Decision-making under spatial uncertainty

Submitted for the degree of
Master of Applied Science
(Geographic Information Systems)

Susannah Jayne Hope

Department of Geomatics
The University of Melbourne

January, 2005

This is to certify that (i) the thesis comprises only my original work except where indicated in the preface, (ii) due acknowledgement has been made in the text to all other material used and (iii) the thesis is approximately 30 000 words in length, exclusive of tables, maps, bibliographies and appendices.

Summary

Errors are inherent to all spatial datasets and give rise to a level of uncertainty in the final product of a geographic information system (GIS). There is growing recognition that the uncertainty associated with spatial information should be represented to users in a comprehensive and unambiguous way. However, the effects on decision-making of such representations have not been thoroughly investigated. Studies from the psychological literature indicate decision-making biases when information is uncertain. This study explores the effects of representing spatial uncertainty, through an examination of how decision-making may be affected by the introduction of thematic uncertainty and an investigation of the effects of different representations of positional uncertainty on decision-making.

Two case studies are presented. The first of these considers the effects on decision-making of including thematic uncertainty information within the context of an airport siting decision task. An extremely significant tendency to select a zone for which the thematic classification is known to be of high certainty was observed. The reluctance to select a zone for which the thematic classification is of low certainty was strong enough to sometimes lead to decision-making that can only be described as irrational.

The second case study investigates how decision-making may be affected by different representations of positional uncertainty within the context of maritime navigation. The same uncertainty information was presented to participants using four different display methods. Significant differences in their decisions were observed. Strong preferences for certain display methods were also exhibited, with some representations being ranked significantly higher than others.

The findings from these preliminary studies demonstrate that the inclusion of uncertainty information does influence decision-making but does not necessarily lead to better decisions. A bias against information of low certainty was observed, sometimes leading to the making of irrational decisions. In addition, the form of uncertainty representation itself may affect decision-making. Further research into the effects on decision-making of representing spatial uncertainty is needed before it can be assumed that the inclusion of such information will lead to more informed decisions being made.

Acknowledgements

Firstly, I would like to thank Associate Professor Gary Hunter for his assistance, support and encouragement as supervisor of this work. I am extremely grateful for the time he so generously volunteered, the invaluable ideas and feedback offered and his all-round enthusiasm for the research.

I am grateful to Dr. Phil Collier and Roger Fraser for their time and patience in explaining the complexities of positioning maritime boundaries, and to Dr. Mark Burgman, Dr. Jane Elith and Brendan Wintle for their ideas relating to thematic uncertainty.

Assistance in the statistical analysis was provided by Dr. Ian Gordon of the Statistical Consultancy Centre at the University of Melbourne. I am indebted to him for the time and help offered, particularly with the REML analysis of unbalanced data.

I would also like to acknowledge the support offered by the Department of Geomatics at the University of Melbourne. The friendliness of staff and students has truly complemented this academic experience.

Contents

List of Figures

List of Tables

List of Abbreviations

ANSI	American National Standards Institute
ASPRS	American Society for Photogrammetry and Remote Sensing
DEM	Digital Elevation Model
DF	Degrees of Freedom
GIS	Geographic Information System
GPS	Global Positioning System
NATO	North Atlantic Treaty Organisation
NCGIA	National Center for Geographic Information and Analysis
NIST	National Institute of Standards and Technology
PCC	Percentage Correctly Classified
REML	Residual Maximum Likelihood
RMSE	Root Mean Square Error
SDTS	Spatial Data Transfer Standard
WADD	Weighted Additive Value Model

1.

Introduction

1.1 Problem Statement and Significance

With the growing use of Geographic Information Systems (GIS), the applications of spatial data have become more sophisticated and diverse. Decisions are increasingly being made on the basis of spatial information that is derived from GIS and the users of these systems may have little knowledge of the processes underlying the GIS output. Although the computer systems themselves may be relatively precise, the output of a GIS can only be as good as the data input. In recent years, it has been recognised as essential that both the quality of this input and its resultant effect on the output are fully understood if the decisions being made are to be informed and robust. However, to date, there has been little consideration of how the inclusion of uncertainty information and the mode of its representation may influence the decision-making process.

It has been estimated that spatial data is now used in over 80% of decisions made by government departments. Indeed, its use is not restricted to such clientele as policy-makers, resource managers and emergency services; spatial information is widely used in the private sector, for example in marketing, insurance and transportation, as well as by the general public in functions such as car navigation systems. With such widespread application, it is in the interest of the entire community that spatial information is used appropriately. Errors are inherent to any spatial data set and it is essential that the resultant uncertainty in the spatial information product be communicated to the users of this information.

The need for a better understanding of spatial uncertainty has been given greater emphasis in recent years, partly as a result of the increased tendency towards litigation. If a party is to be held responsible for the consequences of decisions being made on the basis of spatial data, there is more of a demand for those decisions to be informed and as correct as possible. Decision makers need to know the level of uncertainty within the GIS output that is informing their decisions, to decide if this information is sufficiently accurate to be applicable to the problem at hand. This is particularly important if non-specialised data sets are being used, rather than data specifically created for a nominated purpose. To enable this, data providers need to supply appropriately detailed metadata on the lineage and accuracy of their products.

Many countries have now introduced statutory obligations in the form of spatial data transfer standards to ensure that such metadata is provided. The pioneer amongst these is the Spatial Data Transfer Standard (SDTS) introduced by the United States (NIST, 1992). This mandates that metadata be reported for all datasets that are transferred using the standard. The quality component of this metadata is required to be documented along five themes: the lineage, spatial accuracy, attribute accuracy, logical consistency and completeness of the data set. The inclusion of metadata with spatial datasets has also been propelled by market forces within the private sector that dictate that, as one spatial data provider begins to include metadata as a matter of course, other providers will soon follow suit.

Although the provision of metadata with spatial data sets is now becoming widespread, the decision of whether the data are suitable for the task at hand is, in practice, still left to the user. Many of those users have had little if any exposure to the issue of data quality and, without adequate training, cannot be expected to recognise from metadata reports whether or not a data set is fit for use in the current problem. There has been recognition, particularly within the academic field, that the functionality of GIS needs to be increased to include ways of representing the uncertainty in GIS output resulting from the quality of the data input. These representations need to communicate the uncertainty in a manner that is unambiguous, fully informative and best able to facilitate decision-making.

There has also been substantial research into the means by which uncertainty representations may be included in GIS. Nonetheless, such functionality has been slow to appear in commercial systems, with software developers claiming that there is insufficient demand from the user community. As the importance of spatial uncertainty takes a higher profile, with decisions being made on the basis of spatial information being subject to greater scrutiny and with better education of spatial information users, this should change. Some GIS are already including means of representing the uncertainty in their output. However, since the nature of the representation may affect the manner in which decision-makers apply the uncertainty information, it is essential that further research be conducted into the effects of uncertainty representation on decision-making processes.

1.2 Sources of Error in Spatial Data

As with any data, there will be a degree of error inherent to all spatial data. The accepted definition of error in spatial data is the deviation of the database value from the true field value. Since the true field value is often not known, the observed value is usually compared to a best model value, which represents an accepted truth value. This lack of knowledge of the true value causes the error to become uncertain. Due to the negative semantic implications of the term error, it is often replaced with the concept of accuracy, which is understood to represent the degree of closeness of the database value to the accepted best model value. Thus, errors in spatial data sets are often reported as confidence limits of the accuracy of the data, for example '95% of the values are accurate to within 14m of the true values'.

There are many sources of error in spatial data (for a summary see Hunter *et al*, 2003). Hunter and Beard (1992) propose that these can be considered as arising from one of three stages: data collection and compilation, data processing or data usage. Collection and compilation errors are perhaps the least transparent, as many users are familiar with concepts such as the level of precision of measuring equipment, which will affect the accuracy of both attribute and positional measures. Other examples of how collection and compilation errors arise include the random fluctuations associated with measuring equipment and difficulties in measuring and modelling natural features. For example, the occurrence of random errors, such as fluctuations in platform stability of remote sensors, will affect the resulting images and their subsequent processing. Also, the tendency to represent natural environments as crisply defined polygons within maps, such as soil classification zones, introduces errors by not accurately conveying the fact that they typically display transition zones and a lack of homogeneity.

Processing errors can occur through the digitisation of existing data sources and through the manipulation of data within GIS. These errors are often hidden and inexperienced users of GIS may be unaware that errors are propagating through their systems. Digitising errors arise through generalisation processes, such as line simplification, and feature editing processes, for example node snapping. Manipulation of the data introduces errors as a result of methods such as raster/vector conversion and Boolean overlay within GIS, and through analysis techniques such as interpolation and buffer creation.

The third type of error as classified by Hunter and Beard is that of data usage. Examples of this type of error would be users misusing GIS output through a lack of training or understanding, inappropriate use of a dataset that had been created for another purpose, decision-making without sufficient metadata regarding the lineage of the data set, and use of data that had been collected at a scale unsuitable for the use to which it is being put. While the first two sources (collection and compilation and processing) are to some extent unavoidable, it is possible that this third type of error can be removed, or at least lessened, through an increased understanding of spatial uncertainty among all spatial data users.

1.3 Data Quality and Communicating Uncertainty

Data quality is defined here as the fitness-for-use of a data set for the application at hand. Users need to be able to decide if a data set is of sufficient quality for its application to the current problem to be appropriate. The responsibility for this has traditionally been left to the user, although many novices do not have the experience or training to be able to accomplish this from a metadata report. In addition to the inclusion of appropriately detailed metadata by spatial data providers, correct usage of spatial information requires effective representation of the subsequent uncertainty in GIS output as well as better education of all those using such data.

In recent years the issue of data quality has received a considerable amount of attention. Spatial data are becoming more widely available and are being used for purposes other than that for which they were originally collected. The need for metadata concerning the lineage and accuracy of the data set has been recognized and standards such as the SDTS are being established in many countries. Although many data providers are now including this metadata with their datasets, Hunter (2001) has provided several examples of poor metadata reportage along each of these themes. He has also emphasized the need for greater detail in metadata reporting, particularly in areas such as the level of autocorrelation of errors (Hunter & Goodchild, 1997) and data quality information at local levels as well as the global level (for example, Qiu & Hunter, 2002).

However, metadata are of limited use if the information they provide is restricted to the input data; users need to know what this means with respect to the spatial information output of the GIS upon which they are basing their decisions. In the early 1990s, the National Centre for

Geographic Information and Analysis (NCGIA) established a research initiative focusing on the issue of visualization of spatial data quality (Beard *et al*, 1991). In addition to metadata reporting, research challenges were identified in the areas of communication of data quality to users, error propagation through GIS and the application of data quality information to decision-making. In essence, the conclusion was that users of spatial information need to know if this information is sufficiently accurate for them to make the correct decisions. It is the output of GIS upon which decisions will be made; it is therefore imperative that the quality of this spatial information output is communicated to users in a manner that enables them to apply uncertainty in the data to their decisions in an appropriate manner.

Since the Research Initiative on Visualization of Spatial Data Quality, some progress has been made within these areas. There has been a considerable number of papers written on quantifying spatial data uncertainty and a substantial amount of research into methods of visualizing this uncertainty in GIS. Innovative methods that utilise the advantages of computer systems over paper maps have also been developed, such as interactive displays, dynamic representations and multiple simulations (Hunter, 1999). However, there has been little research conducted into the effectiveness of these methods as communicators of spatial uncertainty and, specifically, how these uncertainty representations may affect the decisions being made.

The few studies that have considered the effectiveness of uncertainty representations have generally been limited to an examination of the extra cognitive demands placed upon decision-makers. They have been concerned with the question of whether or not users can cope with the additional information that is being provided in a display that incorporates uncertainty information together with thematic information (for example Leitner & Buttenfield, 2000). Effective representations have been concluded as being those that do not detract from the accuracy or speed of simple decisions and which users rate as easiest to understand. Whilst this is important, these studies have been run in simple experimental settings that have been limited to binary representations of uncertainty. They have tested whether users can comprehend such representations, without considering how it is that the inclusion of uncertainty information may affect the decisions being made on the basis of spatial data.

Studies in the psychological literature on decision-making have indicated that introducing uncertainty may bias the decisions being made. Many people do not have an intuitive feeling for how to deal with uncertainty and show a tendency to avoid ambiguous situations. Academics within the spatial data sector have identified the need to represent spatial uncertainty, without consideration of how it is that decision-makers may respond to such information. The question of how it is that information about spatial uncertainty may influence the decision-making process requires a lateral step into the realms of psychology, and studies need to be conducted within the framework of behavioural decision research.

1.4 Behavioural Decision Research

The goals of behavioural decision research have been to describe and explain human judgement and choice behaviour so as to aid and improve decision-making performance. Human decision-making often deviates from the rational choice of maximising expected value, especially as decision complexity increases. From an experimental psychology perspective, it has been proposed that this is due to information processing limitations and that decision-makers apply simplified heuristics to lessen the cognitive effort demanded by a problem. Nonetheless, decision-makers often show preference reversals for the same decision under different task conditions. These and other incongruities have been explained through a tendency towards certain biases (Tversky & Kahnemann, 1991).

An understanding of the psychology of human decision-making is important to any assessment of the effectiveness of spatial uncertainty representations. In a groundbreaking experiment, Ellsberg (1961) demonstrated that humans respond differently to information when probabilities are ambiguous (uncertain) than when the same probabilities are represented as fixed, a bias that has strong relevance to studies of decision-making under conditions of spatial uncertainty. However, Ellsberg's study, like much of the early work on human decision-making, focussed on hypothetical gambles. Later studies have shown that similar biases are displayed when humans are making decisions in real life applications such as the purchase of life insurance (Hogarth and Kunreuther, 1985) and the selection of potential employees (Highhouse & Hause, 1995). If the same biases are demonstrated to uncertainty in spatial data, there may be significant consequences on the decisions being made on the basis of uncertainty information being represented. Since decision-making behaviour has also been shown to be task and context dependent within psychological studies, the nature of any spatial uncertainty representation may also affect the decision-making process.

1.5 Research Aim and Objectives

It follows that research is needed to develop an understanding of how the provision of uncertainty information and the mode of its representation may influence decision-making behaviour with spatial information. Within the GIS literature, there has been little testing of the effects on *decision-making* of representing uncertainty. Within the decision-making literature, there has been little consideration of *spatial* uncertainty. This thesis focuses on the ways in which the inclusion of spatial uncertainty and the mode of any such representation may affect the decisions being made within the context of behavioural decision research. Existing research into decision-making under conditions of spatial uncertainty is sparse and this study complements the current literature by introducing a psychological perspective into decision-making on the basis of uncertain spatial information.

Accordingly, the aims of this research are to:

- test the effects of introducing the uncertainty in spatial information to decision makers; and
- investigate the effects of different uncertainty representations on the decisions being made.

The objectives are to:

- examine the different types of spatial data error and uncertainty;
- explore the methods for representing uncertain spatial information;
- investigate how decision-makers deal with uncertainty;
- identify suitable case studies to form the basis for experimental testing; and
- conduct experimental tests and evaluate the results.

The hypotheses that are proposed and tested are that:

- decision-makers will exhibit ambiguity aversion when uncertainty information is included in thematic maps, which may lead to the making of irrational decisions; and
- different decisions will be made when the same positional uncertainty information is displayed, dependent upon the nature of the uncertainty representation.

1.6 Thesis Outline

This thesis begins with a review of the current literature and research relevant to the topic of decision-making under spatial uncertainty. Initially, the types of error in spatial data are outlined and how these errors contribute to uncertainty in GIS output is examined. This is followed by a review of the methods that have been used to quantify and depict the uncertainty in spatial information and the limited research into assessments of the effectiveness of these representations. A lateral step is then taken, into the realm of behavioural decision research, to explore the developments in this field towards an understanding of decision-making and, in particular, decision-making under conditions of uncertainty.

The domains of spatial information and human decision-making are then integrated within an experimental study of the effects of representing spatial uncertainty to decision makers. Two case studies are established. In one of these studies, the uncertainty in the spatial information is a result of thematic uncertainty (airport siting case study), whereas in the second study, the uncertainty relates to position (navigation case study). The airport siting case study examines the effects on spatial decision-making of providing thematic uncertainty information. It investigates how the introduction of this information may affect the relative rankings of different regions as the potential site of a new airport. The hypothesis being tested is that decision-makers will exhibit ambiguity aversion when uncertainty information is provided, which may lead to the making of irrational decisions. The navigation case study considers the effects of different representations of positional uncertainty upon decision-making. The decision as to when a boat should turn away from a restricted zone is examined under different depictions of positional uncertainty. The hypothesis being tested is that different decisions will be made to the same information, dependent upon the nature of the uncertainty representation.

1.7 Scope of Thesis

The title of this thesis, decision-making under spatial uncertainty, is extremely broad, suggesting perhaps more than a single study is able to offer. However, since there has been little research to date in this area, the initial findings from such a study may be general enough to be applicable to all types of spatial decision-making. There are many sources of uncertainty in spatial data; in addition to positional and thematic uncertainty, the logical consistency and completeness of spatial information is uncertain. Temporal uncertainty is also applicable across each of these elements and some academics have made a case for the existence of semantic uncertainty. However, this study is restricted only to consideration of positional and thematic uncertainty in spatial information. The temporal and semantic uncertainty, logical consistency and completeness of the case study data are not considered.

This thesis is concerned with the effects of representing uncertainty in spatial information to decision-makers. However, the diversity of applications of spatial data is huge and continuing to grow. The variety of decisions being made on the basis of spatial information is correspondingly large. A single experimental study can only consider a small number of applications, although if well designed, the results can be generalized across many fields. Accordingly, this study is based upon two such applications; it examines the effects of introducing thematic uncertainty information on siting decisions and explores how different representations of positional uncertainty may also influence decision-making. The findings have potential relevance to all applications of spatial information.

1.8 Chapter Summary

Spatial information is being used in GIS by a diverse group of users and in an increasingly widespread range of applications. Errors from a variety of sources and processes are inherent to spatial data, causing uncertainty in the spatial information output. There is growing recognition of the need for this uncertainty to be represented to users in a comprehensive and unambiguous way, to ensure that their decision-making is fully informed. However, the effects on decision-making of representing uncertain spatial information have not been thoroughly investigated. Studies from the psychological literature indicate decision-making biases when information is uncertain, although these biases have not been examined with spatial data. This study aims to explore the effects of representing spatial uncertainty, through an examination of how decision-making may be affected by the introduction of thematic uncertainty and an investigation of the effects of different representations of positional uncertainty on decision-making.

2.

Errors in Spatial Data and Methods of Visualising Spatial Uncertainty

2.1 Introduction

As Hunter and Beard (1992) note, the manual production of paper maps has traditionally been the domain of experts, specialized in the skills of cartography. These specialists have long been aware of the limitations of the data with which they work and have typically represented the accuracy of their final product using conventional map reliability diagrams and statements of positional accuracy. However, the ease with which spatial data can be used within GIS to produce digital maps has opened up the field of map production to a range of far less specialized users. Many of these users have only a limited awareness of how errors will be inherent to the data sets and propagated through the GIS processes. It is therefore necessary that the limitations of spatial data sets be communicated to their potential users.

In this chapter, the premise that errors are inherent to spatial data is assumed. The sources of these errors are reviewed, under the three-way classification framework of collection errors, application errors and misuse errors. Data sets have an associated level of uncertainty as a result of these errors and current requirements to report data quality along the five elements of spatial, attribute and temporal accuracy, logical consistency and completeness are discussed. The chapter continues with a review of methods aiming to quantify and visualise the uncertainty in spatial data. It concludes with a discussion of the research to date that has considered the effectiveness of visualisation methods to communicate uncertainty to users of spatial data.

2.2 Sources of Error in Spatial Data

Geographic information systems require spatial data input, on which a multitude of operations may be performed, in order to obtain the desired information output. Errors are inherent to all data sets, and these errors will be further propagated throughout any manipulation processes within the GIS. In addition to data acquisition errors and processing errors, Beard (1989) identifies a third source of error in spatial data, that of data usage. Hunter and Beard (1992) propose a model in which these three types of source error lead to two forms of error in the final product, both positional and thematic error. Additional forms of error, those of completeness, logical consistency and temporal error, have since been incorporated into this model (Hunter *et al*, 2003).

2.2.1 Collection and compilation errors

Errors at the stage of acquisition of spatial data are an inherent source of inaccuracy within a data set. Veregin (1989) presents a comprehensive overview of such errors, which include those arising from the technologies and techniques used to measure and record data. For example, remote sensing data is subject to errors arising from factors such as platform tremor and atmospheric variability. The accuracy of photogrammetry and surveying are dependent upon the techniques, geodetic base and equipment used. Different map projections introduce differing distortions in distance, shape or area.

Many natural phenomena do not exist as homogeneous regions, although we often attempt to classify areas such as vegetation or soil type into polygons, and the existence of transition zones between regions is often ignored. Definitions of natural features or class boundaries will affect the accuracy of the resulting data set, as will the method and completeness of sampling. Data currency is also an important factor as the data set attempts to represent a continually changing world.

2.2.2 Processing errors

Errors are further introduced into a data set as it is manipulated within a computer system. If not already in digital form, errors will occur in the digitization process. For example, manual line following will introduce under-shoots, over-shoots and spurious polygons. Feature editing techniques that attempt to rectify these, such as line snapping and the elimination of slivers through the use of tolerance levels, can themselves bring in unwanted results. Generalization methods, for instance line simplification and curve fitting, will also introduce errors.

The manipulation of digital data within the GIS is a hidden source of error, of which inexperienced users may be unaware. Co-ordinate adjustments, for example through rubber sheeting, edge matching or changes of datum will introduce errors, as will techniques such as raster-vector conversions. Errors will further propagate through surface modeling procedures, polygon overlay and spatial analysis techniques, for example network analysis. Spatial statistics and interpolation methods may not adequately account for autocorrelation within the data and the spatial variability of the data set. In addition, the method used to display results, for example the selection of class intervals in a choropleth map, can have a significant effect on the final product.

2.2.3 Application errors

The third source of error identified by Beard is that of data usage. She proposes that inexperience in cartographic principles or lack of knowledge about the limitations of data sets can result in the misuse of data. As an example, Hunter and Beard (1992) suggest that inexperienced users may attempt to add ordinal data values when overlaying raster themes. Such a mathematical operation is not valid on this type of data, although users who are not experienced in handling data may unwittingly perform this form of misuse error and the GIS software will permit them to complete the operation.

Convenience and cost of spatial data are major factors for most users to consider, which may take greater priority than the appropriateness of a data set. Hunter and Beard refer to Blakemore's (1985) example of how data can be inadvertently misused. He describes how a digital data set of administrative districts was collected by the British Department of the Environment for use as a thematic mapping base. Attribute accuracy, rather than positional accuracy, was of greater importance to the intended use of this data set and the position of boundaries was not recorded to a high degree of accuracy. However, as time passed and the data set became popular, agencies began to use it for purposes other than its original intention. Positional accuracy assumed a far greater importance to some of these further applications and it was found that the data set was not of sufficient quality for these uses, with point-in-polygon searches indicating some locations to be several kilometers out into the North Sea. Such errors in the final product are attributable to misuse of a data set that was created for a different purpose.

2.3 Data Quality Elements and Metadata Reporting

The three sources of error lead to errors in the final product of the GIS. However, the true values represented in the final product are not usually known, preventing us from giving an accuracy statement for GIS output. Instead, we refer to the level of uncertainty of the final product. Goodchild (1989) suggests that the term ‘uncertainty’ does not have the negative connotations of the term ‘error’ and is therefore a more appropriate descriptor.

Acquisition errors are inherent to a data set and, as such, cannot be avoided. Similarly, the propagation of errors through a GIS is intrinsic to the manipulation of spatial data and is inescapable. However, if the level of accuracy of a data set is known, it should be possible to track the propagation of these errors through a GIS and to determine the level of uncertainty of the final product. In addition, knowledge of the accuracy of a data set is required if we are to minimise, and possibly eliminate, the third source of error, that of data misuse. If the limitations of a data set could be communicated to the user, with education and training, the frequency of such usage errors would be greatly reduced. Complete knowledge of the quality of a data set is therefore essential to both the assessment of the uncertainty of a GIS product and the reduction of data misuse errors.

The need for reporting of data quality has long been recognized by those working with spatial data. The U.S. Spatial Data Transfer Standard (NIST, 1992; ANSI, 1998) requires that the quality component of spatial data be communicated to users along five themes. These are those of lineage, positional accuracy, attribute accuracy, logical consistency, and completeness of the dataset. Temporal accuracy of the data is one further aspect that must be considered, as this cuts along each of these five themes. Lineage refers to the history of the data set and is, as such, untestable. However, the remaining themes are testable and the quality of a data set with respect to each must be understood by decision-makers, since the erroneous application of a data set can lead to serious consequences.

2.3.1 Positional accuracy

A lack of consideration of the positional accuracy of data led to grave consequences recently in the United Kingdom. Following an outbreak of foot and mouth disease, the government partook in a policy of culling livestock exposed to infected animals. Five hundred sheep were slaughtered at a farm in Cumbria, by soldiers believing that the farm lay within a 3km buffer zone of a confirmed case of the disease. However, the grid reference was wrong by one digit and they should have been slaughtering animals several miles away (Source: the Daily Telegraph newspaper, 21 April 2001).

Reporting of the positional accuracy of a data set is usually done at the global level. Accuracy statements are typically based on the root mean square error, RMSE, (ANSI, 1998) which is calculated from a sample of *n* test points, for which the true *x* values are known and compared to the data set's values, $\hat{x}$, using the rule:

$$RMSE = \sqrt{\frac{\sum (x - \hat{x})^2}{n-1}}$$

The RMSE can then be used to estimate confidence intervals for the data accuracy, a typical positional accuracy statement being that 95% of values are within ± 20m of the true value. Such statements assume that the errors are normally distributed and are independent. However, much spatial data exhibits autocorrelation and the spatial distribution of errors is often not homogeneous.

Hunter *et al* (2002) reflect on the need for more complete data quality descriptions that include local, rather than global, measures and also report the level of autocorrelation within the data. Indeed, Hunter and Goodchild (1997) demonstrate the importance of including information about the spatial autocorrelation of errors in data quality reports for digital elevation models (DEMs). They modeled the spatial uncertainty within a digital elevation model from which slope and aspect calculations were derived, by producing multiple alternative realisations of the DEM. These realisations were produced using different levels of spatial autocorrelation within random error fields applied as a disturbance term to the elevation model. They found that the errors in the calculated slope and aspect values were dependent on the level of autocorrelation used in generating the disturbance field. They concluded that since digital elevation models are generally derived using some form of interpolation, the errors within each must have an associated level of autocorrelation. This spatial autocorrelation will affect subsequent calculations, such as slope and aspect, and must therefore be reported to users of the data set.

2.3.2 Thematic accuracy

Thematic accuracy refers to the correctness of the information concerning attributes of the geo-referenced locations or features. The type of attribute can vary widely, for example, it may be the name of a lake, the soil type of a region or the population of a city. Within a database, many attributes may be recorded for a particular feature, such as the name, length, pH value, salinity and fish population of a river. These attributes may be of different data types, such as categorical or numerical.

Many of those using spatial data further categorize the level of measurement into nominal, ordinal, interval or ratio, following the classification scale of Stevens (1946). However, Chrisman (1995, 1997) proposes that other levels of measurement are relevant to geographic date, such as cyclic ratios for circular data. A further level of measurement that is of particular importance as we consider uncertainty within GIS products is that of absolute ratio. Probability values are within the range from zero to one and meet Ellis' (1966) definition of an absolute scale, that there are no possible transformations that can be applied to these values that retain the original meaning of the measurement.

An understanding of level of measurement is important as it restricts the mathematical operations that are applicable to a data set. For example, it is not appropriate to calculate the mean of a set of ordinal data values, as the values do not necessarily differ by equal intervals. However, as Goodchild (1995) points out, current GIS rarely restrict the use of mathematical operations to those that are appropriate to the data type. Such restrictions should be reasonably straightforward to implement, if the data types were included in the metadata, and would reduce the likelihood of data misuse errors.

The level of measurement of a data set also affects the manner in which thematic accuracy can be reported. For interval or ratio data, measures similar to those used for positional accuracy, such as the RMSE, are applicable. However, different measures are necessary for nominal data. One commonly used statistic is the percent correctly classified (PCC). This requires that the true classification of a set of sample data points is known. A misclassification matrix is then produced, plotting the data set classification against the true classification of each sample point. From this matrix, the PCC is calculated by summing the number of classifications that are in agreement and dividing by the total number of classifications in the sample.

Congalton (1991) has shown that this measure is sensitive to sampling technique, at least in the reporting of classification accuracy for remotely sensed images. He created several sets of sample points, using different sampling techniques, and derived the misclassification matrix for each sample. From these matrices, he calculated the percentage of cells correctly classified within the remotely sensed image. These PCC values differed significantly between the sampling techniques; Congalton concluded that certain techniques were more appropriate than others in assessing classification accuracy.

Furthermore, the PCC value does not take into account the number of correct classifications that would be a result of chance. Goodchild (1995) states that a preferred index is the Kappa statistic, which returns the percentage of cells correctly classified above that expected from chance alone. Two other measures of thematic accuracy for categorical data are producer's accuracy and consumer's accuracy. Producer's accuracy returns the probability that features of a particular class do appear as that class in the database. Consumer's accuracy, on the other hand, returns the probability that features that appear to be of a particular class from the database are actually of that class in reality. These different measures are all useful statistics. However, they are all global, rather than local, measures and do not report the spatial variation in data set accuracy.

These measures of reporting the accuracy of nominal data may sound similar but they return values that can be quite different and have very different meanings. They should all be quoted, as many people do not have an intuitive understanding of the differences between them. For example, the contingency matrix below shows a sample of data that has been classified into two categories: A and B. The true classifications are plotted against the dataset classifications, with the classifications that agree shaded. Thematic accuracy statistics are calculated as shown.

		True classification		
		A	**B**	***Total***
Dataset classification	**A**	60	5	*65*
	B	20	15	*35*
	Total	*80*	*20*	*100*

PCC = 75/100 = 0.75

Kappa = (75-59)/(100-59) = 0.39, where 59 = (80x65 + 20x35)/100

Producer's accuracy for A = 60/80 = 0.75, for B = 15/20 = 0.75

Consumer's accuracy for A = 60/65 = 0.92, for B = 15/35 = 0.43

Table 2.1. Sample misclassification matrix and associated calculations for thematic accuracy statistics.

It can be seen that the Kappa statistic is considerably lower than the PCC, as it does not include the proportion of cells that are correctly classified as a result of chance. Both the PCC value and the producer's accuracy values are 0.75, meaning that three quarters of the cells are correctly identified in the database, for both class A and class B. However, the consumer's accuracy figure for B indicates that only 43% of those cells that appear to be of class B in the database are actually of this class in reality. Unless this figure is explicitly stated, it is understandable that many people would incorrectly interpret the PCC value of 0.75 to mean that three quarters of those cells that appear to be class B are truly of this class.

Although positional and thematic accuracy have been discussed as two distinct measures, several researchers have contended that they often interact and are at times inseparable. Veregin (1989, p.45) states that *'…polygon boundaries are defined in terms of the values of the thematic attribute themselves and thus attribute and positional errors are not independent'*. Goodchild (1995) provides a good introduction to some of the complexities of describing thematic data. These include the different ways that attributes may be related to the earth's surface, such as cities being represented as points at low scales but being composed of a multitude of features, all having their individual attributes, at higher scales. In addition, the accuracy of an attribute such as the area of a feature will be dependent upon the positional accuracy of the dataset and, similarly, attributes such as population density require the use of aggregation areas in their calculation, causing positional accuracy and thematic accuracy of such a data set to be compounded. It is often helpful to consider data quality in terms of the five elements identified in the SDTS, although, in practice, the components may not be easily separable.

2.3.3 Logical consistency

Logical consistency is defined within the SDTS as “the fidelity of relationships” encoded within the data structure. For example, it includes issues such as the conformity of topological relationships; the use of tolerance levels to remove under-shoots and over-shoots and to close polygons must be reported. The importance of communicating the logical consistency component of data quality is illustrated in this example of a German motorist relying on his vehicle’s GPS-based navigation system to select a route between two towns. Approaching a ferry terminal, he drove his car straight into the Harvel River. The navigation system, in which the motorist had shown complete faith, had depicted the ferry crossing as a bridge (Source: The Observer newspaper, 27 Dec. 1998).

2.3.4 Completeness

The completeness report mandated by the SDTS includes information about omissions, selection criteria, definitions used and other rules that may have been applied in deriving the data set. For example, geometric thresholds, such as the minimum width of a bridge for its inclusion in the data set, must be reported. The importance of being aware of the completeness of a spatial data set was demonstrated with disastrous consequences in 1998, when a U.S. military aircraft sliced through an Italian cable car line. The New York Times newspaper reported that the map on which the pilot had been relying did not include the cable car line (source: CNN.com, 1998).

2.3.5 Temporal accuracy

Temporal accuracy is not, in itself, a component that the SDTS identifies as mandatory in the reporting of spatial data quality. However, any representation of a dynamic world needs to be regularly up-dated if it is to remain relevant. The temporal accuracy of a data set relates to each of the previously described components: positional accuracy, thematic accuracy, logical consistency and completeness. The level of temporal accuracy may therefore need to be reported with respect to each of these. The following example illustrates how important an appreciation of temporal accuracy can be.

In 1999, NATO accidentally bombed the embassy of the People's Republic of China in Belgrade. In the subsequent report, apologizing to the president and people of China, the U.S. Under Secretary of State for Political Affairs admitted that it had been a mistake arising from faulty intelligence reports and poorly maintained spatial databases. The intended target of the NATO bombing was the headquarters of the Yugoslav Federal Directorate for Supply and Procurement. However, in attempting to locate this building, officers had relied on imprecise techniques. None of the maps or spatial databases used to verify the target location contained the correct position of the Chinese embassy. Each located the Chinese embassy on the other side of Belgrade, although the embassy had moved to its current location four years earlier. The report from the United States consulate (Source: U.S. State Department Report on Accidental Bombing of Chinese Embassy, 1999) admitted that *'..although database maintenance is one of the basic elements of our intelligence efforts, it has been routinely accorded low priority'*. This failure to update spatial information had catastrophic consequences.

