CREDIBILITY ASSESSMENT AND LABELLING OF MAP MASHUPS

NURUL HAWANI IDRIS, BSc. MSc.

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy

JULY 2014

Abstract

The Web 2.0 revolution has changed the culture of mapping by opening it up to a wider range of users and creators. Map mashups, in particular, are being widely used to map variety of information. There is, however, no gatekeeper to validate the correctness of the information presented. The purpose of this research was to understand better what it is that influence users' perceived credibility and trust within a map mashup presentation and to support the future implementation of automated credibility assessment and labelling of map mashup applications.

This research has been conducted in three stages using mixed method approaches. The objective of the first stage was to examine the influence of metadata related to sources, specifically the map producer and map supplier, on respondents' assessment of the credibility of map mashup information. The findings indicate a low influence of the tested metadata and a high influence of visual cue elements on users' credibility assessment. Only half of the respondents used the metadata whilst the other half did not include it in their assessment.

These findings became the basis of stage two, which was to examine the influence of colour coded traffic light (CCTL) labelling on respondents' assessment of credibility. From the findings, the probability of respondents making informed judgements by choosing a high credibility map based on this rating label (CCTL) was three times higher than where only the metadata was presented.

The third stage was to propose a conceptual framework to support the implementation of automated credibility labelling for map mashup applications. The framework was proposed on the basis of thorough reviews from the literature. The suggested parameters and approaches are not limited to assess credibility of information in the map mashup context, but could be applied to other Web GIS applications.

A list of related publications

IDRIS N H, JACKSON M J, ISHAK M H I. 2014. A Conceptual Model of the Automated Credibility Assessment of the Volunteered Geographic Information. IOP Conference Series: Earth and Environmental Science, vol.18. *(scopus index proceeding)*

IDRIS N H, JACKSON M J, ISHAK M H I, FATHIN N N, MOHAMAD N S, ZAMRI I, MOHAMAD G H. 2013.Evaluating credibility of map mashup through automated metadata assessment. Paper presented at the International Symposium and Exhibition on Geoinformation (ISG), Kuala Lumpur. 24-25 September. *(peer reviewed proceeding)*

IDRIS N H, JACKON M J, ABRAHART, R J. 2011. Mashups: What looks good must be good? GIS Research UK Conference (GISRUK), Portsmouth UK.; pp.27-29. *(peer reviewed proceeding)*

IDRIS N H, JACKSON M J, ABRAHART, R J. 2011, Credibility Labelling in Map Mashup Design: Influence on Users' Judgement, EuroSDR/ISPRS Workshop on Web Cartography, Lund Sweden, 5th-6th May 2011 *(non-peer reviewed)*

IDRIS N H, JACKON M J, ABRAHART, R J. 2011. Colour Coded Traffic Light Labelling: An Approach to Assist Users in Judging Data Credibility in Map Mashup Applications. In Cidalia C. Fonte, Luisa Goncalves & Gil Goncalves (Eds), Proc. of the 7th International Symposium on Spatial Data Quality Coimbra Portugal: INESC Coimbra.pp. 201-206. *(peer reviewed proceeding)*

Acknowledgment

Praise to Allah, the Most Gracious and Most Merciful, for giving me the courage and guidance to proceed and complete this learning process and in putting the last dot in this thesis.

First of all, I would like to express my deepest gratitude to my supervisor Professor Mike Jackson; not only for this thesis, but also for develop some characteristics of how to become a good researcher. His naturally big picture thinking, great efforts to explain things clearly in simple ways, consistent supports, motivation and encouragements have been the driving forces in this journey. I am also indebted to Dr. Robert Abrahart for his guidance particularly at the early stage of this learning process. I'd also like to thanks to the internal and external examiners, Dr. Gary Priestnall (the School of Geography) and Professor Muki Haklay (UCL) for the constructive comments to improve this thesis.

Special thanks to my sponsors, Malaysia Ministry of Education and Universiti Teknologi Malaysia (UTM), for the financial support during my stay in the United Kingdom. I would also like to record my words of thanks to Professor Shahrum Ses (former UTM FGHT dean, 2007) and Assoc. Professor Mohamad Nor Said (Deputy Dean, 2014) for believe in me to complete this journey.

I am also indebted to many students at the University of Nottingham that had accessed to the university intranet portal and the School of Geography student mailing lists and joined in my online surveys dated back in 2010 and 2011. I owe my sincere thanks to my friends in Malaysia and my Malaysian friends in the Nottingham and Loughborough to participate in my series of surveys. Without your contributions, there is nothing to write in this thesis.

