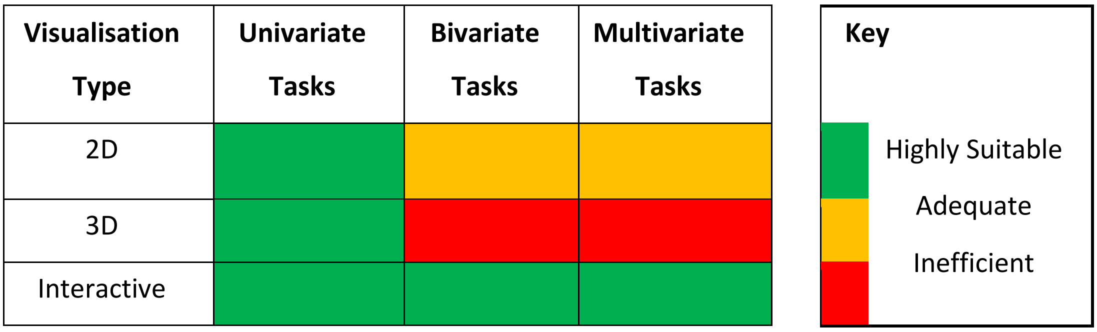
Overall, the conclusions drawn from Aim 2 indicate that data scale does play a role in interpretation of 2D visualisations. The metrics used to evaluate the experiment data indicate a level of perceptual scalability does exist and this is an important factor to note for producers of visual representations of data. This is especially pertinent for geographers and those using geographical or spatial data.

7.2 Guidelines for geovisualisation of GWR output

Geovisual analytics, a field within the discipline of visual analytics, concerns itself with spatial data in particular (Andrienko et al., 2007). From this research thesis, it is now possible to create a summarised guide (of the GWR experiment results coupled with the visual scalability element) for use as a best practise guide for analysis of GWR outputs, or other forms of regression analysis, and appropriate levels for the display of data. This suggested guide is presented in Figure 7.1. The guide is based on a combination of the performance of the visualisations for all tasks, and participant perception and post-experiment comments.

Contributing to the field of geovisual analytics, a basic visualisation guide for GWR outputs has been created to serve as an aid to future researchers in this area. The visualisations were categorised using a simple traffic light system to indicate their effectiveness for interpretation and analysis of GWR out puts.

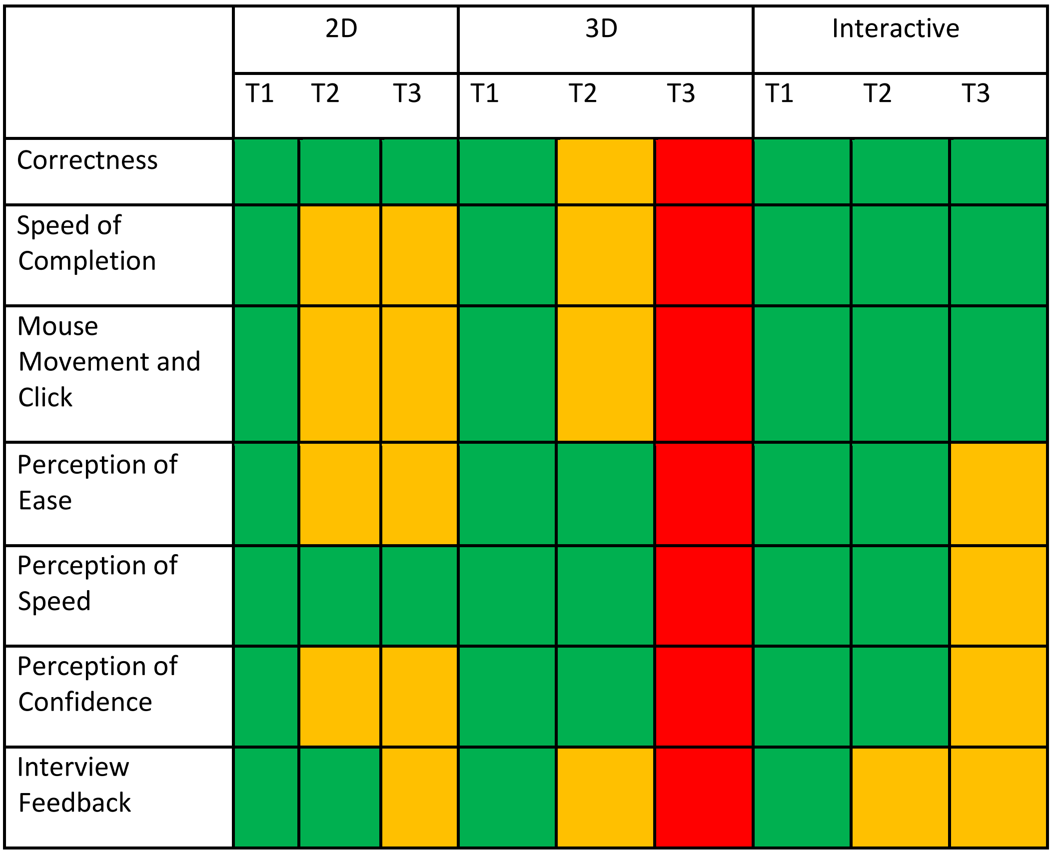
Figure 7.1 Basic Visualisation Guide



In Figure 7.1 the visualisation types are set out as follows; 2D visualisations were represented in Experiment One through ArcMap, the most commonly used tool in ESRIs ArcGIS suite. 3D visualisations were represented in Experiment One by ArcScene, a popular element of ESRI's ArcGIS suite. Interactive visualisations were represented by ProVis, which was created using the Processing scripting language.

For a more specific guide, Figure 7.2 offers an indication of the overall effectiveness of the visualisations with respect to metrics in Experiment One.

Figure 7.2 Visualisation Effectiveness based on Experiment One Metrics. T1 = Task 1, T2 = Task 2 and T3 = Task 3



According to Figure 7.2. The interactive visualisation method is the strongest overall performer. This figure is in line with the results from Experiment One. For more basic tasks, any of the visualisation types are suitable. However, when the most complex tasks (the multivariate tasks) are completed, the interactive visualisation is the strongest. Most GWR analysis will be multivariate in nature so this is why the interactive visualisation should be given preference. The 3D visualisation in this experiment performed worst and is therefore not recommended for standard GWR analysis unless the purpose is to perform basic analysis or to demonstrate some basic GWR outputs in a visually appealing way.

Part of the issue of using 3D visualisation is the difficulty associated with showing GWR outputs in publications. Although there is an increasing number of digitally based publications which offer interactive attachments as a feature of published work, it is not yet widely available. Published visualisations are often displayed in free frame captures or screenshots. Interactive visualisations also suffer from this issue to a degree, but it is possible to display a selected set of values in a dataset for publication purposes using visualisations (including interactive visualisations).

Participant perception as outlined in Figure 7.2 indicates that the more commonly used 2D visualisation type is sufficient for a large part of their GWR analysis. In fact, there are suggestions that the interactive visualisation is less preferred to the 2D visualisation for analysis. This can be attributed to a lack of familiarity with the interactive visualisation. It is recommended that interactive visualisations are assigned a greater level of importance by GWR users as they can be extremely effective in elucidating insights to the data concerned (Keim et al, 2004).

The pattern of fixations is also affected by the spatial distribution of data. “Random” maps generally required a higher number of fixations compared to “Clustered” maps, thus requiring a greater cognitive effort.

Andrienko et al. (2007) highlight the importance of decision support for producers of spatial data such as the GWR explored in this thesis. The above suggested guide for users of GWR methods can therefore act as a decision support tool to maximise conveyance of geo-spatial information.

7.3 Further Discussion

This research thesis has contributed significantly to better understanding the range of visualisations available for GWR data along with considering the effects of perceptual scalability. This has been achieved through the design, implementation and analysis of two key experiments involving real-life users. Kang et al. (2011) highlights the importance of measuring visualisations for their capacity or ability to ‘make sense’ to end users but also acknowledges the difficulty of this task. The research results, as discussed earlier, do act as a guide for researchers on geovisualisation choices along with providing validation for the usability of these visualisations in expert and non-expert audiences.

The literature suggests new insights can be gained from the use of interactive geovisualisations (Keim et al., 2004) and also suggests human perception affects the likelihood of an uptake in the use of interactive systems. In other words, if a researcher is comfortable in the methods they use to analysis complex data or they are reluctant to gain knowledge on the operation of an interactive system then, it is unlikely interactive systems will become the most popular method for analysis. Obviously, this is not to say that they are not being used at all.

As the acquisition of complex data increases, the methods used to display data become more important. Provide too little information and one cannot derive useful meaning. Provide too much, and one can suffer from information overload. Keim et al (2010) suggests visual analytics can provide effective understanding, reasoning and decision making on the basis of very large and complex data sets. The range of data scales tested in this research is broad and the indications for optimal data scales are considered in Experiment Two.

Yost (2006) carried out research into the perceptual scalability of visualisations. In her research she focused on the size of screen a visualisation was scaled upon. Through Experiment Two, there is an observable scalability effect. This is based on the change in participant performance according to the number of data values a human can effectively perceive. As stated in Chapter 1, a human being is capable of processing vast amounts of data but at some point the perceptual load will exceed the cognitive processing ability. At this point, the ability to comprehend data is adversely affected.

The second question on visual scalability attempts to clarify the extent to which a human being experiences a change in performance or behaviour when faced with different levels of data complexity. When data is randomly distributed participants find it more difficult to discern clusters or patterns. An obvious reason for this is the lack of patterns or groups of spatial units. Using smaller sized datasets this does not appear to be a problem, however when participants were presented with a more realistic dataset size it became clear that behaviours change. Visualisations that presented what appears to be more spatially auto-correlated data were straightforward. The change in performance from small to large dataset sizes were to be expected, participants required more time to analyse more complex datasets but their search behaviours largely remained the same. The search patterns of eye movements are a telling sign of the degree of confusion (see Figures 6.4 and 6.5 for a comparison of Clustered and Random spatial distribution search patterns). The degree of confusion could be described as the extent to which a participant struggles to answer an assigned task. For example, an erratic search pattern suggests a greater degree of confusion compared to a uniform like search pattern.

Intelligent design of interactive visualisations would be beneficial (Eick and Karr 2002). Would it be difficult to suggest that the visual scalability results can be applied beyond set visualisation display systems? Yost (2006) focused on size of the screen and resolution data is displayed for example. Google Maps and other frequently used real world mapping systems all incorporate responsive data display algorithms, Kelly and Fotheringham (2011) for example, demonstrate the use of modern online mapping systems. The volume of information presented to the user on these maps changes according to the viewing level, i.e. it depends on how far you have zoomed in on the map. The number of displayed attributes displayed at each level of detail need to be carefully considered to avoid information overload. Would it be difficult to suggest that the visual scalability results can be applied beyond set visualisation display systems?

There is a need to better inform researchers of the potential of interactive visualisations. If people are properly trained to use these systems then more efficient and insightful analysis can occur. It would not be unreasonable to create a set of classes to educate researchers through the suggested guide above. Large bodies of researchers including government based services utilise GIS based systems for the analysis of complex data so there is a place for this. As MacEachren (2004) states, "Cartography is about representation". Visualisation functionality is an important factor in the design of visualisations, not least interactive visualisations. Developing intuitive and highly usable visualisation systems to display complex data is paramount.

Individually, not all measurements taken in the second experiment reveal evidence of a scalability effect. There are mixed perceptions of knowledge groups on their performance for each visualisation set. Participants identified the correct cluster in most of the visualisations too. However, there is a remarkable increase in required cognitive load to complete a task by participants for higher scales, Figures 6.3, 6.10 and 6.19 provide evidence of this. In fact, there is a noticeable change between each spatial unit scale. Participants studied more complex visualisations for longer, and took more time to answer tasks. The correctness levels are still quite high for the most complex visualisation type, so higher cognitive function does not necessarily equate to a drop in the potential to correctly identify patterns within data, at least for Clustered data.

Overall, this thesis was concerned with assessing and evaluating a range of data visualisation techniques to ascertain the optimal way of geographical data presentation with a particular emphasis on the spatial statistical method of GWR and the potential effect of perceptual scalability. Through two in-depth user experiments, it was concluded that whilst 2D, 3D and interactive visualisations all have merits for representing GWR data, the benefits of interactive techniques appears most effective for user interpretation and data analysis. This is a significant finding as effective data representation is a critical element to visual representation of data. As alluded to in Chapter 1, “a picture paints a thousand words” so it is vital for producers of visual data to be aware of optimal types of visualisations. Furthermore, this thesis has uncovered the presence of perceptual scalability when it comes to data scale and related user experiences. Human cognitive load is an important consideration for all producers of visualisations to again ensure the key information can be translated to the user through the visualisation. Through these two aims and their related objectives, one can overall conclude a degree of perceptual scalability does exist and that interactive visualisations may be the most effective way of visualising spatial data, such as GWR output as was used here. These are significant factors to be aware of for all producers of such visualisations.

7.4 Further Research

Stemming from this research, the scope for further research is quite large with several thought-provoking questions emerging. Firstly, how significant are the findings made here? Statistical significance tests could be performed on current metrics. Furthermore, it would be useful to perform a similar experiment with commercially based or governmental institutions that incorporate spatial statistical work into data analysis. GWR workshops to train individuals in the use of GWR are hosted in numerous locations around the world and the inclusion of an additional workshop module on visualisation could be useful to train users and producers of visual data. As stated more than once before, a lack of familiarity with more advanced visualisation systems is evidently one of the problems associated with an uptake in use, even if these approaches can be more effective in information conveyance. Going a step further, it is possible to design a visualisation system to specifically cater for the visualisation of spatial statistical methods, in this case GWR. Looking to the future, it is desirable to increase the accessibility of spatial statistical methods because of their capability to identify relationships within complex data

Complex data present another set of difficulties relating to comprehension, and it is essential to display this data effectively so that is can be understood. Research using advanced visualisation techniques has already proven advantageous in analyse complex data, including data with a temporal aspect (Demsar and Virrantaus, 2011). The second experiment provides evidence of a scalability effect, where participants ability to perceive different scales shows that more complex visualisations of data result in a change in participant performance. Expanding this experiment to include larger data scales (of which there are theoretically infinite scales) would augment this research and potentially further validate the findings. As with the first experiment results, statistical significance tests could reveal relationships between metrics.

Eye movement analysis is a proven effective evaluation tool (Harrie and Stigmar, 2009; Schnur et al., 2010; Coltekin et al., 2010). Additional analysis of eye movements, particularly those associated with the AOIs versus non-AOI attention can provide more insight into the types of patterns most visible to participants. In turn, this could provide an explanation on why participants have decided the areas of focus are important to them. The saccadic movement or eye trajectories could be studied to discover changes in cognitive processes in more detail. Sequential pattern techniques could be applied to further study this eye movement data. It is also possible to combine mouse based movement with eye movement to identify the relationship between mouse movement and cognitive behaviour patterns identified using eye movement data.

The valued geographical concept of perceptual scalability can be further tested. Rather than exclusively focusing on 2D visualisations as were the case in this thesis, a focus on, the deemed most effective, interactive visualisation system would be insightful to research. Returning to the issue of scale, a greater variety of data scales may provide more advanced tangible data on human ability to correctly perceive data using some of the most advanced data visualisation techniques available. Evidently, there is great scope for further research across many areas covered in this thesis. Overall, a more focused approach to study more select techniques would be beneficial to gain a deeper and more refined insight into matters of geovisualisation effectiveness and scale.

7.5 Overall Conclusion

Overall, through the two empirical experiments carried out as part of this research thesis, the research aim was addressed and answered. A range of data visualisation techniques were prepared, assessed and evaluated to gain a deeper insight into the most effective modes of visualising geographical data. The spatial statistical method of GWR was a central component to the research with the concept of perceptual scalability also crucial to understanding how users interpret visual data.

The key findings indicate that popular visualisation methods used to display GWR outputs are useful for analysis and interpretation of GWR results to a certain extent. The results of comprehensibility testing indicated that the largely untested interactive systems have a greater potential for analysis, even while their appearance in published literature is rare. There are two obstacles to overcome in order to increase the users of this visualisation type; the first is familiarity, the second is the modernisation of publication methods to include digital appendices.

The two other key findings of this research are the presence of a scalability effect and the observation indication of perceptual scalability. The scalability effect occurs when faced with increasingly complex data while the perceptual scalability effect offers further evidence on the change in performance, which has been measured to a certain extent in this research, of human beings when they encounter more complex data.