2.4 Communicating Uncertainty in Spatial Information

Acquisition and processing errors that are inherent to the data within a GIS will result in uncertainty in the final product. It is widely accepted that this uncertainty must be communicated to users of spatial information, so that they can be made aware of the limitations of the GIS output on which they may be basing their decisions. However, the current requirements of spatial data standards, such as the SDTS, to report metadata concerning data quality may not be conveying the information in a manner that is comprehensible to many users. Of those users that do understand statistical statements of data quality, such as positional accuracy statements, it would be reasonable to assume that some may experience difficulties in understanding how these relate to the level of uncertainty within a final GIS output.

In light of this, many researchers have proposed that visualisation of uncertainty information may be the most effective means of communicating it to users. Buttenfield and Mackaness

(1991) review the historical use of visualisation to convey information and argue for its effectiveness in communicating complex patterns in both spatial and statistical data, particularly as the volume of this data increases. They contend (p.427) that *'the nature of spatial data and more generally of geographic information mandates the use of visualization for both efficiency and acuity in the analytic process.'*

Buttenfield and Mackaness regard visualisation as the interface between three processes. These include computational analysis, human cognition and graphic design. The latter process, that of graphic design, is concerned with the principles underlying effective communication of information through visual representations. Tufte (1983, 1997, 2001) has demonstrated numerous examples of both graphical excellence and failing in his books on visual communication. His principles of graphical excellence include the need for communicating complex ideas with clarity, precision and efficiency. These same principles would make sensible guidelines underlying the use of visualisation to communicate uncertainty information. However, guidelines regarding the interactions of computational analysis, human cognition and graphical design are, as yet, not so well defined.

Three impediments to the communication of uncertainty are described by Buttenfield (1993). In addition to the problems associated with defining and assessing data quality, she describes the difficulties that arise in attributing data quality elements within a spatial database, particularly when manipulating multiple data sets within a GIS, and goes on to consider the impediments to graphical representation of data quality. She concludes that there is a need for empirical research and cognitive testing of data quality representations, to assess user comprehension of such displays and their ability to convey uncertainty information in a manner that meets the needs of the user.

Beard and Mackaness (1993) identify three levels to the process of communicating uncertainty information through visualisation. The lowest of these levels is notification, which simply alerts the user to the fact there may be a data quality problem. The second level, identification, serves to locate and to identify the nature of this potential problem and the third level, quantification, provides a measure of the problem. They propose that visual displays of uncertainty information should be able to address the issue at any of these levels and also need to take into consideration the use to which the user is to put the information, for example data exploration, assisting decision-making or analyzing patterns in uncertainty.

Reinke and Hunter (2002) propose a theory for communicating uncertainty information to users that incorporates a fourth level beyond these three, that of evaluation. They suggest that the system should also evaluate the significance of the uncertainty measure to the given application. This requires a method that enables users to understand the changes that may occur between the original data and that data when its uncertainty is acknowledged. They conclude that a representation of the quantified uncertainty information needs to be incorporated with the original data.

2.4.1 Quantifying the level of uncertainty in spatial data

To date, much of the research concerning the inclusion of uncertainty information in spatial data has focused on methods of quantifying the level of uncertainty and methods of visualizing this uncertainty. Drummond (1995) describes how the technique of propagation of variances can be used as a means of measuring how positional errors propagate through GIS. For example, if A is a function of variables B, C and D, then the variance of A can be estimated as the sum of the weighted variances of each of the contributors B, C and D, together with their weighted co-variances if they are not independent. In the simplest case:

$$\operatorname{var} A = \frac{dA}{dB}\operatorname{var} B + \frac{dA}{dC}\operatorname{var} C + \frac{dA}{dD}\operatorname{var} D$$

However, this method is computationally extensive and its accuracy is necessarily dependent upon the accuracy of the given variances of contributors B, C and D. Drummond concludes that in many applications it may be more practical to consider simulation as a means of tracking error propagation.

A complete understanding of how error propagates through GIS may be a long way away, although Hunter (1999) demonstrates a method for tracking the positional changes to feature coordinates as they are subject to editing procedures within GIS. In developing a mathematical model to predict probabilities, rather than Boolean results, of a point-in-polygon analysis within GIS, Cheung *et al* (2004) consider the propagation of positional errors of both the point and the polygon vertices. Similarly, Cheung and Shi (2004) propose a model for estimating the positional uncertainty of a line following the propagation of errors associated with line simplification in a GIS.

Error propagation is an extremely complex process as data sets of differing quality may be implemented in a GIS. These can be overlaid and manipulated in many ways before the product of the system appears before the user. Petry *et al* (2003) propose a framework that uses software agents to manage the uncertainty associated with integrating multiple data sources, through the application of fuzzy logic. However, this remains a preliminary framework and they acknowledge that the integration of multiple data sources remains a challenge to researchers.

Rather than understanding the exact nature of error propagation, considerable research has been focused on the effects that the accuracy of spatial data may have on the system output. Openshaw (1989) reflects that users need to learn to live with errors in spatial databases and has proposed that many equally probable realizations of the same data input can be readily output from GIS. An examination of these multiple realizations can give the user an understanding of the influence that errors in the data may have on the product of the GIS and therefore assist in their decision as to whether or not a particular data set is fit for use in their specific application.

Goodchild (1995) expounds this idea in his discussion of using Monte Carlo simulation as a means of modelling error propagation. If the accuracy of a data set is provided, randomly perturbed error fields can be added as distortions, resulting in multiple realizations of the data. The differences in these multiple realizations can be used as a measure of the uncertainty in the dataset, or the multiple realizations can be further manipulated within GIS, providing users with the resultant effects on GIS output. A simple Monte Carlo method simulates an error value for each grid cell by randomly selecting a value from the assumed normal error distribution. Conditional simulation takes into account spatial autocorrelation of errors in generating the simulated error values for each grid cell.

Englund (1993) used simple Monte Carlo simulation to produce three realizations of two surfaces from their variogram models, which he compared to the surfaces obtained from kriging. He found that, when overlaid, the maps from the simulated surfaces were more realistic than that from the kriged surface. The kriging process produces an over-simplified map due to the smoothing nature of the averaging process. The natural spatial variability was better modeled using simple Monte Carlo simulation.

Many researchers have used the Monte Carlo method to examine the effects of data accuracy on output quality. Davis and Keller (1997) used fuzzy surfaces and Monte Carlo procedures to produce multiple realizations to model slope stability. Ehlschlager et al. (1997) applied multiple realizations in their finding of a minimal cost route in modelling road construction through rough terrain and Hunter et al. (1999) demonstrated how random error grids perturbed in two dimensions can be used to produce multiple realizations representing the uncertainty in vector data. In a study of uncertainty and sensitivity analysis, Crosetto and Tarantola (2001) ran multi simulations in a case study of hydrologic modelling to assess flood vulnerability in Italy. These and other similar studies have demonstrated that the effects of error in data input can be represented as uncertainty in the information output, reflected in the differences in the equally probable multiple realizations of the same data.

Holmes *et al* (2000) used conditional simulation to study the effects of error in digital elevation models on terrain modeling. Their exploratory data analysis revealed that, although the global error was relatively small, local errors could be comparatively large and exhibited spatial correlation. They generated 50 realisations of the DEM by applying spatially correlated random error fields to the data set and were able to provide a map showing probability of hillside slope failure, rather than the binary will/will not fail map derived from the original DEM alone. Heo (2003) observes that the generation of spatially correlated error fields is computationally extremely extensive. He compares several methods of computing spatially correlated random error fields for use in Monte Carlo simulations to conclude that the method of steepest descent is the best of those methods with linear computation complexity.

However, Van Niel and Laffan (2003) have demonstrated that the Monte Carlo method of producing multiple realisations of a GIS output is subject to bias, as a result of the biases inherent to random number generators. Depending upon the type of pseudo-random number generator used in the simulation, they showed that significantly different results could be obtained in a Monte Carlo analysis. They concluded that some types of random number generators are more appropriate for spatial analysis than are others and further study into the way that these biases may propagate through such analysis is required. At the least, researchers should report which type of random number generator they have used in a Monte Carlo analysis.

2.4.2 Methods of visualising the uncertainty in spatial data

Although considerable progress has been made into methods of quantifying the level of uncertainty in spatial information, the question still remains as to what is the most effective way to communicate this spatial uncertainty to users. The portrayal of positional uncertainty appears to be the least problematic, particularly with regards to vector data. Typical visualisation methods have placed probabilistic buffer zones around the object of interest, to represent confidence intervals of the object's position.

Often, a circular zone may be placed around a point object, for example, to visualise the 90% Circular Map Accuracy Standard (U.S. Bureau of the Budget, 1947). This assumes that the horizontal positional accuracy is constant in each direction. An error ellipse that divides the horizontal positional error into two components, one in each coordinate direction (ASPRS, 1990), may be more appropriate in some applications (for example, Fraser *et al*, 2003). Similarly, the Perkal epsilon band (Perkal, 1966) is widely used to represent the error band around linear features and this concept can be developed to represent the positional accuracy of polygon boundaries.

These buffer zones tend to be represented as internally homogeneous. However, it would be expected that the object is more likely to be at the centre of the zone than towards the edge if the positional errors are normally distributed. Graduated shading of the buffer zone could be used to represent the probability distribution of the object's position, although this more complicated representation may become confusing. An example of the use of graduated shading to represent positional uncertainty is tested in this study.

Clapham and Beard (1991) propose the potential use of the visual variables identified by Bertin (1983) to display the uncertainty of thematic data. Bertin refers to six retinal variables: size, value, texture, colour, orientation and shape. Colour has been further differentiated into the two variables of hue and saturation (Morrison, 1974). The order of these variables represents the proposed hierarchy of perceptual properties that Bertin suggests these variables hold, with the former variables more suited to numerical data and the latter more suited to nominal data. Reinke (2002) summarises the findings of several researchers who have assessed the suitability of these seven visual variables to portray data of different measurement levels. She reports that these results have generally been in agreement with

Bertin's findings. Her summation is shown in the following table, with each cross indicating a researcher finding the variable unsuitable, each open circle indicating a researcher finding the variable suitable with modification and each closed circle indicating a researcher finding the variable suitable for that level of data.

Visual variable	Nominal data	Ordinal data	Numerical data
Size	✗ ✗ ✗ ●	●●●●●●	●●●●●●●
Value	✗ ✗ ✗ ✗	●●●●●	○○●●●●
Texture	○●●●●●	○○●●	✗○●●
Colour (saturation)	✗ ✗	●●●	○●●
Colour (hue)	●●●●●●●	○○○○●	✗○○
Orientation	●●●●●●	✗○○	✗○○●
Shape	○●●●●●●	✗ ✗ ✗	✗ ✗ ✗

Adapted from Reinke (2002), summarising the findings of Andrienko and Andrienko (1999), Slocum (1999), MacEachren (1995), Senay and Ignatius (1994), McGranaghan (1993), DiBiase *et al* (1992), Mackinlay (1986) and Morrison (1974).

Table 2.2. Summary of appropriateness of using the seven visual variables to portray data of different measurement levels.

It can be seen that size and value are the only two variables generally considered to be suitable for numerical data, whilst hue, orientation and shape are generally considered to be most appropriate for nominal data. The level of uncertainty of a GIS product may be represented as nominal, ordinal or numerical data, depending upon the requirements of the application. For example, it would be possible to produce binary maps to represent slope stability, with regions where the probability of landslide exceeded a set threshold being of one colour and regions below the threshold being another colour. Hue is an appropriate visual variable for this binary representation of the quantified uncertainty. However, hue would not necessarily be appropriate if a choropleth map were used to represent ordered classes of uncertainty levels. In this case, value, saturation or glyph size would be more applicable, as these variables, unlike hue, have an intuitive ordering. Aerts *et al* (2003) used hue in a binary representation of uncertainty but value in an ordinal representation of the same information in a web-based survey assessing the effectiveness of different representation methods. They found that the majority of participants preferred the ordinal representation, although they could see some merit in the binary one.

Clapham and Beard (1991) also argue that the mode of display, such as overlay, inset, adjacency, animation and three-dimensionality, must also be taken into consideration. The possibility of overlaying uncertainty information on a thematic map would only be practical if this information did not interfere with our perception of the underlying thematic information. Several researchers have suggested that overlaying maps with uncertainty information may overload our cognitive abilities and lead to reduced comprehension (for example, Buttenfield & Beard, 1994). MacEachren (1992) suggests that uncertainty information may be overlaid through the use of colour saturation, focus of the image (through blurring or the addition of a fog overlay) or spatial resolution. Similarly, MacEachren *et al* (1998) propose the use of texture as an overlay to represent the level of uncertainty in health statistics information. They found that texture remained visually separable to the hues that were used to represent the statistical data, whereas saturation tended to be integrated with hue and interfered with participants' ability to extract cluster and pattern information from the morbidity rate data.

Fisher (1993) offers the potential use of animation in GIS representations. He proposes that the level of uncertainty in thematic data could be displayed by causing pixels to blink. Using soil maps, he represented greater levels of uncertainty by increasing the frequency with which the pixels blinked, hence lowering the stability of the image. Ehlschlaeger *et al* (1997) utilise a different method of animation to represent the level of uncertainty of costing potential roadways. They made a movie of the different cost realisation maps, causing the less certain cost routes to change more frequently as the movie ran.

An alternative approach to representing the uncertainty in spatial information was adopted by Fisher (1994) and Krygier (1994). They used sound to represent the level of uncertainty within visual representations of thematic information. As the mouse moved over uncertain regions of the map, a high pitch sound was emitted, whereas a lower pitch accompanied more certain regions. Although the use of a second sense to portray the level of uncertainty prevents interference between the uncertain and thematic representations, the use of sound may not be intuitive and may become annoying to users.

Researchers have proposed a variety of different ways to portray the level of uncertainty in spatial information but with little testing of the effectiveness of these methods. Kardos *et al* (2003) conducted a survey amongst subscribers to a GIS user group, to assess how useful these users found a selection of these techniques to be in communicating spatial uncertainty. The assessment scale was from 0 (not useful) to 5 (excellent). Of the nine methods tested, the majority of participants rated all but one of the techniques to be of class 0, not useful. Only the method of blinking pixels was rated as useful by more participants than found it to be of no use. This study highlights the need for empirical research into the effectiveness of techniques for representing spatial uncertainty.

2.5 Effectiveness Research

MacEachren and Kraak (2001) emphasise this need for cognitive testing of representations of uncertainty in spatial data, in both their ability to convey knowledge and to assist decision-making. They also reflect on the need to consider individual differences among users, such as gender, experience, culture and sensory disabilities. Research has not typically been focused in this area. Academics have generally been concerned with quantifying measures of uncertainty and producing innovative methods of visualizing these in spatial data, without necessarily testing the effectiveness of these representations. However, it is imperative that the effectiveness of any method designed to communicate information to users is assessed.

Crossland *et al* (1995) performed one of the first examples of cognitive testing with GIS. They conducted an experiment to empirically assess the effectiveness of using GIS to assist in decision-making. Subjects were presented with a spatial decision task that required them to rank several potential sites for a center developing a new fuel technology. Multiple spatial criteria were provided and the subjects were divided into two groups: those given traditional paper maps of the spatially referenced information and those with access to a GIS. They hypothesized that the subjects using GIS to assist in their decision-making would make faster, more accurate decisions than those without access to a GIS, particularly as the decision complexity increased. The results supported this, although the interaction with task complexity did not reach significance at the 5% level. They were able to conclude that the use of GIS does result in better decision-making.

The few researchers who have attempted to test the effectiveness of representations of spatial uncertainty have taken a similar approach. Leitner and Buttenfield (1997, 2000) introduced uncertainty information into a study of siting decisions for a park and an airport when the complexity of the problem (number of attribute classes) changed. They used static representation of the uncertainty information by value, saturation or texture, which was overlaid on a map using hue to represent attribute information. Subjects were asked to select the best site for the park and the worst site for the airport. Greater accuracy rates, faster response times and greater confidence in decisions were considered to reflect improved decision-making.

Leitner and Buttenfield (2000) found that the uncertainty information did not increase response times, correctness or confidence in the decisions and even decreased response times for the park siting decision. They concluded that users are able to handle the uncertainty information without any cognitive overload and that it may actually serve to help clarify the attribute information. However, this study only introduced a binary level of certainty information and did not account for learning effects.

The nature of the task in the Leitner and Buttenfield study, which was to identify the optimal site for the park and the worst site for the airport, required little interpretation of the uncertainty information. Indeed, it may be that subjects simply used this information to restrict their search to regions that were certain. This would explain the finding that reaction times in the park siting decisions actually decreased when the additional certainty information was supplied. In selecting the worst possible site for the airport, the frequency of correct responses was only around 50 percent, both with and without the inclusion of uncertainty information. The task itself may have been unclear to many participants, masking any effects that the inclusion of uncertainty information may have had.

Additional research into the effectiveness of uncertainty representations has been conducted by Evans (1997). She included binary uncertainty information in land use maps and asked users to interpret the information portrayed. Her representations included adjacent maps depicting thematic and uncertainty information, a toggling of the thematic and uncertainty maps, static maps overlaying thematic and uncertainty information, and dynamic (blinking pixel) maps. She investigated how novice and expert users of both genders used and rated the uncertainty information in interpreting the land use maps. She found that both groups used the uncertainty information and tended to prefer the combination maps, overlay and blinking pixel, to the adjacent depictions and toggling display. However, her subjects were either students or academics. Both of these groups may be expected to be more interested in critical evaluation of represented data, through an analysis of uncertainty information, than the more general user.

Despite this objection, these findings support those of Leitner and Buttenfield in suggesting that users are able to comprehend uncertainty information when it is included in representations of thematic spatial data and that they are able to do this when the uncertainty and thematic information are overlaid on the same map. In both of these studies, the uncertainty information was presented as binary data, either reliable or not. If a threshold level is introduced to uncertainty representations, it may well be that users are able to comprehend this information and incorporate it into their decisions. This requires an *a priori* assumption as to the level at which the uncertainty information is significant, since the data has to be classified into binomial classes of certain or uncertain. Such assumptions and depictions may be more appropriate for some user groups and applications than others. It may be useful to novices to have uncertainty information in this simplistic form. However, analysts may require a more detailed representation of uncertainty if their final decisions are to be truly informed.

Aerts *et al* (2003) conducted a web-based survey for which the participants were GIS users, rather than students or academics. They also represented the uncertainty information as ordinal classes, portrayed with lightness values, rather than binary data. However, possibly due to the complexity of having ordinal, rather than binary, uncertainty information, they did not overlay this on the attribute information but presented subjects with adjacent or toggling displays. They concluded, in agreement with Evans and Leitner & Buttenfield, that subjects were able to understand the uncertainty information included in both the adjacent and toggling displays. However, one question asked subjects whether they were able to make simple approximations (for example, 50%, 75%) about the amount of uncertainty associated with the data. The majority of subjects' responses, on a scale from 1 (not at all) to 5 (completely), were in the 2-3 range. This indicates only a limited understanding of the uncertainty information alone, without any requirement to apply it to decision-making. The confidence of Aerts *et al* in the ability of the display to communicate uncertainty information seems rather optimistic.

These studies have investigated subjects' ability to comprehend uncertainty information when it is presented together with thematic information, without interference or cognitive overload. Similarly, MacEachren *et al* (1998) were able to conclude that texture overlaid on choropleth maps was able to display uncertainty without interfering with the underlying morbidity rate data and Kardos *et al* (2004) are currently assessing survey participants' ability to comprehend the uncertainty information in tessellation overlays. However, with the exception of Leitner and Buttenfield (1997, 2000) none of these studies has required an interpretation of the uncertainty information within the context of a decision task. It can be argued that even the Leitner and Buttenfield study did not require interpretation of the uncertainty information, as the correct decision could have been made by simply using the uncertainty information to restrict the search to certain areas.

It is acknowledged that the uncertainty inherent to spatial data must be communicated to the users of GIS in the manner that is most effective to the decisions being made. However, how the effectiveness of such uncertainty representations is to be assessed is a question that has arisen (for example Slocum et al, 2001) without having been satisfactorily answered. The question is not an easy one, since effectiveness would appear to be specific to the task at hand. Slocum et al suggest the need for analysis of user tasks, software evaluation by experts, formative user-centred evaluation and task-based comparison of alternatives.

It would appear that the logical approach to answering the question of what is the best way to depict uncertainty in spatial information would be to focus on the users and to address their needs. Knapp (1994) has argued for a task analysis approach to the visualization of geographical data, stating Casner's (1989) finding that visual displays are of optimal support only when they directly support user tasks. He contends that closer relationships are required between the needs of users and the output of software systems. Suchan (2001) conducted a study of the usability of geovisualization software through interviewing population census analysts in the workplace. In addition to the application-specific functionality requirements, she identified more general user wishes for representations that are visually easy to work, visually easy to interpret, able to speed up the process of data exploration and able to make the process more interesting. Beard and Mackaness (1993, p.43) agree with this, stating that *'if the visualization techniques for investigating data quality are difficult to use, users will simply not access them'*.

The user-desirability of ‘quick and easy’ representations is of little surprise but is at odds with the nature of uncertainty in GIS output. Indeed, many users are uncomfortable with the concept of uncertainty *per se*, and education in the use of even the most simplistic binary uncertainty representations may well be required. When more complex decisions are being made, researchers are arguing the need for a more detailed depiction of uncertainty information to prevent the mis-interpretation of GIS output. However, for researchers to assume that their job is finished in passing these representations to users seems naive. If pictures are speaking their thousand words, the nature of these thousand words and their application to the decision at hand may be far from clear. Even the most expert users may require assistance in deciding how this uncertainty information is to be applied to decision-making.

There has been some research into the question of how uncertainty information can be incorporated into decision-making. This research has primarily focused on using uncertainty information to assess the fitness for use of the data set. Agumya and Hunter (1999) propose that a risk-based approach to assessing the fitness for use of spatial data is more appropriate than a standards-based approach. Rather than assessing how much uncertainty is acceptable in the decision and taking that back to the data, they argue that it is more appropriate to consider the level of uncertainty in the data and assess how that will affect the level of risk in the decision. A cost-benefit approach can then be used to decide if this level of risk is acceptable.

This approach has been further expanded by De Bruin *et al* (2001). They utilized decision trees and Bayes’ theorem to assess the expected value of information. This theory was then successfully applied to a case study to decide which of two available DEM data sets would be the most cost effective in determining the volume of sand required to build a container port.
Nonetheless, research in this area is relatively sparse and there are no clear guidelines indicating how users of spatial information can utilize the uncertainty information that is provided.

Slocum *et al* (2001) have identified that one of the difficulties that researchers have encountered in assessing the effectiveness of uncertainty representations is that it has not been clearly stated exactly what needs to be assessed. The question remains as to what the depiction of uncertainty is attempting to communicate to the user and how this is applicable to the decision at hand before the effectiveness of the representation can be assessed. It may well be that this is to some extent dependent upon the specific application of the user but it still remains that the provision of some guidelines here is an area in need of much research.

There is a widely held assumption amongst academics working within the GIS community that, if decisions being made on the basis of spatial information are to be as informed and robust as possible, it is essential that the level of certainty of that information be communicated to the decision-makers. It is implied that the inclusion of certainty information in the output of a GIS will lead to better decision-making. For example, Fairbairn *et al* (2001, p.20) state that *'making information available about data uncertainty… is essential, if users are to make informed decisions'*. In a similar vain, within the executive summary to the NCGIA research initiative into visualization of spatial data quality, Beard *et al* (1991, p.iv) argue that:

> *'Information on the quality of data is essential for effective use of GIS data… The credibility of spatial decision support using GIS may indeed depend on the incorporation of quality information within the database and display.'*

However, nobody seems to be asking what it is that decision-makers are going to make of this certainty information. It may well be that the concept of uncertainty is not an intuitive one and the inclusion of such information may not actually assist rational decision-making. Kardos *et al* (2003) report that their subjects, 39% of whom were experts in GIS, 48% advanced and 13% beginners, found that the concept of an uncertainty measure was difficult to grasp. If these subjects, over half of whom were employed in professions using GIS or in government, were hesitant around the notion of uncertainty, one can only speculate as to what users less experienced in GIS are going to make of it.

Is it, then, a fair assumption that the provision of the information *per se* is sufficient to result in better decisions? Mark Harrower (2003, p.1) poses the question: *'…does displaying uncertainty on maps fundamentally change the way people think and problem-solve and ultimately lead to better decisions?'* Studies from the psychological literature indicate that people are averse to the concept of uncertainty and tend to avoid it wherever possible. If this is the case with spatial information, the provision of certainty information may simply lead to bias in decision-making, rather than informed, robust decisions. If this is the case, extensive education may need to accompany the inclusion of certainty information in GIS output.

2.6 Chapter summary

Errors are inherent to spatial data and will lead to uncertainty in the product of GIS. This uncertainty must be communicated to users if the GIS output is to truly inform the decision-making process. Research has primarily focussed on quantifying the level of uncertainty and producing methods of visualising this, without testing the effectiveness of these methods. For example, multiple realization studies, using Monte Carlo simulation to quantify uncertainty, have generally opted for visualization methods to communicate the uncertainty to users. These have included probability maps with colours representing differing susceptibility to landslide as identified from the multiple realizations, dynamic representation of minimal cost and superimposed representations of polygon boundaries complemented with statistical analysis.

Visualisation lends itself to the representation of the spatial distribution of output uncertainty. However, which forms of visualisation are most effective in communicating the level of uncertainty is an area that has received little research. The research to date has generally considered whether or not users are able to comprehend the uncertainty information without it interfering with their perception of the underlying data. When the uncertainty information is binary and presented as overlay or by toggling displays, the findings have generally agreed that users are able to extract this information without interference and without cognitive overload. However, these studies have not questioned what users are to do with the uncertainty information.

There has been little research into what decision-makers actually make of uncertainty information when it is included in GIS output. Academics working with spatial data have assumed that the inclusion of such information will lead to more robust, informed decisions, without testing how decision-makers might use uncertainty information. There appears to be a presumption that decision-makers will respond to uncertainty in a rational manner. However, studies from the psychological literature suggest that most people are averse to uncertainty and tend to avoid it. If such tendencies are apparent with spatial data, the inclusion of uncertainty information may simply lead to bias in decision-making, rather than fully informed decisions.

3.

Behavioural Decision Research

3.1 Introduction

Behavioural decision research is the collective name for studies into human decision-making behaviour under experimental conditions. Its aim is to understand decision-making processes in different choice and judgement tasks, with the goal of assisting decision-makers and enhancing decision-making performance. As such, the findings from behavioural decision research have direct applications to most, if not all, fields. Although generally considered to lie within the realm of psychology, the research focus cuts across many disciplines, including economics, sociology, political science and statistics. In fact, the domain of behavioural decision research has been of keen interest to economists, who have contributed to models of decision-making strategies and found many real life applications to exploit the research findings (for example Einhorn & Hogarth, 1986). However, other fields have been slower to take such findings on board and to apply them to decision-making within their own domains.

If the ultimate goal of studies into decision-making is to improve decision performance, it is necessary that we have a common understanding of what constitutes ‘best practice’ in decision-making. It is widely accepted that, in making the best possible decision, one is required to act rationally. However, the concept of rationality as applied to decision-making has given rise to some debate. It needs to go beyond the standard dictionary definition of rational as being *agreeable to reason* (Oxford English Dictionary, 1989) and to consider the decision outcomes. As such, rational decisions have traditionally been considered to be those that are expected to maximize accepted goals, given a set of premises.

Neoclassical economics, supported by the work of Savage (1954) in his book *Foundations of Statistics*, proposes exactly this and that such decision-making processes are characteristic of human behaviour. It argues that rational decisions are consistent and are made to optimise the expected outcomes, from the given information. Although the neoclassical economists concede to human fallibility and acknowledge that people may sometimes make a decision that is deemed incorrect on such a premise, they believe that decision-makers will recognise such an error when it is brought to their attention and will, on reflection, choose to change their decision to that which would optimise the outcomes. Rationality as applied to decision-making can be considered to be maximising the expected utility.

Simon (1955) contended that, in fact, human decision-making often deviates from such rational behaviour. Simon proposed that information-processing limitations exist, causing decision-makers to exhibit ‘bounded rationality’. The greater the complexity of a decision, through an increased number of conditions or possible outcomes to consider, the greater would be the effect of such limited processing capability. Humans, he suggested, show a tendency to trade-off the ‘correctness’ of a decision for the level of cognitive effort required to make it. In order to understand actual decision behaviour, Simon argued that research was needed to systematically investigate how humans deviate from rational choice.

Simon’s view is not necessarily in opposition to the normative models. Savage accepts that people do make mistaken decisions and these mistakes may well be a result of having employed heuristics of the type that Simon suggests. However, Savage would argue that, given time to reflect on any such mistakes, decision-makers would change their decision to that which is deemed rational in terms of maximising expected utility. Simon’s proposal that decision-makers employ simplifying heuristics may be descriptive of what decision-makers are doing in practice, whereas the normative strategy is more a prescriptive definition of what constitutes rational decision-making behaviour. However, such a definition requires that, given time to reflect, all deliberate decisions would conform to the Savage axioms. The findings from decision research suggest that theories of maximising expected utility do not adequately describe all decision-making behaviour. Some decisions may be reversed according to the task conditions and decision-makers may continue to believe in the correctness of certain ‘irrational’ decisions, even after any contradictions to normative values are brought to their attention.

3.2 Normative Decision-Making Strategies: Prescribing Rational Behaviour

The early studies of human decision-making processes assumed that normative strategies based on rational axioms could be used to model actual decision behaviour. These studies generally concerned two types of decision-making task: choice tasks and decisions under risk. Choice behaviour was investigated by facing subjects with a decision between several alternatives that varied along multiple attributes. No single alternative best met all of the objectives, so subjects had to choose between conflicting values. Decisions under risk involved offering subjects gambles, requiring them to choose between bets with differing payouts and probabilities of occurrence. From both types of task, the resulting models were based on the proposition that decision-makers attempted to optimize the expected value of the decision outcome.

3.2.1 Choice behaviour: Weighted additive value model (WADD) and its variations

In the WADD model, it is assumed that the different attributes (i = 1….n) are each given a weighting, w_i, according to their relative importance to the decision at hand. The value of the attribute for alternative X is then multiplied by its weight to determine its weighted value (w_iX_i). These weighted values are then summed over the n attributes to give the WADD value for alternative X:

$$\mathrm{WADD}(X) = \sum_{i=1}^{n} w_i X_i$$

This summative procedure is repeated for each alternative and the resulting WADD values compared to determine the optimal decision.

A modification to this model was later introduced, constraining the weights to sum to one. This caused the model to become an averaging rather than an additive model. The purely additive nature of the original WADD model meant that if an additional attribute with a positive value were to be included, the overall WADD value for that alternative would improve, even if this value were only mildly positive. However, in practice, we commonly encounter the situation where finding additional information on a particular option lessens its desirability. For example, we may consider a digital camera with high resolution and a high-zoom lens to be extremely desirable. The additional information that its battery life is only moderate to good may dampen its appeal rather than add to it. Research has shown that the averaging modification to the WADD model has improved its application to many decision-making tasks (Anderson, 1981). More recent modifications include variations that allow for first impressions of options and interaction between attributes.

One of the major criticisms of the WADD model has been its assumption that people can reliably evaluate attributes along a common value scale. People do not have pre-defined values for most attributes and objects but appear to construct such values within the experimental constraints of decision tasks. Slovic (1995) argued for the theory of constructive preferences, demonstrating that in decision-making experiments, values are constructed on the spot as needed. This makes expressed choice and judgment open to the effects of task and context factors, such as the presentation order of options. If values are constructed in such an *ad hoc* fashion, it appears difficult to argue that the application of these models to real life decisions is to result in the consistent, rational choices demanded of neoclassical economics.

3.2.2 Decisions under risk: Expected value and expected utility models

Instead of choosing between multi-attribute alternatives each with a known value, some decision tasks, such as gambles, involve the choice between alternatives under conditions of risk. Risk is defined as referring to situations where the consequences of a decision depend upon future outcomes having known probabilities of occurrence. Rational models of decision-making strategies under conditions of risk are the expected value model and its utility variations. The expected value of each outcome, ($i = 1 \ldots n$), for gamble X is the product of its value, $V(X_i,)$ and the probability, P_i, of its occurrence. For each gamble, these products are summed over the n outcomes to give the expected value:

$$\text{EV}\ (X) = \sum_{i=1}^{n} P_i \times V(X_i)$$

Consideration of this expected value therefore enables the decision-maker to evaluate the gamble.

In early studies of risky choice, Kahneman and Tversky (1979) found that most subjects do not behave in a manner that is rational in terms of the expected value model. For example, they asked questions offering choices similar to the following.

> *Which would you prefer: $3000 for sure or an 80% chance of winning $4000?*

The expected value of the latter option is $3200, although the majority of subjects chose the former. However, in a similar question with the outcome representing loss rather than gain:

Which would you prefer: a loss of $3000 for sure or an 80% chance of losing $4000?

Most subjects chose the latter option, despite this resulting in a greater expected loss.

Kahneman and Tversky proposed that these preferences could still be considered rational if the expected value model were modified to reflect expected utility. Bernoulli introduced the concept of utility in 1738, stating that *'any increase in wealth, no matter how insignificant, will always result in utility which is inversely proportional to the quantity of goods already possessed'* (Bernoulli, 1967, p.25). He suggested that a poor man having found a lottery ticket with a 50% chance of winning 20 000 ducats and a 50% chance of winning nothing might be well advised to sell that ticket for 9 000 ducats, despite its expected value being 1000 ducats greater than this. In the same situation, a rich man would be better advised not to sell the ticket for any less than its 10 000 ducats value.