My colleagues and the staff at the Nottingham Geospatial Institute (NGI) (formerly known as Centre for Geospatial (CGS) and at the School of Geography have been a source of inspirations to me. I would like to convey sincere thanks to Ngah Zaifa, Sergiusz Pawlowicz, Suchith Anand, Amir Pourabdollah, Shittu Whanda, Sani Yahya and the Nottingham and Loughborough Malaysian communities for keeping me to act like a normal people throughout this bumpy journey.

Last but not least, special thanks to Stella Fuller, John Miller, Jonathan Walton, Andrea Payne for administrative supports, CELE teachers for language supports, graduate school courses' instructors and all individuals who have directly and indirectly offered help, suggestions and support in bringing towards the completion of this thesis. Thank you so much.

This thesis is dedicated to…

My husband Hafis Izran … it was indeed a very long journey to make our dream come true

And for my 8th years old Melati Aisyah, 7th years old Khalil Imran and 1th year old Addeen…when you are big enough to understand this is for you, darlings!

And to my parents, abah and mama without you I am nothing. Love both of you …for all the things that you have done right.

And to my BFF Zaifa, Izham and Wan Najibah (and family) thanks for saying and make it looks okay

Last but not least, to my little brothers, Syafiq and Dek Pis, thank you for babysitting them. 'Jasamu dikenang'!

Table of Contents

LIST OF TABLES

LIST OF FIGURES

LIST OF CHARTS

LIST OF EQUATIONS

1 INTRODUCTION

1.1 Research Background

The Web has been a medium for the dissemination of information since the time of its invention. The current Web applications are now in the era of Web 2.0. This term was coined by Tim O'Reilly on 30 September 2005 at the first Web 2.0 Conference, to describe the recent trend of innovative and collaborative applications on the Web (Haklay et al., 2008, p.2). Through this revolution, the Web acts as a one stop resource centre that is not limited to information and knowledge discovery. It enables users to connect with communities in social-networking applications, to participate in cyber communities to share intelligence, to be citizen journalists and act as a neogeographer (the term used to describe the trend of amateur citizen geographers on the Web).

This revolution has made big impact on the culture of mapping and the new 'geo' landscape incorporates aspects described by a number of new terms and concepts, including Web Mapping 2.0 (Haklay et al., 2008), neogeography (Turner, 2006) and volunteered geographic information (VGI) (Goodchild, 2007a). Under this revolution, professional geoliterate users are not the only group that are active in mapping activities. Neogeographers, a term used to also include persons that have very little knowledge of the principles of cartography and geographic information sciences and, without any formal training, have discovered the importance of mapping, and hence become both the suppliers and consumers of geographic or location-based data (Goodchild, 2008). This group supplies user-generated spatial content (UGSG) which include contents made publicly available over the Internet and sometimes created outside professional routines and practices (Ochoa and Duval, 2008). The sources of data may draw either from localised individual information or input from the expert domain, including government officials, business owners or environmentalists (Wilson and Graham, 2013, p.13).

Neogeography is not merely about the production of information, but also includes map creation, personal analysis, interactive feedback, collaboration, and the reading and understanding of geographic information. As such, neogeography is not really a reinvention, but a domain of new possibilities that are now approachable by anyone (Wilson and Graham, 2013, p.13). The rise of neogeography activities can be seen by the popularity of map mashup development and the impressive progress of the OpenStreetMap projects. A map mashup is a web mapping application that uses commercial and open source map data such as Google Map for a base map, whilst foreground data are added on the basis of individuals' local information or from other sources including open data services. OpenStreetMap (Haklay and Weber, 2008) is a crowd source based project to develop a free world map for everyone where the major contributors are drawn from neogeographers.

This new mapping landscape has been further motivated by the Digital Earth vision of the former United States (US) Vice President, Al Gore, as delivered in his speech to the California Science Centre in 1998. The vision is to develop a Digital Earth, which is a multi-resolution, three dimensional representation of this planet, into which can be embedded vast quantities of geo-referenced data. The premise of this vision is to capitalise on the advantages of the flood of geospatial information, much of it unused, in order to satisfy the desire for information and knowledge from citizens, including both government and private sectors (Gore, 1998, p.1). The launch of Keyhole's Earth Viewer in 2001 was the first major step by the commercial sector in putting this vision into practice (Craglia et al., 2008). That company was acquired by the Google in October 2004 and rebranded as Google Earth in 2005, which is the free mapping program that successfully achieved 100 million product activations in the first year after being released (Grossner, 2006). Google Earth and Google Maps are the pioneers of the map mashup generation (Goodchild, 2007a).