The general discussion concludes that there is a need to better inform researchers of the potential of interactive visualisations. People do need to be properly trained to use these systems, but the limits of human perceptual processing also need to be considered in order to permit more efficient and insightful analysis. As with all geographical problems, scale is a crucial concept to factor in.

A core component of this research was ultimately concerned with the way in which spatial data is visualised. This research contributes significantly to the wider field of knowledge by demonstrating qualitatively and quantitatively the usefulness of popular and emerging visualisation types for interpretation and analysis of GWR results. The need to be mindful of a scalability effect when presenting data in visual form is also a significant factor to consider.

Reference List

Abbot, A. (1995) Sequence Analysis: New Methods for Old Ideas. *Annual Review of Sociology*, Vol. 21, pp. 93-113.

Abrams, A. R., Meyer, E. D. and Jornblum, S. (1989) Speed and Accuracy of Saccadic Eye Movements: Characteristics of Impulse Variability in the Oculomotor System. *Journal of Experimental Psychology: Human Perception and Performance*, Vol. 1, No. 3, pp. 529-543.

Acevedo, W. and Masuoka, P. (1997) Time-series animation techniques for visualizing urban growth. *Computers and Geosciences Archives*, Vol. 23, No. 4, pp. 423-435.

Agnew, J. (1996) Mapping politics: how context counts in electoral geography. *Political Geography*, Vol. 15, No. 2, pp. 129–146.

Aguilar, G. D. and Farnworth, M. J. (2012) Stray cats in Auckland, New Zealand: Discovering geographic information for exploratory spatial analysis. *Applied Geography*, Vol. 32, pp. 230238.

Andrienko, N. and Andrienko, G. (2005) *Exploratory Analysis of Spatial and Temporal Data: A Systematic Approach*. Berlin: Springer.

Andrienko, N. and Andrienko, G. (2002) Intelligent Visualisation and Information Presentation for Civil Crisis Management. *Transactions in GIS*, Vol. 11, No. 6, pp. 899-909.

Andrienko, G. and Andrienko, N. (1999) Interactive maps for visual data exploration. *International Journal of Geographical Information Science*, Vol. 13, pp. 335-374.

Andrienko, N., Andrienko, G. and Gatalsky, P. (2001a) Exploring changes in census time series with interactive dynamic maps and graphics. Computational Statistics, Vol. 16, pp.417-433.

Andrienko, G., Andrienko, N., Jankowski, P., Keim, D., Kraak, M.-J., MacEachren, A. M., and Wrobel, S. (2007). Geovisual analytics for spatial decision support: Setting the research agenda. *International Journal of Geographical Information Science*, Vol. 21, No. 8, pp.839–857.

Anselin, L. (1999) Interactive techniques and exploratory spatial data analysis. In: Longley, P., Goodchild, M., Maguire, D. and Rhind, D., eds. (1999) *Geographical Information Systems*. 2nd ed., New York: Wiley, pp. 25-264.

Armstrong, M. P., Xiao, N. and Bennett, D. A. (2003) Using genetic algorithms to create multicriteria class intervals for chloropleth maps. *Annals of the Association of American Geographers*, Vol. 93, No. 4, pp. 257-268.

Austin, M. (2007) Species distribution models and ecological theory: A critical assessment and some possible new approaches. *Ecological Modelling,* Vol. 200, No. 1-2, pp. 1-19.

Baeker, R. M. and Buxton, W. A. S. (1987) A historical and intellectual perspective. In: Baecker, M. R. and Buxton, S. A. W., eds. (1987) *Readings in Human Computer Interaction: A Multidisciplinary Approach*, San Mate: Morgan Kaufmann Publishers, pp. 41-54.

Bastin, L., Fisher, P. and Wood, J. (2002) Visualizing uncertainty in multi-spectral remotely sensed imagery. *Computers & Geosciences*, Vol. 28, pp. 337-350.

Bertin, J. (1983) *Semiology of graphics: diagrams, networks, maps*. Madison: University of Wisconsin Press.

Bingham, N. H. and Fry, J., M. (2010) *Regression: Linear Models in Statistics*. London: Springer.

Bitter, C., Mulligan, G. F. and Dall'erba, S. (2007) Incorporating spatial variation in housing attribute prices: a comparison of geographically weighted regression and the spatial expansion method. *Journal of Geographic Information Systems*, Vol. 7, pp. 7- 27.

Blanco-Moreno, J. M., Chamorro, L., Izquierdo, J., Masalles, R. M. and Sans, F. X. (2008) Modelling within-field spatial variability of crop biomass - weed density relationships using geographically weighted regression. *Weed Research*, Vol. 48, pp. 512-522.

Bridgeman, G., Hendry, D. and Stark, L. (1975) Failure to detect displacement of visual world during saccadic eye movements. *Vision Research*, Vol. 15, pp. 719-722.

Brooke, J. (1986) SUS: A quick and dirty usability scale. In: Jordon, W. P., Thomas, B., Weerdmeester, A. B. and McClelland, L. I., eds. (1986) *Usability evaluation in industry.* London: Taylor & Francis Ltd., pp. 189-194.

Brunsdon, C., Fotheringham, A. S. and Charlton, M. (1998) Geographically Weighted Regression-Modelling Spatial. *Journal of the Royal Statistical Society: Series D (The Statistician)*, Vol. 47, No. 3, pp. 431-443.

Brunsdon, C., Fotheringham, A. S. and Charlton, M. (2002) Geographically weighted summary statistics — a framework for localised exploratory data analysis. *Computers, Environment and Urban Systems*, Vol. 26, No. 6, pp. 501-524.

Brunsdon, C., Fotheringham, A. S. and Charlton, M. (2007) Geographically Weighted Discriminant Analysis. *Geographical Analysis*, Vol. 39, No. 4, pp.376-396.

Brunswick, E. (1943) Organismic achievement and environment probability. *Psychological Review*, Vol. 50, pp. 255-272.

Burke, T. (2009) Factors affecting Voter Turnout Levels in County Kildare and County Meath, Unpublished Masters Thesis, National University of Ireland, Maynooth.

Burke, T. and *Demšar, U. (2010) An Evaluation of Geographically Weighted Spatial Statistical Methods*. Unpublished paper presented at: Workshop on Emerging Methods for Studying Use of Spatial Technologies: GIScience Conference. Zurich, Switzerland, September 2010.

Burkhard, R. A. and Meier, M. (2005) Tube Map Visualization: Evaluation of a Novel Knowledge Visualization Application for the Transfer of Knowledge in Long-Term Projects. *Journal of Universal Computer Science*, Vol. 11, No. 4, pp. 473-494.

Bush, V. (1945) As we may think. *The Atlantic Monthly*, Vol. 176, pp. 101-108.

Butler, K. A. (1996) Usability Engineering turns 10. *Interactions*, Vol. 3, pp. 59-75.

Cahill, M. and Mulligan, G. (2007) Using Geographically Weighted Regression to Explore Local Crime Patterns. *Social Science Computer Review*, Vol. 25, pp. 174-193.

Calvo, E. and Escolar, M. (2003) The local Voter; A Geographically Weighted Approach to Ecological Inference. *American Journal of Political Science*, Vol. 47, No. 1, pp 189-204.

Cardozo, D. O., García-Palomares, J. C. and Guitiérrez, J. (2012) Application of geographically weighted regression to the direct forecasting of transit ridership at station-level. *Applied Geography*, Vol. 34, pp. 548-558.

Center for Perceptual Systems (2014) *Space Variant Imaging, Visual Field Simulator* [online]. Available at: http://svi.cps.utexas.edu/software.shtml, accessed 14/11/2013.

Charlton, M., Fotheringham, S. and Brunsdon, C. (2005) Geographically Weighted Regression Research Methods Workshop follow-up, 2005.

Chen, C. and Yu, Y. (2000) Empirical Studies of information visualization: a meta-analysis. *Human-Computer Studies*, Vol. 53, pp. 851-866.

Chen, C. and Rada, R. (1996) Interacting with hypertext: a meta-analysis of experimental studies. *Human-Computer Interaction*, Vol. 11, No. 2, pp. 125-156.

Cho, S-H., Chen, C. Z., Yen, S. T. and English, B. C. (2007) Spatial variation of output-input elasticities: Evidence from Chinese county-level agricultural production data. *Papers in Regional Science*, Vol. 86, No. 1, pp. 139-157.

Cho, S-H., Kim, S. G., Roberts, R. K. and Jung, S. (2009) Amenity values of spatial configurations of forest landscapes over space and time in the Southern Appalachian Highlands. *Ecological Economics*, Vol. 68, pp. 2646-2657.

Chrisman, N.R. (1998) Rethinking levels of measurement for cartography. *Cartography and Geographic Information Systems*, Vol. 25, No. 4, pp. 231-242.

Clement, F., Orange, D., Williams, M., Mulley, C. and Epprecht, M. (2009) Drivers of afforestation in Northern Vietnam: Assessing location variations using geographically weighted regression. *Applied Geography*, Vol. 29, pp. 561 – 576.

Cleveland, W. S. and McGill, R. (1984) Graphical Perception: Theory, experimentation and application to the development of graphical methods. *Journal of American Statistical Association*, Vol. 79, pp. 531-554.

Cogdon, P. (2003) Modelling spatially varying impacts of socioeconomic predictors on mortality outcomes. *Journal of Geographic Systems*, Vol. 5, No. 2, pp. 161–184.

Çöltekin, A., Fabrikant, I. S. and Lacayo, M. (2010) Exploring the efficiency of users' visual analytics strategies based on sequence analysis of eye movement recordings. *International Journal of Geographical Information Science*, Vol. 24, No. 10, pp. 1559-1575.

Çöltekin, A., Heil, B., Garlandini, S. and Fabrikant, S. (2009) Evaluation the Effectiveness of Interactive Map Interface Designs: A Case Study Integrating Usability Metrics with Eyemovement Analysis. *Cartography and Geographic Information Science*, Vol. 36, No. 1, pp. 5-17.

Crespo, R. (2009) *Statistical Extensions of GWR: Spatial Interpolation and a Spatiotemporal Approach*. PhD Thesis, National University of Ireland, Maynooth.

Crespo, R. and Grêt-Regamey, A. (2012) Spatially explicit inverse modelling for urban planning. *Applied Geography*, Vol. 34, pp. 47-56.

Darmofal, D. (2010) Re-examining the Calculus of Voting. *Political Psychology*, Vol. 32, No. 2, pp.149-174.

Darmofal, D., (2008) The Political Geography of the New Realignment. *American Politics Research*, Vol. 36, pp. 934-961.

Davies, C. and Medyckyj-Scott, D. (1996) GIS users observed. *International Journal of Geographical Systems*, Vol. 10, pp. 363-384.

Demšar, U, Fotheringham, A. S. and Charlton, M. (2008) Combining Geovisual Analytics with Spatial Statistics: the example of Geographically Weighted Regression. *The Cartographic Journal*, Vol. 45, No. 3, pp. 182–192.

Demšar, U. and Virrantaus, K. (2010) Space–time density of trajectories: exploring spatiotemporal patterns in movement data International Journal of Geographic Information Science, Vol, 24, No. 10, pp 1527-1542.

Dénes, A. D. and Keedwell, J. (1974) Latin Squares and their applications. New York, Academic Press.

DiBiase, D. (1990) Visualization in the earth sciences. *Bulletin of the College of Earth and Mineral Sciences*, Vol. 59, No. 2, pp. 13–18.

DiBiase, D., Krygier, J., Reeves, C., MacEachren, A. M. and Brenner, A. (1991) Elementary approaches to cartography animation [Video]. University Park, PA: Deasy GeoGraphics Laboratory, Department of Geography, Pennsylvania State University.

Duchowski, A. T. and Çöltekin, A. (2007) Foveated gaze-contingent displays for peripheral LOD management, 3D visualization, and stereo imaging. *ACM Transactions on Multimedia Computing, Communications, and Applications*, Vol. 3, No. 4, pp. 1-18.

Dunbar, K., and Blanchette, I. (2001) The in vivo/in vitro approach to cognition: the case of analogy. *Trends in Cognitive Science*, Vol. 5, No. 8, pp. 334 - 339.

Dykes, J., MacEachren, A. M. and Kraak, M. J. (2005) *Exploring Geovisualization*. Amsterdam: Elsevier Ltd.

Eastman, J. R. (1985) Graphic organization and memory structures for map learning. *Cartographica: The International Journal for Geographic Information and Geovisualization*, Vol. 22, No.1, pp. 1-20.

Edsall, R. M. (2003) Design and Usability of an Enhanced Geographic Information System for Exploration of Multivariate Health Statistics. *The Professional Geographer*, Vol. 55, No. 2, pp. 146-160.

Edsall, R. M., MacEachren, A. M. and Pickle, L. (2001) Case Study: Design and Assessment of an Enchanced Geographic Information System for Exploration of Multivariate Health Statistics. In: Andrews, K., Roth, S. and Wong, P. C., eds. *Proceedings of the IEEE Symposium on Information Visualisation 2001*, October 2001. California: The Institute of Electrical and Electronics Engineers Inc., pp. 159 - 162.

Egenhofer, M. J. and Mark, D. M. (1995), Naïve geography. In: Frank, U. A. and Kuhn, W., eds. *Proceedings of Spatial Information Theory: A Theoretical Basis for GIS, COSIT 1995*, September 1995. Berlin: Springer-Verlag, pp. 1-15.

Eick, S. G. and Karr, A. F. (2002) Visual Scalabilty. *Journal of Computational Graphics and Statistics*, Vol. 11, pp. 22-43.

Engelbart, D. C. (1962) *Augmenting Human Intellect: A conceptual Framework,* Summary Report AFOSR-3223, California: Stanford Research Institute.

Erdmann, B. and Dodge, R. (1898) *Psychologische Untersuchung über das Lesen auf experimenteller Grundlage.* Niemeyer: Halle.

Farber, S. and Paéz, A. (2007) A systematic investigation of cross-validation in GWR model estimation: empirical analysis and Monte Carlo simulations. *Journal of Geographical Systems*, Vol. 9, pp. 371 - 396.

Foley, P, and Demšar, U. (2013) Using geovisual analytics to compare the performance of geographically weighted discriminant analysis versus its global counterpart, linear discriminant analysis. *International Journal of Geographical Information Science*, Vol. 27, No. 4, pp. 633-611.

Fotheringham, A. S., Brunsdon, C., and Charlton, M., (2013) The demographic impacts of the Irish famine: towards a greater geographical understanding. *Transactions of the Institute of British Geographers*, Vol. 38, No. 2, pp. 221-237.

Fotheringham, A. S., Brunsdon, C. and Charlton, M. (2002) *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. Chichester: Wiley.

Fotheringham,A. S., Brunsdon, C. and Charlton, M. (2000) *Quantitative Geography: Perspectives on Spatial Data Analysis*. London: Sage.

Fotheringham, A. S., Charlton, M. and Brunsdon, C. (2001) Spatial Variations in School Performance: a Local Analysis Using Geographically Weighted Regression. *Geographical & Environmental Modelling*, Vol. 5, No. 1, pp. 43- 66.

Fotheringham, A. S., Charlton, M. and Brunsdon, C. (1998) Geographically weighted regression: A natural extension of the expansion method for spatial data analysis. *Environment and Planning A*, Vol. 30, pp. 1905–1928.

Fotheringham, A. S., Charlton, M. and Brunsdon, C. (1997) Two Techniques for Exploring Non-Stationarity in Geographical Data. *Geographical Systems*, Vol. 4, pp. 59-82.

Fuhrmann, S., Ahonen-Rainio, P., Edsall, R. M., Fabrikant, S. I., Koua, E. L., Tobón, C., Ware, C. and Wilson, S. (2005) Making useful and useable Geovisualisation: design and evaluation issues. In: Dykes, J., MacEachren, A. M. and Kraak, J. M., eds. (2005) *Exploring Geovisualisation.* Amsterdam: Elsevier Ltd., pp.553-566.

Gabbard, L. J., Hix, D. and Swan, E. J. (1999) User-Centered Design and Evaluation of Virtual Environments. *IEEE Computer Graphics and Applications*, Vol. 19, No. 6, pp. 51-59.

Gahegan, M. (2001) Visual exploration in geography: analysis with light. In: Miller, J. H. and Han, J., eds. (2001) *Geographic Data Mining and Knowledge Discovery.* London: Taylor and Francis, pp. 260-288.

Gamerman, D., Moreira, B. R. A. and Rue, H. (2003) Space-varying regression models: specifications and simulation. *Computational Statistics & Data Analysis*, Vol. 42, No. 3, pp. 513–533.