Instead of money having absolute value, Bernoulli argued that it has diminishing marginal utility. This means that each additional dollar is appreciated less as the person becomes richer and utility is a subjective measure of value. In the risky choice example of Kahneman and Tversky, the expected utility of the risky $4000 option becomes less than the expected utility of the certain $3000 option, making the choice rational. However, when considering losses, it is the loss of each additional dollar that exhibits diminishing utility, rather than the gain. This makes it rational for the same person to prefer the risky gamble of 80% chance of losing $4000 to a certain loss of $3000.

Kahneman and Tversky also found that most people display loss aversion in risky choice. For example, most people are unwilling to accept a bet where they have an equal chance of losing as winning a given amount and this aversion is greater the higher the stake. People tend to value the loss of the set amount more than its gain. They proposed that the utility function for losses exhibits a steeper gradient than that for gains. Taken together with the diminishing sensitivity of utility, this gives rise to the "S-shaped" utility function, rather than a linear value function, as illustrated below.

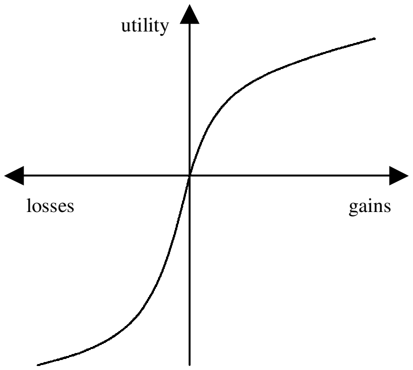


Figure 3.1. A typical utility function.

A model based on expected utility could therefore replace the expected value model to still provide a rational description of choice behaviour under conditions of risk, where the expected utility of gamble *X* is described by the function:

$$\text{EU}\ (X) = \sum_{i=1}^{n} P_i \times U(X_i)$$

However, some risky behaviour is still found to deviate from this model. Many people buy lottery tickets despite their expected value being less than the ticket cost and their expected utility being less still. In their proposed Prospect Theory, Kahneman and Tversky (1979) explained this behaviour by introducing a weighting function (Figure 2) to represent subjects' attitudes to probabilities. Probabilities close to zero are weighted higher, to reflect the risk-seeking behaviour demonstrated when probabilities are low and payoff is high, in situations such as the purchase of lottery tickets.

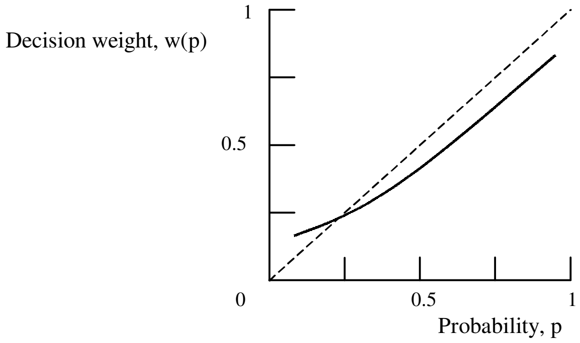


Figure 3.2. A typical weighting function from Prospect Theory.

Prospect theory is therefore an expected utility model that incorporates a weighting function according to subjects' attitude to probabilities.

$$\text{PT}\ (X) = \sum_{i=1}^{n} w(p_i) \times U(X_i)$$

Again, there are aspects of decision behaviour that are not captured by Prospect Theory. Fennema and Wakker (1997) provide an example that illustrates one of its failings. If a gamble has many possible outcomes, all with a low probability, the typical weighting function will cause all of these outcomes to be overweighted. For example, if there are 20 equally likely outcomes of losing \$10 through to winning \$180 in \$10 increments, the probability of each outcome is 0.05. If the weighting function is such that this probability is overweighted, 18 gains will be overweighted and only one loss. Subjects will overestimate the prospect of this gamble, although in reality they tend to prefer its expected value of \$85 for sure.

Tversky and Kahneman (1992) later modified their model to Cumulative Prospect Theory, in which the weighting function is applied to cumulative probabilities. The shape of the function (see figure 3), together with its application to cumulative probabilities ensures that only the extreme values are overweighted. Cumulative Prospect Theory also differs from the original in its use of two different weighting functions, reflecting differing attitudes to probabilities of loss than of gain. The two weighting functions are typically similar in shape, although that for losses is often less curved than that for the gain function.

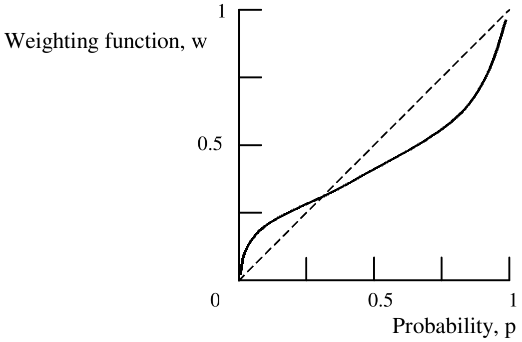


Figure 3.3. A typical weighting function from Cumulative Prospect Theory.

Cumulative Prospect Theory proposes that the expected utility of a gamble is the total of two summations, one for the loss outcomes and one for the gain outcomes. If outcomes *i* = 1…*k* represent losses and outcomes *i* = *k*+1…*n* represent gains, the expected value of taking up the gamble is given by:

$$\text{CPT }(X) = \sum_{i=1}^{k} w_i^- U(x_i) + \sum_{i=k+1}^{n} w_i^+ U(x_i)$$

The decision weights, w_i, are given by:

$$w_1^- = w^-(p_1), \quad w_i^- = w^-(p_1 + ... + p_i) - w^-(p_1 + ... + p_{i-1}) \quad 2 \le i \le k$$

$$w_n^+ = w^+(p_n), \quad w_i^+ = w^+(p_i + ... + p_n) - w^+(p_{i+1} + ... + p_n) \quad k+1 \le i \le n-1$$

Fennema and Wakker demonstrated that Cumulative Prospect Theory and the original Prospect Theory make different predictions in certain conditions. They found that Cumulative Prospect Theory better modeled the actual decision behavior of subjects in two experiments originally conducted by Lopes (1993). Cumulative Prospect Theory has become a universally recognized and widely used model of rational decision-making.

Nonetheless, not all human decision-making follows such a rational model. There are instances when people respond differently to the same decision at different times. The WADD model and expected utility models are deterministic, directing that the single decision that optimises the outcome should always be chosen. Probabilistic versions of these models, such as the multinomial logit model (McFadden, 1981), have been proposed to account for the variance of decision choices in seemingly identical conditions. McFadden's model calculates the probability of choosing alternative *X,* from a set of *m* possible options, as a logistic function of the weighted sum of the *n* attributes that *X* possesses together with an attractiveness value for that particular option.

$$P(X;\{X_1...X_m\} = \frac{e^{V(X)}}{\sum_{j=1}^{m} e^{V(X_j)}}, \quad \text{where} \quad V(X_j) = b_j + \sum_{i=1}^{n} b_i X_i$$

All of these rational models have been shown to be consistent with some choice behaviour, although they all exhibit discrepancies, where decision-making behaviour deviates from their predictions. In attempting to account for the observed idiosyncrasies in actual behaviour, the models have become increasingly complicated. Expected value has been modified to expected utility, a subjective weighting has been introduced to probabilities, which in turn have been made cumulative, and logistic functions have been used to transform deterministic models to probabilistic ones. The resulting models can be used to predict decision behaviour reasonably successfully but it is doubtful that humans are actually using such complex strategies themselves. Rather than relying on complicated calculations to make their decisions, decision behaviour can, instead, be explained through a consideration of some simple heuristics.

3.3 Simplifying Heuristics: Describing Actual Behaviour

Simon's studies of decision-making (1955) led him to propose that, in practice, people do not apply these maximising expected utility models to their day-to-day decisions. He suggested that such models have far too great an information-processing demand and that cognitive limitations lead to simpler heuristics being applied to many decision-making tasks. Simon argued that evidence of such a trade-off between decision-accuracy and cognitive-effort would be greater when tasks become more complex or decisions are made under stress. Under such conditions, non-optimal decisions are more likely to be made and can be predicted through the assumption of certain heuristics being applied. These heuristics are typical of the self-reports that subjects make on the strategies that they are conscious of utilising in decision tasks.

3.3.1 Decision-making heuristics

In his introduction to the theory of bounded rationality, Simon (1955) suggested the heuristic of satisficing. He proposed that people might sometimes choose simply the first option that meets all minimum requirements. As with all heuristics, this strategy would more likely be used for complex decisions, such as those with many alternatives or a high number of attributes to consider. It would also be more likely to be used when decision-makers were under time pressure. In these cases, information-processing demands are high and people might trade-off decision accuracy for the amount of cognitive effort involved.

Tversky (1972) proposed the elimination by aspects heuristic as another means by which people may simplify decision-making. This strategy involves a step-by-step consideration of the different attributes shared by the decision options. Starting with the attribute that is considered to be most important to the decision at hand, all alternatives that do not meet a minimum requirement are eliminated from the option bank. This process continues through the attributes, with all options below a set cut-off value being removed, until a single option remains and this is taken to be the decision choice.

Russo and Dosher (1983) put forward the majority of confirming dimensions heuristic as a strategy to choose between multi-attribute alternatives. They suggested that alternatives are considered two at a time. The values of each attribute are compared across the two options and the alternative with the greater number of favourable attributes is retained. This option is then paired with the next alternative and these pairwise comparisons continue until the most favourable alternative remains. This heuristic is quite different to the elimination by aspects strategy, since it compares all attributes across paired alternatives. It limits its focus to just two options and considers all of the attributes for these two alternatives. The elimination by aspects strategy considers all alternatives but focuses on just one attribute at a time.

Rather than limiting comparisons to a small number of alternatives or consideration of one attribute at a time, the decision-making process can be simplified using the equal weight heuristic. This proposes that all attribute values are simply summed for each alternative, with no attention given to their relative importance or, for outcomes, their probability of occurrence. Dawes (1979) demonstrated that this heuristic is often highly accurate in modelling the decision-making process, although it does assume that all attributes can be evaluated along a common value scale.

Perhaps the most simple of the decision-making heuristics is the lexicographic rule. This proposes that the decision-maker chooses the most important attribute to the task at hand and the alternative with the highest value along this attribute is selected. If two or more alternatives rate similarly highly along that attribute, a comparison between these options will then be performed for the next most important attribute. Fishburn (1991) demonstrated that subjects sometimes exhibit intransitivity among choices, which can be explained through use of this heuristic. For example, if option A is preferred over B and B is preferred over C, in particular circumstances option C may be selected over A. Such intransitivity may be explained if A, B and C are similarly spaced along one attribute but quite differently valued along a less important attribute.

For example, in choosing between computers of similar capabilities, cost may be considered the most important attribute and the cheaper option selected if prices differ by at least $200. If prices are within $200, the warranty may be the next most important attribute to decide between alternatives. Considering the following options:

	Cost	*Warranty*
Option A	$2640	5 years
Option B	$2480	3 years
Option C	$2325	1 year

Option A would be selected over B and option B over C, since their costs are within $200 and warranty becomes the deciding factor. However, option C would be selected over A since it is more than $200 cheaper.

Normative decision theory claims that transitivity among preferences is essential for rational behaviour. It considers binary preferences to construct an order between alternatives, such that if A is preferred to B, and B preferred to C it follows that A will be preferred to C. Examples of intransitivity, such as those provided by Fishburn (1991) are assumed to be mistakes that decision-makers would correct if the intransitivity were made apparent to them. However, the rationale behind decisions like the computer preferences above seem reasonable enough. It would appear to be harsh to label the behaviour of a decision-maker who maintained these preferences, even after the intransitivity were pointed out, as irrational. Fishburn argues that intransitivity is not an essential assumption of rational decision behaviour. He proposes that binary preferences between two alternatives may exhibit intransitivity and do not necessarily tell us anything about how the decision-maker would choose between three alternatives.

Many examples of deviations from normative decision behaviour can be explained in terms of the use of simplifying heuristics. Such strategies may reduce the cognitive demands on decision-makers, although sometimes at the price of making an incorrect decision. Many researchers have argued that the use of particular heuristics may be influenced by task factors, such as presentation mode. It is essential that these effects be fully understood if we are to be aware of how task conditions may influence decision-making behaviour.

3.3.2 Effects of task factors on decision-making behaviour

If the rationale behind decision-making were solely to maximize expected utility, it would be expected that the same decision would invariably be made whenever the same options were presented. However, there have been numerous studies of examples where subjects show preference reversals when factors that should have no effect on the decision are changed. These preference reversals are not random but tend to be exhibited in a systematic, predictable manner. Such factors include the mode by which the subject is to respond; the wording of the decision problem; the response scale presented; the presentation mode; the form of attribute representation and the inclusion of irrelevant alternatives.

3.3.2.1 Response mode effects

Experimental decision tasks generally require one of two response modes: choice or judgment. Choice tasks require that the subject choose between alternatives, whereas judgment tasks require that the subject assess the value of an option. Judgment decisions are often given in the form of a matching task, where subjects are asked to give a value for one alternative that would make it comparable to a given option. The following example is adapted from Payne *et al* (1998) to illustrate the two response modes with decision tasks regarding road safety programs.

Choice task:

Which program do you most prefer?

	Fatalities per year	Annual cost
Program A	570	$12 million
Program B	500	$55 million

Matching task:

Fill in a value for the cost of program B such that it would equate the overall value of the two programs.

	Fatalities per year	Annual cost
Program A	570	$12 million
Program B	500	

Given these particular values, most people prefer Program B to Program A in the choice task but give a value of significantly less than $55 million in the matching task. Their decision in the choice task values the saving of 70 lives as greater than $43 million. Conversely, in the matching task, the saving of 70 lives is valued at less than $43 million.

Tversky *et al* (1988) refer to this effect of response mode on decision-making as the prominence effect. They argue that the differing decisions can be explained on the basis of different heuristics being employed. Since the choice task requires a qualitative response, the lexicographic heuristic is most appropriate. This avoids conflict by selecting the option that is superior on the most salient attribute, in this case the number of fatalities. The matching task requires a quantitative response and a consideration of the values along both attributes, together with their relative weights. The greater level of information processing required for such a decision can therefore result in different decisions to those from the choice task.

A second example of response mode effect is the classic preference reversal phenomenon (Lichtenstein & Slovic, 1971). They offered subjects the choice between two bets, one offering a high probability of winning a small amount, the other offering a low probability of winning a large amount. Lichtenstein and Slovic found that when choosing between two bets, most people preferred the one with the high probability. However, when assigning a value to each of the two bets, most people offered a higher amount of money for the low probability bet, thus exhibiting preference reversal when the response mode was changed. The low probability bet tends to be over-valued in these cases, causing Tversky *et al* (1990) to argue for the concept of scale compatibility. They propose that people may be attempting to reconcile their evaluation of the low probability wager with the size of the potential win.

3.3.2.2 Presentation mode effects

The dramatic effects that the wording of a problem can have on decision-making were demonstrated by Tversky and Kahneman (1981). They presented subjects with the same choice task, framed in two different ways. The choice was between two medical programs that were to be used to combat a new disease expected to kill 600 people. When the decision task was posed in terms of lives saved, A versus B, 72% of subjects preferred the former.

Program A 200 people will be saved
Program B 1/3 chance that 600 people will be saved and 2/3 chance that no one will be saved.

However, when the same decision was posed in terms of lives lost, C versus D, only 22% preferred the former.

Program C 400 people will die
Program D 1/3 chance that no one will die and 2/3 chance that 600 people will die.

Tversky and Kahneman found that such preference reversals were characteristic of undergraduate students, university lecturers and practicing physicians. Several researchers have found similar framing effects, these demonstrations all having the crucial distinction between framing that codes outcomes as losses and that which codes outcomes as gains. The concept of loss aversion (Tversky & Kahneman, 1991) explains these framing effects through its premise that people treat negative and positive consequences differently. The greater impact of a given loss than a corresponding gain causes people to assign different values to losses than to gains.

A similar result that supports the finding that even experts are prone to effects similar to framing was more recently demonstrated by Slovic *et al* (2000). They presented psychiatrists and forensic psychologists with case studies on people with psychological disorders who had been convicted of committing violent crimes. The psychiatrists and psychologists were asked to judge the probability that these people would re-offend within six months of release. Two groups of respondents were given identical case studies but different scales, as below, upon which to mark these judged probabilities:

Low probability	1%	2%	5%	10%	15%	20%	25%	30%	35%	40%	>40%
High probability	1%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%

Slovic *et al* (2000) found that that the group responding on the low probability scale consistently judged the probability of re-offence to be lower than did the group responding on the high probability scale. This was despite the fact that the two groups were experts who frequently made such judgments in their work. Their decisions were significantly influenced by the scale upon which they were to make their response.

In addition to these framing and response scale effects, presentation mode can influence decision-making. Slovic (1972) proposed that decision-makers use information in the same mode as that in which it is presented. Not transforming the displayed information into an alternative mode saves the decision-maker cognitive effort. This 'concreteness' principle explains many of the research findings on presentation effects on decision behaviour. Russo (1977) demonstrated that shoppers are more likely to use unit price information in deciding which brand to buy when this information is displayed in a list. He argued that it is difficult to compare brands in the typical shelf displays of supermarkets.

Jarvenpaa (1989) demonstrated that the display format affects the way that information is processed. Information about the attribute values of the various options can be displayed either by attribute or by alternative and this display format will influence the strategy used to decide between the options. The salience of a particular attribute within the display is also important. For example, cues that are highly prominent but less important may overly-influence decisions, as was demonstrated by MacGregor and Slovic (1986).

The way in which information is represented can also affect decision behaviour. Stone and Schkade (1991) proposed that the heuristic employed to choose between alternatives might depend upon the form of representation of attribute information. They compared numeric and worded representation of computer system attributes in a decision choice task. Since the worded presentation made comparisons within attributes more difficult, in these presentations subjects were found to use qualitative alternative-based information search. However, when the same decision tasks were presented but the attribute information given numerically, direct quantitative attribute comparisons formed the basis of decisions.

Other researchers have found similar effects dependent upon the form by which information is represented. Johnson *et al* (1988) demonstrated preference reversals when probability information was represented by vulgar fractions rather than decimals. They concluded that subjects found the vulgar fractions difficult to directly compare and resorted to alternative decision-making strategies. Erev and Cohen (1990) have demonstrated that subjects prefer to receive numerical probability information, although they tend to convey this probability information verbally. They found that, although people believe that they make decisions more accurately from numerical probabilities, their performance was not significantly improved from that given verbal probability information. Schkade and Kleinmuntz (1994) investigated the effects of display, form of representation and presentation sequence on decision behaviour in choosing between loans. They found that the mode of display, either by matrix (alternative) or by list (attribute) had a strong influence on information acquisition. The form of information representation affected the combination and evaluation of information and the sequence of presentation had only a weak influence on acquisition.

3.3.2.3 Context effects

The inclusion of some alternatives in a choice task may have effects contrary to those expected from rational models of decision-making. The normative model of rational decision behaviour predicts that the inclusion of an additional alternative to a choice task cannot increase the probability of selecting one of the original alternatives. However, in the case of asymmetric dominance, exactly this does occur. An option is said to be asymmetrically dominated if it is dominated by at least one alternative in the choice set but not dominated by at least one other. For example, option A may be greatly preferred to option B, whereas there may be little difference in preference between options B and C. Option B is said to be asymmetrically dominated by option A. If option B is added to the choice set of options A and C, the probability of selecting option A actually increases. This effect is observed despite there now being an additional alternative to A, option B as well as option C, so if anything the probability of selecting A would be expected to decrease.

3.4 Decision-making Biases

A considerable amount of behavioural decision research has been performed under conditions of risk, where subjects evaluate gambles according to their payoff and probability of occurrence. These experimental conditions, where the probability of a future outcome is given are not true to many real life decision tasks. In reality, people have to make decisions under uncertainty, in which the probability of future outcomes is not known but has to be estimated. In making these subjective estimates, many people exhibit decision-making biases that appear to be a result of their misperceptions of probability. These cognitive failings have considerable effect on decision-making behaviour and lead to predictable deviations from the expectations of normative decision theory. Tversky and Kahneman (1982a) have identified several such biases, which they attribute to the use of heuristics in estimating the likelihood of an event occurring. Three of these heuristics and the types of bias to which they give rise are discussed below.

3.4.1 Availability

The availability heuristic proposes that the frequency of a given event is estimated using the ease with which examples come to mind. It would be expected that such availability would be related to the class size of the event and therefore a reasonable measure of frequency. However, other factors will also affect availability, giving rise to a bias in the estimated frequency when this heuristic is employed. Tversky and Kahneman (1973) offer several examples of an availability bias. In one experiment, they asked subjects to estimate how many committees of either two or eight members could be made from a group of 10 people. The median estimate for the number of two-person committees was 70, whereas that for eight-member committees was 20 (Tversky & Kahneman, 1973). The correct number is the same for both sized committees, ${}^{10}C_2 = 45$. They proposed that it is far easier to imagine groups of two members than groups of eight members, hence the vastly different estimates.

Slovic *et al* (1982) demonstrated an example of the availability bias that can be readily applied to real life decision-making tasks. They asked people to estimate the probability of death from various causes. Estimates of highly publicized causes of death, such as tornado and flood, were over-estimated, whereas those of less “glamorous” causes, such as stomach cancer and diabetes, were under-estimated. The ease with which examples of a class come to mind is related to the class size but is also open to the influence of factors such as recency and salience, hence the bias that arises from using the availability heuristic.

3.4.2 Representativeness

Sometimes a decision-maker has to judge the likelihood that an object belongs to a certain class or will generate a certain event. In such decisions, Tversky and Kahneman (1982a) suggest that the decision-maker will employ the representativeness heuristic. This proposes that the judged likelihood will depend upon the level of similarity between the object and the proposed member class. Although this sounds reasonable, it fails to take into account the base rate probabilities of the possible member classes and leads to a bias known as the base rate fallacy. This bias produces a systematic deviation from the normative conditional probability rule of Bayes' theorem:

$$p(A|B) = \frac{p(B|A)p(A)}{p(B)}$$

Since the similarity between A and B is symmetrical, the representativeness heuristic assumes that $p(A|B) = p(B|A)$, which only occurs if $p(A) = p(B)$.

Examples of the base rate fallacy were provided by Kahneman and Tversky (1973). They presented two groups of subjects with short personality profiles of professional people and asked them to judge whether each profile was more likely to belong to an engineer or a lawyer. One group were told that the sample profiles had been randomly drawn from a collection of 70 engineers and 30 lawyers. A second group were told that they had been drawn from a sample of 30 engineers and 70 lawyers. However, this base rate information did not affect the probability judgments; both groups made essentially the same assessments. When no profiles were provided, or when the profiles were no more similar to one stereotype than the other, the two groups did use base rate information and made different probability judgments of approximately 0.7 and 0.3.

The base rate fallacy is so strong that it often leads to decision-makers judging the probability of two events both occurring as being greater than the probability of the occurrence of just one of those events. The following example of this conjunction fallacy is from a similar experiment to the personality profiles described above and is taken from Tversky and Kahneman (1982b, p.92).

> Bill is 34 years old. He is intelligent, but unimaginative, compulsive and generally lifeless. In school, he was strong in mathematics but weak in social studies and humanities.

Please rank in order the following statements by their probability, using 1 for the most probable and 8 for the least probable:

A Bill is a physician who plays poker for a hobby.
B Bill is an architect.
C Bill is an accountant.
D Bill plays jazz for a hobby.
E Bill surfs for a hobby.
F Bill is a reporter.
G Bill is an accountant who plays jazz for a hobby.
H Bill climbs mountains for a hobby.

Tversky and Kahneman found that 87% of subjects rated the probability of G as being greater than D. This means that they rated the likelihood of Bill being an accountant who plays jazz as greater than the chance that he plays jazz (regardless of his profession). When repeated on students who had taken several advanced courses in probability and statistics, they found that 80% of these subjects still made the conjunction fallacy.

Eddy (1982) reports how prevalent the base rate fallacy can be, even amongst experienced professionals. He discusses the example of breast cancer and mammography, for which many physicians do not adequately consider the low base rate of malignancy. The following example illustrates the sort of information that was provided to doctors.

In patients complaining of painful lumps in breast tissue, the frequency of malignant cancer is one in 100. The accuracy of mammography is 90%, meaning that the probability of a positive x-ray given a malignant cancer is 0.9 and the probability of a negative x-ray given no cancer is 0.9.

The physicians were asked to judge the probability that a woman complaining of painful breast tissue and having a positive mammogram result did actually have a malignant cancer. This can be calculated using Bayes' theorem:

$$p(C|+) = \frac{p(+|C)p(C)}{p(+|C)p(C) + p(+|C')p(C')} = \frac{(0.9)(0.01)}{(0.9)(0.01) + (0.1)(0.99)} = 0.083$$

The probability of the woman having cancer given a positive x-ray is approximately 8%. However, the vast majority of the physicians informally interviewed estimated the probability as being around 75%. The physicians had made an estimate that was too close to that of the probability of a positive x-ray given that there was a malignant cancer, assuming that $p(C|+)$ was similar in value to $p(+|C)$. They had not sufficiently accounted for the effect of the low base rate of malignant cancer in their estimates.

3.4.3 Anchoring and adjustment

In certain instances when asked to estimate a probability, people may start from an initial value and make adjustments. The initial value may arise as a part computation of the required value or it may be suggested by the problem. However, this strategy of anchoring and adjustment is prone to mistakes as the adjustment from the initial anchor value is typically insufficient. The anchoring and adjustment heuristic therefore gives rise to a bias towards the initial anchor value.

An illustration of the use of part computation as an anchor value was provided by Tversky and Kahneman (1973). They asked high school students to mentally evaluate the product of the numbers from 1 to 8. For one group of students the product was presented in ascending order, $1\times2\times3\times4\times5\times6\times7\times8$, whereas for a second group it was presented in descending order, $8\times7\times6\times5\times4\times3\times2\times1$. The two groups had to make their estimates within five seconds. In such a short timeframe, Tversky and Kahneman proposed that the students would evaluate only the first three or four steps of the calculation. Their estimates would then be extrapolated from these initial anchors. They therefore predicted that both groups would under-estimate the true value of 40320, this under-estimation being greater for the ascending group who would have a lower anchor value. Both of these predictions were realised. Both groups under-estimated the true value and the ascending group tended to make estimates that were considerably lower than those of the descending group, the median value being 512 as opposed to 2215.

Tversky and Kahneman (1982a) report of an example where the initial anchor values were randomly generated but still produced a bias in the estimated values. Subjects were asked questions such as the percentage of the member countries of the United Nations that are

African. Before replying, a wheel with the numbers from 0 to 100 was spun. The subjects had to answer firstly whether the percentage of African nations was higher or lower than this random value and then to make an estimate of the percentage. The estimates varied greatly, dependent upon the number that was randomly generated and regardless of payoffs for accuracy. For example, the median answer to the above question was 25 when the number 10 was spun but 45 when the number 65 was spun.

Some lapses from the normative view of rational decision-making behaviour may be examples of task factors influencing the decision-maker. Others may be a result of biases resulting from the use of heuristics that misperceive probability. Normative theorists assume that any such deviations are unwitting and will be corrected when they are brought to the attention of a rational decision-maker. The examples above demonstrate the extent of such 'irrationality' in actual decision behaviour. Although many of these decisions are undoubtedly unwitting mistakes, some may be deliberate choices by decision-makers who would not consider themselves to be acting irrationally. The following section considers decisions made under conditions of ambiguity and offers examples of decision-making behaviour that are difficult to reconcile with the normative view of rationality.

3.5 Decision-making and Ambiguity

Much behavioural decision research has been performed under conditions of risk, where gambles are assessed according to their payoff and probability of occurrence, or conditions of uncertainty, where the likelihood of future outcomes has to be subjectively estimated. Ellsberg (1961) introduced an additional experimental condition under which to perform decision-making research, that of ambiguity. This can be considered as the uncertainty associated with an outcome's probability, due to incomplete or vague knowledge of the task conditions or contradictory information.

It can be argued that many decisions made in real life are decisions made under the condition of ambiguity, since the facts informing decisions are typically incomplete, vague or contradictory. For example, in tossing a standard coin, it is accepted that the probability of it landing on Heads is one half, since we know that there are two equally likely outcomes: Heads or Tails. However, if we are told that the names of all of the students in a particular class are to be put into a hat and one drawn at random, the probability that the name drawn belongs to a male student is ambiguous. Our knowledge of the facts is incomplete, as we do not know the proportion of males in the class. This ambiguity might be lessened slightly if we were told, perhaps, that it was an engineering class, although the probability would remain ambiguous until we were told the group's exact composition. From the initial problem, the probability that the drawn name belongs to a male student is best described as one half, but there is uncertainty attached to this probability.

Ellsberg (1961) demonstrated that many people appear to demonstrate aversion to such ambiguity. He described to subjects two urns, each filled with one hundred balls. In the first urn, U_1, fifty of these balls were red and the other fifty were black. In the second urn, U_2, it was known that the hundred balls were made up of red and/or black balls only, but the numbers of each colour were unknown. He offered subjects a hypothetical $100 prize if the outcome were to match their bet and presented the following four questions:

> Would you prefer to bet on a red or a black ball being drawn from U_1?
> Would you prefer to bet on a red or a black ball being drawn from U_2?
> Would you prefer to bet on a red ball being drawn from U_1 or U_2?
> Would you prefer to bet on a black ball being drawn from U_1 or U_2?

Not surprisingly, Ellsberg found that most people were indifferent to the choice of red or black in the first two questions. However, the majority of people preferred to bet on U_1 rather than U_2 in questions 3 and 4. This is inconsistent with the use of the normative principle of maximizing expected utility. A preference for a red ball from U_1 over a red ball from U_2 would imply that the subjective rating of the probability of drawing a red ball from U_1 is greater than that from U_2, since the prize is the same in each condition. However, the concurrent preference for a black ball from U_1 over a black ball from U_2 implies that the probability of drawing a not-red ball from U_1 is also considered greater than that from U_2. Clearly, in terms of maximizing expected utility, these preferences are self-contradictory.

Ellsberg found that most of his subjects maintained their preferences even after this contradiction was made explicit. Normative theorists claim that any violations of the maximizing expected utility principle are unwitting mistakes and will be corrected by decision-makers who are given time to reflect on their decisions. However, the finding that many ‘reasonable’ people persisted in this violation caused Ellsberg to question the normative definition of rational behaviour.

In a further experiment, Ellsberg presented his subjects with choice questions based on an urn filled with balls of three colours. The urn was known to contain 30 red balls, with the remaining 60 balls being in an unknown ratio of black to yellow. Ellsberg asked his subjects whether they would prefer to bet:

(I) on a red ball
or (II) on a black ball

being drawn from the urn, with again a hypothetical $100 prize at stake. Most people preferred to bet on the red ball, which had a known probability of 1/3, than the black ball, whose probability was ambiguous but best considered as 1/3.

Ellsberg then asked his subjects which bet they would prefer if the prize were to be offered:

(III) on drawing either a red or a yellow ball
or (IV) on drawing either a yellow or a black ball.

Most people preferred the latter bet, with a known probability of 2/3, than the former, with its ambiguous probability best described as 2/3.

The choice questions are summarized below, adapted from Ellsberg (2001, p.137).

	30	**60**	
	Red	**Black**	**Yellow**
I	$100	0	0
II	0	$100	0
III	$100	0	$100
IV	0	$100	$100

The preference pattern for I over II and for IV over III is at odds with normative theory. Specifically it violates Postulate 2 of Savage’s theory, known as the ‘Sure-thing Principle’ (Savage, 1954). The ‘Yellow’ column remains constant over conditions I and II and also over conditions III and IV. It should therefore not influence preferences between I and II, neither should it influence preferences between III and IV. However, if this third column can effectively be ignored, it would be expected that anyone selecting I over II would also have to select III over IV. Even when this was pointed out to Ellsberg's subjects, the majority persisted in their contradictory preferences for I over II, but for IV over III.

This behaviour pattern is also inconsistent with the principle of maximizing expected utility. There is no ratio of black to yellow balls that would simultaneously support the preference for I over II and for IV over III. The proportion of yellow balls in the urn could vary between 0 and 2/3, with the corresponding expected utility for each preference varying as illustrated below: (Adapted from Ellsberg, 2001, p.166)

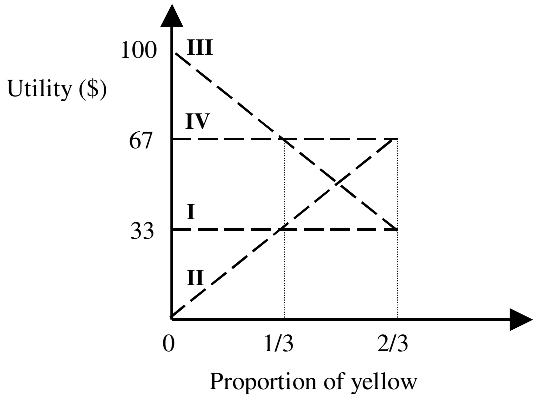


Figure 3.4. Graph showing expected utility for each of the bets I to IV, as the proportion of yellow balls in the urn varies.

Ellsberg argued that the tendency for ambiguity aversion in decisions is deliberate and in violation of Savage's axioms. He proposed that the normative model needed revision to account for this paradox and suggested a decision rule that included consideration of the minimum expected utility of each alternative, in addition to the actual expected utility. The actual expected utilities of both decision choices in the three-colour urn example are equal. However, the minimum utility of II is less than that of I, as is the minimum utility of III compared to that of IV. This decision rule would therefore support the preference of I over II and IV over III. Ellsberg's decision rule uses p to represent the degree of confidence in the probability function, est_x to represent the actual expected utility and $\min_x$ to represent the minimum expected utility. He then proposes that the expression below is evaluated for each alternative, x, and the option with the highest value is selected:

$$p \times \text{est}_x + (1 - p) \times \min_x$$

This is a conservative rule, since the minimum expected utility receives increased salience, being used in the calculation of the actual estimated utility in addition to its separate representation in the expression. However, most people do tend to be conservative in their decision-making behaviour, as demonstrated by Tversky and Kahneman (1991) in their loss aversion experiments. Ellsberg suggested that for some people, it might be more appropriate to replace the minimum expected utility term with maximum expected utility. These people would be ambiguity prone, showing the reverse preference pattern for II over I and for III over IV. Such preferences were exhibited by a minority of respondents.

Ellsberg (2001) also provided an example where most people prefer an ambiguous option to a more definite one. He asked subjects to imagine two urns, each filled with 100 coloured balls. The first urn, U_1, contained equal numbers of ten different colours. The second urn, U_2, contained no balls of colour different to these ten, although the exact composition was unknown. He asked subjects whether they would prefer to bet on a green ball being drawn from U_1 or U_2. The majority of people preferred the ambiguous second urn, although the expected utility of the two options is again equal. Ellsberg's explanation for this ambiguity prone behaviour is easiest understood with reference to the following diagram: (Adapted from Ellsberg, 2001, p. 203).

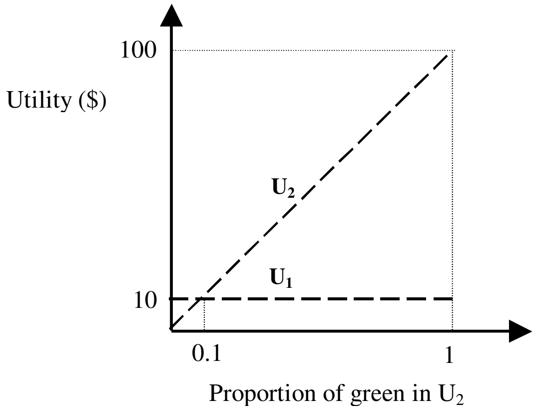


Figure 3.5. Graph showing expected utility for the two bets U_1 and U_2, as the proportion of green balls in the urn varies.

The expected utility of U_2 is ambiguous. However, it can be seen from figure 3.5 that the maximum value of this expected utility is considerably greater than the actual expected utility of $10, whereas the minimum value is only slightly less. Expected utilities that are higher than the actual expected value may be more available to the decision-maker. If the decision rule is modified to consider both the minimum and maximum expected utility values, as well as the actual expected utility, such preferences can be easily understood.

Ellsberg's experiments demonstrate that many people display ambiguity averse behaviour, tending to over-emphasise the minimum expected utility of an option for which the probability is ambiguous. Mukerji and Tallon extend the concept of ambiguity aversion to portfolio inertia (2004a) and wage indexation (2004b). Hogarth and Kunreuther (1985) develop Ellsberg's work into a subjective model of human decision-making behaviour that they apply to insurance decisions. They found that people were more likely to pay for insurance in low probability, high risk scenarios where the expressed probabilities were ambiguous. Einhorn and Hogarth (1986) develop this model to incorporate the addition of new information over time and it has been used to successfully model decision-making behaviour in other contexts, such as the recruitment of potential employees (Highhouse & Hause, 1995).

Ellsberg's decision rule accounts for his findings that most people exhibit ambiguity aversion except at low probabilities. However, the rule is similar in complexity to the modified expected utility models of normative theorists. Ellsberg was also unclear as to whether the decision rule was to be prescriptive of 'rational' behaviour or was simply to model actual decision behaviour. Is it rational to be conservative in decision-making? The normative theorists were willing to accept the concepts of subjective utility and loss aversion, should ambiguity aversion be similarly assimilated into a theory of what constitutes rational decision-making? Alternatively, both loss aversion and ambiguity aversion may be considered as human biases, of which decision-makers need to be aware if their decisions are to be truly rational.

It is generally acknowledged that uncertainty should be included in spatial information (for example, Zhang & Goodchild, 2002). However, representation of the uncertainty inherent to spatial data is tantamount to emphasizing the ambiguity of the information. It would therefore be expected that the effects of ambiguity aversion would be prevalent in spatial decision-making when uncertainty information is included. In an experimental study, Van Dijk and Zeelenberg (2003) found that participants tended to discount ambiguous information in making decisions about costs and benefits. If uncertainty information is included in spatial data, a similar bias to avoid regions for which the information is labeled uncertain may be apparent. Rather than informing the decision-making process, the representation of uncertainty information may simply lead to biased decisions.

3.6 Chapter Summary

In attempting to describe human decision-making behaviour, models have been developed that are based on the strategy of maximising expected utility. However, these have introduced subjective measures of both probability and value, making them difficult to disprove and less applicable to predicting behaviour without prior knowledge of the individual decision-maker. In addition, as they have been refined to accommodate some of the idiosyncrasies in actual behaviour, they have evolved into complicated mathematical functions. Particularly as decisions become more complex, it would appear to be highly unlikely that people are actually utilising such strategies in practice.

The use of simpler heuristics has been proposed, particularly as decisions become more complex. These do not place such high information-processing demands on decision-makers and predict actual behaviour with a similar level of accuracy. They are also more similar to subjective self-reports of conscious decision-making strategies. Some of the preference reversals that are inconsistent with normative strategies can be systematically predicted in terms of the use of heuristics. This gives strong support for the theory that heuristics may be employed in practice and particularly when processing demands are high.

People do not always make the same decision when presented with the same information. They can be sensitive to the effects of presentation, task and context. Humans are also prone to decision-making biases, particularly in subjectively judging probabilities. It is assumed by normative theorists that any decisions that deviate from the normative ideal of maximising expected utility are unwitting mistakes and would be corrected if decision-makers were able to reflect on their decisions.

Ellsberg demonstrated that most people display ambiguity aversion, tending to prefer options for which the probabilities are known to those for which the probabilities are ambiguous. This behaviour can be modelled by giving greater emphasis to the minimum expected utility of an ambiguous option. This conservatism reflects human behaviour, even after reflection, although it does not necessarily lead to a decision that maximises expected utility. It can be argued that the concept of a rational decision needs to incorporate ambiguity aversion in addition to the already assimilated loss aversion. Otherwise, ambiguity aversion can be considered as another human decision-making bias.

Most real life decisions are made under conditions of vague or incomplete knowledge and can therefore be considered to be made under ambiguity. An explicit acknowledgement of the uncertainty in our information would be expected to accentuate the effects of ambiguity aversion. This would lead to the decision-maker tending to under-value options that are labelled as uncertain and not necessarily making a rational decision, as defined in terms of maximising expected utility. It is essential that we appreciate these effects of ambiguity aversion on decision-making behaviour if we are to represent the uncertainty in spatial data. In the next section, an experiment is conducted to test the effects on decision-making of representing the uncertainty associated with thematic information.

4.

Methodology

4.1 Introduction

Two case studies were designed to investigate decision-making under spatial uncertainty. The first of these tested the effects of introducing thematic uncertainty information on decision-making. The second study investigated how different representations of positional uncertainty may affect decision-making.

4.2 Thematic Uncertainty: Airport Siting Case Study

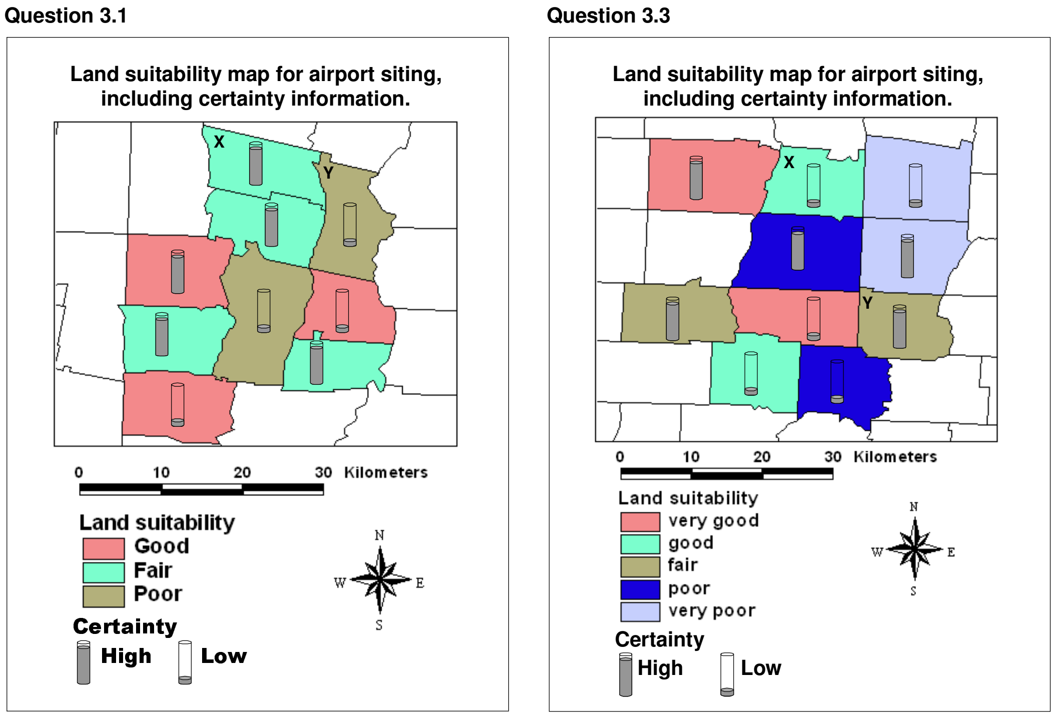
4.2.1 Background

GIS are being widely used to support decision-making. However, the uncertainty that is inherent to GIS output is not usually communicated to the decision-maker. This study aims to explore the effects of introducing spatial uncertainty information within a decision support context. It examines how decision-making may be affected by the introduction of thematic uncertainty into the GIS output. Specifically, it tests the hypothesis that the inclusion of thematic uncertainty information may lead to the making of irrational decisions as a result of decision-makers exhibiting ambiguity aversion.

4.2.2 Case study

Similar to the research by Leitner and Buttenfield (2000), this case study investigated decision-making in the context of siting a new airport. Subjects were presented with thematic maps depicting several zones that were classified in terms of their potential land suitability for the new airport. The land suitability classification was represented using block colouring. The level of certainty associated with this classification was represented using glyphs overlaid within each zone. A full cylinder represented high certainty and an empty cylinder represented low certainty.

In each of six questions, subjects selected which of two zones, marked X and Y, they would prefer for the new airport. They were also able to give a ‘no preference’ response. Both zones were of similar area. In some questions the zones were of equal land suitability and in others the land suitability classes differed. However, in all questions the classification of one of the two zones was high certainty whereas that of the other zone was low certainty. Examples of the maps presented for a simple (3 land suitability classes) and a complex (5 land suitability classes) question are shown below:


Question 3.1

Question 3.3

Select whether you would choose Zone X or Zone Y as the airport site.

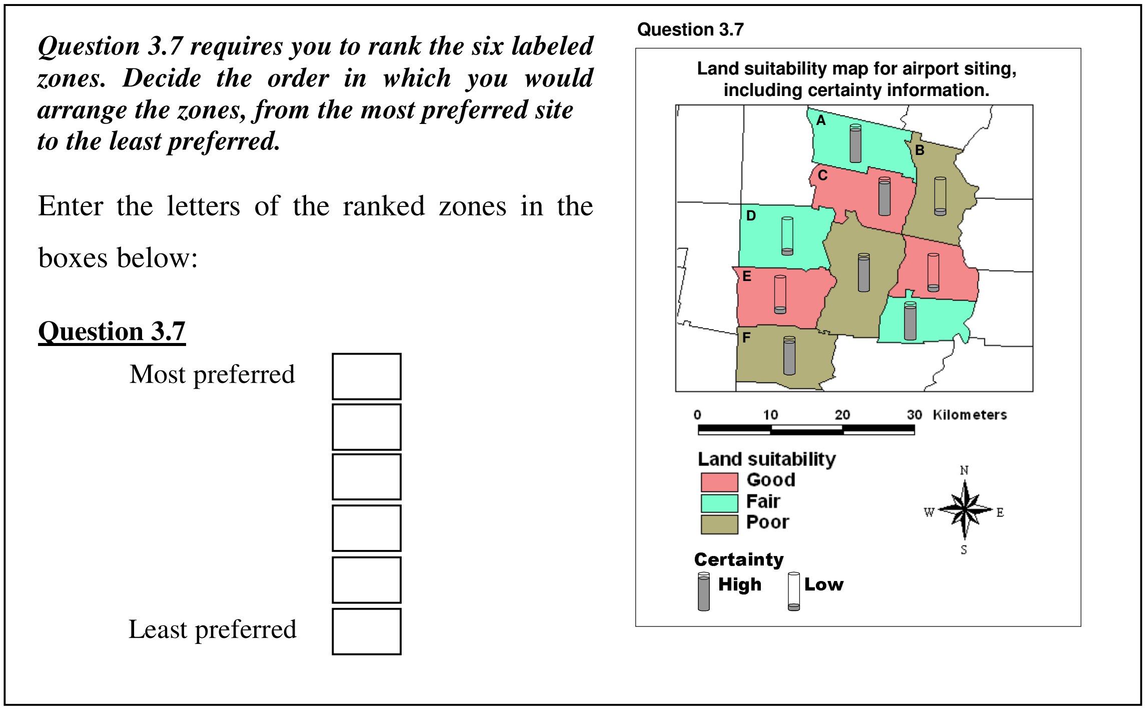
For each question, tick the box for the zone you would choose in the table below:

	Zone X	Zone Y	No preference
Question 3.1			
Question 3.2			

Figure 4.1. Example of a simple and complex pairwise comparison for testing of thematic uncertainty representation.

These pairwise comparisons required a level of decision-making that went beyond the decisions made in the Leitner and Buttenfield study. In their research, they only considered selection of the optimal zone and the worst zone for siting the airport. These simple decision tasks do not require much interpretation of the certainty information. Indeed, it may be that subjects responded faster when certainty information was included as they could simply narrow their search by only considering zones that were classified as certain. In this study however, subjects were selecting between zones that may not be optimal in terms of land suitability. It therefore required an interpretation of the certainty information in light of the level of land suitability. Sometimes the higher level of certainty was advantageous but sometimes the lower level of certainty was associated with the more preferable land.

Similarly, an interpretation of the certainty information with regards to the level of land suitability was required in a final question. In this, a simple map was presented, with six zones labelled A to F. These comprised one zone of each of the six suitability-by-certainty combinations. Subjects were asked to rank the six zones in terms of their preference as the potential site of the new airport. Again, it was ensured that all six zones were of similar area. An example of this ranking question is shown in figure 4.2:



Question 3.7 requires you to rank the six labeled zones. Decide the order in which you would arrange the zones, from the most preferred site to the least preferred.

Enter the letters of the ranked zones in the boxes below:

<u>Question 3.7</u>

Most preferred ☐
☐
☐
☐
☐
Least preferred ☐

Figure 4.2. Example of a ranking question for testing of thematic uncertainty representation.

4.3 Positional Uncertainty: Navigation Case Study

4.3.1 Background

In this next case study, an investigation was conducted into how different representations of positional uncertainty may impact on decision-making. This was examined within a marine navigation context. Positional uncertainty of maritime cadastre boundaries is a topical issue and is a subject of current research (Fraser *et al*, 2003). However, this research is primarily concerned with quantifying the positional uncertainty of nautical boundaries and has not really addressed the issue of how this uncertainty can most usefully be represented. The current study aims to investigate how decisions may vary according to the nature of the positional uncertainty representation and to assess both participants' understanding of, and subjective preferences for, the different representations. Specifically, it tests the hypothesis that the same people may make different decisions when presented with the same information, dependent upon the nature of the positional uncertainty representation.

4.3.2 Case study

The study was divided into two sections: testing of dynamic representations of positional uncertainty and testing of static representations. The former component investigated how the four different representations may have different effects on decision-making. The latter component investigated participants' understanding of the information portrayed and their subjective preferences for the four representations.

The dynamic testing stage required subjects to respond to animations depicting a hand-held GPS receiver that displayed a boat advancing from Zone A towards Zone B. Subjects were told that Zone B represented a restricted region and they were asked to make a decision as to when they would turn away from the restricted zone. Information concerning the positional uncertainty of both the boat and the inter-zone boundary location was included in a legend for each animation. Subjects were shown four animations, one for each of the four different methods used to depict the positional uncertainty.

The Limits representation used dotted lines to depict the outer limits of the likely position (99%) of both boat and boundary. The Scale and Probability representations both displayed the most likely boat and boundary positions. However, the Scale representation included a written statement in the legend as to the positional accuracy of the displayed boat and boundary positions, whereas the Probability representation provided an on-screen probability statement as to the likelihood that the boat was in Zone B. The Graduated representation portrayed the positional uncertainty of the boundary using graduated shading which became darker as the probability of being in Zone B increased. The outer limit of the boat's likely position was depicted with a dotted circle. Examples of these four representations are shown below:

Limits representation:

Legend:

Possible boat position, within a circle that contains the true boat position 99% of the time.

Possible boundary position between Zones A and B, within two outer lines that contain the true boundary position 99% of the time.

Slide 30

Figure 4.3. Example slide showing the Limits representation of positional uncertainty in the navigation case study.

Scale representation:

Legend:

● Boat position. Positional accuracy of +/- 10m (99% confidence limit).

B
- - -
A
Boundary position between Zones A and B. Positional accuracy of +/- 20m (99% confidence limit).

Slide 30

Figure 4.4. Example slide showing the Scale representation of positional uncertainty in the navigation case study.

Probability representation:

Legend:

● Possible boat position.

B
– – – Possible boundary position between Zones A and B.
A

Probability of boat being in Zone B = **36%**

Slide 30

Figure 4.5. Example slide showing the Probability representation of positional uncertainty in the navigation case study.

Graduated representation:


Legend:

Circle within which the true boat position should lie 99% of the time.

Probability of being in Zone B.

Slide 30

Figure 4.6. Example slide showing the Graduated representation of positional uncertainty in the navigation case study.

Towards the end of the testing phase, one participant pointed out a mistake in the Graduated representation. Only the transition zone between Zones A and B was shaded. The region above this, which was intended to depict ‘definitely in Zone B’, should have been coloured purple rather than white. This inadvertent mistake may have led to some bias in the results. However, only this one participant questioned the representation and stated that he understood what the display had been intended to portray. All other participants responded to this representation without question. Nonetheless, it would be worthwhile repeating the testing with the appropriate correction being made to the Graduated representation.

The static testing stage investigated subjects' comprehension of the information portrayed within these four representations of positional uncertainty. Five forms of multiple choice question were generated for each representation. These five forms were considered to be typical examples of displays representing the boat in each of five possible locations: Definitely in Zone A; Probably in Zone A; Equal chance of being in either zone; Probably in Zone B; Definitely in Zone B.

Participants were required to select which of these alternatives best described the boat's position for each of eight multiple choice questions. The option 'Cannot understand the diagram' was also available. The example below shows the multiple choice question for the Graduated representation, for which the expected response was 'Equal chance of being in either zone':

Figure 4.7. Example multiple choice question for the Graduated representation, for which the expected response was 'Equal chance of being in either zone'.

Participants were also asked to rank the four representation methods from most preferred to least preferred.

4.4 Experimental design

The testing of both thematic and positional uncertainty representations was conducted in a single experimental session, composed of three sections: testing of thematic uncertainty, dynamic testing of positional uncertainty and static testing of positional uncertainty.

4.4.1 Testing of thematic uncertainty

To test the effects of introducing the level of uncertainty associated with thematic information, participants were asked to select between two regions for the potential siting of a new airport. Twelve maps were prepared, using ArcView software. In six of these, the two zones had equal land suitability classifications and, in the other six, the land suitability classes differed. The ‘minerals’ colour ramp was used to block colour the zones according to their land suitability classification. In all twelve maps, one of the two zones had a high level of certainty associated with this classification, whereas the other zone had a low level of certainty. This was depicted using glyphs overlaid within each zone. A filled cylinder was used to represent high certainty and an empty cylinder depicted low certainty. No numerical value was assigned to the high and low certainties.

A second group of twelve questions was prepared by reversing the land suitability and certainty information assigned to the two zones, to account for any semantic or positional effects that may be associated with the zones labelled X and Y. A total pool of 24 questions was available. From these, two sets of six questions were generated, with a further two (1b and 2b) being question reverals of these. In each set, the first two questions were simple, in that the land suitability classification was along three levels: good, fair or poor. The other four questions were complex, having five levels of land suitability classification, from very good to very poor. The order of the two simple questions and four complex questions was also reversed between sets a and b, in an attempt to compensate for any order effects. Three questions of each set were comparisons between zones of equal land suitability and the remaining three questions compared zones of differing land suitability.

The four question sets are shown below, using 1-3 (3 being good) or 1-5 (5 being very good) to represent land suitability and H (high certainty) or L (low certainty) to represent the associated certainty level:

Set 1a: 2H1L 2L2H 4L3H 4L4H 2L4H 1H1L
Set 1b: 2H2L 1L2H 1L1H 4H2L 4H4L 3H4L
Set 2a: 3L3H 3L2H 2H2L 3L4H 3H3L 3H5L
Set 2b: 2H3L 3H3L 5L3H 3L3H 4H3L 2L2H

A set of six questions was randomly allocated to each of the subjects within a test group, under the constraint of approximately equal numbers.

A final question required subjects to rank 6 zones that differed in land suitability classification and level of certainty. Two questions were prepared and randomly allocated within each group. The certainty information was reversed between the two questions, again to account for any semantic or positional effects between the labelled zones.

4.4.2 Dynamic testing of positional uncertainty

Participants were initially shown four animations depicting a boat advancing from Zone A towards the restricted region, Zone B. They were required to make a judgment as to when they would turn away from the restricted zone. Each animation was prepared in Microsoft PowerPoint, with a background positional change of 1mm between slides and slide transition of 0.6 seconds. This gave a realistic portrayal of movement as depicted in a typical hand-held GPS receiver. In total, 70 slides were prepared for each animation. The slides were numbered, to give a measure of relative position at which subjects chose to turn away from Zone B.

Each of the four representations depicted the same information but used different methods. Although this was not made explicit to subjects, the starting slide of each of the four animations was offset, so that the same slide number for each representation depicted different relative positions. This offset was taken into account when analysing the slide numbers at which subjects responded.

Four presentation orders were generated using a Latin Square design, with the letters A-D used to signify the four representations. This design ensured that each representation was placed once in each ordered position and that each representation was subsequently followed once each by the other three representations: CBDA; DCAB; ADBC; BACD. A different presentation order was made to each of the four groups tested. The dynamic testing was thus designed to be balanced for both order and carry-over effects.

4.4.3 Static testing of positional uncertainty

Twenty static multiple choice questions were prepared to test subjects' understanding of the positional uncertainty representations. These comprised one of each of the five forms for each of the four representations. The same slide, considered to be a typical example of its form, was used for each representation, as below:

Form	**Slide**
1. Definitely in Zone A	15
2. Probably in Zone A	33
3. Equal chance of being in either zone	38
4. Probably in Zone B	55
5. Definitely in Zone B	61

Participants were each presented with eight multiple-choice questions. Five question sets were generated and allocated randomly within each experimental group, under the constraint of approximately equal numbers. Each question therefore appeared twice within the five question sets. The 4 representations were ordered using a Latin Square design for the first four questions of the first four sets, then reversed for the last 4 of these sets. The order of the first four questions of the final set was randomly selected, then reversed for the remaining questions. Each representation and each form were equally, as far as possible, placed in each of the eight possible positions:

Set 1:	C1 B4 D3 A2 A5 D1 B3 C2
Set 2:	B2 A1 C4 D5 D3 C5 A4 B1
Set 3:	A3 D2 B5 C1 C4 B2 D5 A4
Set 4:	D4 C5 A2 B3 B1 A5 C3 D1
Set 5:	B5 C3 A1 D4 D2 A3 C2 B4

The question sets were thus balanced, as far as possible, for both order and carry-over effects.

4.4.4 Question and answer booklets

Question booklets were prepared, using colour printed A4 paper (see appendix B). These comprised of the eight static positional uncertainty questions, the positional uncertainty survey question, the six pairwise comparisons of thematic uncertainty representations and the thematic uncertainty ranking question. Twenty different booklets were prepared, using the five question sets for the static positional uncertainty testing and the four question sets for the thematic uncertainty testing. These booklets were allocated randomly among the participants within each test group. An answer booklet was also prepared (see appendix C).

4.4.5 Pre-testing

The whole experiment (comprising of both the thematic and positional uncertainty components) was initially pre-tested on twelve subjects in groups of two or three. Following a brief introduction to the nature of the experiment, the dynamic representations were viewed on a desktop computer. Subjects recorded their answers in individual answer booklets. They were then each issued with a question booklet and completed the static testing at separate desks. The pre-testing sessions concluded with group interviews, where subjects were asked to comment on the clarity of the tasks and instructions and to give suggestions as to how the testing may be improved.

Most subjects stated that the instructions were easy to understand and that at all stages they felt that they knew what was required of them. However, one subject had not initially understood what was required in the dynamic representations and had not been able to respond to the first animation. It was suggested that a dynamic demonstration be presented before the first animation, to clarify the nature of the first task.

Three survey questions had initially been included to test for subjects' preferred method of testing positional uncertainty. These assessed ease of use, level of informativeness and overall preference. Several subjects stated that they had not really understood the difference between the three survey questions and had responded in the same way to them all. They proposed that only a single survey, that for overall preference was necessary. It was also suggested that the written instructions for the third stage, testing of thematic uncertainty, be broken down into bullet points as it was difficult to take in all of the information as presented. These recommendations were all adopted before the commencement of the actual experimental testing.

4.5 The Experiment

4.5.1 Participants

Participants were undergraduate and postgraduate students of the Department of Geomatics at the University of Melbourne. They voluntarily participated in the research at the conclusion of a lecture. Although there are methodological arguments against using students as a sample representative of the population at large (Gordon *et al*, 1986), other researchers have countered these arguments (Greenberg, 1987). In this study, it was considered that the students should represent a random sample of potential users of spatial information. Many of those using GIS to support their decision-making are not experts in the field of spatial information. The participants in this study had varying levels of experience in GIS and spatial information in general, reflective of the diverse user group that they were considered to represent.

Participants' experience in dealing with spatial data varied from novices to postgraduate students employed in positions where they were using spatial information to assist in decision-making on a daily basis. The students subjectively classified themselves into one of three experience levels, which were described as:

1: Novice (less than six months experience of using GIS)
2: Some experience (between six months and two years experience of using GIS)
3: Experienced (more than two years experience of using GIS)

Gender was also recorded.

In total, 100 students participated. The following table shows their numbers by experience level and by gender:

	Experience level				
		1	**2**	**3**	**total**
Gender	**male**	33	23	13	**69**
	female	14	14	3	**31**
	total	**47**	**37**	**16**	**100**

Table 4.1. Participant numbers, by experience level and by gender.

The subjects were tested in four groups. It had been assumed that these groups would consist of approximately equal numbers, with similar distributions of experience and gender between the four groups. However, the actual composition of the four groups varied considerably, as shown in the table below. This caused the dynamic testing phase of the experiment to be unbalanced. This was accounted for in the data analysis by completing a residual maximum likelihood (REML) analysis of the data from this stage.

Group	**Experience 1**		**Experience 2**		**Experience 3**		**Total**	
	Male	**Female**	**Male**	**Female**	**Male**	**Female**	**Male**	**Female**
1	23	5	8	4	0	0	**31**	**9**
2	8	7	3	3	0	0	**11**	**10**
3	0	0	8	3	4	2	**12**	**5**
4	2	2	4	4	9	1	**15**	**7**

Table 4.2. Group participant numbers, by experience level and by gender.

4.5.2 Data variables

There were two between-subjects variables, relevant to each of the three stages of testing:

Experience was an ordinal variable representing subjects' level of experience in using GIS (1-3, 3 being highest).

Gender was a binary variable denoting subjects' gender.

4.5.2.1 Testing of thematic uncertainty

The test variable, ***Pr(Expected)***, was a continuous variable [0,1] representing the sample proportion selecting the expected zone.

Land suitability class was an ordinal variable denoting the suitability of a zone as a potential airport site. (Simple case, 1(poor)-3(good); Complex case, 1(very poor)-5(very good)).

Certainty level was a binary variable, signifying the two certainty levels (H = high certainty, L = low certainty).
Ranking was an ordinal variable representing assigned rank (1-6, 6 being highest).

4.5.2.2 Dynamic testing of positional uncertainty

The test variable, ***slide number***, was a discrete variable representing the slide number at which subjects decided to turn away from Zone B.
Representation was a categorical variable, denoting the four representations of positional uncertainty.

4.5.2.3 Static testing of positional uncertainty

The test variable, ***frequency***, was a discrete variable representing the frequency with which choices were selected.
Form was a categorical variable, denoting the five question forms (1-5).
Rank was an ordinal variable representing the assigned rank (1-4, 4 being highest).

4.5.3 Conduct of the experiment

The students chose to voluntarily participate in the study, after a brief introductory explanation of the nature of the experiment. At the end of a lecture, approximately one week prior to testing, they were issued with a Plain Language Statement (see appendix A) that introduced the purpose of the study and explained what would be expected of participants. It was emphasized that there were no correct answers to the test questions, but that the study was concerned with their individual responses to the uncertainty representations. They were invited to contact the researchers if they required further information, although nobody took up this offer. The actual testing was conducted at the conclusion of their subsequent lecture in that subject.

The four groups of subjects were tested in separate sessions of approximately thirty minutes duration. Initially, consent forms were signed and again the voluntary nature of their participation and the study's concern with individual responses rather than correct answers were emphasized. Participants were each issued with an Answer Booklet (see appendix C). They were asked to complete the details on the front of this booklet, concerning their group number, their experience in using GIS and their gender. It was then briefly explained that there would be three stages to the experiment, beginning with the testing of dynamic representations of positional uncertainty.

A demonstration animation was displayed on the screen at the front of the lecture theatre and it was explained that this represented a boat moving through a channel of water from Zone A towards Zone B. It was pointed out that the slide number, on screen to the right of the GPS receiver, continually ticked over to represent time as the boat advanced towards Zone B. Participants' attention was then brought to the instructions on the Answer Booklet, which stated that Zone B was a restricted zone and that they were to decide as to when they would choose to turn away from this zone. They were told that they would see four different representations and that, for each, the method of representing the boat and boundary positions would be shown in the legend. They were to firstly familiarize themselves with this legend, then the animation would run and they were to record the slide number at which they would turn away from the restricted Zone B. Participants were asked at this point if they had any questions, as they would only have the opportunity to see each animation once.

The four animations were presented on the front screen one after the other, with approximately three minutes between presentations to allow subjects to record their responses and familiarise themselves with the subsequent legend. After the fourth animation, participants were issued with a Question Booklet (see appendix B). They were asked to record the number of their question booklet in the appropriate box on their answer booklet. They were then told that they could work through the remaining two stages of the experiment, static testing of the positional uncertainty representations and testing of thematic uncertainty representation, at their own pace, following the written instructions in their answer booklets. They were asked to ensure that they recorded all answers in the appropriate spaces in this booklet and were told that no writing in the question booklet was required. They were free to ask questions throughout these two stages, although few participants actually did so.

4.6 Chapter Summary

The effects of introducing thematic uncertainty information on decision-making were investigated within the context of an airport siting case study. An investigation of how different representations of positional uncertainty may affect decision-making was performed within a marine navigation context. Both case studies were carefully designed to be balanced for order and carry-over effects. The experiment was pre-tested on 12 volunteers and refinements were made as a result of their feedback. However, a mistake on the Graduated representation of positional uncertainty was not picked up.

Four groups of students, numbering 100 in total, participated in the experiment. The two case studies were conducted within a single experimental session for each group. Each session began with the dynamic representations of positional uncertainty being presented on a screen at the front of the lecture theatre. Participants responded to each of the four animations in turn, before moving on to individual question booklets. They proceeded to answer the static positional uncertainty questions and survey. They also completed the pairwise comparisons and ranking task that comprised the testing of thematic uncertainty.

5.

Results and Discussion: Airport Siting Case Study

5.1 Introduction

In this chapter, the results of the airport siting case study are presented. This study explored the effects on decision-making of including thematic uncertainty information in GIS output. The study had two components: pairwise comparisons of two regions and a task requiring six regions to be ranked. The results from each of these components are presented. The chapter continues with a discussion of the two sets of results and finishes with a section concluding the findings.

5.2 Results

In the following results sections, land suitability is coded from 1 (poor) to 3 (good) for the simple classification scale and from 1 (very poor) to 5 (very good) for the complex scale. The level of certainty is coded as H (high) or L (low). Hence the comparison *2H2L simple*, for example, refers to a pairwise comparison of fair, high certainty land with fair, low certainty land.

5.2.1 Pairwise testing of thematic uncertainty

Analysis of the pairwise data was performed through an examination of the proportions selecting the high or low certainty zone in each comparison. Two-tailed tests were performed against the null hypothesis that there was no difference between the proportion selecting the high or low certainty zones. The analysis was divided into two conditions: those where the high and low certainty zones were rated equally in terms of land suitability and those where the two zones differed in land suitability classification. There were few effects of gender or experience on individual comparisons (these are considered in section 5.2.1.3, which considers the overall data).

5.2.1.1 Equal land suitability classes

Six pairwise comparisons in which the two regions were of equal land suitability classes were tested. In only two of these, *2H2L simple* and *3H3L complex*, were the two zones of equal expected value. These results are presented first. The two regions in the comparisons *2H2L complex* and *4H4L complex* were not of equal expected value, despite being of equal land suitability. It was expected that participants would exhibit a preference for the low certainty land in the former comparison but for the high certainty land in the latter case. These results are considered second. Finally, the results from the comparisons *1H1L complex* and *3H3L simple* are presented. It is argued that the only logical response to these comparisons is to select the low certainty land in the former case and land of high certainty in the latter.