Gao, J. and Li, S. (2011) Detecting spatially non-stationary and scale-dependent relationships between urban landscape fragmentation and related factors using Geographically Weighted Regression. *Applied Geography*, Vol. 31, pp. 292-302.

Gebreab, Y. S. and Diez Roux, V. A. (2012) Exploring racial disparities in CHD mortality between blacks and whites across United States: A geographically weighted regression approach. *Health and Place*, Vol. 18, pp. 1006-1014.

Geisler, W. S. and Perry, J. S. (2002) Real-time simulation of arbitrary visual fields, In: *Proceedings of the 2002 Symposium on Eye Tracking Research and Applications*, March 2002, New York: AMC, pp.83-87.

GEOVISTA Centre (2015) *GEOVISTA Studio Application [online*], Available at: http://www.geovistastudio.psu.edu/jsp/whatCanYouDo.jsp, accessed 5th May 2015.

Gibson, J. J. (1979) *The Ecological Approach to Visual Perception*, Michigan: Houghton-Mifflin.

Gilbert, A. and Chakraborty, J. (2011) Using geographically weighted regression for environmental justice analysis: Cumulative cancer risks from air toxics in Florida. *Social Science Research*, Vol. 40, pp. 273-28.

Goldberg, J. H., Stimson, M. J., Lewenstein, M., Scott, N. and Wichansky, A. M. (2002) Eye Tracking in Web Search Tasks: Design Implications. In: *Proceedings of the 2002 Symposium on Eye Tracking Research and Applications*, March 2002, New York: AMC, pp. 51-58.

Goldstein, J., Roth, S. F., Kolojejchick, J. and Mattis, J. (1994) A framework for knowledge-based interactive data exploration. *Journal of Visual Languages and Computing*, Vol. 5, pp. 139-155.

Graif, C. and Sampson, R. J. (2009) Spatial Heterogeneity in the Effects of Immigration and Diversity on Neighbourhood Homicide Rates. *Homicide Studies*, Vol. 13, No. 3, pp. 242-260.

Griffith, D. (2008) Spatial-filtering-based contributions to a critique of geographically weighted regression (GWR). *Environment and Planning A*, Vol. 40, pp. 2571-2769.

Guo, L., Ma., Z. and Zhang, L. (2008) Comparison of bandwidth selection in application of geographically weighted regression: a case study. *Canadian Journal of Forest Research*, Vol. 38, pp. 2526 - 2534.

Haklay, M., and Tobón, C. (2003) Usability evaluation and PPGIS: towards a user-centred design approach. *International Journal for Geographical Information Science*, Vol. 17, No. 6, pp. 577-592.

Harris, P. and Brunsdon, C. (2010) Exploring spatial variation and spatial relationships in a fresh water acidification critical load data set for Great Britain using geographically weighted summary statistics. *Computer and Geosciences*, Vol. 36, pp. 54-70.

Harrower, M., MacEachren, A. M. and Griffin, A. (2000) Developing a geographic visualisation tool to support Earth Science learning. *Cartography and Geographic Information Science*, Vol. 27, pp. 279-294.

Harrower, M. and Fabrikant, S. (2009) The role of map animation for geographic visualization. In Dodge, M., McDerby, M. and Turner, M., eds. (2009) Geographic Visualization: Concepts, Tools and Applications, pp.49-65, Chichester: John Wiley & Sons Ltd.

Hawkins, T. (2003) Geostatistical Analysis of Arizona Summertime Precipitation. *Journal of the Arizona-Nevada Academy of Science*, Vol. 36, No. 1, pp. 9-17.

Hayhoe, M. and Ballard, D. (2005) Eye movements in natural behaviour. *Trends in Cognitive Science*, Vol. 9, No. 4, pp. 188-194.

Heinrich, J. and Weiskopf, D. (2012) State of the Art of Parallel Coordinates, *Proceedings of the European Associated for Computer Graphics (Eurographics)*, Cagliari, Italy.

Hewett, T. T., Baecker, R. R., Card, S. K., Carey, T., Gasen, J., Mantei, M., Perlman, G., Strong, G. and Verplank, W. (2002) *Curricula for Human Computer Interaction* [online]. Available at: http://old.sigchi.org/cdg/cdg2.html#2_1, accessed 12/08/2010.

Holt, J. B. and Lo, C. P. (2008) The geography of mortality in the Atlanta metropolitan area, Computers. *Environment and Urban Systems*, Vol. 32, pp. 149-164.

Hsueh, H. Y., Lee, J. and Beltz, L. (2012) Spatio-temporal patterns of dengue fever cases in Kaoshiung City, Taiwan, 2003-2008. *Applied Geography*, Vol. 34, pp. 587-594.

Huang, Y. and Leung Y. (2002) Analysing regional industrialisation in Jiangsu province using geographically weighted regression. *Journal of Geographical Systems*, Vol. 4, pp. 233-249.

Hunter, J. E. and Schmidt, F. L. (1990) *Methods of meta-analysis: correcting error and bias in research findings*. London: Sage.

Hunter, J. E., Schmidt, F. L. and Jackson, G. B. (1982) *Meta-analysis: cumulating research findings across studies*. California: Sage Publications.

Hurley, C. and Buja, A. (1990) Analyzing high-dimensional data with motion graphics. *SIAM Journal of Statistical Computing*, Vol. 11, No. 6, pp. 1193–211.

Jaimes, P. B. N., Sendra, B. J., Delgado, G. M. and Plata, F. R. (2010) Exploring the driving forces behind deforestation in the state of Mexico using geographically weighted regression. *Applied Geography*, Vol. 30, pp.576-591.

Jenks, G. (1953) An improved cartographic curriculum for cartographic training at the college and university level. *Annals of the Association of American Geographers*, Vol. 43, No. 4, pp. 317-31. Reprinted in *Surveying and Mapping*, Vol. 14, No. 3, pp. 319-330.

Kahraman, H. E. Z. (2010) Using user-centered design approach in course design. *Procedia Social and Behavioural Sciences*, Vol. 2, pp. 2071-2076.

Kang, Y.-A., Gorg, C., and Stasko, J. (2011). How Can Visual Analytics Assist Investigative Analysis? Design Implications from an Evaluation. *IEEE Transactions on Visualization and Computer Graphics*, 17(5):570–583.

Kara, L. B., Shimada, K., and Marmalefsky, S. D. (2007) An evaluation of user experience with a sketch-based 3D modelling system, Computers and Graphics, Vol. 31, pp. 580 -597.

Karat, J. (1997) User Centered software evaluation methodologies. In: Helander, M., Landauer, K. T. and Prabhu, P., eds. (1997) *Handbook of human-computer interaction*. Amsterdam: Elsevier Ltd., pp. 689-704.

Kavanagh, A., Mills, G. and Sinnott, R. (2004) The geography of Irish voter turnout: A case study of the 2002 General Election. *Irish Geography*, Vol. 37, pp.177-186.

Kavanagh, A., Sinnot, R., Fotheringham, A. S. and Charlton, M. (2006) A Geographically Weighted Regression Analysis of General Election Turnout in the Republic of Ireland. Paper presented at the *Political Studies Association of Ireland Conference*, October 2006. University College Cork.

Keim, D. A., Hao, M. C., Dayal, U., Janetzko, H. and Bak, P. (2010) Generalized scatter plots, Information Visualization, Vol. 9, No. 4, pp. 301-311.

Keim, D., Kohlhammer, J., Ellis, G., and Mansmann, F., editors (2010). *Mastering the Information Age: Solving Problems with Visual Analytics*. Eurographics Association. Available from : http://www.vismaster.eu/wpcontent/uploads/2010/11/VisMaster-book-lowres.pdf.

Keim, D., Panse, C., Sips, M., and North, S. C. (2004). Pixel based visual data mining of geo-spatial data. *Computers & Graphics*, Vol. 28, No. 3, pp. 327–344.

Kelly, M. and Fotheringham, A.S. (2011) The online atlas of Irish population change 1841-2002: a new resource for analysing national trends and local variations in Irish population dynamics. *Irish Geography*, Vol. 44, No. 2-3, pp.215-244.

Knapp, L. (1995) A task analysis approach to the visualisation of geographic data. In: Nyerges, L. T., Mark, M. D., Laurini, R. and Egenhofer, J. M., eds. (1995) *Cognitive aspects of humancomputer interaction for geographic information systems*. The Netherlands: Springer, pp. 355-72.

Kok, V.T.K and Van Liere, R. (2007) A multimodal virtual reality interface for 3D interaction with VTK. *Knowledge and Information Systems*, Vol. 11, No. 3, pp.197-219.

Koláčný. A. (1969) Cartographic Information—a Fundamental Concept and Term in Modern Cartography. *The Cartographic Journal,* Vol. 6, No. 1, pp. 47-49.

Koua, E. L., MacEachren, A. M. and Kraak, M. J. (2006) Evaluating the usability of visualisation methods in an exploratory geovisualisation environment. *International Journal of Geographical Information Science*, Vol. 20, No. 4, pp. 425-448.

Koua, E. L. and Kraak, M. J. (2005) Evaluating self-organising maps for geovisualisation. In: Dykes, J., MacEachren, A. M. and Kraak, J. M., eds. (2005) *Exploring Geovisualisation.* Amsterdam: Elsevier Ltd., pp . 627-643.

Kraak, M. J. and Klomp, A. (1995) A classification of cartographic animations: towards a tool for the designing of dynamic maps in a GIS environment. In: Ormeling, F., Kobben, B. and PerezGomez, R., eds. (1995) *Proceedings of the Seminar on Teaching Animated Cartography*. August/September 1995, Madrid: p. 29-37.

Kraak, M-J., and Omreling F. J. (1996) *Cartography, Visualization of Spatial Data*. Essex: Addison Wesley Longman.

Kupfer, J. A. and Farris, C. A. (2007) Incorporating spatial non-stationarity of regression coefficients into predictive vegetation models. *Landscape Ecology*, Vol. 22, pp. 837-852.

Laffan, S. W. (1999) Spatially assessing model error using geographically weighted regression. In: Diaz, J., Tynes, R., Caldwell, D. and Ehlen, J., eds. (1999) *Proceedings of the 4th International Conference on Geocomputation*, July 1999. Virginia: Geocomputation CDROM. Available at: http://www.geocomputation.org/1999/, (accessed 30/03/2014).

Lai, M-L., Tsai, J-M., Yang, Y-F., Hsu, C-Y., Liu, T-C., Lee, S., W-L., Lee, M-H., Chiou, G-L., Liang, JC. and Tsai, C-C. (2013) A review of using eye-tracking technology in exploring learning from 2000-2012. *Educational Research Review*, Vol. 10, pp. 90-115.

Landolt, E. (1879) *The artificial eye*, London: Trübner and Co.

Lave, J. and Wegner, E. (1991) *Situated Learning Legitimate Peripheral Participation*, Cambridge: Cambridge University Press.

Leung, Y., Mei, L. and Zhang X. W. (2000) Testing for spatial autocorrelation among the residuals of the geographically weighted regression. *Environment and Planning A*, Vol. 32, pp. 871-890.

Leyk, S., Maclaurin, G. J., Hunter, L. M., Nawrotzki, R., Twine, W., Collinson, M and Erasmus, B. (2012) Spatially and Temporally Varying Associations between Temporary Outmigration and Natural Resource Availability in Resource-Dependent Rural Communities in South Africa: A Modelling Framework. *Applied Geography*, Vol. 34, pp. 559-568.

Li, F., Li, M. and Liang, J. (2007) Study on disparity of regional economic development based on geoinformatic Tupu and GWR model - A case of growth of GDP per capita in China from 1999 to 2003. In: *Proceedings of SPIE 6754, Geoinformatics 2007: Geospatial Information Technology and Applications*, Vol. 6754, 3A.

Licklider, J. C. R. (1960) Man-computer symbiosis. *IRE Transactions on Human Factors in Electronics*, Vol. 1, pp. 4-11.

Liversedge, P. and Findlay, J. M. (2000) Saccadic eye movements and cognition. *Trends in Cognitive Sciences*, Vol. 4, No. 1, pp. 6-14.

Lobo, M. J., Pietriga, E., and Appert, C. (2015) An Evaluation of Interactive Map Comparison Techniques. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, Soeul, South Korea, 2015.

Lopéz-Carr, D., Davis, J., Jankowska, M. M., Grant, L., López-Carr, C. A. and Clark., M. (2012) Space versus place in complex human-natural systems: Spatial and multi-level models of tropical land use and cover change (LUCC) in Guatemala. *Ecological Modelling*, Vol. 229, pp. 64-75.

Lu, B., Charlton, M. and Fotheringham, A. S. (2011) Geographically Weighted Regression Using a Non-Euclidean Distance Metric with a Study on London House Price Data. *Procedia Environmental Sciences,* Vol. 7, pp. 92-97.

Lucieer, A. and Kraak, M-J. (2004) Interactive and visual fuzzy classification of remotely sensed imagery for exploration of uncertainty. *International Journal of Geographical Information Science*, Vol. 18, No. 5, pp. 491-512.

Luo, J. and Wei, D. Y. H. (2009) Modeling spatial variations of urban growth patterns in Chinese cities: The case of Nanjing. *Landscape and Urban Planning*, Vol. 91, pp. 51-64.

Lyndsay, P. and Norman, D. A. (1977) *Human Information Processing: An Introduction to Psychology*. New York: New York Academic Press Inc.

Ma, Z., Zuckerberg, B., Porter, F. and Zhang, L. (2012) Use of localized descriptive statistics for exploring the spatial pattern changes of bird species richness at multiple scales. *Applied Geography*, Vol. 32, pp. 185-194.

MacEachren, A. M. (1995 and 2004) *How Maps Work: Representation, Visualisation and Design*, New York: The Guilford Press.

MacEachren, A. M., (1994a) *Some Truth with Maps: A Primer on Symbolization & Design.* Washington DC: Association of American Geographers.

MacEachren, A. M. (1994b) Visualisation in modern cartography: Setting the Agenda. In: MacEachren, A. M. and Taylor, D. R . K., eds. (1994) *Visualisation in Modern Cartography*, Oxford: Pergamon, pp. 1-12.

MacEachren, A. M. (1992) Visualizing Uncertain Information. *Cartographic Perspective*, Vol. 13, pp. 10-19.

MacEachren, A. M. (1982) Map complexity: Comparison and measurement. *The American Cartographer*, Vol. 9, No. 1, pp. 31-46.

MacEachren, A. M., Buttenfield, B., Campbell, J., DiBiase, D. and Monmonier, M. (1992) Visualization. In: Abler, R., Marcus, M. and Olson, J., eds. (1992) *Geography's Inner Worlds: Pervasive Themes in Contemporary American Geography*, New Brunswick, New Jersey: Rutgers University Press, pp. 99–137.

MacEachren, A. M. and DiBiase, D. (1991) Animated maps of aggregate data: Conceptual and practical problems. *Cartography and Geographic Information Systems*, Vol. 18, No. 4, pp. 221-229.

MacEachren, A. M. and Kraak, M. J. (2001) Research challenges in geovisualisation. *Cartography and GeoInformation Science*, Vol. 28, pp. 3-12.

MacEachren, A. M. and Kraak, M. J. (1997) Exploratory cartographic visualization: Advancing the agenda. *Computers and Geosciences*, Vol. 23, No. 4, pp. 335-43.

MacEachren, A. M. and Taylor, D. R. F., eds. (1994) *Visualization in Modern Cartography*. Oxford: Pergamon Press.

MacEachren, A. M., Wachowicz, M., Edsall, R. and Haug, D. (1999) Constructing knowledge from multivariate spatiotemporal data: Integrating geographical visualization with knowledge discovery in database methods. *International Journal of Geographic Information Science*, Vol. 13, No. 4, pp. 311-334.

Mandal, R., St-Hilaire, S., Kie, J. G. and Derryberry, D. (2009) Spatial trends of breast and prostate cancers in the United States between 2000 and 2005. *International Journal of Health Geographics*, Vol. 8, No. 53.

Matthews, S. A. and Yang, C. T. (2012) Mapping the results of local statistics: Using Geographically Weighted Regression. *Demographic Research*, Vol. 26, No. 6, pp.151-166.

McArdle, G. and Demšar U. (2011) investigating similarity of trajectories through physical decomposition of movement in space-time cube. *AGILE 2011*, Utrecht, The Netherlands.

Mittal, V., Kamakura, A. W. and Govind, R. (2004) Geographic Patterns in Customer Service and Satisfaction: An Empirical Investigation. *Journal of Marketing*, Vol. 68, No. 3, pp. 48-62.