5.2.1.1.1 Pairwise comparisons: *2H2L simple, 3H3L complex*

The data from the two comparisons *2H2L simple* and *3H3L complex* were combined, as these two comparisons were equivalent. The two zones being compared were of equal expected value in each case, since they were rated equally in terms of land suitability and this classification was at the mid-class level. As there are an equal number of suitability classes above and below the assigned value, it would be expected that the low certainty zone would have as great a chance of being in a better land suitability class as it would have of being lower in land suitability. From a normalist point of view, we would therefore predict that subjects exhibit no preference for either zone in these comparisons. The table below summarises the aggregated responses to the two comparisons by gender and by experience level, with the columns **H, L** and **N** representing the number of subjects selecting the high certainty zone, low certainty zone or indicating no preference respectively:

	Experience									**Total**		
	1			**2**			**3**					
Gender	**H**	**L**	**N**	**H**	**L**	**N**	**H**	**L**	**N**	**H**	**L**	**N**
Male	31	0	2	18	1	4	12	1	0	**61**	**2**	**6**
Female	11	0	1	13	0	1	2	0	0	**26**	**0**	**2**
Total	**42**	**0**	**3**	**31**	**1**	**5**	**14**	**1**	**0**	**87**	**2**	**8**

Table 5.1. Aggregated responses, by gender and experience level, to *2H2L simple* and *3H3L complex* pairwise comparisons.

The null hypothesis was that the probability of selecting the high certainty zone was equal to that of selecting the low certainty zone, Pr(High) = 0.5, with the test statistic being the proportion H/(H+L) and two-tailed tests being performed. The ‘no preference’ responses were ignored in calculating this sample proportion.

There was no significant difference between the proportion selecting the high certainty zone by either gender or experience. The overall proportion selecting this zone was therefore tested against Pr(High) = 0.5 and a 95% confidence interval for this proportion was calculated:

Sample Pr(High)	**95% confidence interval**	**p (Fisher's)**
87/89 = 97.8%	(92.1%, 99.7%)	<0.001

Table 5.2. Two-tailed test results of null hypothesis that Pr(High) = 0.5 for *2H2L simple* and *3H3L complex* pairwise comparisons.

Regardless of gender and experience, the proportion selecting the high certainty zone was much greater than 0.5. In fact, all but two of the 89 subjects who expressed a preference between the two zones in these comparisons selected the high certainty zone.

Only eight of the 97 subjects indicated 'no preference' between the two zones. Even though this was the predicted response, since both zones were of equal expected value, this sample proportion was significantly lower than would be predicted from the null hypothesis that Pr(No pref) = 0.5:

Sample Pr(No pref)	**95% confidence interval**	**p (Fisher's)**
8/97 = 8.2%	(3.6%, 15.6%)	<0.001

Table 5.3. Two-tailed test results of null hypothesis that Pr(No pref) = 0.5 for *2H2L simple* and *3H3L complex* pairwise comparisons.

Subjects tended to indicate a preference and this preference strongly favoured the high certainty zone over the low certainty zone when the two zones were of equal expected value.

5.2.1.1.2 Pairwise comparisons: *2H2L complex, 4H4L complex*

The comparisons *2H2L complex* and *4H4L complex* were a little different to those above. Although both zones were classed equally in terms of land suitability, it might be expected that subjects would show a preference for the low certainty zone in the *2H2L* comparison, and a preference for the high certainty zone in the *4H4L* comparison. This is due to the ease with which a better scanario may be imagined for the low certainty zone. If it is classed as 2 for land suitability, it could only be worse if it were in reality class 1, whereas it would be better if it were any of classes 3, 4 or 5. The expected value of the low certainty zone is therefore higher than that of the high certainty zone in the *2H2L* comparison, despite the assigned land suitability classes being equal in each case. Conversely, there would be only one potentially better class when comparing land of class 4, whereas there would be three potentially worse classes. The expected value of the low certainty zone is therefore less than that of the high certainty zone in the *4H4L* comparison. The results for these two comparisons are summarised in the tables below:

2H2L	Experience									Total		
	1			2			3					
Gender	H	L	N	H	L	N	H	L	N	H	L	N
Male	6	5	1	3	5	3	7	3	0	**16**	**13**	**4**
Female	3	3	1	2	3	2	0	0	0	**5**	**6**	**3**
Total	**9**	**8**	**2**	**5**	**8**	**5**	**7**	**3**	**0**	**21**	**19**	**7**

4H4L	Experience									Total		
	1			2			3					
Gender	H	L	N	H	L	N	H	L	N	H	L	N
Male	20	1	0	12	0	0	3	0	0	**35**	**1**	**0**
Female	4	1	0	6	1	0	2	0	0	**12**	**2**	**0**
Total	**24**	**2**	**0**	**18**	**1**	**0**	**5**	**0**	**0**	**47**	**3**	**0**

Table 5.4. Summarised responses, by gender and experience level, to *2H2L complex* and *4H4L complex* pairwise comparisons.

For these two comparisons, the proportion selecting the expected zone was calculated and tested against the null hypothesis Pr(Expected) = 0.5. For the *2H2L* comparison, the sample proportion was L/(H+L) whereas for the *4H4L* comparison it was H/(H+L). Again, there were no significant effects of gender or experience, so the overall totals were used to calculate the sample proportions:

Comparison	Sample Pr(Expected)	95% confidence interval	p (Fisher's)
2H2L	19/40 = 47.5%	(31.5%, 63.9%)	0.88
4H4L	47/50 = 94%	(83.5%, 98.7%)	<0.001

Table 5.5. Two-tailed test results of null hypothesis that Pr(Expected) = 0.5 for *2H2L complex* and *4H4L complex* pairwise comparisons.

When the expected zone was of low certainty, the sample proportion was not significantly different to the null hypothesis, suggesting that there is no difference between the probability of selecting either zone in this case. However, when the expected zone was of high certainty, the sample proportion was significantly higher than that predicted from the null hypothesis. In this comparison, subjects showed a strong tendency to select the expected zone.

A two-proportion test was also performed on these results to test the difference in the proportions selecting the expected zone. This showed an extremely significant difference, with the proportion selecting the expected zone being much greater when it was of high certainty as opposed to low certainty:

Difference in sample proportions	95% confidence interval	p (Fisher's)
46.5%	(29.7%, 63.3%)	<0.001

Table 5.6. Two-tailed test results of difference in proportions selecting the expected zone in *2H2L complex* and *4H4L complex* comparisons.

In both comparisons, few subjects indicated 'no preference' between the two zones, the sample proportions being significantly lower than would be predicted from the null hypothesis that Pr(No pref) = 0.5:

Comparison	Sample Pr(No pref)	95% confidence interval	p (Fisher's)
2H2L	7/47 = 14.9%	(6.2%, 28.3%)	<0.001
4H4L	0/50 = 0.0%	(0%, 5.8%)	<0.001

Table 5.7. Two-tailed test results of null hypothesis that Pr(No pref) = 0.5 for *2H2L complex* and *4H4L complex* pairwise comparisons.

Although subjects did tend to indicate a preference in each of these two comparisons, the proportion indicating 'no preference' was significantly higher when the low certainty zone was the expected choice as opposed to the high certainty zone.

5.2.1.1.3 Pairwise comparisons: *1H1L complex, 3H3L simple*

There were also expected preferences for the *1H1L complex* and *3H3L simple* comparisons. In these, we would expect a strong tendency for subjects to select the low certainty zone when the land is in the lowest possible land suitability class (*1H1L*) and a strong tendency towards the high certainty zone when it is in the highest possible land suitability class (*3H3L*). In fact, for the *1H1L* comparison it would be irrational to select the high certainty zone, seeing as this is land of the lowest suitability class. The logical preference must be for the low certainty zone in this case, as there is some potential that the suitability of this land will actually be better than recorded. Similarly, choosing the low certainty zone in the *3H3L* comparison would defy logic and the rational choice would have to be for the high certainty zone. The results for these comparisons are summarized in the following tables:

1H1L	**Experience**									**Total**		
	1			**2**			**3**					
Gender	**H**	**L**	**N**	**H**	**L**	**N**	**H**	**L**	**N**	**H**	**L**	**N**
Male	7	9	5	5	5	2	1	2	0	**13**	**16**	**7**
Female	2	1	2	1	4	2	2	0	0	**5**	**5**	**4**
Total	**9**	**10**	**7**	**6**	**9**	**4**	**3**	**2**	**0**	**18**	**21**	**11**

3H3L	**Experience**									**Total**		
	1			**2**			**3**					
Gender	**H**	**L**	**N**	**H**	**L**	**N**	**H**	**L**	**N**	**H**	**L**	**N**
Male	12	0	0	10	1	0	10	0	0	**32**	**1**	**0**
Female	7	0	0	7	0	0	0	0	0	**14**	**0**	**0**
Total	**19**	**0**	**0**	**17**	**1**	**0**	**10**	**0**	**0**	**46**	**1**	**0**

Table 5.8. Summarised responses, by gender and experience level, to *1H1L complex* and *3H3L simple* pairwise comparisons.

Again, the sample proportion for the expected zone was tested against the null hypothesis Pr(Expected) = 0.5. There were no significant effects of gender or experience, so the overall totals were used to calculate the sample proportions.

Comparison	**Sample Pr(Expected)**	**95% confidence interval**	**p (Fisher's)**
1H1L	21/39 = 53.8%	(37.2%, 69.9%)	0.75
3H3L	46/47 = 97.9%	(88.7%, 99.9%)	<0.001

Table 5.9. Two-tailed test results of null hypothesis that Pr(Expected) = 0.5 for 1*H1L complex* and *3H3L simple* pairwise comparisons.

For the *1H1L* comparison, the sample proportion was not significantly different to the null hypothesis, indicating no preference for the low certainty zone over the high certainty zone. However, for the *3H3L* comparison, the sample proportion selecting the expected zone was extremely significant, revealing a strong preference for the high certainty zone. A two-proportion test showed that the proportion selecting the expected zone was significantly higher when this zone was of high certainty rather than low certainty:

Difference in sample proportions	95% confidence interval	p (Fisher's)
44.0%	(27.8%, 60.2%)	<0.001

Table 5.10. Two-tailed test results of difference in proportions selecting the expected zone in *1H1L complex* and *3H3L simple* comparisons.

Again, in both comparisons few subjects indicated 'no preference' between the two zones, the proportions that did so being significantly lower than would be predicted from the null hypothesis that Pr(No pref) = 0.5:

Comparison	Sample Pr(No pref)	95% confidence interval	p (Fisher's)
1H1L	11/50 = 22%	(11.5%, 36.0%)	<0.001
3H3L	0/47 = 0%	(0%, 6.2%)	<0.001

Table 5.11. Two-tailed test results of null hypothesis that Pr(No pref) = 0.5 for *1H1L complex* and *3H3L simple* pairwise comparisons.

However, as before, although subjects did tend to indicate a preference in each of these two comparisons, the proportion indicating 'no preference' was significantly higher when the low certainty zone was the expected choice.

5.2.1.1.4 Summary of results for equal land suitability classes

The following bar graph summarises the results from the pairwise comparisons between regions of equal land suitability classes. It also includes, in dashed boxes, the responses expected from a consideration of expected utility. It is clear that the actual responses only follow those predicted when the expected response is the high certainty zone. When the low certainty zone or no preference between the two zones was expeced, the tendency to select the high certainty zone was much greater then predicted from a consideration of expected utility.

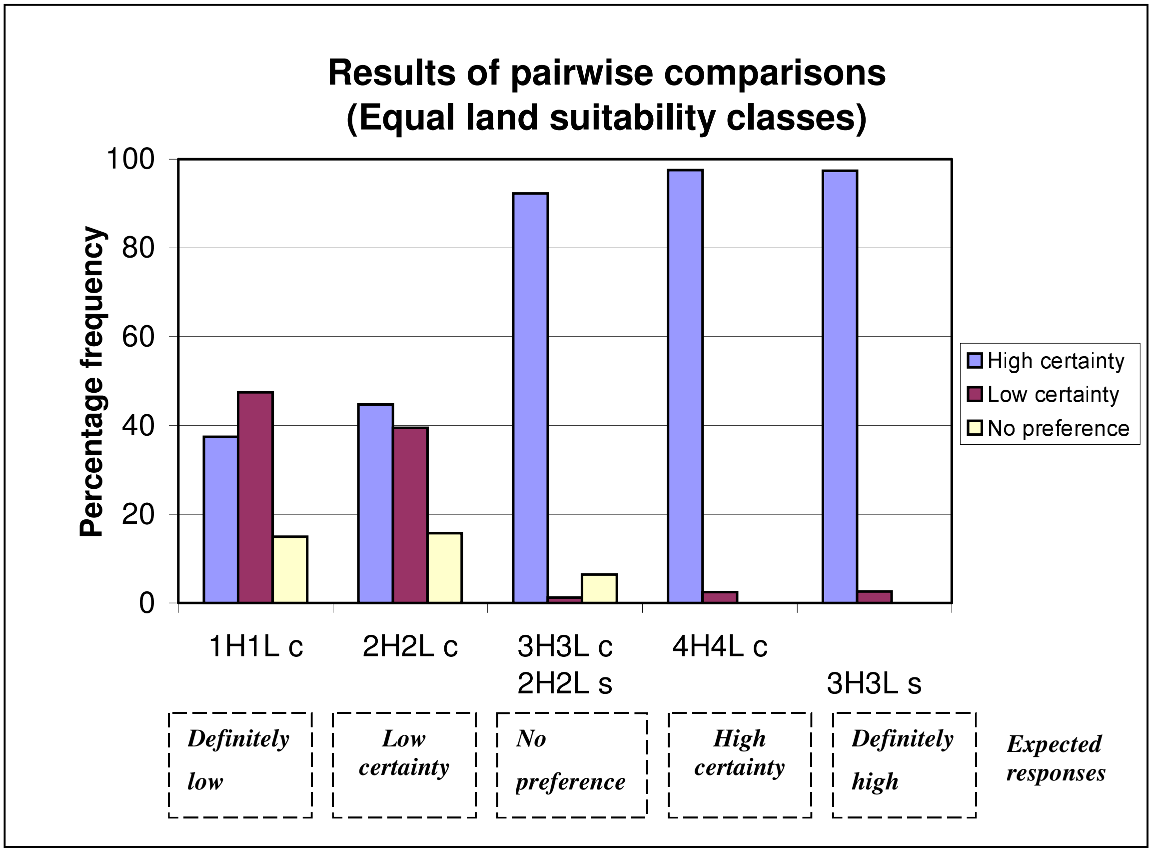


Figure 5.1. Bar graph showing actual responses to pairwise comparisons between zones of equal land suitability classes, together with expected responses from a consideration of expected utility.

5.2.1.2 Different land suitability classes

In each of the six pairwise comparisons between zones differing in land suitability classification, it was expected that subjects would select the zone of higher land suitability. In the analysis, the appropriate sample proportion was tested against Pr(Expected) = 0.5.

5.2.1.2.1 Pairwise comparisons: *3H4L complex, 4H3L complex*

The pairwise comparisons *3H4L* and *4H3L* are analysed together as, in these conditions, the certainty information is reversed. In the former comparison, the expected preference was for the low certainty zone, as this is of higher land suitability. In the latter comparison, the high certainty zone was the expected preference. The results are summarized in the following tables:

3H4L	**Experience**									**Total**		
	1			**2**			**3**					
Gender	**H**	**L**	**N**	**H**	**L**	**N**	**H**	**L**	**N**	**H**	**L**	**N**
Male	15	5	1	3	6	3	2	1	0	**20**	**12**	**4**
Female	2	0	3	2	3	2	2	0	0	**6**	**3**	**5**
Total	**17**	**5**	**4**	**5**	**9**	**5**	**4**	**1**	**0**	**26**	**15**	**9**

4H3L	**Experience**									**Total**		
	1			**2**			**3**					
Gender	**H**	**L**	**N**	**H**	**L**	**N**	**H**	**L**	**N**	**H**	**L**	**N**
Male	11	1	0	11	0	0	10	0	0	**32**	**1**	**0**
Female	7	0	0	7	0	0	0	0	0	**14**	**0**	**0**
Total	**18**	**1**	**0**	**18**	**0**	**0**	**10**	**0**	**0**	**46**	**1**	**0**

Table 5.12. Summarised responses, by gender and experience level, to *3H4L complex* and *4H3L complex* pairwise comparisons.

The results of the one-proportion test on the null hypothesis that Pr(Expected) = 0.5 are shown in the table below, together with 95% confidence intervals for this probability:

Comparison	**Sample Pr(Expected)**	**95% confidence interval**	**p (Fisher's)**
3H4L	15/41 = 36.6%	(22.1%, 53.1%)	0.117
4H3L	46/47 = 97.9%	(88.7%, 99.9%)	<0.001

Table 5.13. Two-tailed test results of null hypothesis that Pr(Expected) = 0.5 for 3*H4L complex* and *4H3L complex* pairwise comparisons.

Again, it can be seen that when the expected zone is of low certainty, the sample proportion is not significantly different from the null hypothesis, whereas when the expected zone is of high certainty, the sample proportion was much higher than would be expected from the null hypothesis. A two-proportion test confirms that the difference in the sample proportions is significantly different according to the level of certainty associated with the expected zone:

Difference in sample proportions	**95% confidence interval**	**p (Fisher's)**
61.3%	(46.0%, 76.6%)	<0.001

Table 5.14. Two-tailed test results of difference in proportions selecting the expected zone in *3H4L complex* and *4H3L complex* comparisons.

In both comparisons, few subjects indicated 'no preference' between the two zones, causing us to reject the null hypothesis that Pr(No pref) = 0.5:

Comparison	Sample Pr(No pref)	95% confidence interval	p (Fisher's)
3H4L	9/50 = 18%	(8.6%, 31.4%)	<0.001
4H3L	0/47 = 0%	(0%, 6.2%)	<0.001

Table 5.15. Two-tailed test results of null hypothesis that Pr(No pref) = 0.5 for *3H4L complex* and *4H3L complex* pairwise comparisons.

Similar to previous findings, subjects did tend to indicate a preference in each of these two comparisons, although the proportion indicating 'no preference' was significantly higher when the low certainty zone was the expected choice.

5.2.1.2.2 Pairwise comparisons: *2H1L simple, 2H3L simple*

The results for these two comparisons are summarized in the following tables:

2H1L	Experience									Total		
	1			2			3					
Gender	H	L	N	H	L	N	H	L	N	H	L	N
Male	20	1	0	10	1	1	3	0	0	**33**	**2**	**1**
Female	3	1	1	6	1	0	2	0	0	**11**	**2**	**1**
Total	**23**	**2**	**1**	**16**	**2**	**1**	**5**	**0**	**0**	**44**	**4**	**2**

2H3L	Experience									Total		
	1			2			3					
Gender	H	L	N	H	L	N	H	L	N	H	L	N
Male	8	2	2	9	1	1	8	1	1	**25**	**4**	**4**
Female	4	1	2	7	0	0	0	0	0	**11**	**1**	**2**
Total	**12**	**3**	**4**	**16**	**1**	**1**	**8**	**1**	**1**	**36**	**5**	**6**

Table 5.16. Summarised responses, by gender and experience level, to *2H1L simple* and *2H3L simple* pairwise comparisons.

Again, there were no significant differences in the sample proportions by gender or by experience, so the overall totals were used in the sample proportions tested against the null hypothesis that Pr(Expected) = 0.5.

Comparison	Sample Pr(Expected)	95% confidence interval	p (Fisher's)
2H1L	44/48 = 91.7%	(80.0%, 97.7%)	<0.001
2H3L	5/41 = 12.2%	(4.1%, 26.2%)	<0.001

Table 5.17. Two-tailed test results of null hypothesis that Pr(Expected) = 0.5 for 2*H1L simple* and *2H3L simple* pairwise comparisons.

There is a clear difference in the way that subjects responded to these two comparisons. In both cases, the null hypothesis is to be rejected. However, when the expected zone is of high certainty, this is because the sample proportion is much higher than predicted by the null hypothesis. In contrast, when the expected zone is of low certainty, the proportion selecting it is much lower than would be expected. A two-proportion test shows an extremely significant difference between the two sample proportions:

Difference in sample proportions	95% confidence interval	p (Fisher's)
79.5%	(66.8%, 92.2%)	<0.001

Table 5.18. Two-tailed test results of difference in proportions selecting the expected zone in *2H1L simple* and *2H3L simple* comparisons.

In fact, we can see that the proportion selecting the high certainty zone is almost as high in the *2H3L* comparison as it is in the *2H1L* comparison. A two-proportion test on the proportions selecting the high certainty zone in each of these conditions reveals no significant difference. Subjects are tending to select the high certainty zone in both comparisons, regardless of the difference in land suitability class of the low certainty zone.

Difference in sample proportions	95% confidence interval	p (Fisher's)
3.9%	(-8.8%, 16.6%)	0.73

Table 5.19. Two-tailed test results of difference in proportions selecting the high certainty zone in *2H1L simple* and *2H3L simple* comparisons.

As before, few subjects indicated 'no preference' between the two zones, causing us to reject the null hypothesis that Pr(No pref) = 0.5:

Comparison	Sample Pr(No pref)	95% confidence interval	p (Fisher's)
2H1L	2/50 = 4%	(0.5%, 13.7%)	<0.001
2H3L	6/47 = 12.8%	(4.8%, 25.7%)	<0.001

Table 5.20. Two-tailed test results of null hypothesis that Pr(No pref) = 0.5 for *2H1L simple* and *2H3L simple* pairwise comparisons.

In these two comparisons, a greater proportion of subjects indicated 'no preference' when the low certainty zone rather than the high certainty zone was the expected choice, although this difference did not reach significance.

5.2.1.2.3 Pairwise comparisons: *4H2L complex, 3H5L complex*

The following tables summarise the results for these two comparisons, in which the two zones differed by two land suitability classes:

4H2L	**Experience**									**Total**		
	1			**2**			**3**					
Gender	**H**	**L**	**N**	**H**	**L**	**N**	**H**	**L**	**N**	**H**	**L**	**N**
Male	20	0	1	11	0	1	3	0	0	**34**	**0**	**2**
Female	5	0	0	6	0	1	2	0	0	**13**	**0**	**1**
Total	**25**	**0**	**1**	**17**	**0**	**2**	**5**	**0**	**0**	**47**	**0**	**3**

3H5L	**Experience**									**Total**		
	1			**2**			**3**					
Gender	**H**	**L**	**N**	**H**	**L**	**N**	**H**	**L**	**N**	**H**	**L**	**N**
Male	3	8	1	5	4	2	6	3	1	**14**	**15**	**4**
Female	1	4	2	3	2	2	0	0	0	**4**	**6**	**4**
Total	**4**	**12**	**3**	**8**	**6**	**4**	**6**	**3**	**1**	**18**	**21**	**8**

Table 5.21. Summarised responses, by gender and experience level, to *4H2L complex* and *3H5L complex* pairwise comparisons.

The overall proportion selecting the expected zone in each comparison was tested against the null hypothesis that Pr(Expected) = 0.5 and 95% confidence intervals for these sample proportions were calculated:

Comparison	Sample Pr(Expected)	95% confidence interval	p (Fisher's)
4H2L	47/47 = 100%	(93.8%, 100%)	<0.001
3H5L	21/39 = 53.8%	(37.2%, 69.9%))	0.75

Table 5.22. Two-tailed test results of null hypothesis that Pr(Expected) = 0.5 for 4*H2L complex* and *3H5L complex* pairwise comparisons.

Similar to earlier comparisons, it can be seen that when the expected zone is of low certainty, the sample proportion is not significantly different from the null hypothesis, whereas when the expected zone is of high certainty, the sample proportion was much higher than would be expected from the null hypothesis. A two-proportion test confirms that the difference in the sample proportions is extremely significant according to the certainty associated with the expected zone.

Difference in sample proportions	95% confidence interval	p (Fisher's)
46.2%	(30.5%, 61.8%)	<0.001

Table 5.23. Two-tailed test results of difference in proportions selecting the expected zone in *4H2L complex* and *3H5L complex* comparisons.

As in previous findings, the null hypothesis that Pr(No pref) = 0.5 was rejected for both of these comparisons:

Comparison	Sample Pr(No pref)	95% confidence interval	p (Fisher's)
4H2L	3/50 = 6%	(1.3%, 16.5%)	<0.001
3H5L	8/47 = 17.0%	(7.6%, 30.8%)	<0.001

Table 5.24. Two-tailed test results of null hypothesis that Pr(No pref) = 0.5 for *4H2L complex* and *3H5L complex* pairwise comparisons.

Again, there was a greater tendency for subjects to indicate 'no preference' when the low certainty zone was the expected choice, although this difference did not reach significance.

5.2.1.2.4 Summary of results for different land suitability classes

The bar graph below summarises the results from the pairwise comparisons between regions of differing land suitability classes. It also includes, in dashed boxes, the responses expected from a consideration of expected utility. Again, it can be seen that the actual responses differed from those predicted whenever it was expected that the low certainty zone would be selected.

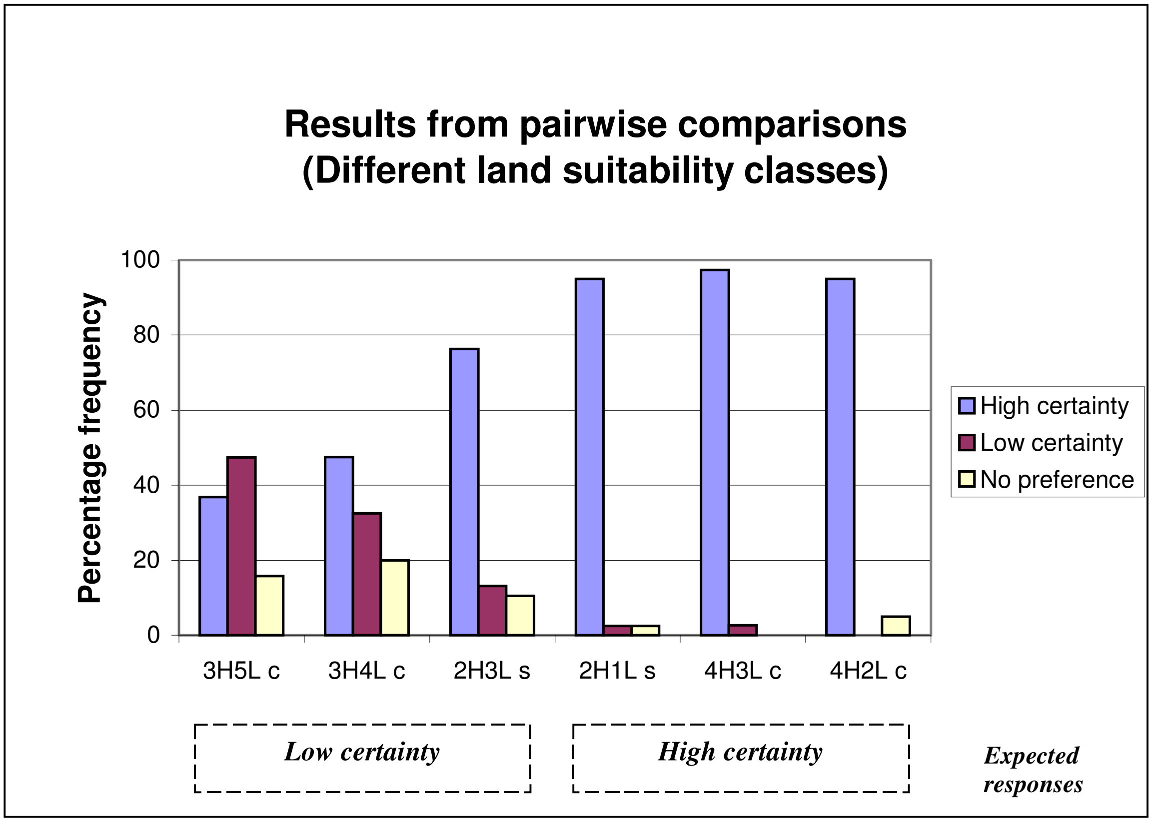


Figure 5.2. Bar graph showing actual responses to pairwise comparisons between zones of different land suitability classes, together with expected responses from a consideration of expected utility.

5.2.1.3 Overall gender and experience effects

The following table summarises the overall responses to the 12 comparisons by experience and gender:

Overall	**Experience**									**Total**		
	1			**2**			**3**					
Gender	**H**	**L**	**N**	**H**	**L**	**N**	**H**	**L**	**N**	**H**	**L**	**N**
Male	153	32	13	97	24	17	65	11	2	**315**	**67**	**32**
Female	49	11	12	60	14	10	12	0	0	**121**	**25**	**22**
Total	**202**	**43**	**25**	**157**	**38**	**27**	**77**	**11**	**2**	**436**	**92**	**54**

Table 5.25. Aggregated responses, by gender and experience level, to the 12 pairwise comparisons.

There were no significant differences between the responses of the two genders to each of the individual comparisions. Overall, it was found that females showed a higher tendency to indicate ‘no preference’ than did males, although this did not reach significance at the 5% level (estimated difference = 5.4%, p=0.066).

However, some differences were evident in the way that subjects of different experience levels responded. In the *3H4L* comparison, the sample proportion selecting the low certainty zone was significantly greater for experience level 2 than the other subjects (estimated difference = 42.1%, Fisher’s p = 0.015). In the *3H5L* comparison, the sample proportion selecting the low certainty zone was significantly greater for experience level 1 than the other subjects (estimated difference = 35.9%, Fisher’s p = 0.049). Overall, subjects of experience level 3 (most experienced) showed a significantly higher tendency to select the high certainty zone than did other subjects (estimated difference = 12.6%, p = 0.003) and a significantly lower tendency to indicate ‘no preference’ (estimated difference = -8.3%, Fisher’s p = 0.009).

Looking at the aggregated totals, subjects exhibited an extremely strong tendency to select the high certainty zone over the low certainty zone. This was evident despite the fact that the land suitability classes were balanced over the 12 comparisons and was typical of both genders and all experience levels. The overall sample proportion H/(H+L) was tested against the null hypothesis that Pr(High) = 0.5:

	Sample Pr(H)	**95% confidence interval**	**p (Fisher’s)**
Overall	436/528 = 82.6%	(79.1%, 85.7%)	<0.001

Table 5.26. Two-tailed test results of null hypothesis that Pr(High) = 0.5 for the 12 pairwise comparisons.

5.2.2 Ranking of suitability/certainty zones

The ranking rask was only performed for the simple (3 class) land suitability classification. In the following results and discussion, the six zones have been coded from 1 (poor) to 3 (good) for land suitability, and using H for high certainty and L for low certainty. For example, 1H represents poor land suitability, high certainty. The boxplot below shows the distribution of rankings against zone, experience and gender:

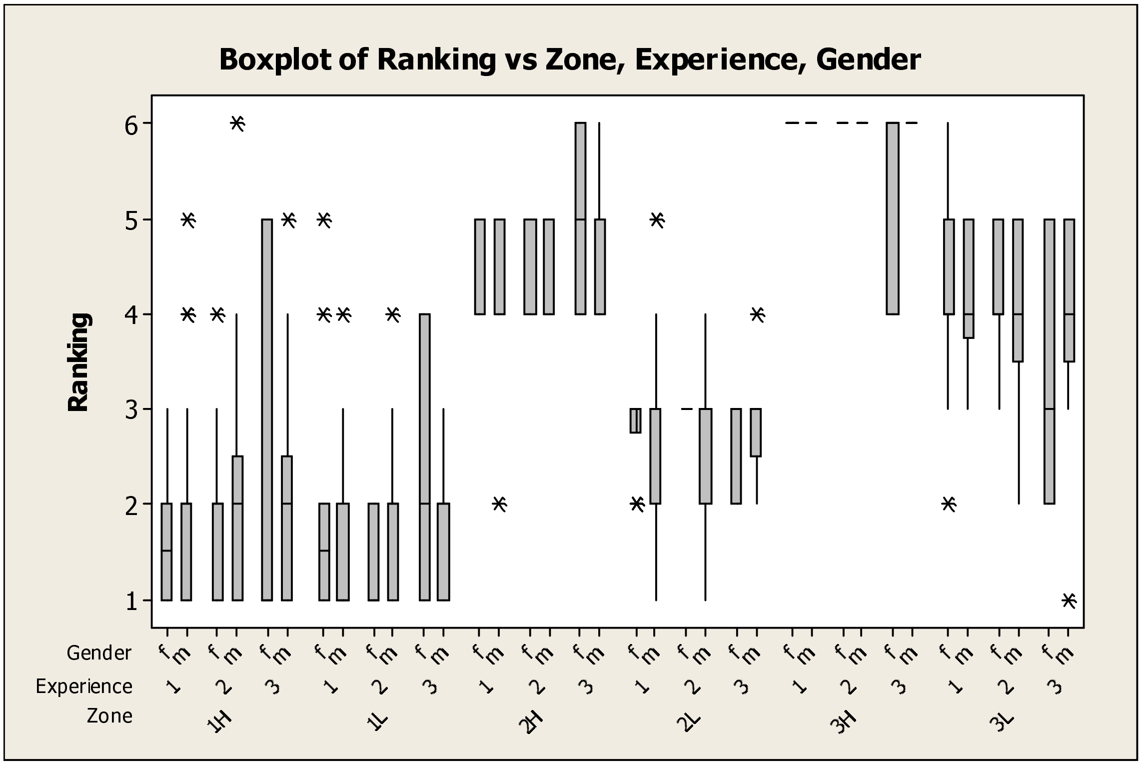


Figure 5.3. Boxplot showing distribution of rankings for each suitability/certainty zone against experience level and gender.

It can be seen that there are few apparent differences in the way that participants of both genders and all experience levels are ranking the six zones. It is clear that the zones of lowest land suitability, 1H and 1L, generally receive the lowest rankings, although with some variability. Zone 2L tends to be assigned the next lowest ranking, followed by zone 3L. Zone 2H appears to be ranked second highest on average, with only a single outlier ranking it below 4. Almost without exception, participants assigned the highest ranking to zone 3H.

The following matrix plot indicates no apparent interactions between zone and experience or between zone and gender:

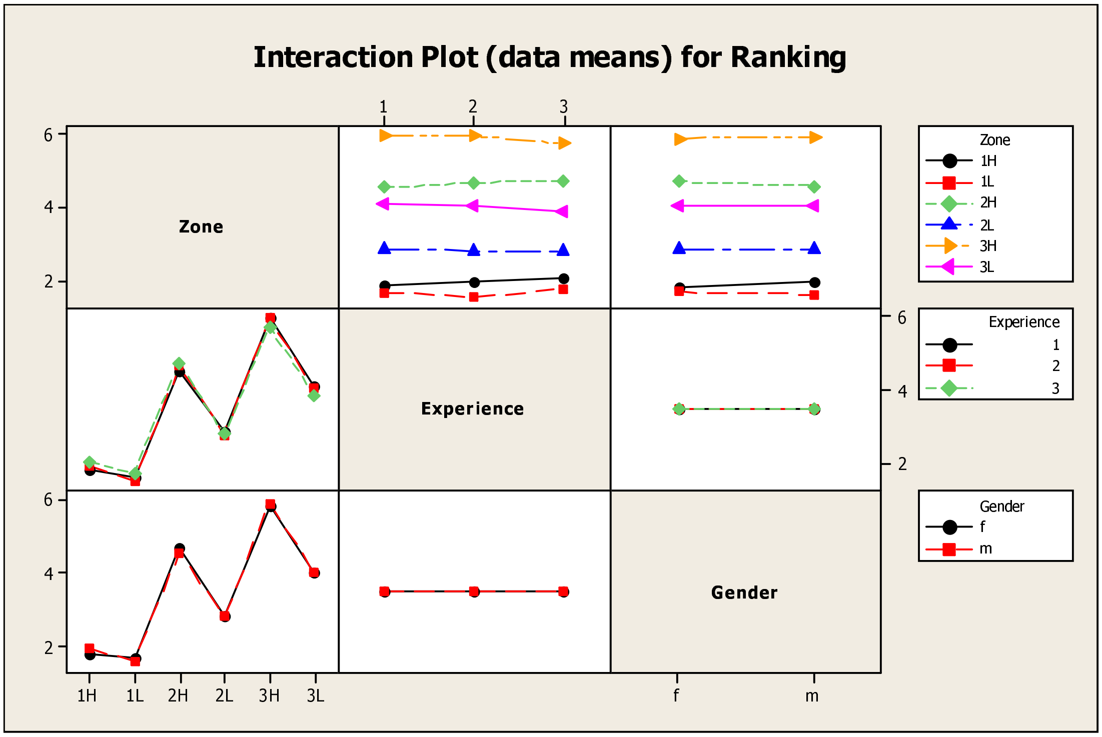


Figure 5.4. Matrix plot showing ranking interactions between zone, experience level and gender.

The interactions plot suggests that subjects of different experience levels and different genders appear to be ranking the six zones in similar fashions. An analysis of variance was not appropriate for this data as the assumption of equal variances is violated (Bartlett's test statistic = 146.38, p<0.001) and the rankings are not independent or normally distributed. A nonparametric Kruskal-Wallis test was therefore performed to test for differences in the rankings that subjects assigned to the six zones. This showed extremely significant differences (DF = 5, H = 454.28, p<0.001, adjusted for ties). Wilcoxon signed rank tests were performed to test for the significance of pairwise differences in the median rankings. These tests were significant at the 0.1% level for all pairwise comparisons except that between 1L and 1H. This latter comparison did not reach a significant level of difference. The estimated median difference, together with the 95% confidence interval and associated p-values for adjacent preferences are shown in the following table:

		Wilcoxon		Estimated	95% confidence interval	
	n	Statistic	p	Median	Lower	Upper
3H - 2H	95	4405	<0.001	1.50	1.00	1.50
2H - 3L	95	3305	<0.001	0.50	0.00	1.00
3L - 2L	95	4162	<0.001	1.50	1.00	1.50
2L - 1H	95	3582	<0.001	1.00	1.00	1.50
1H - 1L	95	2691	0.128	0.00	0.00	0.50

Table 5.27. Results of Wilcoxon signed rank tests, including estimated median differences between adjacent ranked zones and associated p-values.

The order of preference for the six zones is shown below, with the non-significant difference indicated by a line:

Most preferred Least preferred

3H 2H 3L 2L 1H 1L

5.3 Discussion

5.3.1 Pairwise testing of thematic uncertainty

Equal land suitability classes

When the high and low certainty zones were equally classed in terms of land suitability and this classification was at the mid-class level, the two zones were of equal expected value. It was not expected that subjects would exhibit a preference for either zone. However, an extremely significant tendency to select the high certainty zone over the low certainty zone was observed. Only 8.2% of subjects indicated ‘no preference’ as their response. Of the 89 subjects who did indicate a preference, all but two chose the land with the high certainty classification as their preferred site. This tendency was similarly displayed across experience levels and both genders, with the 95% confidence interval for the proportion selecting the high over the low certainty zone being (92.1%, 99.7%).

Although not expected from a normalist point of view, this finding is not surprising in the light of Ellsberg’s work on ambiguity. People generally dislike working ‘in the dark’ and prefer to work with data that is ‘certain’, despite it not necessarily having a greater intrinsic expected value. The observed responses in the *2H2L simple* and *3H3L complex* comparisons demonstrate the strength of this bias. The inclusion of certainty information in a simple binary form can lead to extremely strong preferences that appear to be irrational in terms of expected value.

However, it could be argued that from a loss aversion viewpoint, this behaviour is rational. People are generally conservative and suffer from a loss more strongly than they value the corresponding gain. The potential loss from the low certainty zone being of inferior land suitability is seen as more damaging than the possible reward would be beneficial if the land were to be of greater suitability. Effectively, loss aversion may decrease the subjective expected value of the low certainty zone by giving a greater weight to the possible result that the land is of lower suitability than its current classification. It would therefore make more sense to choose the high certainty zone.

The possible effect of loss aversion provides a plausible explanation for the strong bias towards selecting the high certainty zone observed in the *2H2L simple* and *3H3L complex* comparisons. It would also provide an explanation for the pattern of subject responses in the *2H2L complex* and *4H4L complex* comparisons. We would expect subjects to select the low certainty zone in the *2H2L* comparison, since there are more classes of potentially higher land suitability than lower ones for this zone. Conversely, we would expect a preference for the high certainty zone in the *4H4L* comparison, since there are a greater number of potentially lower classes for the low certainty zone. The results were not entirely consistent with this logic. There was no significant preference for either zone in the *2H2L* comparison (p=0.88), although a strong preference for the high certainty zone in the *4H4L* comparison was observed (p<0.001).

If the loss aversion effect is introduced into subjects' reasoning, these results can be explained. Although there are more potentially higher than lower land suitability classes for the low certainty zone in the *2H2L* comparison, if greater weighting were given to the lower class, some subjects may still, rationally, choose the high certainty zone. Approximately equal numbers of subjects chose the high and low certainty zone in this case. However, in the *4H4L* comparison, increasing the weight given to the potential lower land suitability classes of the low certainty zone would serve to further lessen its attraction and therefore strengthen the preference for the high certainty zone.

However, loss aversion does not provide an explanation for the observed responses when zones of the extreme suitability classes were compared. In the *1H1L* comparison, the land is of the lowest possible land suitability. The rational choice would therfore be for the low certainty zone, as this has potentially higher land suitability but no potentially worse outcomes. In this condition, loss aversion would have no effect and there can be no logical reason for selecting the high certainty zone. However, 18 of the 39 subjects who expressed a preference did exactly that. Only 53.8% of subjects expressing a preference responded in a logical manner and chose the low certainty zone. It was to this comparison that the greatest proportion of subjects, 22%, indicated no preference between the two zones. Logic dictates that preference for the low certainty zone is the only rational response, but many subjects appear to be loath to select this option.

There was no such unexpected result when the high certainty zone was the expected preference in the *3H3L* comparison. All but one of the 47 subjects who expressed a preference did select the logical option. In contrast to the 22% who indicated no preference in the *1H1L* comparison, no subjects indicated no preference in this condition. It appears that subjects have no problems with selecting the high certainty zone when it is the rational choice but are significantly less likely to select the low certainty zone in the converse condition ($p<0.001$). This pattern of behaviour is similar to that observed in the previous comparisons, although loss aversion cannot be used to explain the reluctance to select the low certainty zone in this case.

Different land suitability classes

When comparing zones of differing land suitability classes, it was expected that subjects would select the zone of greater land suitability, regardless of certainty information. It was assumed that greater expected value would be attributed to the zone with the higher land suitability class. However, the observed pattern of subjects' behaviour was again inconsistent with this. In the *3H4L complex* and *4H3L complex* comparisons, the certainty information was reversed between the two conditions. A preference for the low certainty zone was expected in the *3H4L* comparison and a preference for the high certainty zone was expected in the *4H3L* comparison. However, only 36.6% of subjects expressing a preference responded as expected in the former comparison, while 97.9% responded as expected in the latter comparison.

It could again be argued that loss aversion can provide an explanation for the reluctance to select the low certainty zone in the *3H4L* comparison. It was expected that subjects would choose the 4L zone. If the three potentially lower land suitability classes are weighted more heavily than the one potentially higher class, the expected value of this zone would decrease. Nonetheless, to adequately justify the responses of the 63.4% of subjects selecting the high certainty zone, this expected value would have to drop by more than one land suitability class. This would require that all of these subjects exhibit extremely strong loss aversion. Whether or not such extreme loss aversion can be considered rational would depend upon the context of the decision being made. Little context was provided in this example; for instance, no social, environmental, political or economic effects of the airport were discussed. If the loss aversion explanation is upheld, the level of conservatism exhibited was beyond that expected.

A similar, but more pronounced, pattern of results was observed in the *2H1L simple* and *2H3L simple* comparisons. The strong preference for the high certainty zone in the former comparison was expected, with 91.7% of those subjects expressing a preference responding in this manner. However, only 12.2% of subjects selected the low certainty zone in the *2H3L* comparison. This caused us to reject the null hypothesis of no preference between the two zones. However, the significant preference was observed towards the high certainty zone and not, as expected, towards the low certainty zone. In both comparisons, subjects were selecting the high certainty zone, regardless of the land suitability class assigned to the low certainty zone.

Again, this can be explained in terms of loss aversion. In the *2H1L* comparison, the expected value of the low certainty zone would be expected to rise above one, but not by as much as the expected value of the low certainty zone in the *2H3L* comparison would be expected to fall below three. If strong loss aversion were exhibited, it would be possible for the expected value of the low certainty zone to be similar in the two comparisons. This could explain the fact that there was no significant difference betweeen the proportion selecting the high certainty zone across the two comparisons (p=0.73). As before, such an explanation requires that extremely strong weights are applied to the potentially lower land suitability classes.

Similar results were observed when the zones differed by two land suitability classes. In the *4H2L complex* comparison, all 47 subjects expressing a preference selected the high certainty zone. However, in the *3H5L complex* comparison, only 53.8% of subjects expresssing a preference responded as expected and selected the low certainty zone. An explanation in terms of loss aversion requires that the expected value of the low certainty zone in this latter comparison is actually lowered by more than two land suitability classes.

Throughout these comparisons, subjects have shown a reluctance to select the low certainty zone when it was the expected choice but have shown no such deviation from expectation when selecting the high certainty zone. Ignoring the possible effects of subjective loss aversion, in the two comparisons when the zones were of equal expected value, an overwhelming majority of subjects selected the high certainty zone. Of the five comparisons in which the low certainty zone had the higher expected value, no significant preference was observed in four comparisons and in the fifth, the observed preference was for the high certainty, not the low certainty, zone. In all five comparisons when the high certainty zone had the higher expected value, an extremely strong preference ($p<0.001$) was observed for this zone.

Although this strong reluctance to select the low certainty zone can be explained in terms of loss aversion for many of the comparisons, such an argument would not be applicable to the *1H1L* comparison. In this condition, only 21 of the 50 subjects correctly responded that the low certainty zone was the preferred option. There is no rational justification for the 11 responses that indicated no preference and the 18 responses that selected the high certainty zone. It appears that subjects exhibit an aversion to selecting a zone labelled as low certainty, even when this zone has the higher expected value. This can lead to the making of irrational decisions when certainty information is included.

Overall, 82.6% of expressed preferences favoured the high certainty zone, despite the comparisons being balanced so that the high and low certainty zones were expected with equal frequency. This strong bias toward the high certainty zone was evident across both genders and all experience levels. Although there were no significant differences between the genders, there were some differences between subjects of different experience levels. Two of the individual comparisons were significant at the 5% level, although this may be a chance result of making multiple comparisons. However, the overall proportions of experience level 3 subject responses were significantly different to those of the other subjects. The level 3 subjects were more likely to select the high certainty zone than the other subjects ($p=0.003$) and were less likely to indicate no preference ($p=0.009$). This result is somewhat surprising as it might be expected that the more experienced subjects would be more familiar with the concept of uncertainty and would show less of a bias towards the high certainty zone. Conversely, this preliminary finding indicates that the more experienced subjects are no better, and may in fact be slightly worse, than other subjects in dealing with uncertainty.

5.3.2 Ranking of suitability/certainty zones

Irrespective of gender and experience, there were extremely significant differences in the rankings asssigned to the six zones. It was expected that zone 3H would be the most preferred zone, as this land is classed as most suitable and the classification is of high certainty. Of the 95 subjects, 91 selected this as their most preferred zone. This result is consistent with the findings of other researchers (Leitner & Buttenfield, 2000) that people are able to apply certainty information in a manner that assists in decision-making. Nevertheless, when subjects were asked to go beyond the optimal zone and rank the remaining land, irrational decisions became apparent. In agreement with the findings from the pairwise comparisons, zone 2H was preferred to zone 3L, despite the higher suitability classification of the latter zone. Although it could be argued that this finding is a result of loss aversion, subjects also repeated their illogical treatment of zones 1H and 1L. In fact, in this ranking task the poor, high certainty zone was actually preferred to the poor, low certainty zone, although this preference did not reach significance (p=0.128).

These findings support the conclusion from the pairwise comparisons that people do not handle certainty information in a logical manner. Most people can identify the optimal zone when certainty information is incorporated into the attribute display, but beyond this many people respond in an irrational fashion. Several of the subjects who ranked zone 1H higher than 1L were incredulous as to their mistake when it was pointed out to them. When it was explained in terms of the 1H zone being definitely poor land but the 1L zone having some chance of better (but none of worse) suitability, they stated that they had not considered the information like this and could not believe that they had not done so. They appeared to quickly grasp the concept after a brief explanation, despite this not being intuitive from the display itself. Indeed, six of the 95 subjects (6.3%) ranked the data primarily by certainty and then by the apparently subordinate quality of land suitability, as below:

Most preferred Least preferred

3H 2H 1H 3L 2L 1L

This demonstrates how strong the aversion to a low certainty classification can be.

5.4 Conclusion

The introduction of thematic uncertainty can have significant effects on decision-making, although these may not always be rational. Within a siting task, an extremely significant tendency was observed for subjects to prefer the high over the low certainty zone when the zones had equal expected value in terms of land suitability. This preference was exhibited by subjects of both genders and all experience levels. When comparing zones that differed in land suitability, subjects did not show a preference for low certainty zones that were classified higher, despite their greater expected value. Although this could be explained in terms of loss aversion, the reluctance to select the low certainty zone in the 1H1L comparison can only be described as irrational.

Overall, participants showed a strong tendency to select the high certainty zone more readily than the low certainty zone. The observed reluctance to select the low certainty zone is consistent with Ellsberg's finding of ambiguity aversion. It implies that many people have an irrational bias against information labelled as low certainty, beyond that which can be explained through loss aversion. The inclusion of certainty information may therefore lead to irrational decisions, rather than promoting informed, robust decision-making as intended. The bias towards selecting the high certainty zone was greatest in subjects of the highest experience level. This indicates that even experienced GIS users may not be able to intuitively handle uncertainty information, although with a little education, they appeared to understand its relevance with regards to this simple task.

6.

Results and Discussion – Navigation Case Study

6.1 Introduction

In this section, the results from the navigation case study are presented. This study investigated how different representations of positional uncertainty may affect decision-making. The results are presented in three sections. The first of these is concerned with participants' responses to the four dynamic representations of positional uncertainty. The slide numbers at which they chose to turn the boat away from the restricted zone are analysed. In section 6.2.2, the responses to the static multiple-choice questions are presented. This section aims to demonstrate whether or not the participants were able to comprehend each of the four types of representation. Finally, the results from the survey question asking participants to rank their preference for the four representation types are analysed. A discussion of the results from the three sections follows and the chapter concludes with a summary of the findings.

6.2 Results

The four methods used to represent positional uncertainty in this study are referred to as Limits, Scale, Probability and Graduated (see figures 4.3-4.6, p.66-69). As with the airport siting case study, responses are analysed with respect to the gender and experience level of the participant.

6.2.1 Dynamic testing of positional uncertainty

The response variable, the slide number at which participants chose to turn the boat away from the restricted zone, was analysed against the treatment variables of representation, order, gender and experience. Representation and order were treated as variables that vary within subjects, whereas gender and experience varied between subjects. The following boxplot shows the distribution of responses (slide number) against the two within-subject variables of representation and order:

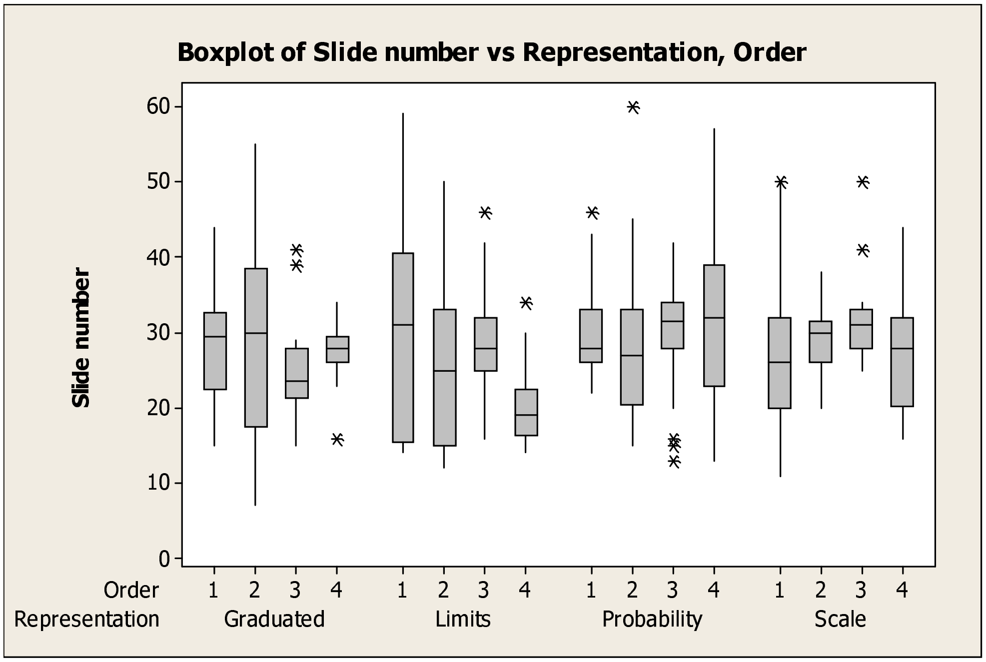


Figure 6.1. Boxplot showing the distribution of responses (slide number) against representation and order.

The boxplot below shows the distribution of responses against the between-subject variables of gender and experience:

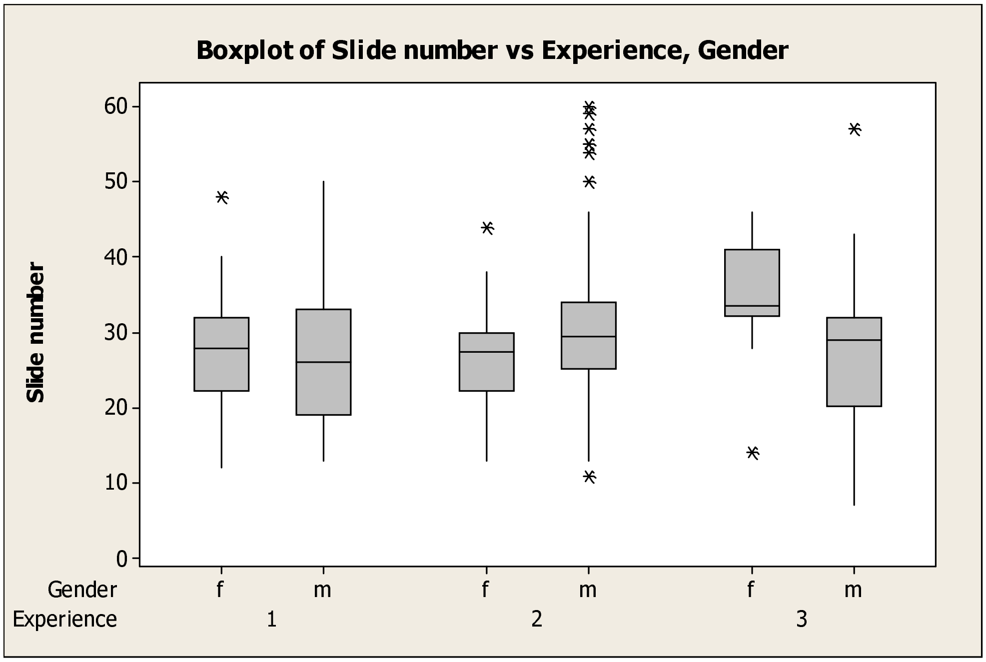


Figure 6.2. Boxplot showing the distribution of responses (slide number) against experience and gender.

From both boxplots, it can be seen that the participants' responses were quite variable, although it is difficult to compare spreads as the sample sizes differed between these conditions. Several outliers were identified in the data. These were fairly evenly distributed between the four representation types but appeared to be more common when the representations were placed at order 3 than at the other orders. The outliers also appeared to be clustered around the male subjects of experience level 2, with six responses being unusually high in this group. There is no apparent reason for this cluster of outliers.

Since the experiment was not balanced across all variables, a residual maximum likelihood (REML) analysis was performed to investigate the significance of any of the treatment variable effects or any interactions. A logarithmic transformation was applied to the response variable, slide number, as this provided a more satisfactory plot of residuals against the fitted values. Subject number 82 was also omitted from the analysis as two of this participant's four responses exhibited abnormally large residuals. On further inspection, it was observed that subject 82 had responded in an atypical manner. His responses including the minimum slide number at which any participant decided to turn away from the restricted zone as well as the third highest response for slide number for any subject. It may be that this participant had not fully understood the task, as his responses were far more variable than the average.

The following matrix plot indicates any pairwise interactions between the four treatment variables of representation, order, experience and gender:

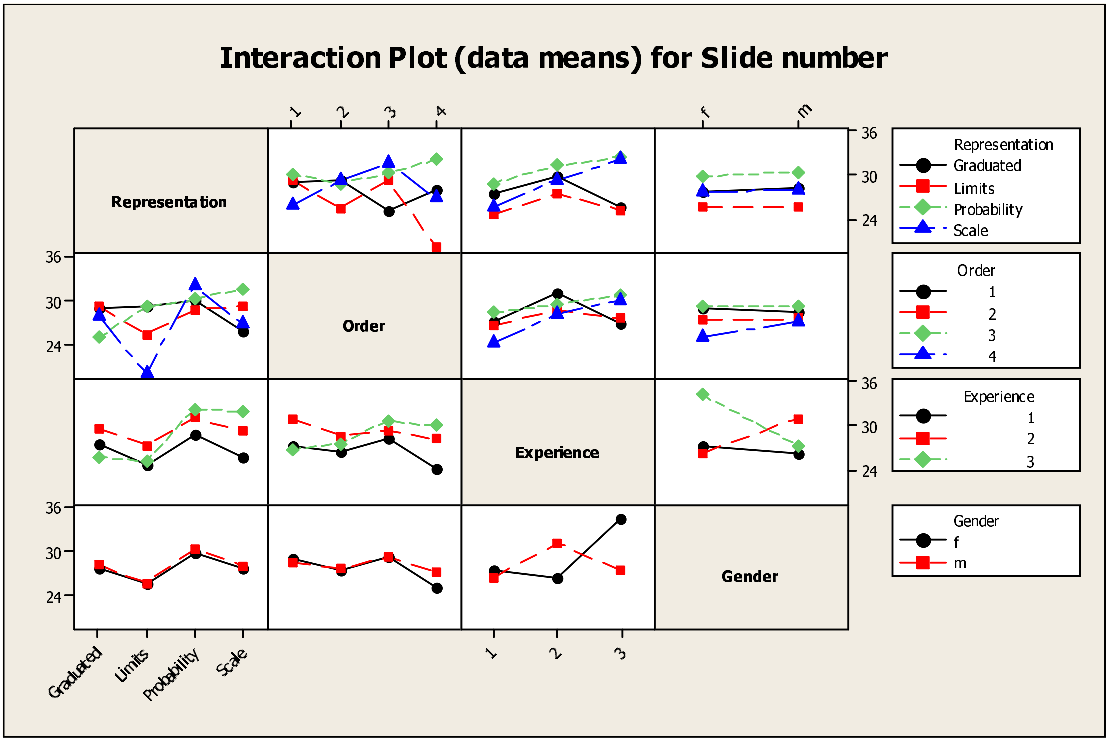


Figure 6.3. Matrix plot indicating any pairwise interactions between the four treatment variables.

The plot reveals little interaction between gender and the two variables of representation and order, since the lines are approximately parallel for these two pairings. There is some evidence that there may be an interaction between gender and experience, with the more experienced females showing a tendency to select higher slide numbers than less experienced females; a tendency that is not evident amongst male subjects. There is also some indication that the more experienced subjects respond differently to the displays, depending upon the order and type of representation. However, the most noticeable interaction appears to be between representation and order. The plot suggests that responses to the Limits representation are more prone to order effects than those to the other representations, this being particularly noticeable when the Limits representation is displayed last in the sequence of four animations.

The results of the REML analysis of variance are summarised in the following table:

Fixed term	**Wald stat.**	**d.f.**	**Wald/d.f.**	**p**
Experience	3.76	2	1.88	0.152
Gender	0.01	1	0.01	0.909
Representation	37.18	3	12.39	<0.001
Order	10.66	3	3.55	0.014
Representation*Order	20.98	9	2.33	0.013
Gender*Order	3.40	3	1.13	0.333
Experience*Order	4.12	6	0.69	0.661
Gender*Representation	0.66	3	0.22	0.884
Experience*Representation	13.48	6	2.25	0.036
Experience*Gender	6.36	2	3.18	0.042

Table 6.1. Results of REML analysis of variance for slide number against the four treatment variables and pairwise interactions.

The analysis indicates significant interactions of representation*order, experience*gender and experience*representation. The two interactions involving experience may be a chance result, as the number of highly experienced subjects, especially females, was small. The REML analysis can also under-estimate p-values, so the calculated values of 0.036 and 0.042 may not actually reflect interactions that are significant at the 5% level. Nonetheless, the interaction of representation*order (W=20.98, DF=9, p=0.013) does appear to be significant and implies that the effects of order vary between the different representations of the uncertainty information.

The REML analysis also indicates significant main effects of order and representation. The effect of order (W=10.66, DF=3, p=0.014) suggests that subjects are making different decisions dependent upon the order in which the display is viewed. Such an effect would be expected if subjects were learning throughout the duration of the experiment and is of no surprise, considering that they were given no prior training in tasks of this nature. Most importantly to this study, the analysis reveals an extremely significant effect of representation (W=37.18, DF=3, p < 0.001). Subjects' responses to the different representations of uncertainty varied significantly, despite the four representations depicting the same information, albeit in different ways.

The significant effects were further examined through an analysis of pairwise comparisons. There were 120 such comparisons for the representation*order interaction, so the analysis was restricted to a comparison of the four orders separately within each representation. Inspection of the interactions plot indicated that subjects might be responding differently to order within the Limits representation than within the other three representations. The pairwise comparisons confirmed this. The responses to the Limits representation differed significantly, according to the order in which the representation was viewed. The only non-significant difference was that between orders 3 and 1. However, no order effects were significant at the 5% level when each of the other three representations was considered individually. The figure below arranges the predicted means for log(slide number) for the four orders of each representation, with the non-significant differences indicated by lines:

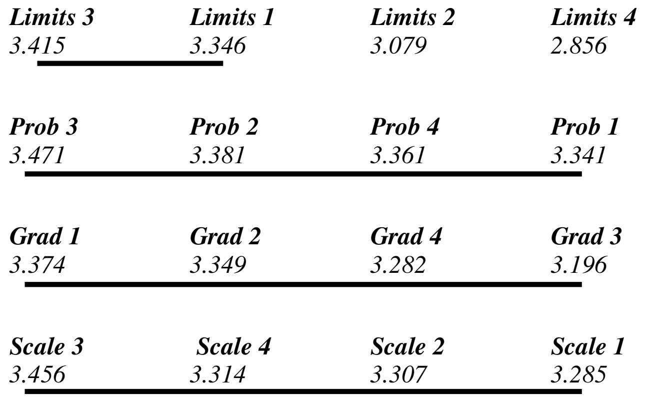


Figure 6.4. Predicted means for log(slide number), arranged for the four orders for each representation and with non-significant pairwise differences indicated by lines.

Although this representation*order interaction appeared to be significant, a main effect of order was also indicated by the REML analysis. Pairwise comparisons of the four orders revealed that the slide number when a representation was presented third was significantly greater than when it was presented second or fourth. The slide number for a representation presented first was also significantly greater than that when the representation was presented fourth. The following figure arranges the predicted mean log(slide number) for the four orders, with non-significant differences indicated by lines:

Figure 6.5. Predicted means for log(slide number) for the four orders with non-significant pairwise differences indicated by lines.

The main effect of representation type on the response variable was found to be extremely significant (p<0.001). As this effect was of primary concern to this study, descriptive statistics summarizing participants' responses by representation type were calculated. These are shown below, together with a boxplot showing the distribution of responses against type of representation:

Representation	**n**	**Mean**	**StDev**	**Min**	**Q1**	**Med**	**Q3**	**Max**
Graduated	94	28.01	8.457	7	22	28	32	55
Limits	94	25.67	10.38	12	16	25	33	59
Probability	94	30.21	8.753	13	25	30	34	60
Scale	94	27.93	7.170	11	23	28	32	50

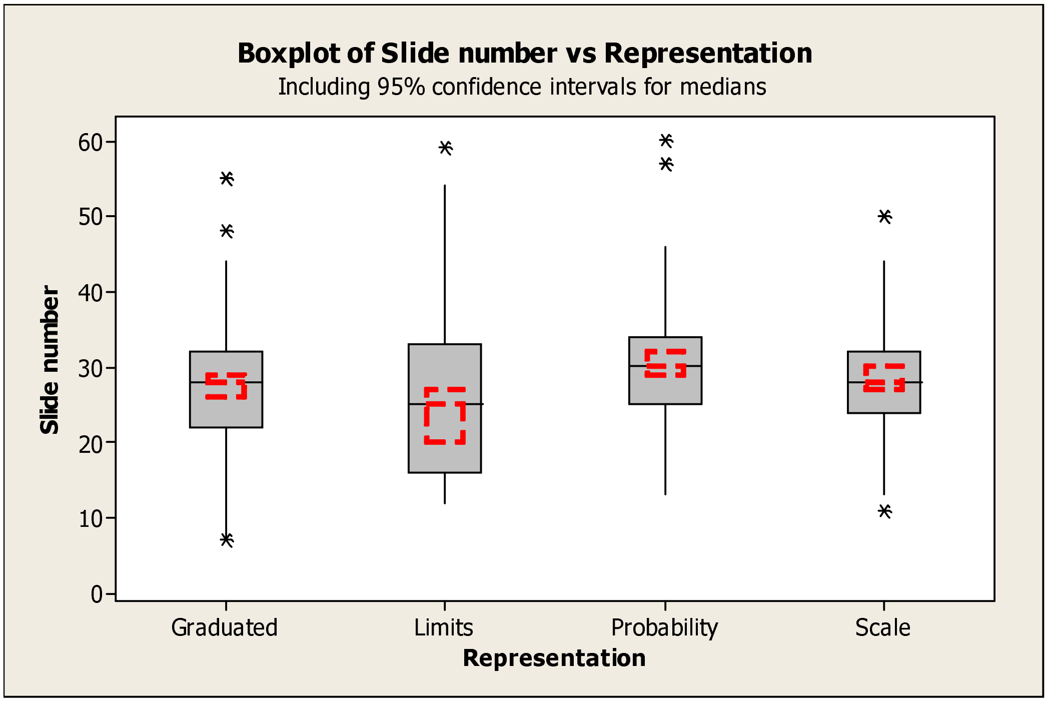


Figure 6.6. Boxplot showing distribution of responses (slide number) for the four types of representation.

In order to determine where the significant differences in responses lie, comparisons were performed between each pair of representations. The results of testing the null hypothesis that the difference in the predicted means of log(slide number) equalled zero are summarized below:

Two-tailed test of difference in predicted means = 0.00

	Estimated difference	**Standard Error**	**Effective D.F.**	**p**
Prob-Scale	0.048	0.057	84	0.402
Prob-Grad	0.088	0.049	84	0.076
Prob-Limits	0.214	0.049	84	<0.001
Scale-Grad	0.040	0.049	84	0.416
Scale-Limits	0.166	0.049	84	0.001
Grad-Limits	0.126	0.057	84	0.030

Table 6.2. Results of pairwise comparisons testing for significant differences in responses to the four representations, including estimated mean differences of log(slide number) and associated p-values.