Monmonier, M. (1989) Geographical brushing: Enhancing exploratory analysis of the scatterplot matrix. *Geographical Analysis*, Vol. 21, No. 1, pp. 81-84.

Montello, D. R., Fabrikant, S. I., Ruocco, M. and Middleston, R. S. (2003) Testing the First Law of Cognitive Geography on Point-Display Spatializations. *Spatial Information Theory: Foundations of Geographic Information Science, Lecture Notes in Computer Science*, Vol. 2825, pp. 316-331.

Moore, A. and Drecki, I. (eds.) (2013) *Geospatial Visualisation*. Berlin, Springer-Verlag.

Morgan, D. and Krueger, R. A. (1997) *The Focus Group Kit*. London: Sage.

Morrison, J. L. (1997) Topographic mapping in the twenty-first century. In: Rhind, D. W., ed. (1997) *Framework for the World*, Cambridge: GeoInformation, pp. 14-28.

Muntz, R. R., Barclay, T., Dozier, J., Faloutsos, C., MacEachren, A. M., Martin, J. L., Pancake, C. M. and Satyanarynan, M. (2003) *IT Road Map to a Geospatial Future*. Report of the Committee on Intersections between Geospatial Information and Information Technology, National Research Council of the National Academies, Washington, DC: The National Academies Press.

Nakaya, T. (2001) Local spatial interaction modelling based on the geographically weighted regression approach. *GeoJournal*, Vol. 53, pp. 347-358.

Nakaya, T., Fotheringham, A. S., Brunsdon, C. and Charlton, M. (2005) Geographically weighted Poisson regression for disease association mapping. *Statistics in Medicine*, Vol. 24. pp. 2695–2717.

National Eye Institute (2014) *Structure of the Eye* [online]. Available at: http://www.nei.nih.gov/health/coloboma/images/eye_with_labels.jpg, accessed 29/10/2013.

Nelson, A. (2001) Analysing data across geographic scales in Honduras: Detecting levels of organisation within systems. *Agriculture, Ecosystems and Environment*, Vol. 85, pp. 107–131.

Nelson, T., H. (1965) A file structure for the complex, the changing, and the indeterminate. In: Winner, L., ed. (1965) *Proceedings of the ACM 20th National Conference 1965*, August 1965. New York: AMC New York, pp. 84-100.

Nielsen, J. (1994) Heuristic Evaluation. In: Nielsen, J. and Mack, R. L., eds., (1994) *Usability Inspection Methods*. New York: John Wiley and Sons, pp. 25-62.

Nielsen, J. (1993) *Usability Engineering*. London: Academic Press Ltd.

Nielsen, J. (1993a) Guerilla HCI: using discount usability engineering to penetrate the intimidation barrier. In: Bias, R. and Mayhew, D., eds. (1993) *Cost-Justifying Usability.* Boston: Academic Press, pp. 245-272.

Norton, D. and Stark, I. (1971) Eye movements and visual perception. *Scientific American*, Vol. 224, pp. 35–43.

Novick, L. R. (1988) Analogical transfer, problem similarity, and expertise. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, Vol., 14, pp. 510-520.

Nyerges, T. L., Mark, D. M., Laurini, R. and Egenhofer, M. J., eds. (1995a), *Cognitive Aspects of Human-Computer Interaction for Geographic Information Systems*. Proceedings of the NATO Advanced Research Workshop, March 1994, NATO Science Series D, Vol. 83, The Netherlands: Kluwer Academic Publishers.

Ogao, P. J. and Kraak, M. J. (2002) Defining Visualization operations for temporal cartographic animation design. *International Journal of Applied Earth Observation and GeoInformation*, Vol. 4, pp. 22-31.

Ogneva-Himmelberger, Y., Pearsall, H. and Rakshit, R. (2009) Concrete evidence & geographically weighted regression: A regional analysis of wealth and the land cover in Massachusetts. *Applied Geography*, Vol. 29, pp. 478 - 487.

Osborne, P. E., Foody, G. M. and Suárez-Seoane, S. (2007) Non-stationarity and local approaches to modelling the distributions of wildlife. *Diversity and Distributions*, Vol. 13, pp. 313-323.

O' Sullivan, D. and Unwin, D. J. (2010) *Geographic Information Analysis*. 2nd ed. New Jersey, John Wiley and Sons Inc.

Oxford Dictionaries (2014) *Visual Acuity* [online]. Available at: http://www.oxforddictionaries.com/definition/english/visual-acuity?q=visual+acuity, (accessed 30/1/2014).

Páez, A. (2006) Exploring contextual variations in land use and transport analysis using a probit model with geographical weights. *Journal of Transport Geography*, Vol. 14, pp. 167-176.

Páez, A., Long, F. and Farber, S. (2008) Moving Window Approaches for Hedonic Price Estimation: An Empirical Comparison of Modelling Techniques. *Urban Studies*, Vol. 45, No. 8, pp. 1565-1581.

Páez, A., Takashi, U. and Miyamoto, K. (2002) A general framework for estimation and inference of geographically weighted regression models: Location-specific kernel bandwidths and a test for locational heterogeneity. *Environment and Planning A*, Vol. 34, pp. 733-754.

Pearsall, H. and Chirstman, Z. (2012) Tree-lined lanes or vacant lots? Evaluating nonstationarity between urban greenness and socio-economic conditions in Philadelphia, Pennsylvania, USA at multiple scales. *Applied Geography*, Vol. 35, No. 1-2, pp. 257-24.

Peterson, M. P. (1995) *Interactive and Animated Cartography*. New Jersey: Prentice Hall.

Pieters, R., Rosbergen, E. and Wedel, M. (1999) Visual Attention to Repeated Print Advertising: A Test of Scanpath Theory. *Journal of Marketing Research*, Vol. 36, No. 4, pp. 424-438.

Pinker, S. (1990). A theory of graph comprehension. In: Freedle, R., ed. (1990) *Artificial intelligence and the future of testing*, New Jersey: Lawrence Erlbaum Associates Inc., pp. 73-126.

Plaisant, C. (2004) The Challenge of Information Visualisation Evaluation. In: Costabile, M. F., ed. (2004) *Proceedings of the Working Conference on Advanced Visual Interfaces*, May 2004. New York: ACM New York, pp. 109-116.

Preece, J., Rogers, Y., Sharp, H., Benyon, D., Holland, S., and Carey T. (1994) Human-Computer *Interaction.* Harlow: Addison-Wesley.

Preece, J., Benyon, D., Davies, G., Keller, L. and Rogers, Y. (1993) *A Guide to Usability: Human Factors in Computing*, Dorset: Addison Wesley.

Propastin, P. A. (2009) Spatial non-stationarity and scale-dependency of prediction accuracy in the remote estimation of LAI over a tropical rainforest in Sulawesi, Indonesia. *Remote Sensing of Environment*, Vol. 113, pp. 2234-2242.

Pullar, D. V. and Tidey, M. E. (2001) Coupling 3D visualisation to qualitative assessment of built environment designs. *Landscape and Urban Planning,* Vo. 55, No. 1, pp. 29-40.

Rayner, K. (1998) Eye Movements in Reading and Information Processing: 20 Years of Research. *Psychological Bulletin*, Vol. 124, No. 3, pp. 372-422.

Robins, N.S., Rutter, H.K., Dumpleton, S. and Peach D.W. (2005) The role of 2D visualisation as an analytical tool preparatory to numerical modelling, Journal of Hydrology, Vol. 201, pp. 287295.

Robinson, P. D., Lloyd, D. C. and McKinley, M. J. (2013) Increasing the accuracy of nitrogen dioxide (NO2) pollution mapping using geographically weighted regression (GWR) and geostatistics. *International Journal of Applied Earth Observation and Geoinformation*, Vol. 21, pp 374–383.

Robinson, A.C., Kersi, J., Long, E.C., Luo, H, DiBiase, D. and Lee, A. (2015) Maps and the geospatial revolution: teaching a massive open online course (MOOC) in geography. *Journal of Geography in Higher Education*, Vol. 39, No. 1, pp.62-82.

Rogers, Y. (1999) What is different about interactive graphical representations? *Learning and Instruction,* Vol. 9, pp. 419-425.

Rogerson, P. A. (2006) *Statistical Methods for Geography: A Students Handbook*, 2nd ed., London: SAGE Publications.

Roth, S., F. and Mattis, J. (1990) Data characterization for intelligent graphics presentation. In: Carrasco-Chew, J. and Whiteside, J., eds. (1990) *Proceedings of CHI 1990, SIGCHI Conference on Human Factors in Computing Systems*. Seattle, April 1990. New York: ACM New York, pp. 193-200.

Sage, L. J. and Goldberger, R. J. (2012) Decisions to direct market: Geographic influences on conventions in organic production. *Applied Geography*, Vol. 34, pp. 57-65.

Salom, P., Megret, R., Donias, M. and Berthoumieu Y. (2009) Dynamic picking system for 3D seismic data: Design and evaluation. *International Journal of Human-computer Studies*, Vol. 67, pp. 551-560.

Schnur, S., Bektas K., Salahi M. and Çöltekin A. (2010) A Comparison of Measured and Perceived Visual Complexity for Dynamic Web Maps. *Proceedings of GIScience 2010, 6th International Conference on Geographic Information Science*. Zurich, September 2010. pp. 1-4.

Shearmur, R., Apparicio, P., Lizion, P. and Polése, M. (2007) Space, Time, and Local Employment Growth: An Application of Spatial Regression Analysis. *Growth and Change*, Vol. 38, No. 4, 696-722.

Shekhar, S., Evans, M. R., Kang, J. M. and Mohan, P. (2011) Identifying patterns in spatial information: A survey of methods. *Data Mining and Knowledge Discovery*, Vol 1, pp. 193–214.

Shneiderman, B. (1996) The eyes have it: A task by data type taxonomy of information visualizations. In: *Proceedings of the 1996 IEEE Symposium on Visual Languages*. September 1996. Washington DC: IEEE Computer Society, pp. 336-343.

Slocum. T. A., McMaster, R. B., Kessler, F. C. and Howard, H. H. (2009) *Thematic Cartography and Geovisualization*. New Jersey: Pearson Prentice Hall.

Stanger, N. (2008) Scalability of techniques for online geographic visualization of web site hits. In: Moore, A. and Drecki, I., eds. (2008) *Geospatial Vision: New Dimensions in Cartography.* Berlin: Springer, pp. 193-217.

Su, S., Xiao, R. and Zhang, Y. (2011) Multi-scale analysis of spatially varying relationships between agricultural landscape patterns and urbanization using geographically weighted regression. *Applied Geography*, Vol. 32, pp. 360-375.

Suchman, L. (1987) *Plans and Situated Actions: the Problem of Human-Machine Communication*. Cambridge: Cambridge University Press.

Svenning, J-C., Normand, S. and Skov, F. (2009) Plio-Pleistocene climate change and geographic heterogeneity in plant diversity-environment relationships. *Ecography*, Vol. 32, pp. 13-21.

Swienty, O. and Reichenbacher, T. (2008) Evaluating the visual scanning efficiency of geovisualisation displays. In: *Proceedings of Geospatial Visual Analytics Workshop: GIScience 2008*, September 2008. Park City, USA, pp.23-26.

Szymanowski, M. and Kryza, M. (2011) Application of geographically weighted regression for modelling the spatial structure of urban heat island in the city of Wroclaw (SW Poland). *Procedia Environmental Sciences*, Vol. 3, pp. 87-92.

Talter, W. B., Wade, J. N., Kwan, H., Findlay, M. J. and Velichkovsky, M. B. (2010) Yarbus, eye movements, and vision. *i-Perception*, Vol. 1, pp. 7-27.

Taylor, D. R. F. (1997) Maps and Mapping in the Information Era, *Keynote address to the 18th International Cartographic Association Conference*. 12 June 1997, Stockholm, Sweden.

Tech Smith (2010) *Camtasia Studio: Screen Recording and Video Editing Software* [online]. Available at: http://www.techsmith.com/camtasia/ (accessed 10th November 2010).

Terribile, l. C. and Diniz-Filho, J. A. F. (2009) Spatial Patterns of species richness in New World coral snakes and the metabolic theory of ecology. *Acta Oecologica*, Vol. 35, pp. 163-173.

Thomas, P. and Macredie, R. D. (2002) Introduction to new usability. *Journal of ACM Transactions on Computer-Human Interaction (TOCHI)*, Vol. 9, No. 2, pp. 69-73.

Tobler, W.R. (1987) Experiments in migration mapping by computer. *The American Cartographer*, Vol. 14, No. 2, pp. 155-163.

Tobler, W. R. (1976) Analytical Cartography. *The American Cartographer*, Vol. 3, No.1, pp. 21-31.

Tobler W.R. (1970) A computer movie simulating urban growth in the Detroit region. *Economic Geography*, Vol. 46, No. 2, pp. 234-240.

Tobón, C. (2005) (Evaluating geographic visualisation tools and methods: an approach and experiment based upon user tasks. In: Dykes, J., MacEachren, A. M. and Kraak, J. M., eds. (2005) *Exploring Geovisualisation.* Amsterdam: Elsevier Ltd., pp. 546-666.

Traynor, C. and Williams, M. G. (1995) Why are Geographic Information Systems hard to use? In: Katz, I., Mack, R. and Marks, L., eds. (1995) *CHI '95 Conference Companion on Human Factors in Computing Systems*. Denver, May 1995. New York: ACM New York, pp. 288-289.

Tu, J. (2011) Spatially varying relationships between land use and water quality across an urbanization gradient explored by geographically weighted regression. *Applied Geography*, Vol. 31, pp. 376-392.

Tu, J. and Xia, G. Z. (2008) Examining spatially varying relationships between land use and water quality using geographically weighted regression I: Model design and evaluation. *Science of the Environment*, Vol. 407, pp. 358-378.

Tu, J., Tu. W. and Tedders, H. S. (2012) Spatial Variation in the associations of birth weight with socioeconomic environmental, and behavioral factors in Georgia, USA. *Applied Geography*, Vol 34, pp. 331-344.

Tufte, E. R. (2007) *The Visual Display of Quantitative Information.* 2nd ed. Connecticut: Graphic Press.

Tukey, J. W. (1977) *Exploratory Data Analysis*. Massachusetts: Addison-Wesley Publishing Company.

Tulloch, D. (2008) Public Participation GIS (PPGIS). In Kemp, K. ed., (2008) *Encyclopedia of geographic information science*, California: Sage Publications, pp.352-355.

Tutmez, B., Kaymak, U., Erhan-Tercan, A. and Llyod, D. C. (2012) Evaluating geo-environmental variables using clustering based area model. *Computers and Geosciences*, Vol. 43, pp. 34-41.

UAUUG (2012) *What are the advantages of data visualisation.* Available at: http://www.uauug.org.uk/what-are-the-advantages-of-data-visualisation.html.

Vessy, I. (1991) Cognitive fit: a theory based analysis of the graphs versus table literature. *Decision Sciences*, Vol. *22, No.* 2, pp.219-240.

Virrantuas, K., Fairbairn, D., and Kraak, M., J., (2009) ICA Research Agenda on Cartography and GIScience. *Cartography and Geographic Information Science*, Vol. 36, No. 2, pp. 209-222.

Wade, J. N. and Tatler W. B. (2005) T*he moving tablet of the Eye: Origins of Modern Eye Movement Research*. Oxford: University Press.

Wang, N., Mei, C-L. and Yan, X-D. (2008) Local linear estimation of spatially varying coefficient models: an improvement on the geographically weighted regression technique. *Environment and Planning A*, Vol. 40, pp. 986-1005.

Wang, Q., Zhao, P., Ren, H. and Kakubari, Y. (2008) Spatiotemporal dynamics of forest net primary production in China over the past two decades. *Global and Planetary Change*, Vol. 61, pp. 267-274.

Ware, C. (2008) *Visual Thinking for Design*. Burlington, MA: Morgan Kauffmann Publishers.

Wegman, E. J. (1990) Hyperdimensional data analysis using parallel coordinates. *Journal of the American Statistical Association*, Vol. 85, No. 411, pp. 664–75.

Wehrend, S. and Lewis, C. (1990) A Problem-Oriented Classification of Visualisation

Techniques. In: Kaufman, A. (1990) *Proceedings of the 1st Conference on Visualisation*. October 1990, California: IEEE Computer Society Press, pp.139-143.