The slide number for the representation Limits was found to be significantly lower than that for all other representations. Subjects tended to turn away from the restricted zone sooner when the uncertainty was represented using Limits. The slide number for the representation Probability was higher than that for the representations Scale and Graduated, although these differences were not found to be significant at the 5% level (p=0.402 and p=0.076 respectively). The predicted means for log(slide number) at which subjects turned away are illustrated below, with the line indicating non-significant differences:

Figure 6.7. Predicted means for log(slide number) for the four types of representation, with non-significant pairwise differences indicated by lines.

6.2.2 Static multiple-choice tests

The following tables summarise subject responses to the 20 representation-by-form multiple-choice questions, the values representing the number of participants making that particular response, using the legend:

Code	**Response**
1	Definitely in Zone A
2	Probably in Zone A
3	Equal chance of being in either Zone
4	Probably in Zone B
5	Definitely in Zone B
N	Do not understand the diagram

Form 1 **Expected response: 1 (Definitely in A)**

	1	2	3	4	5	N
scale	34	4	0	0	0	0
grad	33	3	1	0	0	1
limits	38	3	0	0	0	0
prob	36	2	0	0	2	0

Form 2 **Expected response: 2 (Probably in A)**

	1	2	3	4	5	N
scale	3	27	10	0	0	0
grad	4	19	11	3	1	1
limits	1	29	9	1	0	0
prob	7	27	4	1	0	0

Form 3 **Expected response: 3 (Equal chance of either Zone)**

	1	2	3	4	5	N
scale	0	2	33	3	1	0
grad	1	10	26	2	0	1
limits	0	0	34	3	1	0
prob	0	0	37	1	1	1

Form 4 **Expected response: 4 (Probably in B)**

	1	2	3	4	5	N
scale	0	0	1	15	23	0
grad	0	1	0	30	8	1
limits	0	2	0	25	12	0
prob	0	0	0	21	17	0

Form 5 **Expected response: 5 (Definitely in B)**

	1	2	3	4	5	N
scale	1	0	0	3	36	0
grad	1	0	0	2	35	1
limits	0	0	0	2	36	0
prob	0	0	0	1	38	0

Table 6.3. Responses to the 20 representation-by-form questions, organised by expected response.

These results were aggregated across the four representation types and tabulated against the expected responses. An attribute agreement analysis was then performed. The following table shows, in the shaded boxes, the number of responses that are in agreement with the expected response. The overall agreement percentage is also given:

Cumulative responses:

Actual response

Expected response

	1	2	3	4	5	N	*total*
1	141	12	1	0	2	1	***157***
2	15	102	34	5	1	1	***158***
3	1	12	130	9	3	2	***157***
4	0	3	1	91	60	1	***156***
5	2	0	0	8	145	1	***156***
total	***159***	***129***	***166***	***113***	***211***	***6***	***784***

Agreement = 609/784 = 77.7%

Table 6.4. Aggregated responses to the 20 representation-by-form questions.

To enable an analysis of the level of agreement between subjects' responses and the expected responses for the different types of representation, the data was then tabulated by representation type. A similar attribute agreement analysis was then performed separately for each. The following tables summarise the results by type of representation:

Scale representation:

Actual response

		1	**2**	**3**	**4**	**5**	**N**	***total***
Expected response	**1**	34	4	0	0	0	0	***38***
	2	3	27	10	0	0	0	***40***
	3	0	2	33	3	1	0	***39***
	4	0	0	1	15	23	0	***39***
	5	1	0	0	3	36	0	***40***
	total	***38***	***33***	***44***	***21***	***60***	***0***	***196***

Agreement = 145/196 = 74.0%

Graduated representation:

Actual response

		1	**2**	**3**	**4**	**5**	**N**	***total***
Expected response	**1**	33	3	1	0	0	1	***38***
	2	4	19	11	3	1	1	***39***
	3	1	10	26	2	0	1	***40***
	4	0	1	0	30	8	1	***40***
	5	1	0	0	2	35	1	***39***
	total	***39***	***33***	***38***	***37***	***44***	***5***	***196***

Agreement = 143/196 = 73.0%

Limits representation:

Actual response

		1	**2**	**3**	**4**	**5**	**N**	***total***
Expected response	**1**	38	3	0	0	0	0	***41***
	2	1	29	9	1	0	0	***40***
	3	0	0	34	3	1	0	***38***
	4	0	2	0	25	12	0	***39***
	5	0	0	0	2	36	0	***38***
	total	***39***	***34***	***43***	***31***	***49***	***0***	***196***

Agreement = 162/196 = 82.7%

Probability representation:

Actual response

		1	**2**	**3**	**4**	**5**	**N**	***total***
Expected response	**1**	36	2	0	0	2	0	***40***
	2	7	27	4	1	0	0	***39***
	3	0	0	37	1	1	1	***40***
	4	0	0	0	21	17	0	***38***
	5	0	0	0	1	38	0	***39***
	total	***43***	***29***	***41***	***24***	***58***	***1***	***196***

Agreement = 159/196 = 81.1%

Table 6.5. Agreement analysis for the four types of representation.

A version of McNemar's test was used to determine whether the levels of subject agreement varied significantly between the four representation types. The test had to be modified as each subject responded to a total of eight questions, two of each of the four representation types. The responses were therefore not independent. This was accounted for by comparing the number of times each subject was in total agreement with the expected response for each representation, with the number of times they were in some disagreement (either one or two discrepancies). Two-way tables were calculated for each pair of representations and two-tailed sign tests were performed on the off-diagonals to test the null hypothesis of no difference between the agreement levels. The results are summarised below:

		Graduated display		
		Agree	**Disagree**	**Total**
Scale display	**Agree**	30	23	53
	Disagree	21	24	45
	Total	51	47	98

		Limits display		
		Agree	**Disagree**	**Total**
Graduated display	**Agree**	38	13	51
	Disagree	28	19	47
	Total	66	32	98

		Limits display		
		Agree	**Disagree**	**Total**
Scale display	**Agree**	39	14	53
	Disagree	27	18	45
	Total	66	32	98

		Probability display		
		Agree	**Disagree**	**Total**
Graduated display	**Agree**	35	16	51
	Disagree	29	18	47
	Total	64	34	98

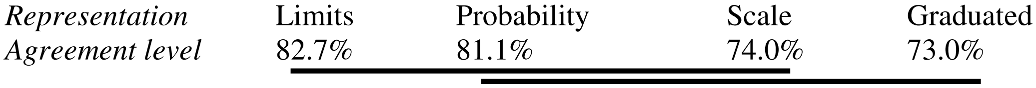
		Probability display		
		Agree	**Disagree**	**Total**
Scale display	**Agree**	38	15	53
	Disagree	26	19	45
	Total	64	34	98

		Probability display		
		Agree	**Disagree**	**Total**
Limits display	**Agree**	44	22	66
	Disagree	20	12	32
	Total	64	34	98

	test proportion	**p**
Scale display vs Graduated display	21/44	0.880
Scale display vs Limits display	14/41	0.060
Scale display vs Probability display	15/41	0.117
Graduated display vs Limits display	13/41	0.028
Graduated display vs Probability display	16/45	0.072
Limits display vs Probability display	20/42	0.878

Table 6.6. Modified McNemar's test of differences between levels of participant agreement for the four types of representation.

These two-tailed tests only reveal one pairing to differ significantly at the 5% level; subject agreement for Limits was significantly higher than that for Graduated (p=0.028). The greater subject agreement for Limits than for Scale approached, but did not reach, significance (p=0.060), as did the greater subject agreement for Probability than for Graduated (p=0.072). These findings are summarised below, with the lines indicating non-significant differences:



Representation	Limits	Probability	Scale	Graduated
Agreement level	82.7%	81.1%	74.0%	73.0%

Figure 6.8. Agreement levels for the four representation types, with non-significant differences indicated by lines.

These agreement levels do not account for the level of agreement that would be expected from chance alone. In light of this, Cohen's kappa values were also calculated for each type of representation and for the aggregated data. The following table gives the kappa values for each of the five forms, together with the overall kappa value, for the four representations and for the cumulative responses.

	Scale	**Graduated**	**Limits**	**Probability**	Cumulative
Form 1	0.869	0.822	0.937	0.832	**0.865**
Form 2	0.681	0.422	0.734	0.752	**0.647**
Form 3	0.740	0.584	0.798	0.891	**0.754**
Form 4	0.419	0.725	0.653	0.620	**0.612**
Form 5	0.629	0.801	0.779	0.716	**0.728**
Overall K	**0.675**	**0.664**	**0.783**	**0.764**	**0.722**

Table 6.7. Cohen's kappa values, by form, for each of the four representation types and aggregated values.

The data was also organised by experience and by gender and an attribute agreement analysis was performed on each. The following tables summarise the kappa values from these analyses:

Cohen's kappa values by experience level:

	Experience 1	**Experience 2**	**Experience 3**
Form 1	0.888	0.876	0.770
Form 2	0.613	0.683	0.658
Form 3	0.706	0.787	0.821
Form 4	0.589	0.720	0.407
Form 5	0.696	0.814	0.621
Overall K	**0.698**	**0.776**	**0.658**

Table 6.8. Cohen's kappa values, by form, for each of the three experience levels and aggregated values.

Cohen's kappa values by gender:

	Female	**Male**
Form 1	0.803	0.892
Form 2	0.657	0.641
Form 3	0.790	0.738
Form 4	0.624	0.606
Form 5	0.707	0.736
Overall K	**0.719**	**0.723**

Table 6.9. Cohen's kappa values, by form, for each gender and aggregated values.

It can be seen that the kappa values are extremely similar for both genders but appear to be more variable across the three experience levels.

6.2.3 Survey of participants' preferred representation type

The boxplot below shows the distribution of rankings against the treatment variables of representation, experience and gender:

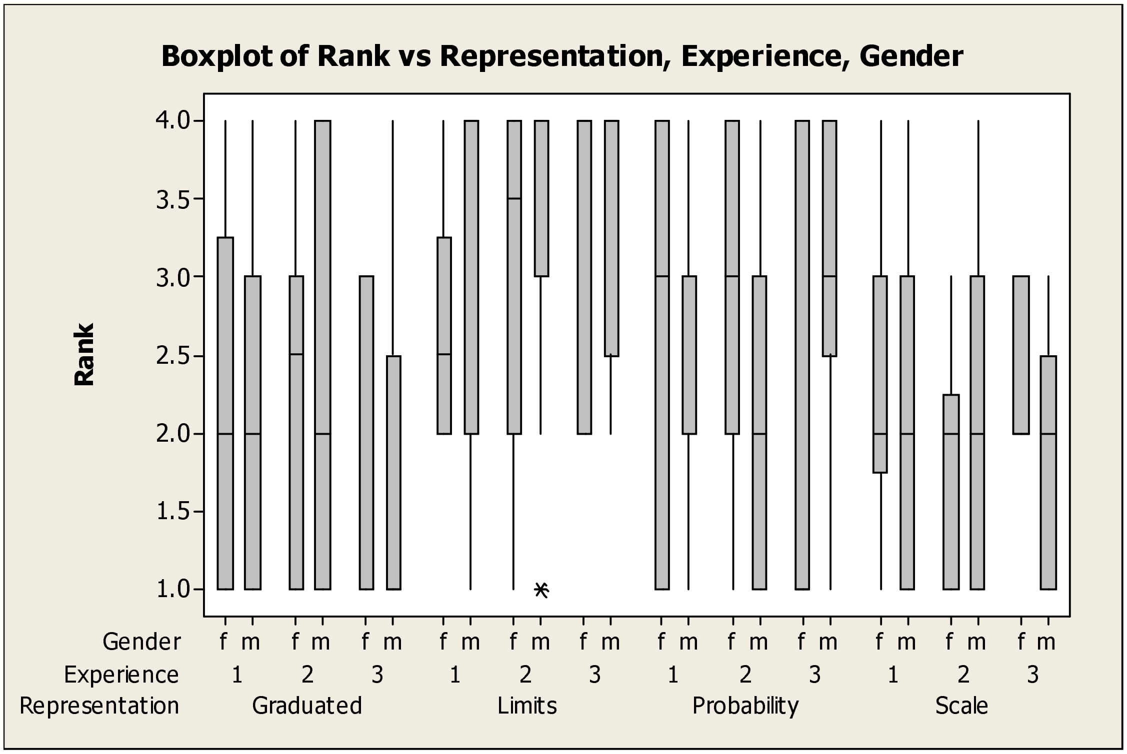


Figure 6.9. Boxplot showing distribution of subjective preference rankings for the four types of representation by experience level and gender.

From the boxplot it is clear that the Limits representation tended to receive higher rankings than the other three representation types, with participants of all experience levels and both genders responding similarly. The Scale representation tended to be the least preferred, again irrespective of experience level and gender. It appears that the Probability representation is the second most preferred, although the rankings it receives are quite variable. Similarly, the rankings assigned to the Graduated representation are quite variable, although this seems to be the third most preferred overall.

The following matrix plot reveals little interaction between the three treatment variables of representation, experience and gender, since the lines are approximately parallel:

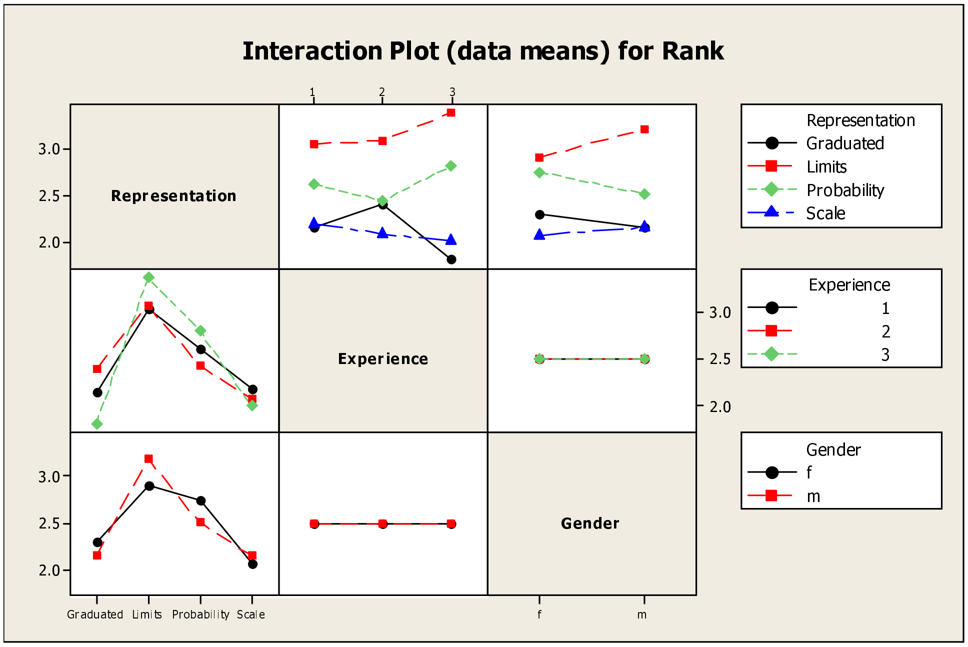


Figure 6.10. Matrix plot indicating any pairwise interactions between the three treatment variables.

In light of the apparent lack of interaction between the variables, the data were aggregated and the following boxplot summarises the overall preference rankings assigned to the four representations:

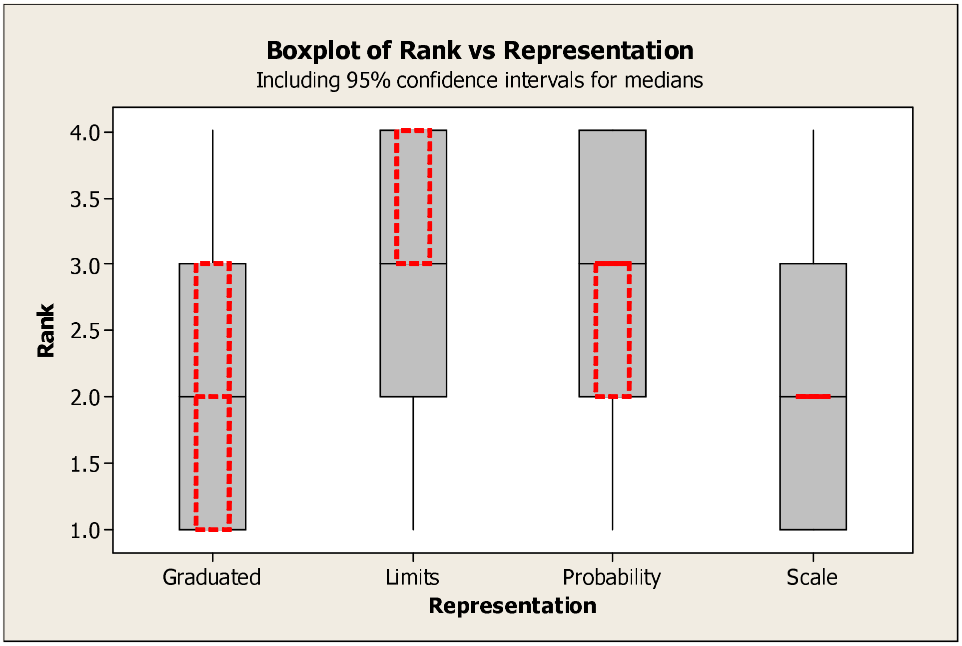


Figure 6.11. Boxplot showing the preference rankings assigned to the four types of representation, aggregated across experience levels and gender.

Since the ranked values are dependent, ordinal data and are not normally distributed, an analysis of variance is not appropriate for this data. A nonparametric Kruskal-Wallis test was therefore performed on the rankings, to investigate whether any of the representations were preferred over the others. This confirmed that there were significant differences in the rankings assigned to the four representations (DF = 3, H = 49.4, p<0.001, adjusted for ties). Wilcoxon signed rank tests between each pair of representations confirmed that Limits was significantly preferred to all other forms of uncertainty representation and also revealed a significant preference for Probability over Scale. Subjects' preference for Probability over Graduated did not quite reach significance at the 5% level:

	n	**Wilcoxon Statistic**	**P**	**Estimated Median**	**95% confidence interval**	
					Lower	**Upper**
lim-prob	100	3309.0	0.007	0.5000	0.00	1.00
lim-grad	100	3899.0	<0.001	1.000	0.50	1.50
lim-scale	100	4088.0	<0.001	1.000	0.50	1.50
prob-grad	100	3085.0	0.054	0.5000	0.00	1.00
prob-scale	100	3343.0	0.005	0.5000	0.00	1.00
grad-scale	100	2623.0	0.737	0.0000	0.00	0.50

Table 6.10. Results of Wilcoxon signed rank tests, including estimated median differences between each pair of representations and associated p-values.

These findings are summarized in the following figure, with the representations ordered by preference and the lines indicating non-significant differences:

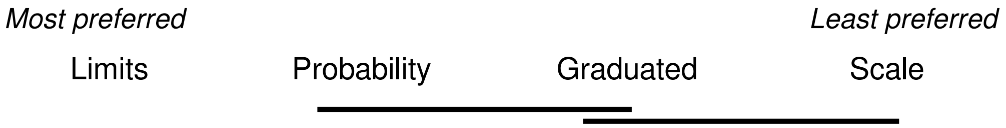


Figure 6.12. The four representation types ordered by preference, with non-significant differences indicated by lines.

6.3 Discussion

6.3.1 Dynamic testing of positional uncertainty

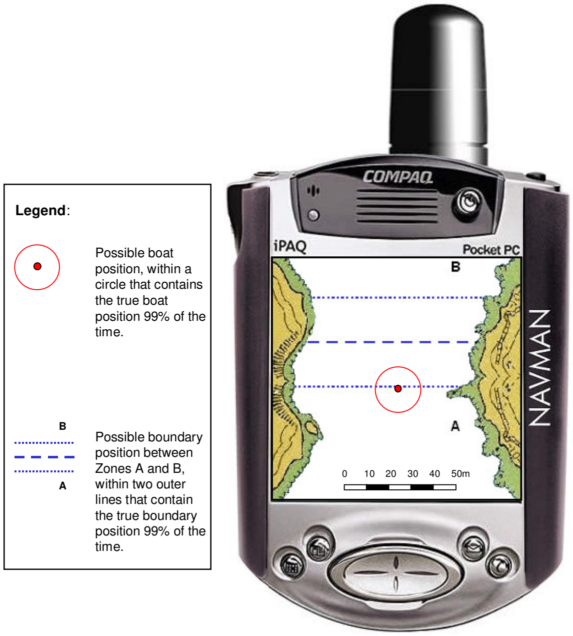
The interactions of experience*gender, experience*representation and representation*order were found to be significant at the 5% level. However, the experience*gender (p=0.042) and experience*representation (p=0.036) interactions may be chance results, partly due to the REML analysis under-estimating p-values. Study of the interactions plot shows that the interactions appear to be due to subjects of experience level 3 responding in a different manner to subjects of the other experience levels, interacting with both gender and representation. It could therefore be that these are chance interactions due to the small numbers, especially females, in the highly experienced group. Nonetheless, there is some evidence that subjects of differing experience levels may be responding to the four representations in different ways. Further research is required to confirm the presence of such an interaction.

The representation*order (p=0.013) interaction is less likely to be a chance result and suggests that the effect of representation on response differed according to the order in which the representations were viewed. Responses to the Limits representation varied significantly with order, subjects choosing to turn away from the restricted zone soonest when the representation was viewed fourth in the sequence of animations. There was no significant order effect for any of the other three representation types when considered individually. It could be considered that the prior animations set some context to the task, or provide an opportunity for learning, and the interaction indicates that this context has differential effects on the four representation types. This is an interesting possibility and study of the effects of context on spatial decision-making is an area requiring further research.

There was no main effect of either gender or experience on slide number, indicating that subjects responded in similar fashions irrespective of these two variables. However, there was a significant effect of order on subjects' responses (p=0.014). Other than the demonstration animation, to which subjects were not asked to respond, no prior training had been given to the subjects before they viewed the four animations. It is therefore not surprising to see an effect of order, as this may be a consequence of subject learning throughout the task. However, the order effect is not systematic. Subjects responding soonest to order 3, followed by 1, then 2 and finally order 4. This effect may also be a result of prior animations setting a context to the task and may be confounded with the representation*order interaction. It is difficult to provide a full explanation without further information.

There was extremely strong evidence that the type of positional uncertainty representation affects people's judgments as to when to turn away from the restricted zone (p<0.001). The predicted mean log(slide number) for the Limits representation was 3.174, corresponding to a slide number of 24, which was significantly lower than that for all other representations. The fact that subjects tended to turn away soonest when the uncertainty was represented using

Limits may be due to the physical nature of the boundary depiction in this case. The predicted mean slide number at which subjects turned away was at the point where they could be considered to be ‘crossing a line’ (see figure below); something which subjects appeared to be reluctant to do. The other three representations had no physical line drawn at this boundary limit and subjects were more willing to advance further towards Zone B in these cases.



Slide 24

Figure 6.13. Predicted mean slide for Limits representation.

One subject remarked that he had not wanted to cross the line when the uncertainty was represented in this manner, although he was happy to advance further toward Zone B when the boundary was ‘fuzzy’, as it was in the other forms of representation. Interestingly, subjects tended to advance farthest towards Zone B when the uncertainty was represented using Probability. For this representation, the predicted mean log(slide number) was 3.388, corresponding to a slide number of 30. This form of representation attempts to interpret the uncertainty information for the decision-maker and it could be argued that this would be the easiest to understand in terms of responsibility and potential consequences. However, subjects advanced furthest in this case. It may be that they do not fully comprehend how the other forms of representation correspond to the probability of being in Zone B and are therefore responding differently, particularly to the Limits representation of uncertainty.

It can be concluded that the form by which uncertainty information is portrayed can have an effect on the decisions being made. The fact that subjects tended to exhibit reluctance to cross a bounday line, albeit an uncertain limit, may be used by agencies attempting to enforce restricted zones. The use of 'fuzzy' representations of uncertain boundaries does not appear to act as such a strong deterrent, with subjects advancing further towards the restricted zone in these cases. This preliminary finding suggests that further research into the ways in which decision makers interact with different uncertainty representations is required.

6.3.2 Static multiple-choice tests

Participants' responses to the static representations of positional uncertainty were generally in agreement with the expected responses. Overall, 77.7% of responses were in agreement, indicating that, on the whole, subjects were able to understand the information portrayed in the positional uncertainty displays. Cohen's kappa statistic for the overall proportion of agreement, above that expected by chance alone, is 0.722. However, the kappa statistics for the cumulative data show that some forms were easier to understand than others. Agreement was extremely strong for Form 1 (K = 0.865), followed by Form 3 (K = 0.754) and Form 5 (K = 0.728). These values indicate that subjects' responses tended to be as expected when the boat was definitely in one of the two zones or had equal chance of being in either zone. However, the kappa values were lower for Form 2 (K = 0.647) and Form 4 (K = 0.612), indicating that subjects were not reliably responding as expected when the boat was probably in one of the two zones.

This might be due to a problem with the imprecision of language concerning probability rather than subjects misinterpreting the information conveyed by the displays. Karelitz and Budescu (2004) report on the potential for miscommunication when uncertain events are described using verbal probability terms and this may have been a factor in these findings. An example of a Form 2 question is shown in figure 6.14 for the Scale representation. The expected response for this question was *Probably in Zone A*. The most common response that did not agree with this was response 3, *Equal chance of being in either Zone*. This response could be considered reasonable, dependent on how rigidly we apply the term 'equal chance'. The boat was fairly close to the most likely boundary position in the Form 2 displays. It may be that participants considered the boat to be close enough to the expected boundary position that they believe the equal chance response to be the most appropriate.

Question 2.4

Figure 6.14. Example of a Form 2 question for Scale representation.

Similarly, the expected response for Form 4 was *Probably in Zone B*. The most common 'incorrect' response for this form was response 5, *Definitely in Zone B*. Inspection of the Scale representation for Form 4, shown below, reveals why this might be a reasonable answer. The boat is approximately 25 metres from the most likely boundary position, with a total positional uncertainty of both boat and boundary of up to 30m. Dependent on our understanding of the term 'definitely', the boat could be considered to be certainly in Zone B.

Question 2.7

Figure 6.15. Example of a Form 4 question for Scale representation.

Examination of the Cohen's kappa values by representation reveals that, not only are some question forms easier to understand than others, but so are some representation types. Subjects' responses were in stronger agreement with the expected responses when the Limits representation (K = 0.783) or Probability representation (K = 0.764) were used. The level of agreement was less when the representations of Scale (K = 0.675) or Graduated (K = 0.644) were used to depict the positional uncertainty. The results of a modified McNemar's test to determine the significance of these differences in agreement levels indicated that only the difference between the Limits representation and the Graduated representation was significant at the 5% level. However, the difference between the Limits and Scale representations and between the Probability and Graduated representations also approached significance (p=0.060 and p=0.072 respectively). It may be that the modification to the McNemar's test to account for the dependence in the data may have reduced the sensitivity of the test.

The different types of representation appeared to differ in their reliability of depicting uncertain information according to the form of the question. Subjects' responses were in least agreement with the expected responses for Form 4 when the Limits, Scale and Probability representations were used. As discussed above, this may be partly due to the vagueness surrounding probabilistic language. However, subjects' responses were in least agreement with Form 2, followed by Form 3, when the Graduated representation was used. In fact, closer inspection of the results tabulated by representation type reveals that the responses to the Graduated representation were quite disparate compared to those of the other three types of representation. The Graduated representation appears to be the most confusing to subjects. In support of this, it is also notable that in five of the six occurences that the subject responded that they did not understand the diagram, the Graduated representation was being used to depict the information (p=0.005). It should be noted that there was a mistake in the Graduated display, with the shading not extending into the region considered to be definitely in Zone B. It may be that this mistake caused some confusion, although only one participant questioned this display.

An attribute agreement analysis was also performed on the data when it was tabulated by subjects' experience level. The Cohen's kappa values indicate differences in the levels of agreement, according to subject experience. The responses of experience level 2 subjects were generally in better agreement with the expected responses than those of the other two experience levels. This would be expected when comparing experience level 2 (k = 0.776) with the novice subjects of experience level 1 (k = 0.698). However, it was also evident when comparing level 2 with the more experienced level 3 subjects (k = 0.658). These most experienced subjects had particularly low agreement with the expected responses for questions of form 4 (k = 0.407). As previously argued, this may be attributable to the imprecision of language concerning probability. It could also be that subjects with the greatest experience with GIS have incorporated uncertainty into their subjective definitions of terms like 'probably' and 'definitely'. This is an area that requires further research before any definite conclusions can be drawn.

A possible effect of gender on the agreement between subjects' responses and the expected responses was also investigated. However, it was found that both genders exhibited extremely similar levels of agreement. Cohen's kappa value for female subjects was 0.719 and that for male subjects was 0.723. As observed in the responses to the animations, females and males reacted in similar fashions to the static representations of positional uncertainty.

6.3.3 Survey of participants' preferred representation type

There was strong evidence that subjects preferred some representations over others, irrespective of their experience level or gender. The Limits representation was ranked significantly higher than its nearest contender, Probability. The Probability representation was not significantly preferred to the Graduated representation, although it was ranked significantly higher than the least-preferred Scale representation. The preferences assigned to the Probability and Graduated representations were more variable than those assigned to Limits and Scale.

It is interesting to note that the least preferred representation, Scale, is the one that offers no attempt to either visualise the uncertainty information or to interpret it. This form of representation is effectively the provision of metadata concerning positional uncertainty and is typical of current practice. The strong preferences displayed for other forms of uncertainty representation reflect the need for an improved method of handling uncertainty in GIS. One subject reported that he liked both the Limits and the Probability displays and would be interested to see both offered to the decision-maker. Such a display would provide both a visualisation of the uncertainty information and an attempt to interpret it. However, it would be interesting to see how subjects would respond to an integrated approach, seeing as the responses to these two forms of representation were the most disparate.

It is also noteworthy that this order of subjective preferences is very similar to the ordering obtained using agreement between subjects' response and the expected response. Subjects prefer the Limits representation and are also in greatest agreement with the expected response when the Limits representation is used. The Probability representation is second best on both scales; whilst there is little difference between subjective preferences for Scale and Graduated, despite the lowest agreement evident when the Graduated rerpesentation is used.

6.4 Conclusion

The results of this study provide evidence that the form of representation of positional uncertainty can have significant effects on the decisions being made. The observed differences in subjects' responses to the dynamic displays were strongest between the representations of Limits and Probability, despite there being no significant differences in subjects' apparent ability to comprehend the information provided by these two representations. Users also report strong subjective preferences for certain representations of positional uncertainty over others, which generally corresponded to the level of agreement noted between subjects' responses and expected responses.

7.

Conclusions and Recommendations

This thesis has assumed the premise that all spatial datasets are representations of a reality and, as such, will contain inherent errors. In accordance with the established classification of data errors into those of collection, processing and usage, the potential sources of such errors have been discussed. Recently, standards such as the SDTS have been introduced for reporting metadata of spatial datasets along the quality dimensions of positional, thematic and temporal accuracy, completeness and logical consistency. However, it is argued that these do not adequately convey the level of uncertainty associated with a dataset to many potential users.

Academics have argued for visualization as a means of communicating the uncertainty associated with spatial information, contending that it particularly lends itself to the portrayal of the spatial variability of uncertainty. The NCGIA research initiative, principally concerned with the use of visualisation methods to display the level of uncertainty associated with a dataset, aimed to prioritise a research agenda into the associated issues. In addition, of equal importance to the decision-maker is the representation of uncertainty in the final product of a GIS. The research arising from the NCGIA initiative has primarily focused on quantifying the level of uncertainty in GIS output and in developing visualisation methods to represent this uncertainty. However, there has been little research that has empirically tested the effectiveness of these visualisation methods and, more recently, this has been identified as a major research challenge.

Most of the cognitive testing of uncertainty representations to date has considered whether or not the inclusion of uncertainty information interferes with the user's ability to comprehend the underlying thematic information. Although researchers have concluded that several visualisation methods can be integrated with thematic displays without interference, these have tended to use a binary classification of uncertainty level which is rather simplistic in nature. These studies have also required little interpretation of the uncertainty information. It may be that users can read uncertainty displays but we must also ask what they are to make of this information and, most pertinently, how it is to affect their decision-making.

If researchers are claiming that it is essential that the level of uncertainty associated with a GIS product be communicated to the decision-maker, it follows that we must investigate how decision-makers might use this information. A recent survey of GIS users indicated that the majority of uncertainty visualisation methods are considered to be of no use whatsoever. This finding calls into question the assumption that the inclusion of uncertainty information will lead to improved, fully informed decisions. Instead, it invites us to ask exactly what it is that we are expecting decision-makers to do with uncertainty information when it is included in a GIS output.

Studies from the psychological literature indicate that most people display ambiguity aversion and tend to under-value options that are classified as uncertain. This is a powerful bias that often remains even when the apparent irrationality of such behaviour is overtly pointed out to the decision-maker. It could be argued that such conservatism is acceptable and should be incorporated into our definition of a rational decision, just as expected utility has been used to replace expected value. However, if this is the case, we still need to be aware of this bias in the options that decision-makers may select. By explicitly including a representation of the uncertainty associated with a spatial information, it would be expected that the tendency towards ambiguity aversion will be exacerbated and those classifications that are labelled as uncertain will be under-valued by decision-makers.

This hypothesis was tested within the first of two experiments designed as part of the current research. The case study investigated the effects of introducing thematic uncertainty information on decision-making. Participants were asked to judge which of two land parcels they would choose as the potential site of a new airport, given a land suitability classification and the level of uncertainty associated with this classification for each parcel. They were also asked to rank six land parcels that differed in both land suitability and the associated level of uncertainty. The results showed an extremely significant tendency for participants to select the land for which the classification was of high certainty. This finding is consistent with the concept of ambiguity aversion and confirms that many people under-value options that are labelled as low certainty in spatial data.