Wei, D. Y. H. and Ye, X. (2009) Beyond Convergence: Space, Scale, and Regional Inequality in China. *Tijdschrift voor Economische en Sociale Geografie*, Vol. 100, No. 1, pp. 59-80.

Weldon, J. L. (1996) Data mining and visualization. *Database Programming and Design*, Vol. 9, pp. 21-24

Wentz E. A. and Gober, P. (2007) Determinants of Small-Area Water Consumption for the City of Phoenix, Arizona. *Water Resource Management*, Vol. 21, pp. 1849-1863.

Wheeler, D. C. (2009) Simultaneous coefficient penalization and model selection in geographically weighted regression: the geographically weighted lasso. *Environment and Planning A*, Vol. 41, pp. 722-742.

Wheeler, D. C. and Calder, C. A. (2007) An assessment of coefficient accuracy in linear regression models with spatially varying coefficients. *Journal of Geographical Systems*, Vol. 9, pp. 145-166.

Wheeler, D. C. and Páez, A. (2009) Geographically weighted regression. In: Fischer, M. M. and Getis, A., eds. (2009) *Handbook of Spatial Analysis.* Oxford: Elsevier Ltd., pp. 461-486.

Wheeler, D. C. and Waller, L. A. (2009) Comparing spatially varying coefficient models: a case study examining violent crime rates and their relationships to alcohol outlets and illegal drug arrests. *Journal of Geographical Systems*, Vol. 11, pp. 1-22.

Wimberly, M. C., Yabsley, M., Baer, A. D., Dugan, V. G. and Davidson, R. (2008) Spatial heterogeneity of climate and land-cover constraints on distributions of tick-borne pathogens. *Global Ecology and Biogeography*, Vol. 17, pp. 189-202.

Xiao, N. and Armstrong, M. P. (2006) Choroware: A software toolkit for chloropleth map classification. *Geographical Analysis*, Vol. 38, No. 1, pp. 102-120.

Yang, T. C. and Matthews, S. A. (2012) Understanding the non-stationarity associations between distrust of the health care system, health conditions and self-rated health in the elderly: A geographically weighted regression approach. *Health and Place*, Vol. 18, pp. 576-585.

Yarbus, L. A. (1967) *Eye Movement and Vision*. New York: Plenum Press.

Yoo, D. (2012) Height and death in the Antebellum United States: A view through the lens of geographically weighted regression. *Economics and Human Biology*, Vol. 10, pp. 43-53.

Yost, B., Haciahmetogly, Y. and North, C. (2007) Beyond Visual Acuity: The Perceptual Scalability of Information Visualisations for Large Displays. In: Gilmore, D., ed. (2007) *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. California, April 2007. New York: ACM New York, pp. 101-110.*

Yost B. and North C. (2006) The Perceptual Scalability of Visualization. *IEEE Transaction on Visualization and Computer Graphics*, Vol. 12, No. 5, pp. 837-844.

Zhang, L., Bi, H., Cheng, P. and Davis, J. C. (2004) Modelling spatial variation in tree diameter-height relationships. *Forest Ecology and Management*, Vol. 189, No 1-3, pp. 317-329.

Zhang, L., Gove, H. G. and Heath, S. L. (2005) Spatial residual analysis of six modelling techniques. *Ecological Modelling*, Vol. 186, pp. 154-177.

Zhang, L., Ma, Z. and Guo, L. (2008) Spatially assessing model errors of four regression techniques for three types of forest stands. *Forestry*, Vol. 81, No. 2, pp. 209-225.

Zhang, P., Wong., W. D., So., L. K. B. and Lin, H. (2012) An exploratory spatial analysis of western medical services in Republic Beijing. *Applied Geography*, Vol. 32, pp. 556-565.

Zhou, M. X. and Feiner, S. K. (1998) Visual Tasks Characterisation for Automated Visual Discourse Synthesis. In: Karat, M. C., Lund, A., Coutaz, J. and Karat, J., eds. (1998) *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* Los Angeles, April 1998. pp. 392-399.

Zudilova-Seinstra, E.V., de Koning P.J.H., Suinesiaputra, A., van Schooten, B.W., van der Geest, R.J., Reiber, R.H.C., and Sloot, P.M.A (2010) Evaluation of 2D and 3D glove input applied to medical image analysis, *International Journal of Human-Computer Studies*, Vol. 68, pp. 355-369.

Appendix

Appendix 1: Experiment One

a. Ethical Application Form

National University of Ireland Maynooth - Application Forms
National University of Ireland Maynooth Ethics Committee

This section contains the following forms:

1. Application Form involving research with Human participants (with the exception of the medical or clinical based projects)

2. Application Form for Ethical Approval of a Research Project involving participation of Humans or Human Derived Material

Please complete sections 1-16 (where applicable) Applications should be emailed to research.ethics@nuim.ie

1. **Name of applicant:** Tommy Burke
2. **Appointment or position held:** PhD Student
3. **Qualifications:** Bachelor of Arts Double Honours in History and Geography, Masters in Geographical Analysis.
4. **Department:** National Centre for Geocomputation (NCG)
5. **Contact details:** tommy.burke@nuim.ie, 086 3166815.

6. **If the applicant is a postgraduate student:**
   Name of Supervisor: Dr. Urška Demšar
   Position held: Lecturer
   Department: National Centre for Geocomputation
   Contact details: urska.demsar@nuim.ie

A detailed letter from the supervisor must be included outlining how the applicant is suitably qualified/prepared for the type of work proposed.

7. **Will the research be carried out with:**
   i. Human Participant(s)
   ii. Will the human participant(s) be children or individuals with mental disabilities? No
   iii. If yes will the sessions be supervised by a guardian or a person responsible for the individual? N/A
   iv. If the sessions are to be unsupervised we are required to carry out Garda Vetting. Do you agree to NUI Maynooth carrying out this procedure? Yes
   v. Human derived material
   vi. Please describe the source of material

8. Why does this project require ethical approval?
Section 3.1 of ethical policy dictates that any interaction with living individuals requires ethical approval. Data will be collected from participants through a computer based experiment and short interview. The University Ethics Committee must approve of this work before it can commence.

9. Brief title of research project:
An evaluation of visualisations of Geographically Weighted Spatial Statistical Methods.

10. Please submit a copy of the research application which has been offered funding (if applicable)
N/A

11. Describe the purpose of the research (c. 150 words)
The purpose of this research is to discover the most appropriate visualisations which facilitate interpretation and analysis of Geographically Weighted Spatial Statistical Methods. Geographically Weighted Spatial Statistical Methods are employed in a wide range of disciplines to analyse and interpret data where they are used to detect significant patterns or relationships. One of these methods, Geographically Weighted Regression (GWR), is used to examine processes that vary over space and time. There is little variation in the types of visualisations which are used to analyse the results of GWR. 2D Univariate maps, statistical summary tables and graphs of residual values are primarily used. Consequently, it is unclear whether other visualisation methods could be more effective for displaying the results. The research we are conducting focuses on evaluating different visualisation techniques for GWR. This will be achieved through user trials with different visualisation techniques to ascertain their effectiveness for a given set of tasks. The initial goal is to discover the most appropriate way to facilitate interpretation and analysis of Geographically Weighted Regression. These results will be expanded upon for use with other methods such as Geographically Weighted Discriminant Analysis and Spatio-Temporal Geographically Weighted Regression.

12. Describe the methods and procedures to be used, showing how they adhere to the NUI MAYNOOTH Ethical Policy (*See Section 3*), under the following headings

a. Techniques to be used.
An experiment involving up to 30 individuals will be carried out. In accordance with NUI Maynooth Ethical policy each research participant will be asked to give informed consent before the experiment begins. An information sheet with details of the aims, objectives and general outline of the experiment will be provided. If the participant has any questions they will be answered to their satisfaction. They also have the option to opt out at any point.
Any data collected from the experiments will be managed according to the NUI Maynooth Ethical Policy on data storage. All participant information will be

treated as confidential. Each participant will be given an alias to ensure they are not identifiable, and the data will be stored in a secure environment.

b. Personal questions, interview schedules and questionnaires (for interviews please give a list of potential/possible questions that will be asked):
The first phase of the experiment will be computer based. A general participant profile will be established through several questions including:
- Persons Position (e.g. student).
- Persons knowledge of Geographically Weighted Regression.
- Persons knowledge of visualisation software.

During the experiment the participant will be asked to complete a series of tasks using visualisation. These visualisations will show the spatial distribution of voting data. An example of these questions are as follows:
- Identify the area where the influence of Owner Occupied housing has a significant effect on voter turnout levels.
- Locate the areas where Third Level Education had a positive influence on voter turnout levels.
- What is the relationship between the population aged 65 and over and male population?

A short semi-structured interview will be carried out after the computer based experiment. Questions will be asked to gain a sense of the participants perception of the different visualisation types presented. A sense of user satisfaction and ease of completion will also be obtained from these interviews. Sample interview questions are as follows:
- How difficult was it for you to complete the tasks?
- How happy were you with the time it took to complete each task?
- Do you think one visualisation was better to use than another?

c. Duration and frequency of sessions:
Each participant will complete one experiment. The sessions will last no longer than 2 hours to avoid participant fatigue. The computer based phase will require 90 minutes of this time.

13. Describe any discomfort or inconvenience to which participants may be subjected (if applicable), for example:

a. Procedures that for some people may be physically stressful or might impinge on the safety of the participants e.g. noise levels
The participants will be monitored by Eye-gaze tracking software through a webcam and experiment progress will be recorded via another ‘observer’ computer by tracking mouse movement, keyboard interactions, face recording and voice recording. All of this will be explained to the participant before they begin and they should be relatively relaxed. Stress should not be encountered.

b. Procedures that for some people could be psychologically stressful, e.g. tasks with high failure rate
The user will be aware that they are being timed. The participants however, will

be clearly informed in instructions that the test is not exclusively based on the time taken to complete the tasks. It is the visualisations being evaluated and not them. This should relax them and prevent any psychological stress.

14. Participants.

a. Who will be participants be?

The participants will be people with a knowledge of Geographically Weighted Regression, and an adequate working knowledge of the visualisation software which will be used in the experiment. Participants will mainly consist of PhD students, MSc students, and interested individuals from NCG that have completed a GWR workshop.

b. How will they be recruited?

They will be recruited on the basis of their knowledge as outlined above since it is a requirement. In most cases I will verbally invite them to participate in the experiment. In the event I can not conveniently invite them in person, a telephone call will be place, or an email shall be sent.

c. Will the participants be paid, if so how much?

The participants will not be paid on the basis that it may skew the experiment results. Participants may change their process on thought on certain aspects of the experiment if their time was compensated financially.

15. What will the participants be told about the study?

The participants will be presented with a A4 sheet detailing the aims and objectives of the experiment. They will be informed that the goal is to discover the visualisations which best facilitate the interpretation and analysis of GWR. They will also be informed that it is the visualisations that are being evaluated, not them. Details of how the experiment will benefit their own use of GWR will be provided. Since the most appropriate visualisations will be identified, the participants can utilize these in their own work.

16. What information, if any, will be withheld about the research procedure or the purposes of the investigation?

No information will be withheld about the procedure or the purposes of the investigation.

17. Consent:

a. When will consent be obtained? Prior to or at the time of the investigation.

Consent from participants will be obtained before the start of the experiment.

b. Will consent be verbal or written? (If not written, please justify)

The consent will be written. I will provide a consent form for each participant to sign.

c. Will consent be personal or third party on behalf of the participant?
The consent will be directly from the participant.

d. Will personally identifiable information be made available beyond the research team? If so, to whom and how will consent be obtained.
Participants information will not be identifiable. Alias names/numbers will be used.

e. Please include a consent form (See following pages for the required form)
The consent form is attached.

18. When the research is completed, outline how the participants will be debriefed as well as ways of alleviating and/or dealing with any distress or other problems that may arise.
The participant will be thanked and they will be informed of what I hope to achieve by carrying out the experiments. I would stress that the tests are no reflection on the person in any way and that the information gathered in the experiment will remain confidential. I will offer to provide keep the participant informed of results of the experiment.

19. If researchers are proposing to refer to a professional code of ethics governing research in their area, this must be specified and the appropriate part of the code appended to this application.
N/A

Note: Based on the Council of the School of the Biological Sciences, Cambridge Human Biology Research Ethics Committee, application for ethical approval of a research project form.

b. Experiment Profiler

Background Profile Participant Number: ☐

Gender Male ☐ Female ☐

English Knowledge poor ☐ basic ☐ good ☐ excellent ☐ native ☐

Native language (if not English)____

Describe your academic background (check one or more fields):

☐ Mathematics and Computer Science
☐ Geocomputation
☐ GIS and Remote Sensing
☐ Geography
☐ Other (Please Specifiy)____

What is your current job/position: ____

How long have you approximately been using Geographical Weighted Regression (GWR) in your work (please check just one box)

☐ <1 Year
☐ 1-3 Years
☐ 3-10 Years
☐ >10 Years

Describe your knowledge of GWR (Please check one box)

☐ Poor
☐ Basic
☐ Good
☐ Excellent
☐ Expert

Which types of visualisation software do you have experience with?
(Please check one or more boxes)

☐ ArcMap
☐ ArcScene
☐ Interactive Visualisation Software
☐ Other (Please Specify)____
☐ None of the above

c. Experiment Introduction Sheet

Title of Research:

An Evaluation of Visualisations of Geographically Weighted Regression.

Aim:

To evaluate the visualisations of Geographically Weighted Regression (GWR), and to discover those which best facilitate the interpretation and analysis of Geographically Weighted Regression results.

Experiment schedule:

The experiment will be split up into three parts, A, B and C.

Part A will have you working with a small dataset. **Part B** will have involve work with a large dataset. Both A and B contain three sections, detailed below:

1 – Working with Static 2D visualisation software - ArcMap. This software is capable of producing static 2D maps.

2 – Working with 3D visualisation software – ArcScene. This software is capable of producing 3D models.

3 – Working with Interactive visualisation software –Processor. This software is capable of producing an interactive 2D map, a parallel coordinate plot, and scatterplots.

In **Part C** you will participate in a short interview. Questions on your overall experience with the different visualisations will be covered here.

Experiment length:

Part A will be approximately 50 minutes, as will Part B. Part C will last for approximately 20 minutes. In total, the experiment will take approximately two hours.

Experiment Hardware and Software.

Morae is a usability testing software package which can collect data, log and observe important moments, and analyse and visualise the results.

Mousetracker is designed to follow your mouse movements to help with analysis.

The **HDcamcorder** will provide a high quality recording of the participant's facial expressions to help with analysis.

A **digital voice recorder** will record the participants' voice in Part C.

If you have any questions regarding the dataset, the software, the tasks or the instructions feel free to ask the observer. If you think you are unable to continue the experiment for any reason at any time you are free do to so.

d. Experiment Consent Form

Title of the Study: An Evaluation of Visualisations of Geographically Weighted Regression.

Researcher details: Tommy Burke
National Centre for Geocomputation
Íontas Building, North Campus,
NUI Maynooth,
Maynooth, Co.Kildare.

tommy.burke@nuim.ie
01 6286731

Supervisor details: Dr. Urška Demšar
National Centre for Geocomputation
Íontas Building, North Campus,
NUI Maynooth,
Maynooth,
Co.Kildare.

urska.demsar@nuim.ie
01 6286178

* * *

The purpose of this research is to discover the most appropriate visualisations which facilitate interpretation and analysis of Geographically Weighted Spatial Statistical Methods. Geographically Weighted Spatial Statistical Methods are employed in a wide range of disciplines to analyse and interpret data where they are used to detect significant patterns or relationships. One of these methods, Geographically Weighted Regression (GWR), is used to examine processes that vary over space and time. There is little variation in the types of visualisations which are used to analyse the results of GWR. 2D Univariate maps, statistical summary tables and graphs of residual values are primarily used. Consequently, it is unclear whether other visualisation methods could be more effective for displaying the results. The research we are conducting focuses on evaluating different visualisation techniques for GWR. This will be achieved through user trials with different visualisation techniques to ascertain their effectiveness for a given set of tasks. The initial goal is to discover the most appropriate way to facilitate interpretation and analysis of Geographically Weighted Regression. These results will be expanded upon for use with other methods such as Geographically Weighted Discriminant Analysis and Spatio-Temporal Geographically Weighted Regression.

A camcorder will be used to record the participants facial expressions, and a Dictaphone will be used to record their speech. Any data obtained during the course of the experiment will be stored in a secure environment, i.e. it will be kept in a locked cabinet at work. The data is available to the subjects at their discretion. Any digital voice records, video records, transcripts or other recorded data can be accessed at any time. The identity of each participant will be anonymous.