Although a tendency to display ambiguity aversion could be considered rational in terms of conservatism, one of the results obtained in the airport siting experiment can only be deemed to report irrational behaviour. In selecting between land parcels of the lowest suitability, which differed only in terms of the uncertainty level associated with this classification, over 45 percent of those participants exhibiting a preference selected the high certainty zone. In this comparison, the low certainty zone must be the rational choice since no parcel could be worse than land that is definitely of the lowest land suitability. However, the reluctance to select the low certainty zone exhibited by nearly one half of participants demonstrates that many people may not respond rationally to uncertainty information. The widely held assumption that the inclusion of such information will lead to better decision-making may be fundamentally flawed.

The second experiment implemented within the current research considered positional uncertainty and had the aim of assessing whether or not participants respond in a consistent manner to different representations of the same uncertainty information. The study was designed within the context of GPS-guided navigation, with decision-makers deciding at which point they would turn away as a boat approached a restricted zone. The positional uncertainties of both the boat and boundary positions were represented in different ways, including (i) *Graduated*: graduated shading of a transition zone, (ii) *Limits*: portrayal of the outer-most boundaries of the 99% positional confidence interval, (iii) *Probability*: a probability statement concerning positional uncertainty and (iv) *Scale*: a metadata statement concerning positional uncertainty.

The results indicated that there were significant differences in participants' responses to the representations, although the same information was being portrayed in the four displays. Participants tended to turn away significantly sooner when the Limits representation was used to display the positional uncertainty than when any of the other three representations were used. This tendency was evident across the three experience levels and both genders. It is noteworthy that the median response time to the Limits representation corresponded to the position at which the boat first crossed the outer boundary line. It appears that participants exhibited a reluctance to cross a physically displayed line; this reluctance was less evident when the boundary demarcation was fuzzy.

It could be argued that the differences in subjects' responses to the various representation methods are a consequence of a lack of understanding of some of the display methods. However, an assessment of the agreement between participants' responses and the expected responses indicated that agreement was relatively strong for all four display methods, indicating that participants generally understood the information that was being presented. The only significant difference was the greater level of agreement for the Limits representation than the Graduated representation, which may be attributable to the confusion caused by the mistaken colouring of the 100% zone in the Graduated representation. Participants were also asked to rank their preferences for the four representation methods. The results indicated a strong overall preference for the Limits representation and a significant preference for the Probability representation over the other least preferred Scale display. This preference pattern corresponded to the trend in the kappa results assessing comprehension of the displays, suggesting that participants tended to most easily understand those representations for which they also showed the greatest preference.

The results from both the experiment testing the effects of introducing thematic uncertainty information and the experiment investigating different representations of positional uncertainty information were extremely significant. They suggest that great care must be taken in incorporating the level of uncertainty associated with spatial information. Decision-makers do not necessarily respond in a rational manner to the inclusion of uncertainty information and can make significantly different decisions dependent upon the form by which the information is portrayed. Academics may be guilty of having made too great an assumption in claiming that the visualisation of uncertainty will lead to better decision-making.

However, these are preliminary results from two experiments. The decision-making tasks were highly specific and the findings need to be verified through additional cognitive testing of many different tasks if the results are to be generalised to decision-making with spatial information as a whole. A decision-making context was provided for both experiments but it is not clear whether this was sufficient to immerse participants in the tasks. The effects noted may be relatively bottom-up perception effects rather than being characteristic of considered decision-making. In addition, the participants in these experiments were students. Cognitive testing of GIS users in their workplace needs to be conducted, to investigate whether or not the same findings are typical of experienced decision-makers.

Nonetheless, these results raise some important questions that are worthy of further research. The inclusion of thematic uncertainty information resulted in participants under-valuing land that was labelled as low certainty and sometimes making irrational decisions. Can thematic uncertainty information be provided in a manner that does not lead to this bias? Perhaps a bar chart or histogram showing the probability of a zone belonging to each class would be more useful to decision-makers than a binary representation of uncertainty. Alternatively, an animation in which the zone's classification changes over time, according to this probability could replace the bar chart. Such representations require testing to see whether they are effective in reducing the bias against uncertainty.

These displays also require testing as to whether or not they remain comprehensible. It may be that more complicated representations of uncertainty information are required if the decision-maker is to be truly informed. However, it is probable that such displays will become confusing if incorporated into thematic maps and it may be that they will need to remain separate from the underlying information. If this is the case, the question remains as to how the user is to be alerted to the issue of uncertainty in GIS output. Perhaps this issue can be solved through a greater emphasis on educating spatial data users in the concept of uncertainty and including message boxes or similar means in GIS interfaces to attract users' attention to uncertainty information.

The effects on decision-making of including types of uncertainty information other than thematic also need to be investigated. In this study, the manner in which positional uncertainty information was displayed had significant effects on the decisions being made. Further testing is required to see if the physical depiction of a boundary line consistently has effects that differ to more fuzzy representations of boundaries. If this is the case, the appropriateness of the different styles for particular applications is an area in need of research. The possible effects of wording and context on spatial decision-making also warrant further investigation.

A strong preference was stated in this study for the representation of uncertainty information using the visualisation method *Limits* or the interpretative method *Probability*. The least preferred representation was the *Scale* method, which was most similar to the metadata reports in current use. This suggests that visualisation and interpretative representations are of greater potential use in communicating uncertainty information than metadata statements. It may be possible to combine both visualisation and interpretative methods into a single, optimum representation, although further research is required to investigate this possibility.

In conclusion, the research conducted within this thesis has indicated that the inclusion of uncertainty information in spatial data may affect decision-making in ways that have not generally been given a great deal of consideration by academics. Participants exhibited a strong tendency towards ambiguity aversion when thematic uncertainty was explicitly represented in a GIS output. This tendency to select a *high certainty* region over a *low certainty* one was so powerful that many participants made decisions that can only be described as irrational. This finding is contrary to the assumption that the inclusion of uncertainty information in spatial data will lead to better decision-making.

In addition, the study concerning different representations of positional uncertainty indicated that decision-makers may behave differently to the same information, according to the nature of its representation. This finding raises questions as to how uncertainty may be most appropriately represented for the application at hand. Participants also exhibited strong preferences for certain styles of representation over others. The least preferred style was most similar to the metadata statements currently used to communicate uncertainty. Subjective preferences for the more visual and interpretative representation methods indicate that further research is required to identify the most effective means of communicating the uncertainty associated with GIS output in a manner that is truly informative to decision-makers.

Bibliography

Aerts, J.C.J.H., Clarke, K. C. and Keuper, A. D. (2003) ‘Testing popular visualization techniques for representing model uncertainty’, *Cartography and Geographic Information Science*, Vol.30(3), pp.249-262.

Agumya, A. and Hunter, G. J. (1999) ‘A risk-based approach to assessing the ‘fitness for use’ of spatial data’, *Journal of the Urban and Regional Information Systems Association*, Vol.11(1), pp.33-44.

ANSI (American National Standards Institute) (1998) *National Committee on Information Technology Standards 320 (Spatial Data Transfer Standard)*, US Department of Commerce, Washington D.C.

ASPRS (American Society for Photogrammetry and Remote Sensing) (1990) Specifications and Standards Committee, ‘ASPRS Accuracy Standards for Large-Scale Maps’, *Photogrammetric Engineering and Remote Sensing*, Vol.56(7), pp.1068-1070.

Anderson, N. H. (1981) *Foundations of information integration theory,* New York: Academic Press.

Andrienko, G. L. and Andrienko, N. V. (1999) ‘Interactive maps for visual data exploration’, *International Journal of Geographical Information Science*, Vol.13(4), pp.355-374.

Beard, K. (1989) ‘Dimensions of use and value of geographic information’, in *Use and Value of Geographic Information Initiative Four Specialist Meeting Report and Proceedings 89-7*, eds. Onsrud, H. J., Calkins, H. W. and Obermeyer, N. J., NCGIA, University of California, Santa Barbara.

Beard, M. K., Buttenfield, B. P. and Clapham, S. B. (1991) *Visualization of Spatial Data Quality, Scientific Report for the Specialist Meeting*, National Centre for Geographic Information and Analysis, Castine, Maine, NCGIA Research Initiative 7.

Beard, K. and Mackaness, W. (1993) ‘Visual access to data quality in geographic information systems’ *Cartographica*, Vol.30(2-3), pp.37-45.

Bernoulli, D. (1967) *Exposition of a new theory on the measurement of risk*, Farnsborough Hants: Gregg Press (original work published in 1738).

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

Blakemore, M. (1985) ‘High or low resolution? Conflicts of accuracy, cost, quality and application in computer mapping’, *Computers and Geosciences*, Vol.11, pp.345-348.

Burrough, P. A. and McDonnell R. A. (1998) *Principles of Geographical Information Systems*, New York: Oxford University Press.

Buttenfield, B. and Beard, M. K. (1994) ‘Graphical and geographical components of data quality’, in *Visualization in Geographical Information Systems*, eds. Hearnshaw, H. M. and Unwin, D. J., New York: J. Wiley & Sons, Chap.16, pp.150-157.

Buttenfield, B. P. (1993) ‘Representing data quality’, *Cartographica*, Vol.30(2), pp.1-7.

Buttenfield, B. P. and Ganter, G. H. (1990) ‘Visualization and GIS: What should we see? What might we miss?’, in *Proceedings of the 4th International Symposium on Spatial Data Handling*, Zurich, Switzerland, Vol.1, pp.307-316.

Buttenfield, B.P. and Mackaness, W.A. (1991) ‘Visualization’, in *Geographical Information Systems: Principles and Applications*, eds. Maguire, D., Goodchild, M. and Rhind, D.W., New York: Wiley, Chap.28, pp. 427-443.

Casner, S. (1989) ‘A task-analytic approach to the automated design of information graphics’, *Carnegie Mellon University Technical Report AIP-82*.

Cheung, C. K. and Shi, W. Z. (2004) ‘Estimation of the positional uncertainty in line simplification in GIS’, *Cartographic Journal*, Vol.41(1), pp.37-45.

Cheung, C. K., Shi, W. Z. and Zhou, X., (2004) ‘A probability-based uncertainty model for point-in-polygon analysis in GIS’, *GeoInformatica*, Vol.8(1), pp.71-98.

Chrisman, N. (1997) *Exploring Geographic Information Systems*, New York: J. Wiley & Sons.

Chrisman, N. R. (1995) ‘Beyond Stevens: A revised approach to measurement for geographic information’, in *Proceedings of AUTO-CARTO 12*, Vol.4, Charlotte, North Carolina, pp.271-280.

Clapham, S. B. and Beard, K. (1991) ‘The development of an initial framework for the visualization of spatial data quality’ in *Proceedings of AUTO-CARTO 10*, Baltimore, Maryland, pp.73-82.

CNN.com (1998) ‘Report: U.S. military lapses caused Italian cable car accident’, URL http://www.cnn.com/WORLD/9802/18/italy.cable.car/ last accessed 29/11/04.

Congalton, R. G. (1991) ‘A review of assessing the accuracy of classifications of remotely sensed data’, *Remote Sensing of Environment*, Vol.37(1), pp.35-46.

Crosetto, M. and Tarantola, S. (2001) ‘Uncertainty and sensitivity analysis: tools for GIS-based model implementation’, *International Journal of Geographical Information Science*, Vol.15(5), pp.415-437.

Crossland, M. D., Wynne, B. E. and Perkins, W. C. (1995) ‘Spatial decision support systems: An overview of technology and a test of efficacy’, *Decision Support Systems*, Vol.14(3), pp.219-235.

Davis, T. J. and Keller, C. P. (1997) ‘Modelling and visualizing multiple spatial uncertainties’, *Computers & Geosciences*, Vol.23(4), pp.397-408.

Dawes, R. M. (1979) ‘The robust beauty of improper linear models in decision making’, *American Psychologist*, Vol.34, pp.571-582.

De Bruin S., Bregt, A. and Van De Ven, M. (2001) 'Assessing fitness for use: the expected value of spatial data sets', *International Journal of Geographical Information Science*, Vol.15(5), pp.457-471.

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

Drummond, J. (1995) 'Positional accuracy', in *Elements of Spatial Data Quality*, eds. Guptill, S. C. and Morrison, J. L., Oxford: Elsevier Science, Chap.3, pp.31-58.

Eddy, D. (1982) 'Probabilistic reasoning in clinical medicine: Problems and opportunities', in *Judgment under Uncertainty: Heuristics and Biases,* eds. Kahneman, D., Slovic, P. and Tversky, A., Cambridge: Cambridge University Press, Chap.18, pp.249-267.

Ehlschlaeger, C. R., Shortridge, A. M. and Goodchild, M. F. (1997) 'Visualising spatial data uncertainty using animation', *Computers & Geosciences*, Vol.23(4), pp.387-395.

Einhorn, H. J. and Hogarth, R. M. (1985) 'Ambiguity and uncertainty in probabilistic inference', *Psychological Review,* Vol.92, pp.433-461.

Einhorn, H. J. and Hogarth, R. M. (1986) 'Decision making under ambiguity', *Journal of Business,* Vol.59(4), pp.S225-S250.

Elith, J., Burgman, M. A. and Regan, H. M. (2002) 'Mapping epistemic uncertainties and vague concepts in predictions of species distribution', *Ecological Modelling*, Vol.157(2-3), pp.313-329.

Ellis, B. D. (1966) *Basic Concepts of Measurement*, Cambridge: Cambridge University Press.

Ellsberg, D. (1961) 'Risk, ambiguity and the Savage axioms', *Quarterly Journal of Economics*, Vol.75, pp.643-669.

Ellsberg, D. (2001) *Risk, Ambiguity and Decision*, New York: Garland Publishing.

Englund, E. (1993) 'Spatial simulation: Environmental applications' in *Environmental Modelling with GIS* eds. Goodchild, M. F., Parks, B. O. and Steyaert, L. T., New York: Oxford University Press, Chap.43, pp.432-437.

Erev, I. And Cohen, B. L. (1990) 'Verbal versus numerical probabilities: Efficiency, biases and the preference paradox', *Organizational Behavior and Human Decision Processes*, Vol.45, pp.1-18.

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

Fairbairn, D. G., Andrienko, G., Andrienko, N., Buziek, G. and Dykes, J. (2001) 'Representation and its relationship with cartographic visualization', *Cartography and Geographic Information Science*, Vol.28(1), pp.13-28.

Fennema, H. and Wakker, P. (1997) 'Original and Cumulative Prospect Theory: A discussion of empirical differences', *Journal of Behavioral Decision Making*, Vol.10, pp.53-64.

Fishburn, P. (1991) 'Nontransitive preferences in decision theory', *Journal of Risk and Uncertainty*, Vol.4(2), pp.113-124.

Fisher, P. (1993) 'Visualizing uncertainty in soil maps by animation', *Cartographica*, Vol.30(2-3), pp.20-27.

Fisher, P. (1994) 'Hearing the reliability in classified remotely sensed images', *Cartography and Geographical Information Systems*, Vol.21(1), pp.31-36.

Fraser, R., Collier, P. and Leahy, F. (2003) 'Positioning maritime boundaries with certainty - a rigorous approach', in *Proceedings of the International Conference on Addressing Difficult Issues in UNCLOS*, International Hydrographic Bureau, Monaco.

Frisch, D. and Baron, J. (1988) 'Ambiguity and rationality', *Journal of Behavioral Decision Making*, Vol.1(3), pp.149-157.

Goodchild, M. F. (1989) 'Modeling error in objects and fields', in *Accuracy of Spatial Databases*, eds. Goodchild, M. and Gopal, S., London: Taylor & Francis, pp.107-113.

Goodchild, M. F. (1991) 'Issues of quality and uncertainty', in *Advances in Cartography*, ed. Muller, J. C., New York: Elsevier Science, Chap.6, pp.113-139.

Goodchild, M. F. (1995) 'Attribute accuracy', in *Elements of Spatial Data Quality*, eds. Guptill, S. C. and Morrison, J. L., Oxford: Elsevier Science, Chap.4, pp.59-79.

Goodchild, M., Buttenfield, B. and Wood, J. (1994) 'Introduction to visualizing data validity', in *Visualization in Geographic Information Systems*, eds. Hearnshaw, H. M. and Unwin, D. J., New York: Wiley & Sons, Chap.15, pp.141-149.

Gordon, M. E., Slade, L. A. and Schmitt, N. (1986) 'The "science of the sophomore" revisited: From conjecture to empiricism' *Academy of Management Review*, Vol.11(1), pp.191-207.

Greenberg, J. (1987) 'The college sophomore as guinea pig: Setting the record straight', *Academy of Management Review*, Vol.12(1), pp.157-159.

Harrower, M. (2003) 'Representing uncertainty: Does it help people make better decisions?', URL http://www.ucgis.org/visualization/whitepapers/Harrower.pdf last accessed 29/11/2004.

Harrower, M., MacEachren, A. and Griffin, A. L. (2000) 'Developing a geographic visualization tool to support earth science learning', *Cartography and Geographic Information Science*, Vol.27(4), pp.279-294.

Hearnshaw, H. (1994) 'Psychology and displays in GIS', in *Visualization in Geographic Information Systems*, eds. Hearnshaw, H. M. and Unwin, D. J., New York: Wiley & Sons, Chap.21, pp.193-199.

Heo, J. (2003) 'Steepest descent method for representing spatially correlated uncertainty in GIS', *Journal of Surveying Engineering-ASCE*, Vol.129(4), pp.151-157.

Highhouse, S. and Hause, E. (1995) ‘Missing information in selection: An application of the Einhorn-Hogarth Ambiguity Model’, *Journal of Applied Psychology,* Vol.80, pp.86-93.

Ho, J. L. Y., Keller, L. R. and Keltyka, P. (2002) ‘Effects of outcome and probabilistic ambiguity on managerial choices’, *The Journal of Risk and Uncertainty*, Vol.24(1), pp.47-74.

Hogarth, R. M. and Kunreuther, H. (1985) ‘Ambiguity and insurance decisions’, *The American Economic Review*, Vol.75(2), pp.386-390.

Holmes, K. W., Chadwick, O. A. and Kyriakidis, P. C. (2000) ‘Error in a USGS 30-meter digital elevation model and its impact on terrain modeling’, *Journal of Hydrology*, Vol.233(1-4), pp.154-173.

Hunter, G. J. (1999) ‘Managing uncertainty in GIS’, in *Geographical Information Systems Volume 2: Management Issues and Applications*, eds. Longley, P. A., Goodchild, M. F., Maguire, D. J. and Rhind, D. W., New York: John Wiley & Sons, Chap.45, pp. 633-641.

Hunter, G. J. (1999) ‘New tools for handling spatial data quality: Moving from academic concepts to practical reality’, *URISA Journal*, Vol.11(2), pp.25-34.

Hunter, G. J. (2001) ‘Keynote address: Spatial data quality revisited’ *GeoInfo 2001 Workshop Brasileiro de Geoinformatica*.

Hunter, G. J. and Beard, K. (1992) ‘Understanding error in spatial databases’, *The Australian Surveyor*, Vol.37(2), pp.108-119.

Hunter, G. J. and Goodchild, M. F. (1995) ‘Dealing with error in spatial databases: A simple case study’, *Photogrammetric Engineering & Remote Sensing*, Vol.61(5), pp.529-537.

Hunter, G. J. and Goodchild, M. F. (1996) ‘Communicating uncertainty in spatial databases’, *Transactions in GIS*, Vol.1(1), pp.13-24.

Hunter, G. J. and Goodchild, M. F. (1997) ‘Modelling the uncertainty of slope and aspect estimates derived from spatial databases’, *Geographical Analysis*, Vol.29(1), pp.35-49.

Hunter, G. J., Qiu, J. and Goodchild, M. F. (1999) ‘Application of a new model of vector data uncertainty’, in *Spatial Accuracy Assessment: Land Information Uncertainty in Natural Resources*, eds. Lowell, K. and Jaton, A., Chelsea, Michigan: Ann Arbor Press, Chap.25, pp.203-208.

Jarvenpaa, S. L. (1989) ‘The effect of task demands and graphical format on information processing strategies’, *Management Science*, Vol.35, pp.285-303.

Johnson, E. J., Payne, J. W. and Bettman, J. R. (1988) ‘Information displays and preference reversals’, *Organizational Behavior and Human Decision Processes*, Vol.42, pp.1-21.

Kahneman, D. and Tversky, A. (1973) ‘On the psychology of prediction’, *Psychological Review*, Vol.80, pp.237-251.

Kahneman, D. and Tversky, A. (1979) ‘Prospect theory: an analysis of decision under risk’, *Econometrica*, Vol.47, pp.263-291.

Kardos, J., Benwell, G. L. and Moore, A. (2004) ‘Assessing different approaches to visualising spatial and attribute uncertainty in socioeconomic data using the hexagonal or rhombus (HoR) quadtree’, in *Proceedings of GIS Research UK 2004,* ed. Lovett, A., University of East Anglia, Norwich.

Kardos, J. D., Moore, A. and Benwell, G. L. (2003) ‘Visualising uncertainty in spatially-referenced attribute data using hierarchial spatial data structures’, in *Proceedings of the 7th International Conference on GeoComputation*, University of Southampton, United Kingdom. CD-ROM produced by Martin, D., Publisher: GeoComputation CD-ROM.

Karelitz, T. M. and Budescu, D. V. (2004) ‘You say “probable” and I say “likely”: Improving interpersonal communication with verbal probability phrases’, *Journal of Experimental Psychology-Applied*, Vol.10(1), pp.25-41.

Knapp, L. (1995) 'A task analysis approach to the visualization of geographic data', in *Cognitive Aspects of Human-Computer Interaction for Geographic Information Systems*, eds. Nyerges, T. L., Mark, D. M., Laurini, R. and Egenhofer, M. J., Boston: Kluwer Academic Publishers, pp.355-371.

Krygier, J. B. (1994) 'Sound and geographic visualization' in *Visualization in Modern Cartography*, eds. MacEachren, A. M. and Fraser Taylor, D. R., New York: Permagon, pp.149-166.

Leitner, M. and Buttenfield, B. P. (1997) 'Cartographic guidelines for visualizing attribute accuracy' in *Proceedings of AUTO-CARTO 13,* Seattle, Washington, pp.184-194.

Leitner, M. and Buttenfield, B. P. (2000) 'Guidelines for the display of attribute certainty', *Cartography and Geographic Information Science,* Vol.27(1), pp.3-14.

Lichtenstein, S. and Slovic, P. (1971) 'Reversals of Preferences Between Bids and Choices in Gambling Decisions', *Journal of Experimental Psychology,* Vol.89, pp.46-55.

Liu, W. G., Gopal, S. and Woodcock, C. E. (2004) ' Uncertainty and confidence in land cover classification using a hybrid classifier approach', *Photogrammetric Engineering and Remote Sensing*, Vol.70(8), pp.963-971.

Lloyd, R. (1997) *Spatial Cognition: Geographic Environments*, Boston: Kluwer Academic Publishers.

Lopes, L. L. (1993) 'Reasons and resources: The human side of risk taking', in *Adolescent Risk Taking*, eds. Bell, N. J. and Bell, R. W., Newbury Park, Calif.: Sage Publications, pp.29-54.

MacEachren, A. M. (1992) 'Visualizing uncertain information', *Cartographic Perspective*, Vol.13, pp.10-19.

MacEachren, A. M. (1995) *How Maps Work: Representation, Visualization and Design*, New York: Guilford Press.

MacEachren, A. M., Brewer, C. A. and Pickle, L. W. (1998) ‘Visualizing georeferenced data: representing reliability of health statistics’, *Environment and Planning A*, Vol.30(9), pp.1547-1561.

MacEachren, A. M. and Kraak, M.-J. (2001) ‘Research challenges in geovisualisation’, *Cartography and Geographic information Science*, Vol.28(1), pp.3-12.

MacGregor, D. and Slovic, P. (1986) ‘Graphical representation of judgmental information’, *Human-Computer Interaction*, Vol.2, pp.179-200.

Mackinlay, J. (1986) ‘Automating the design of graphical presentations of relational information’, *ACM Transactions in Graphics*, Vol.5(2), pp.110-141.

McFadden, D. (1981) ‘Econometric models of probabilistic choice’, in *Structural analysis of discrete data with econometric applications*, eds. Manski, C. F. and McFadden, D., Cambridge, MA: MIT Press, pp.198-272.

McGranaghan, M. (1993) ‘A cartographic view of spatial data quality’, *Cartographica*, Vol.30(2-3), pp.8-19.

Montello, D. R. (2002) ‘Cognitive map-design research in the twentieth century: Theoretical and empirical approaches’, *Cartography and Geographic Information Science*, Vol.29(3), pp.283-304.

Morrison, J. L. (1974) ‘A theoretical framework for cartographic generalization with the emphasis on the process of symbolization’, *International Yearbook of Cartography*, Vol.14, pp.115-127.

Mukerji, S. and Tallon, J. M. (2004a) ‘Ellsberg’s two-colour experiment, portfolio inertia and ambiguity’, *Journal of Mathematical Economics*, Vol.39(3-4), pp.299-315.

Mukerji, S. and Tallon, J. M. (2004b) ‘Ambiguity aversion and the absence of wage indexation’, *Journal of Monetary Economics*, Vol.51(3), pp.653-670.

NIST (National Institute of Standards and Technology) (1992) *Federal Information Processing Standard Publication 173 (Spatial Data Transfer Standard)*, U.S. Department of Commerce, Washington D.C.

Openshaw, S. (1989) 'Learning to live with errors in spatial databases' in *Accuracy of Spatial Databases*, eds. Goodchild, M. and Gopal, S., London: Taylor & Francis, Chap.23, pp.263-276.

Openshaw, S., Waugh, D. and Cross, A. (1994) 'Some ideas about the use of map animation as a spatial analysis tool', in *Visualization in Geographic Information Systems*, eds. Hearnshaw, H. M. and Unwin, D. J., New York: Wiley & Sons, Chap.14, pp.131-138.

Payne, J. W., Bettman, J. R. and Luce M. F. (1998) 'Behavioral decision research: An overview', in *Measurement, Judgment and Decision Making,* ed. Birnbaum, M. H., London: Academic Press, Chapter 5, pp.303-359.

Perkal, J. (1966) 'On the length of empirical curves', *Discussion Paper 10*, Michigan Inter-University Community of Mathematical Geographers, University of Michigan, Ann Arbor Press.

Petry, F. E., Cobb, M. A., Paprzycki, M. and Ali, D. (2003) 'An agent system for managing uncertainty in the integration of spatio-environmental data', *Soft Computing-A Fusion of Foundations, Methodologies and Applications*, Vol.7(6), pp.402-411.

Qiu, J. and Hunter, G. J. (2002) 'A GIS with the capacity for managing data quality information', in *Spatial Data Quality*, eds. Shi, W., Fisher, P. F. and Goodchild, M. F., London: Taylor & Francis, pp.230-250.

Reinke, K. (2002) 'Communicating thematic uncertainty in spatial information', unpublished Ph.D. thesis, The University of Melbourne, Australia.

Reinke, K. and Hunter, G. J. (2002) 'A theory for communicating uncertainty in spatial databases', in *Spatial Data Quality*, eds. Shi, W. Fisher, P. F. and Goodchild, M. F., London: Taylor & Francis, Chap.6, pp.76-101.

Russo, J. E. (1977) ‘The value of unit price information’, *Journal of Marketing Research*, Vol.14, pp.193-201.

Russo, J. E. and Dosher, B. A. (1983) ‘Strategies for multiattribute binary choice’, *Journal of Experimental Psychology: Learning, Memory and Cognition*, Vol.9, pp.676-696.

Savage, L. J. (1954) *The Foundations of Statistics*, New York: Wiley, pp.21-23.

Schkade, D. A. and Kleinmuntz, D. N. (1994) ‘Information displays and choice processes: Differential effects of organization, form and sequence’, *Organizational Behavior and Human Decision Processes*, Vol.57, pp.319-337.

Senay, H. and Ignatius, E. A. (1994) ‘A knowledge-based system for visualization design’, *IEEE Computer Graphics and Applications*, Vol.14(6), pp.36-47.

Simon, H. A. (1955) ‘A behavioural model of rational choice’, *Quarterly Journal of Economics,* Vol.69, pp.99-118.

Slocum, T. A., Blok, C., Jiang, B., Koussoulakou, A., Montello, D. R., Fuhrmann, S. and Hedley, N. R. (2001) ‘Cognitive and usability issues in geovisualization’, *Cartography and Geographic Information Science*, Vol.28(1), pp.61-75.

Slovic, P. (1972) ‘From Shakespeare to Simon: Speculations and some evidence about man’s ability to process information’, *Oregon Research Institute Research Monograph*, Vol.12(2).

Slovic, P. (1995) ‘The construction of preference’, *American Psychologist*, Vol.50, pp.364-371.

Slovic, P., Fischhoff, B. and Lichtenstein, S. (1982) ‘Facts versus fears: Understanding perceived risk’, in *Judgment under Uncertainty: Heuristics and Biases,* eds. Kahneman, D., Slovic, P. and Tversky, A., Cambridge: Cambridge University Press, Chap.33, pp.463-489.

Slovic, P., Monahan, J. and MacGregor, D. G. (2000) 'Violence risk assessment and risk communication: The effects of using actual cases, providing instruction, and employing probability versus frequency formats', *Law and Human Behavior*, Vol.24(3), pp.271-296.

Stevens, S. S. (1946) 'On the theory of scales of measurement', *Science*, Vol.103, pp.677-680.

Stone, D. N. and Schkade, D. A. (1991) 'Numeric and linguistic information representation in multiattribute choice', *Organizational Behavior and Human Decision Processes*, Vol.49, pp.42-59.

Suchan, T. A. (2001) 'Usability Studies of Geovisualization Software in the Workplace', URL http://www.digitalgovernment.org/library/library/pdf/suchan.pdf last accessed 30/08/2004.

Tufte, E. R. (1983) *The Visual Display of Quantitative Information*, Cheshire: Graphics Press.

Tufte, E. R. (1990) *Envisioning Information*, Cheshire: Graphics Press.

Tufte, E. R. (1997) *Visual Explanations: Images and Quantities, Evidence and Narrative*, Cheshire: Graphics Press.

Tversky, A. (1972) 'Elimination by aspects: A theory of choice', *Psychological Review*, Vol.79, pp.281-299.

Tversky, A. and Kahneman, D. (1973) 'Availability: A heuristic for judging frequency and probability', *Cognitive Psychology,* Vol.5, pp.207-232.

Tversky, A. and Kahneman, D. (1981) 'The framing of decisions and the psychology of choice', *Science*, Vol.211, pp.453-458.

Tversky, A. and Kahneman, D. (1982a) 'Judgment under uncertainty: Heuristics and biases', in *Judgment under Uncertainty: Heuristics and Biases,* eds. Kahneman, D., Slovic, P. and Tversky, A., Cambridge: Cambridge University Press, Chap.1, pp.3-20.

Tversky, A. and Kahneman, D. (1982b) 'Judgments of and by representativeness', in *Judgment under Uncertainty: Heuristics and Biases,* eds. Kahneman, D., Slovic, P. and Tversky, A., Cambridge: Cambridge University Press, Chap.6, pp.84-98.

Tversky, A. and Kahneman, D. (1991) 'Loss aversion in riskless choice: A reference-dependent model', *Quarterly Journal of Economics*, Vol.106, pp.1039-1062.

Tversky, A. and Kahneman, D. (1992) 'Advances in prospect theory: Cumulative representation of uncertainty', *Journal of Risk and Uncertainty*, Vol.5, pp.297-323.

Tversky, A., Sattath, S. and Slovic, P. (1988) 'Contingent weighting in judgment and choice', *Psychological Review*, Vol.95, pp.371-384.

Tversky, A., Slovic, P. and Kahneman, D. (1990) 'The determinants of preference reversal', *American Economic Review*, Vol.80, pp.204-217.

U.S. Bureau of the Budget (1947) *United States National Map Accuracy Standards,* U.S. Bureau of the Budget, Washington, D.C.

U.S. State Department Report on Accidental Bombing of Chinese Embassy (1999) URL http://hongkong.usconsulate.gov/uscn/state/1999/0706.htm last accessed 6/8/04.

Van Dijk, E. and Zeelenberg, M. (2003) 'The discounting of ambiguous information in economic decision making', *Journal of Behavioral Decision Making*, Vol.16(5), pp.341-352.

Van Niel, K. and Laffan, S. W. (2003) 'Gambling with randomness: the use of pseudo-random number generators in GIS', *International Journal of Geographical Information Science*, Vol.17(1), pp.49-68.

Veregin, H. (1989) *A Taxonomy of Error in Spatial Databases*. Technical Paper 89-12, U.S. National Centre for Geographic Information and Analysis, University of California, Santa Barbara.

Von Mises, L. (1949) ‘Rationality and irrationality; Subjectivism and objectivity of praxeological research’, in *Human action: A treatise on economics*, Chap.1, Sec.4, p.19, London: Hodge.

West, L. A. and Hess, T. J. (2002) ‘Metadata as a knowledge management tool: Supporting intelligent agent end user access to spatial data’, *Decision Support Systems*, Vol.32(3), pp.247-264.

Zhang, J. and Goodchild, M. F. (2002) *Uncertainty in Geographic Information*, New York: Taylor & Francis.

Author/s:
Hope, Susannah Jayne

Title:
Decision-making under spatial uncertainty

Date:
2005

Citation:
Hope, S. J. (2005). Decision-making under spatial uncertainty. Masters Research thesis, Department of Geomatics, The University of Melbourne.

Publication Status:
Unpublished

Persistent Link:
http://hdl.handle.net/11343/38960

File Description:
Decision-making under spatial uncertainty