The results will be analysed using special analysis software to help answer my research question. They will form part of my PhD thesis.

The computer based experiment or subsequent interview do not pose any risk to your person. Your participation is voluntary, and should you wish to discontinue the experiment at any point you may do so.

If during your participation in this study you feel the information and guidelines that you were given have been neglected or disregarded in any way, or if you are unhappy about the process please contact the Secretary of the National University of Ireland Maynooth Ethics Committee at research.ethics@nuim.ie

Please be assured that you concerns will be dealt with in a sensitive manner.

"I have read and understood this consent form, and agree to participate"

Signed:

____

Date:

____

e. Experiment Script

The Dataset:

The dataset for this experiment is a combination of voter level turnout and census information*. We will use two different data set sizes. The large dataset comprises of electoral division (ED) level data for Ireland. The small dataset contains data for the province of Leinster. The values in the dataset can be either negative or positive, remember this when you are using the visualisations to answer the tasks.

You will see the word relationship often in the in the task questions. This word is asking to you compare attributes to see if there are any patterns which stand out. For example, the attributes could be affecting voter turnout in a similar way.

Table 1: The Experiment Dataset

Abbreviated Parameter Name	Full Parameter Name	Parameter Details
Males	Total Male Population.	Proportion of the population whose gender is male.
Soc 1and2	Social Classes One and Two.	Proportion of the population that occupies social classes one and two.
ThirdLevel	Third Level Education.	Proportion of the population with third level qualification.
Over65	Populated Over 65.	Proportion of the population that is aged 65 and over.
Unemp	Unemployed Population	Proportion of the population that is unemployed.

The level of voter turnout is the dependent variable, and the significant attributes are the independent variables.

Geographically Weighted Regression (GWR) Refresh.

GWR works on the Nearest Neighbour principle which weighs a point in the dataset against other data points which are nearest to it (the neighbours) and assigns it a value based on this weighting.

GWR has a number of outputs including Parameter Estimates, T-Values, Standard error values (S-Values) and Local R-Squared Values.

Parameter estimate values explain the influence an attribute has on the dependent variable. In the case of the dataset you will be using the dependent variable is of course, voter turnout.
T-Values help to highlight areas where their parameter estimate values are significant, particularly if the values are more than 2, or less than -2.
S-Values are a measure of the accuracy of predictions of the parameter estimates. The closer these values are to zero, the better the estimate.
Local R-Squared values measure the variation of the dependent variable which is explained by the independent variables.

In the centre at the top of the screen you will notice a start and 'Exit Session' button panel. This is the experiment 'progress panel'. Please, do not press this button as it will stop the test. This 'progress panel' is designed to assist in the analysis of the experiment results after it is complete.

If any of the visualisation software crashes inform the observer to get them working again.

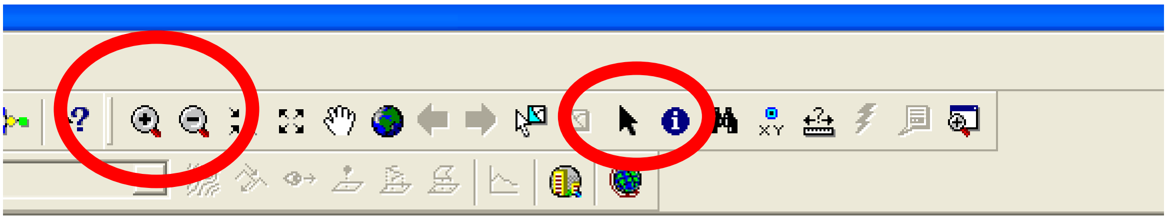
PART A

First you will use ArcMap, a GIS software created by the Economic and Social Research Institute (ESRI). It is the principal component of ESRI's ArcGIS suite of geospatial processing programs. It allows the user to explore data within a data set and produce maps.

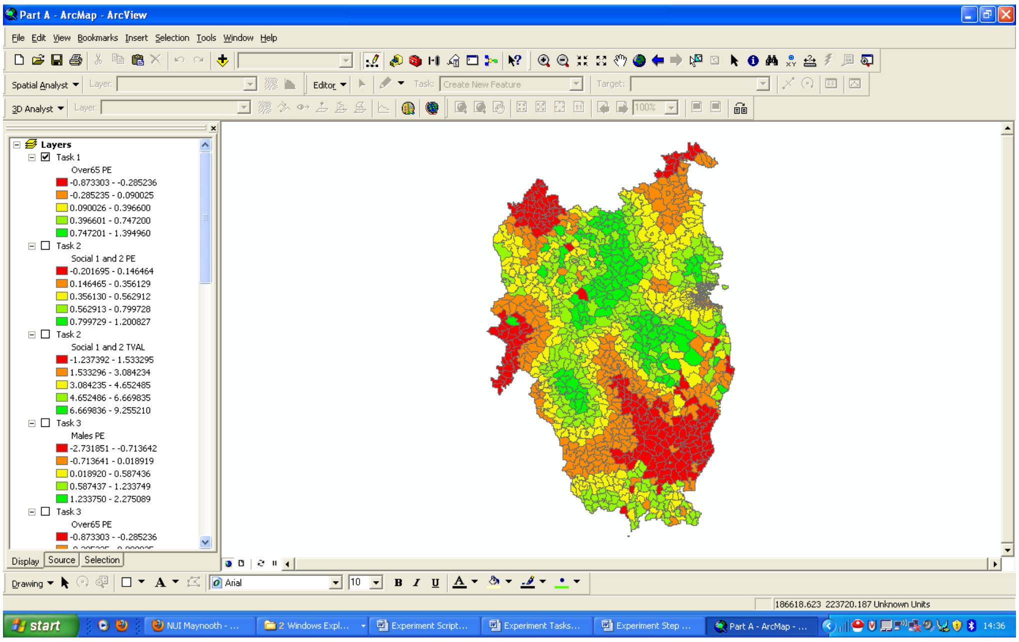
Interacting with ArcMap:

On the screen in front of you, you can see ArcMap open with a Map. The map information is contained in the Table of Contents (TOC) to the left of the screen. You may display any of these maps by ticking their blank box at any time. If you want to examine a map you must un-tick all the maps above it. Try to limit the number of maps you have ticked at once to five or six as the speed of the computer is affected. Please do not delete any map from the table of contents.

As you can see in TOC the maps are coloured and ranked from most negative (red) to most positive (green). Each map visualisation displays one attribute/parameter. You can use the zoom in and zoom out buttons at the top of your screen in the menu bars. You can also use the 'identification button' to discover the county names. Select the 'black arrow' button to return to normal browsing. These buttons are highlighted in the screenshot below:



The screen should now look like this:



If you have any questions please ask them now to avoid any disruptions during the course of the experiment.

Tasks

Select the ‘Start’ button from the progress panel when you are ready, next hit the ‘Start Task’ button to begin answering the tasks. When you have answered a task, hit the ‘End Task’ button. Your mouse cursor will begin to flicker but this is normal.

Task 1:

Rank in order for “Male” parameter estimates, the five counties which have the most positive effect on voter turnout, and the five counties which have the most negative effect on voter turnout.

A:

____

____

____
____
____

Task 2:

What is the relationship between the parameter estimates for the population of Social 1+2 and the parameter estimates for Third Level on voter turnout levels?

A:

____
____
____
____
____

Task 3:

What is the relationship between Over 65's parameter estimates, Over 65's T-Values, and Over 65's S-Values?

A:

____
____
____
____
____

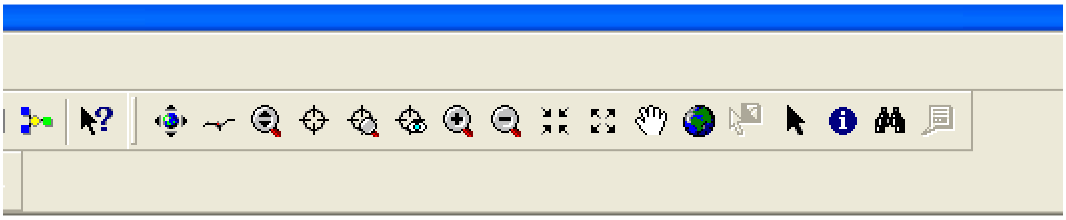
When you have completed task 3 you can minimise the ArcMap program

---

The second set of tasks will be completed using visualisations produced by ArcScene. ArcScene is a 3D visualisation application which allows you to view GIS data in three dimensional visualisations. It is fully integrated with the geo-processing environment, providing access to many different analytical functions and tools.

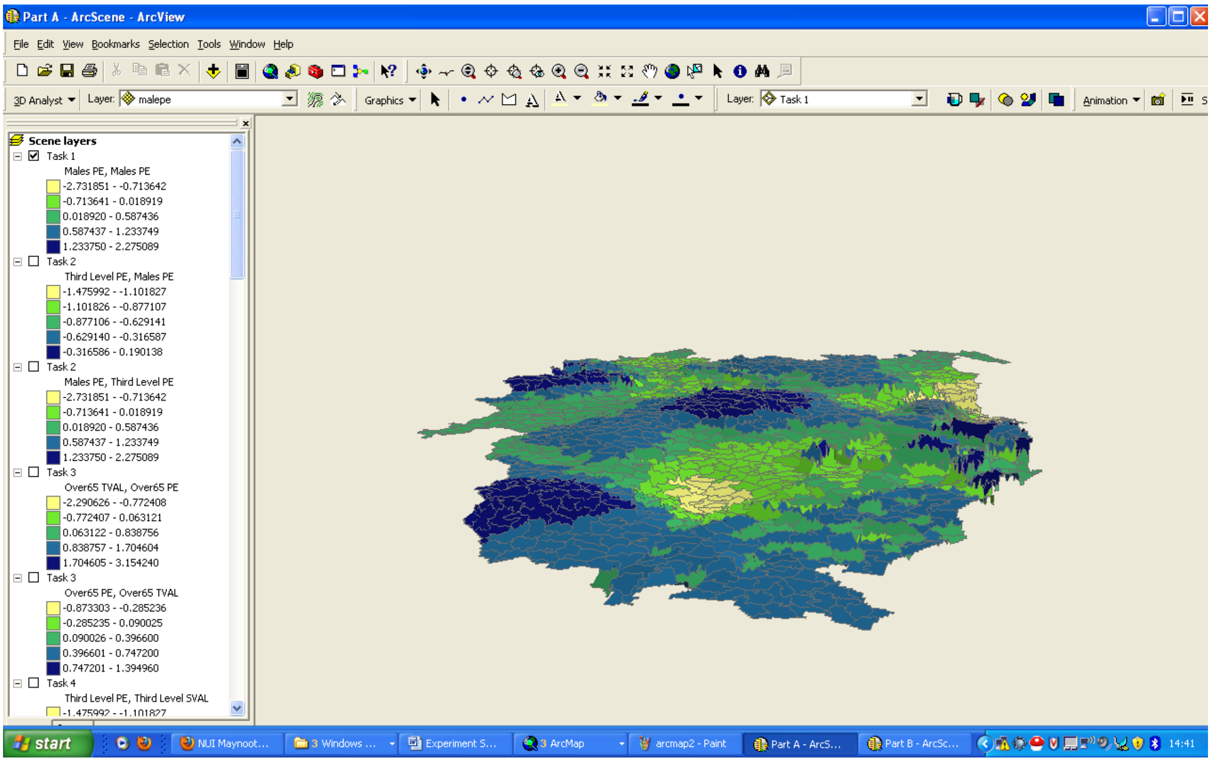
<u>Interacting with ArcScene:</u>

Maximise the 'ArcScene A' from the tabs toolbar at the bottom of the screen. You will notice there is a list of visualisations to the left of your screen in the table of contents. You can zoom in and out using the magnifying buttons with the ' + ' and ' - ' signs. You can also rotate the 3D visualisations by holding down with the left mouse button on the image in the viewing window and moving the mouse, alternatively you can use the 'rotate' icon. You can use the 'identification button' to find out the county names, and 'black arrow' button to return to normal browsing of the 3D visualisation. These buttons are highlighted in the screenshot below:



You can adjust the height of the visualisation by right clicking the 'scene layer' button at the top of the Table of Contents', and entering a larger or smaller number in the 'vertical exaggeration' box. At the top of each 3D visualisation you will see a title describing what it displays. The first parameter name provide the colour, the second parameter name provides the height. These are coloured and ranked from most negative (yellow) to most positive (blue). To help you answer each task you should tick the appropriate county boundary file. Please do not delete any visualisations.

The screen should now look something like this:



Tasks

Don't forget to select the 'Start Task' button when you are ready to answer a question, and 'End Task' when you have completed a question.

Task 1:

Locate the counties where the parameter estimates for the Unemployment population have a negative effect on voter turnout levels.

A:

____
____
____
____
____

Task 2:

What is the relationship between the parameter estimates and T-Values Males?

A:

____
____
____
____
____

Task 3:

For this task, please correct the height exaggeration so that your will become 3D. To do this, right click 'Scene Layers' at the top of the TOC. Select 'Properties', Under 'General' in the height exaggeration box, type in 50000. This will provide your 3D model with an appropriate height to help you answer the final ArcScene question.

In which areas does the model not fit well (have the lowest local R-squared values)? What are the influences of all parameter estimates in these areas?

A:

____
____
____
____
____

When you have completed task 3 you can minimise the ArcScene program.

Interacting with Processor:**

Processor is a program used to develop a set of interactive visualisations. Several different visualisation types are displayed at once, and they are integrated. This means if you highlight something or hover over an attribute value in any of the visualisations, the same attribute will be automatically highlighted in all of the other visualisations. You can deselect any highlights you have made by clicking anywhere on the screen that is not taken by one of the visualisations. This highlighting technique is known as interactive brushing.

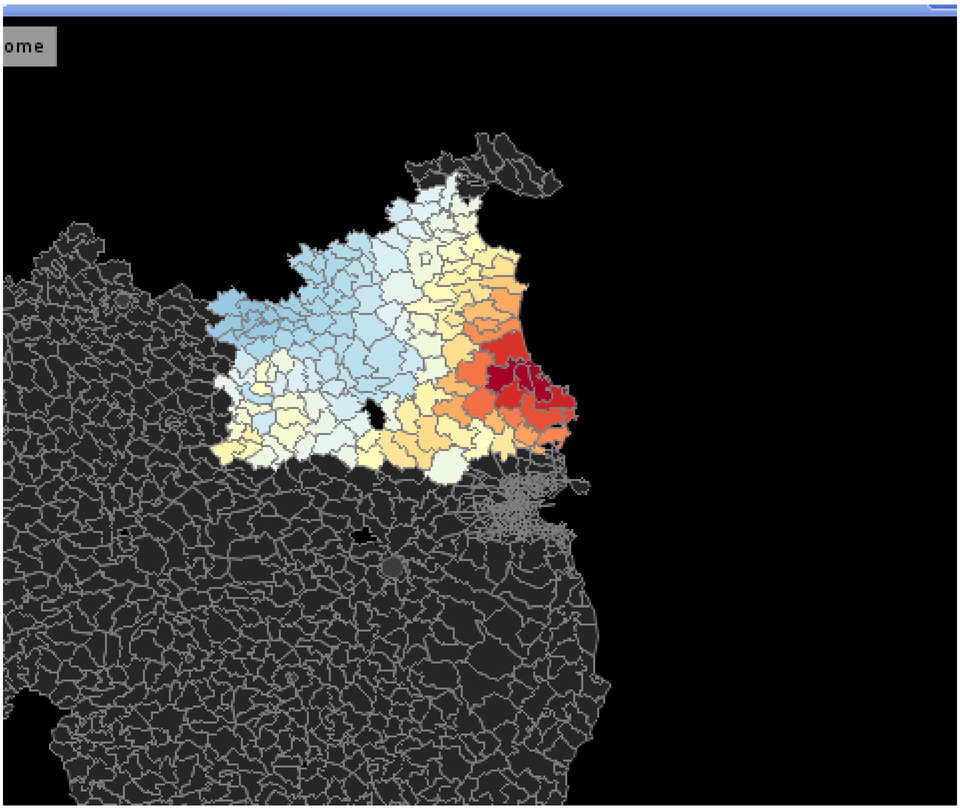
The visualisations on the right are scatterplots. They are useful for displaying groups of data and for examining the relationship between one variable and another. Each axis is related to one attribute with each dot representing a point of data.

At the bottom you will see a Parallel Co-ordinates Plot. Each attribute of a dataset are mapped onto a series of vertical axes. Each data object is represented by a continuous linear line which appears to connect the axes together. These lines intersect the vertical axes at the points which correspond to the value for that attribute.

Let us try this highlighting and brushing technique now. First, click on the colour of the first legend titled 'Males'. Next, click on the interactive map window to make it active. Now, browse over the Leinster region. You will notice the name of the ED and the county will appear if you hover your mouse over an ED. Now click on the PCP window to make it active, hover your mouse over the values in the visualisation. You can also zoom in on a particular area by holding the shift button, and using the scroll button on your mouse. To return to the original viewing distance, just click the home button in the visualisation window. Remember to click on the visualisation window to make it active. If you want to change which attribute is displayed, click on the 'colour' button in the colour legend of that attribute.

Lastly, let us use the select attributes feature. This feature allows you to highlight a certain number of attribute values, and examine them on their own. Click on the map, and hold and drag your mouse over a county sized area (just as you would if you would highlighting text in a word document for example). Release the mouse button when you are happy with the area you have selected. Any attribute values within this box will now be visible with all others blanked out.

Here is a screenshot showing an example of an area that has been highlighted:



Do the same with the PCP, click on the window to activate it, and highlight a number of values. To deselect the values and return to normal viewing, perform the click hold and drag operation in any area that does not have attributes. Please do not close any of the visualisations.

<u>**Tasks**</u>

Task 1:

Identify the counties where parameter estimates for Third Level Education have the least positive effect on voter turnout levels.

A:

____
____
____
____
____

Task 2:

Identify two parameter estimates from the five you have available that exhibit the least similar behaviour. Which two have you selected? Why did you select these two?

A:

____
____
____
____
____

Task 3:

Identify the areas of least stability, where the behaviour of all parameter estimates is least similar. Do such areas exist?

A:

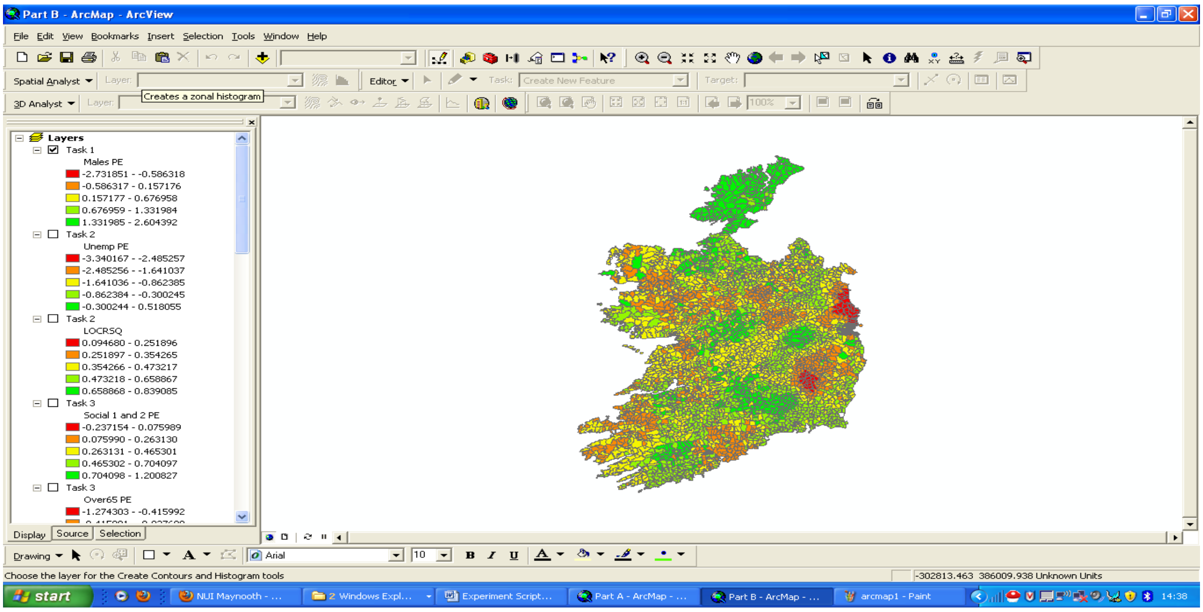
____
____
____
____
____

When you have completed task 3, you can minimise the processor visualisations using the box in the centre of the screen.

You are now half way through the computer based part of this experiment!

PART B

First, maximise the ArcMap window named 'Part B' which can be found in the toolbar list at the bottom of your screen. The ArcMap window should pop up and cover the whole screen, make sure it is maximised. Your screen should now look like this:



You should already have an understanding of how to interact with the different visualisations to complete the tasks required so let us continue onto the tasks.

Tasks

Task 1:

Rank in order for "Unemployment" parameter estimates, the counties which have the most negative to most positive effect on voter turnout?

A:

____

____

____

____

____

Task 2:

What is the relationship between parameter estimates for Third Level Education Levels and parameter estimates for Males on turnout levels?

A:

____

____

____

____

____

Task 3:

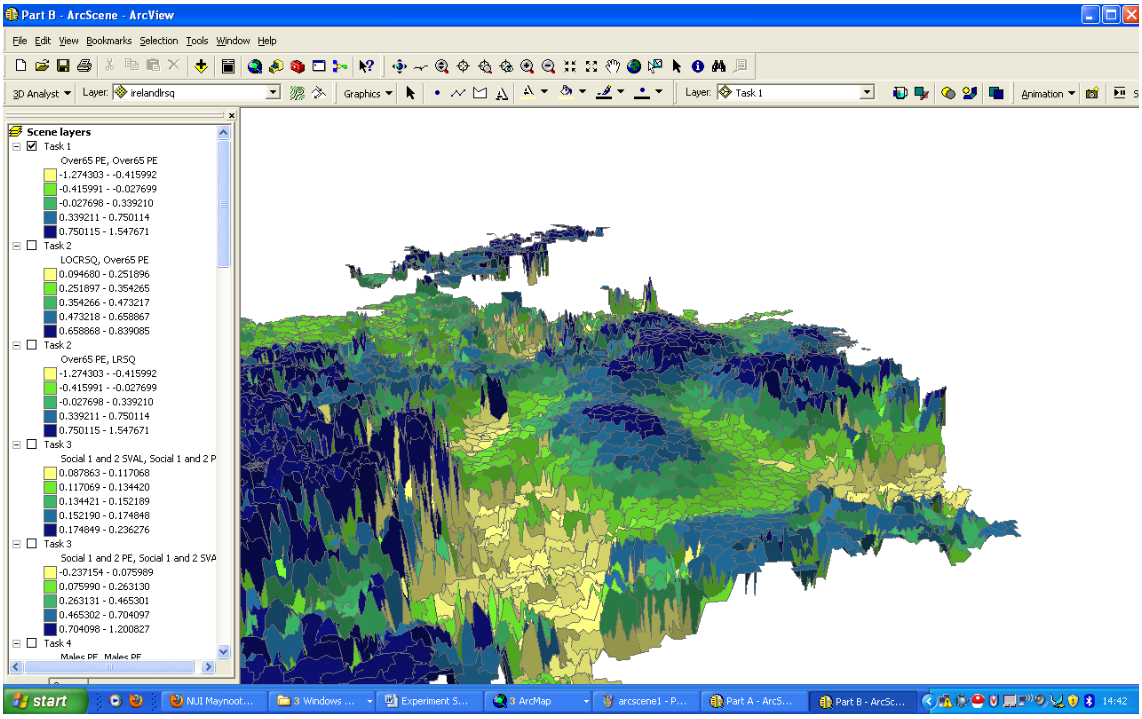
Rank the parameter estimates in order for which they exhibit the most negative to least negative influence on voter turnout levels. The order should be from 1 to 5, with 5 being the least negative.

A:

____
____
____
____
____

When you have completed task 3, answer the System Usability Survey, and then you can minimise the ArcMap program.

Maximise the ArcScene program called 'Part B' by selecting it from the programs in the toolbar list at the bottom of the screen. The screen should now look like this:



You already have an understanding of how to interact with ArcScene, so let us begin answering the tasks.

<u>Tasks</u>

Task 1:

Identify the counties where parameter estimates for population over the age of 65 have a positive influence on voter turnout levels.

A:

____
____
____
____
____

Task 2

What is the relationship between Social Classes 1 and 2 S-Value and Social Classes 1 and 2? In the counties where the S-Values are most stable, how do parameter estimates for Social Classes 1 and 2 affect turnout levels?

A:

____
____
____
____
____

Task 3

For this task, please correct the height exaggeration so that your will become 3D. To do this, right click 'Scene Layers' at the top of the TOC. Select 'Properties', Under 'General' in the height exaggeration box, type in 50000. This will provide your 3D model with an appropriate height to help you answer the final ArcScene question.

In which areas does the model best fit best (have the highest local R-Squared values)? What are the influences of all parameter estimates in these areas?

A:

____
____
____
____
____

When you have completed task 3, answer the System Usability Survey, and then you can minimise the ArcScene program.

The final visualisation program we will again work with is Processor. Double click the program called "Ireland" in the 'deploy' folder which can be found in the tab menu at the bottom of your screen.

From Part A you should already have an understanding of how to interact with the interactive visualisations so let us begin answering the set of tasks.

<u>Tasks</u>

Task 1:

Identify the counties where parameter estimates for Social Classes 1 and 2 have the most positive effect on voter turnout levels.

A:

____
____
____
____
____

Task 2

What is the relationship between Local R-Squared values and the areas where parameter estimates for Males positively affect turnout levels? How do parameter estimates for Males affect turnout levels in areas where Local R-Squared values are highest?

A:

____
____
____
____
____

Task 3

Identify the areas of stability, where all parameter estimates behave the same. Do such areas exist?

A:

____
____
____
____
____

When you have completed task 3, answer the System Usability Survey, and then you can minimise the processor visualisations.

When you have completed all of the tasks in Part A and Part B you can select the 'Exit Session' button in the Morae 'Progress Panel' at the top of your screen. I know you were informed not to do this before but since you are finished the computer based aspect of the experiment you can do this now.

That's it for Part A and Part B! Part C is next. Please inform the observer that you are ready to proceed to this part of the experiment now.

**The voter turnout data is courtesy of Dr. Adrian Kavanagh from the Department of Geography, NUI Maynooth. The census information was retrieved from the Central Statistics Office (CSO) online database.*

*** The Processor visualisation software in this experiment was kindly designed by Peter Foley, National Centre for Geocomputation (NCG).*

Appendix 2: Experiment Two

a. Ethical Application form

Approval Code:

1.1 University of St Andrews

1.2 Teaching and Research ethics committee (UTREC)

Please Tick: (*click on the box then click 'Checked' for a cross to appear in the box)*

Undergraduate ☐ Postgraduate Research ☒ Postgraduate Taught ☐ Staff ☐

Lecturer/Course Controller on behalf of Taught module ☐ Module Code:

Researchers Name(s):	Tommy Burke		
Project Title:	An Evaluation of Perceptual Scalability of Visualisations		
School/Unit: (Please indicate)	Centre for GeoInformatics, Geography and GeoSciences,	**Supervisor:**	Dr. Urška Demšar
Emails	tb44@st-andrews.ac.uk urska.demsar@st-andrews.ac.uk	**Date Submitted**	5/10/2012

Rationale: Please detail the project in 'lay language'. *This summary will be reviewed by UTREC and may be published as part of the reporting procedures. DO NOT exceed 75 Words (for database reasons).* Elucidation, if required can be given in Q.28

We will use eye movement trajectories and other usability techniques to investigate how the perceptual behavior of individuals changes with increased data set sizes through choropleth map visualisations. Eye trajectories are recorded using non-invasive eye tracking technology. Other usability techniques include task times and mouse movement trajectories.

Ethical Considerations: Please detail the main ethical considerations raised by the project, concentrating on any issues raised specifically in the red sections, and addressing, where appropriate, the issue of whether basic ethical criteria has been met in all supporting documentation and if not why not. *This summary will be reviewed by UTREC and may be published as part of its reporting procedures. DO NOT exceed 75 words (for database reasons).* Elucidation, if required can be given in Q.28

All ethics concerns have been addressed. Participants will be provided with a consent form to sign, they will be aware that their participation is voluntary and that they may stop the experiment at any time. Participants will be briefed on all aspects of the experiment before it begins, and will be debriefed afterwards. Participant identities will be anonymous, and any data obtained during the experiment will be stored in a secure location.

1.4

1.5 APPLICATIONS MUST BE SUBMITTED TO THE RELEVANT SCHOOL ETHICS COMMITTEE

1.6 HTTPS://WWW.ST-ANDREWS.AC.UK/UTREC/SEC/SECMEMBERS/

Please DO NOT submit directly to UTREC.

1.7

- Please submit an electronic copy and one hard copy (with signatures) to the Secretary/Administrator. In the absence of a Secretary please submit to the SEC Convener.
- Applicants must be accompanied by the relevant supporting documents without which a full ethical assessment cannot be made.
- Please do not type out with the text boxes provided, note that the Text Boxes are fixed in size and will not allow any viewing beyond the word limit permitted.

If ethical approval has been obtained from the University of St Andrews for research so similar to this project that a new review process may not be required, please give details of the application and the date of its approval.
Approval Code:
Date Approved:
Project Title:
Researchers Name(s):

RESEARCH INFORMATION

1. Estimated Start Date:
5th November 2012

2. Estimated Duration of Project: Up to six months including research and analysis.

3. Is this research funded by any external sponsor or agency? **YES** ☒ **NO** ☐

If YES please give details:

For projects funded by ESRC please be aware of the Ethical and Legal Considerations found at http://www.esds.ac.uk/aandp/create/ethical.asp

ESRC Funded Studentships (postgraduate Students) please be aware of the requirements as outlined at (in particular in relation to Submission of data to the Economic and Social Data Service, ESDS

4. Does this research entail collaboration with researchers from other institutions and/or across other University Schools/Units? **YES** ☒ **NO** ☐

If YES state names and institutions of collaborators: Dr. Arzu Çöltekin, Department of Geography, Geographic Information Visualization & Analysis, University of Zurich.

5. If the research is collaborative has a framework been devised to ensure that all collaborators, including all University Staff, External Researchers, and Students, are given appropriate recognition in any outputs? **N/A** ☐ **YES** ☒ **NO** ☐

6. Where projects raise ethical considerations to do with roles in research, Intellectual property, publication strategies/authorship, responsibilities to funders, research with policy or other implications etc., have you taken appropriate steps to address these issues? **N/A** ☐ **YES** ☒ **NO** ☐

7. Location of Research Fieldwork to be conducted: Geographic Information Visualization & Analysis Unit, Department of Geography, University of Zurich, Zurich, Switzerland.

RESEARCH INFORMATION

8. Are you using only library, internet sources or unpublished data *(with appropriate licenses and permissions)* and so have no human involvement such as interviewing of people?

YES ☐ **NO** ☒

9. a. Who are the intended Participants (e.g. students aged 18-21) and how will your recruit them (e.g. advertisement)	Ages: 18-25 Advertisement: Through Emails, and Posters around the department, contact through potential participant modules.
b. Estimated duration of Participant Involvement.	30-40 minutes.

ETHICAL CHECKLIST

If you have answered YES to Q8 but the project has other Ethical Considerations
please go to Q.28. If there are no other Ethical Considerations please sign and submit.

10. Have you obtained permission to access the site of research?

N/A ☐ **YES** ☒ **NO** ☐

If YES please state agency/authority etc. & provide documentation.
If NO please indicate why in Q.28

Prof. Dr. Sara I. Fabrikant (Head of Unit). Dr. Arzu Çöltekin (Senior Research)

11. Will inducement i.e. other than expenses, be offered to participants?

:S ☐ **NO** ☒

12. Has ethical approval been sought and obtained from any external body e.g., REC(NHS)/LEA and or including other UK Universities? If YES, please attach a copy of the external application and approval.

N/A ☒ **YES** ☐ **NO** ☒

13. Will you tell participants that their participation is voluntary?

YES ☒ **NO** ☐

14. Will you describe the main project/experimental procedures to participants in advance so that they can make an informed decision about whether or not to participate?

YES ☒ **NO** ☐

15. Will you tell participants that they may withdraw from the research at any time and for any reason, without having to give an explanation?

YES ☒ **NO** ☐

16. Please answer either a. or b.
   a. Will you obtain written consent from participants?

YES ☒ **NO** ☐

   b. *(ONLY: Social Anthropology, Geography/Geoscience, International Relations & Biology)*

Will you obtain written consent from participants, in those cases where it is appropriate?

YES ☒ **NO** ☐

17. Please answer either a. or b.
   a. If the research is photographed or videoed or taped or observational, will you ask participants for their consent to being photographed, videoed, taped or observed?

N/A ☐ **YES** ☒ **NO** ☐

   b. *(Social Anthropology & Biology ONLY)*
   Will participants be free to reject the use of intrusive research methods such as audio-visual recorders and photography?

N/A ☒ **YES** ☐ **NO** ☐

18. Please answer either a. or b.
   a. Will you tell participants that their data will be treated with full confidentiality and that if published, it will not be identifiable as theirs?

YES ☒ **NO** ☐

   b. Will you tell participants their work /contribution will be credited unless they specifically request anonymity?

YES ☐ **NO** ☐

19. Will participants be clearly informed of how the data will be stored, who will have access to it, and when the data will be destroyed?

YES ☒ **NO** ☐

20. Will you give participants a brief explanation in writing of the study? i.e. a debrief

YES ☒ **NO** ☐

21. With questionnaires and/or interviews, will you give participants the option of omitting questions they do not want to answer?

N/A ☐ YES ☒ NO ☐

If you have answered NO to any question 12- 21, please give a brief explanation in the statement of
Ethical Considerations on Page 1 and expand in Q28 if necessary.
If you have answered YES, it must be clearly illustrated in
the relevant paperwork which must be attached i.e.
Participants Information Sheet, Consent Form, Debriefing
Form, Questionnaire, Letters etc……

WORKING WITH CHILDREN AND OR VULNERABLE PEOPLE

Do participants fall into any of the following special groups?

22. a. Children (under the age of 16 in Scotland or 18 in England/Wales) **YES** ☐ **NO** ☒

b. Vulnerable Adult, receiving care or welfare services **YES** ☐ **NO** ☒

c. People with learning or communicative difficulties **YES** ☐ **NO** ☒

d. Residents/Carers in a specific location, e.g. Care Home **YES** ☐ **NO** ☒

If you answer YES to Q.22 a. – d., you may be required to obtain Protection of Vulnerable Groups [PVG] *Disclosure* approval (for all research being conducted in Scotland and elsewhere). Except, England / Wales please obtain the England/Wales equivalent – Police Check. Please check with the relevant people in Student Support (student applications) or Human Resources (staff applications) for clarification.

Refer to the UTREC Working with Children and Vulnerable Groups webpage for further guidance

e. NHS Patients or Staff **YES** ☐ **NO** ☐

f. Institutionalised persons **YES** ☐ **NO** ☐

If you answer YES to Q 22.,e. or f., it is likely you will be required to obtain approval from the NHS. This should be sought prior to approval from the relevant SEC or

g. People in custody

YES ☐ NO ☐

h. People engaged in illegal activities, e.g., drug-taking

YES ☐ NO ☐

If YES to Q22. g. or h., you should ensure that the relevant Risk Assessment Checklist has

23. As an adult have you lived/worked outside the UK in the last 12 months?

YES ☐ NO ☐

If you have answered YES you may be required to provided, in addition to PVG approval, a police reference/check from the country of residence during this period. Please check with the relevant people in Student Support (student applications) or Human Resources (staff applications) for clarification. https://www.st-andrews.ac.uk/utrec/EthicalApplication/children/

ETHICAL RISK

This section is for ethical use only and does not replace the University official procedures on Risk and Safety measures. In addition to completing this section you must review the following https://www.st-andrews.ac.uk/utrec/EthicalApplication/riskassessment/ and http://www.st-andrews.ac.uk/staff/policy/Healthandsafety/Publications/Fieldwork/ and follow the relevant procedures.

24. Are any of the participants in a dependant relationship with the investigator e.g. lecturer/student? If YES, give explanation in Q.28.

YES ☐ NO ☒

25. Will your project involve deliberately misleading participants in any way?
If YES, give details in Q.28 and state why it is necessary and explain how debriefing will occur

YES ☐ NO ☒

26. Is there any significant risk to any paid or unpaid participant(s), field assistant(s), helper(s) or student(s), involved in the project, experiencing either physical or psychological distress or discomfort?
If Yes, give details in Q.28 and state what you will do if they should experience any problems e.g. who to contact for help.

YES ☐ **NO** ☒

27. Do you think the processes, including any results, of your research have the potential to cause any damage, harm or other problems for people in your study area? If YES, please explain in Q.28 and indicate how you will seek to obviate the effects.

YES ☐ **NO** ☒

There is an obligation on the Lead Researcher & Supervisor to bring to the attention of the School Ethics Committee (SEC) any issues with ethical implications not clearly covered by the above

ETHICAL STATEMENT

28. Write a clear but concise statement of the ethical considerations raised by the project and how you intend to deal with them. It may be that in order to do this you need to expand on the Ethical Considerations section on page 1. (continue on additional pages if necessary)

An experiment with up to 50 participants will be carried out. In accordance with the University of St Andrews (UTREC) policies, each participant will be asked to give informed consent before the experiment begins. An information sheet with details of aims, objectives and general outline of the experiment will be provided. If the participant has any questions they will be answered o their satisfaction. They also have the option to opt out at any point and they will be aware of this. Any data collected from the experiment will be managed according to UTREC policies, and will be stored in a secure location and encrypted.

The participants will be monitored by Eye-gaze tracking software through non-invasive Tobii eye tracking technology, and experiment progress will be recorded via another 'observer' computer. Mouse movement, keyboard interactions, and face and voice recording will also be carried out. All of this will be explained to the participant before they being and they should be relatively relaxed. Stress should not be encountered at any point. The user will be aware that they are being timed, but will be clearly informed through experiment instructions that the test is not exclusively based on the time taken to complete the tasks.

The eye tracking technology is non-invasive. The hardware is positioned below the computer participants sit in front of for the duration of the experiment. No information will be withheld about the procedure or the purposes of the investigation. No participant information will be identifiable, random identification numbers or alias names will be used. All participant information will be treated as confidential. Any link between a persons' video and voice records will be security protected and kept in a secure location at all times. It will be stored for up to six months so that it can be analysed fully. No aspect of the experiment puts any participant at risk at any time. Participants will be thanked at the end of the experiment, and I would offer to keep them informed of the results of the experiment when they have been analysed.

DOCUMENTATION CHECKLIST

Ethical Application Form	YES	☒	NO	☐
Participant Information Sheet	YES	☒	NO	☐
Consent Form	YES	☒	NO	☐
Debriefing Form	YES	☒	NO	☐
External Permissions	YES	☒	NO	☐
Letters to Parents / Children / Head Teachers etc…..	YES	☐	NO	☐
PVG (Scotland) or Police Check (England/Wales) Approval *(as necessary)*	YES	☐	NO	☐
Advertisement	YES	☐	NO	☐
Other (please list):				

DECLARATION

I am familiar with the UTREC Guidelines for Ethical Research http://www.st-andrews.ac.uk/utrec/guidelines/ and *BPS, *ESRC, *MRC and *ASA (*please delete the guidelines not appropriate to your discipline) Guidelines for Research practices, and have discussed them with other

STUDENTS ONLY
My Supervisor has seen and agreed all relevant paperwork linked to this project YES ☐ NO ☐

Print Name:	Tommy Burke
Signature	Tommy Burke
Date:	5-10-2012

SUPERVISOR(S)
The Supervisor must ensure they have read both the application and the guidelines, and also has approved the project and application, before signing below, with clear regard for the balance between risk and the value of the research to the School/Student. (Supervisors should provide this on a separate sheet or supply to the student to insert below) Please, if you wish, add comments in no more than 200 words:

Print Name:

Signature	
Date:	

STAFF RESEARCHER ONLY YES ☐ NO ☐

Print Name:	
Signature	
Date:	

SCHOOL ETHICS COMMITTEE OFFICIAL USE ONLY

STATEMENT OF ETHICAL APPROVAL

This project has been considered using agreed University Procedures and has been:

☐ Approved

☐ Not Approved pending:

☐ More Clarification Required

☐ New Submission Recommended

☐ Discussed with Supervisor

☐ Referred to UTREC

☐ Referred to Fieldwork Subcommittee

Convenor's Name	
Signature	
Date:	

Please use the space below and additional pages to attach any supporting documents i.e. Participant Information Sheets, Consent Forms, Debriefing Forms, Questionnaires, Letter to Parents etc.

We recommend you refer to the sample documents provided at

b. Briefing Script

<u>Experiment Information:</u>

<u>Title of the Study:</u>

An Evaluation of the Perceptual Scalability of Visualisations.

We conduct a usability experiment to acquire and analyse trajectories of eye movement data in two visual exploration settings that differ in the size of geographic data. We couple eye movement analysis in the context of usability evaluation with analysis of spatial trajectories. We investigate how the type of perceptual behaviour of analysis of individuals with regards to scalability is linked with particular patterns in eye movement trajectories.

The experiment is split into two parts, A (1 & 2) and B (1 & 2). In the first part, the units in the dataset remain the same size on screen, while the number of spatial units in the dataset gets bigger. In the second part the size of dataset units gets smaller on screen, while the number of spatial units in the data set increases. A script will be provided showing clear instructions on how to complete the experiment, and there will be a research observing the experiment that can answer any questions a participant may have at any time.

This duration of the experiment will be approximately 40 minutes including pre experiment and post experiment briefings. It will take place in the Eye Tracking Labs of the Geographic Information Visualisation & Analysis Unit. Participants will interact with a computer screen to complete a set of tasks using varied 2D choropleth map visualisations, and datasets containing varying numbers of spatial units.

Any data obtained during the course of the experiment will be security protected and stored in a secure environment initially at in the Geographic Information Visualisation & Analysis unit and at the Centre for GeoInformatics in a locked drawer. The data is available to the subjects at their discretion. Any digital voice records, video records, transcripts or other recorded data can be accessed at any time. The identity of each participant will be anonymous, and data will be stored for up to 12 months after the experiment to carry out analysis.

We have permission to carry out this research with the full approval of the School of Geography and GeoSciences at the University of St Andrews, and the Geographic Information Visualisation & Analysis unit, which is part of the Department of Geography in the University of Zurich.

The content of this experiment presents no risk what so ever to participants. If a participant wishes to withdraw from the experiment at any time and for any reason you may do so. If you would like to express any concerns or have any questions about this experiment please contact one of the following persons:

Tommy Burke – PhD Student
Centre for GeoInformatics,
Geography and GeoSciences,

University of St Andrews,
Scotland.
Email: tb44@st-andrews.ac.uk

Dr. Urška Demšar – PhD Supervisor
Centre for GeoInformatics,
Geography and GeoSciences,
University of St Andrews,
Scotland.
Email: urska.demsar@st-andrews.ac.uk

Dr. Arzu Çöltekin – Research Collaborator
Geographic Information Visualisation & Analysis,
Department of Geography,
University of Zurich,
Switzerland.
Email: arzu.coltekin@geo.uzh.ch

c. Participant Consent Form

Participant Consent Form

Aim: To discover the perceptual scalability of 2D visualisations.

Objectives:

1. To collect data of users task completion on increasingly complex datasets using a 2D visualisation.
2. To detect the threshold of users change in performance due to perceptual scalability.
3. To develop specific methods for pattern detection in eye movement trajectories to detect this scalability.
4. To detect the difference (if any) in performance between the two different display sizes of spatial units.

Any data obtained during the course of the experiment will only be accessible by the experiment researchers listed below. The data will be security protected and stored in a secure environment at all times in a locked drawer, both in the University of Zurich and the University of St Andrews. The data will be used to answer the research aim and objectives stated above, and may be used to answer further questions that may arise from conducting this experiment in future research. All participant information will be anonymous, and at no time will participants be identifiable. The data may be stored for up to six months for the purposes of research and analysis.

Experiment Reserachers:

Tommy Burke – PhD Student
Email: tb44@st-andrews.ac.uk

Dr. Urška Demšar – PhD Supervisor
Email: urska.demsar@st-andrews.ac.uk

Dr. Arzu Çöltekin – Research Collaborator
Email: arzu.coltekin@geo.uzh.ch

"I have read and understood this consent form, and agree to participate"

Signed: **Date:**

____

d. Participant Information Sheet

University of St Andrews

Participant Information Sheet

Project Title

An Evaluation of the Perceptual Scalability of a 2D Visualisation.

What is the study about?

You are invited to take part in a usability experiment to acquire and analyse trajectories of eye movement data in two visual exploration settings that differ in the size of geographic data. We couple eye movement analysis in the context of usability evaluation with analysis of spatial trajectories. We investigate how the type of perceptual behaviour of analysis of individuals with regards to scalability is linked with particular patterns in eye movement trajectories.

This study is being conducted as part of Tommy Burkes PhD thesis in the Centre for GeoInformatics from the School of Geography and GeoSciences at the University of St Andrews, and is in collaboration with Dr. Urška Demšar, Centre for GeoInformatics from the School of Geography and GeoSciences at the University of St Andrews and Dr. Arzu Çöltekin from the Department of Geography, Geographic Information Visualization & Analysis of the University of Zurich.

Do I have to take Part?

This information sheet has been written to help you decide if you would like to take part. It is up to you and you alone whether or not to take part. If you do decide to take part you will be free to withdraw at any time without providing a reason.

What would I be required to do?

You will be asked to complete an experiment which contains two parts (A and B). You will complete a series of simple tasks using the Geographical Information System (GIS) Software, ArcMap. The Arcmap visualisations will be varied 2D choropleth maps, and the datasets used to produce these maps will contain spatial units varying in both number and size as you move from task to task, and from part A to part B.

We anticipate the experiment will not take more than 40 minutes to complete, including pre and post experiment briefs.

Will my participation be Anonymous and Confidential?

The identity of each participant will be anonymous, and data will be stored for up to 12 months after the experiment to carry out analysis. Only the researcher(s) and supervisor(s) will have access to the data which will be kept strictly confidential. Your permission may be sought in the Participant Consent form for the data you provide, which will be anonymised, to be used for future scholarly purposes.

Storage and Destruction of Data Collected

The data we collect will be accessible by the researcher(s) and supervisor(s) involved in this study only Any data obtained during the course of the experiment will be security protected and stored in a secure environment initially at in the Geographic Information Visualisation & Analysis unit and then at the Centre for GeoInformatics on a computer system hard drive. All data will be stored in an anonymised format. Your data will be stored for at least 12 months after the experiment to carry out analysis before being destroyed.

What will happen to the results of the research study?

The results will be finalised by the end of 2013 and written up as part of my/our PhD Thesis. We will also seek to produce at least one paper from the results of the experiment, and it is hoped that it will provide a basis to carry out research in the future.

Are there any potential risks to taking part?

The content of this experiment presents no risk what so ever to participants. If a participant wishes to withdraw from the experiment at any time and for any reason you may do so

Questions

You will have the opportunity to ask any questions in relation to this project before giving completing a Consent Form.

Consent and Approval

This research proposal has been scrutinised and been granted Ethical Approval through the University of St Andrews ethical approval process, and is hosted in the University of Zurich with the approval of Prof. Dr. Sara I. Fabrikant, head of the Geographic Information Visualization & Analysis Unit.

What should I do if I have concerns about this study?

A full outline of the procedures governed by the University Teaching and Research Ethical Committee is available at http://www.st-andrews.ac.uk/utrec/Guidelines/complaints/

Contact Details

Researcher: Tommy Burke (email: tb44@st-andrews.ac.uk, Phone:+44 77 333 44 5 66)
Supervisor: Dr. Urška Demšar (email: urska.demsar@st-andrews.ac.uk)
Host Collaborator: Dr. Arzu Çöltekin (email: arzu.coltekin@geo.uzh.ch)

e. Debriefing sheet

Debriefing Sheet (Verbal Debriefing)

First I will thank the individual for their participation. I will ask them how they found the experiment. I will explain we hope to discover the differences in patterns of eye movements on different sizes of datasets, and on different variations of the same type of visualisations. I will mention how the researchers mentioned in the Information sheet and Consent form hope to discover the point at which a participants search pattern changes, and to discover if there is a pattern in the type of eye movement trajectories for the different data sizes and the different visualisations.

I will emphasise that any data obtained during the experiment will remain confidential, and that if they would like to be kept updated on the results of the research to let me know. They will be informed that they are free to contact any of the researchers mentioned on the consent and participant information sheet if they have any concerns or questions regarding the experiment that just took place. I will also let them know that they can withdraw their data at any time for any reason.

Reason for verbal debriefing: I want to verbally debrief the participants because I believe they will have already had enough to read before the experiment began, and verbal debriefings can help to relax participants after the experiment has been complete because of the informality introduced.