The emergence of map mashup technology is one of the outputs driven by this vision. The term mashup was originally used to describe the blending of musical tracks to create new forms of song; the term now refers to websites that weave data from different sources into new integrated user services (Batty et al., 2010, p.2). At the time of writing, more than 2000 map mashup applications are identified by the Programmableweb (2013) website portal; mapping is one of the dominant themes that deploy mashup technology as shown in Figure 1.1. Google Map APIs have become the most frequently used APIs to deploy map mashup applications.

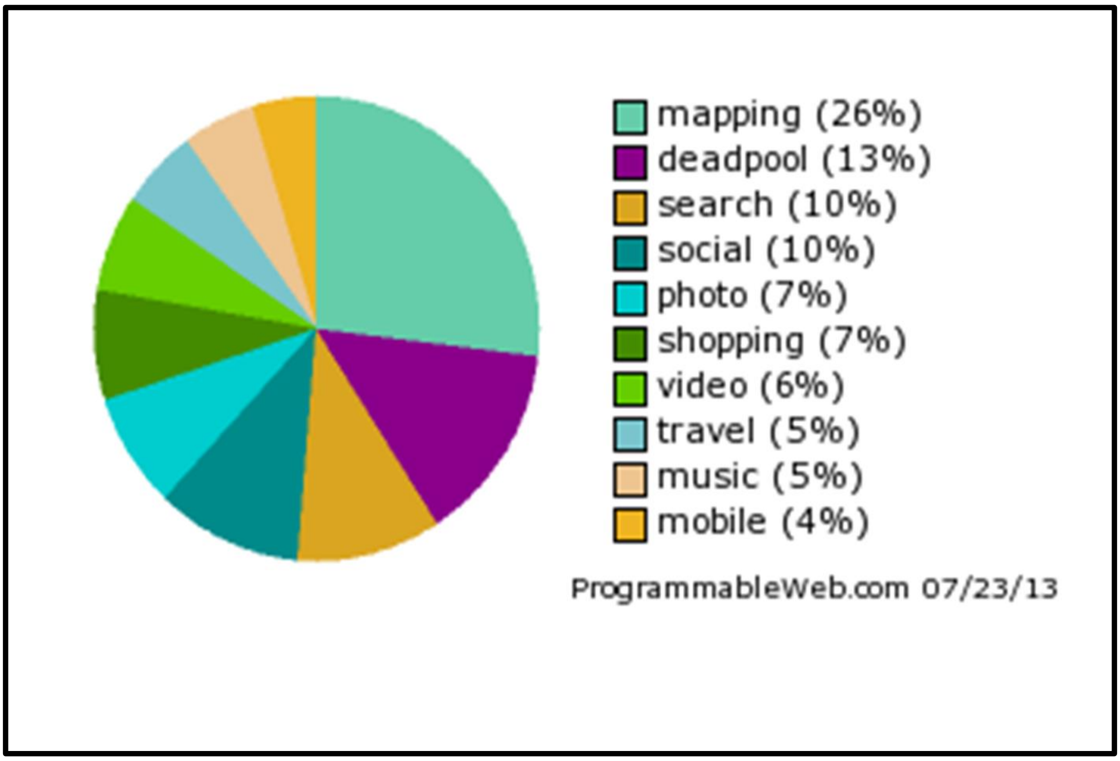


Figure 1-1 The themes that deploy mashup

The adoption of mashup technology by professional and amateur developers to easily disseminate and share geospatial data has increased the availability of location based information on the Web. Moreover, the Pew Research Centre (Fox, 2006) has indicated that online users turn to online news ‘to get information that is unavailable elsewhere, for convenience and for the ability to search for news on a particular topic’. The availability and the needs might explain why map users turn to a map mashup for location-based information. The incapacity of the national mapping providers in providing relevant and timely data to its citizen and to make all of its data publicly available might one of the reasons of this evolution (Caquard, 2014). Although there is no quality standards or gatekeepers to control and guarantee the correctness or accuracy of the data and information presented on the medium, the applications of map mashups include the dissemination of trivia information (e.g. celebrity mashups), consumer based and enterprise linked information (e.g. store locator mashups), to that relating to the news and current events (e.g. crisis mashups).

1.2 Research Problem

In this new mapping landscape, the nature of users' perceived credibility towards information presented on online maps is changing. Users now not only use authorised data sources but also use user-generated spatial content (UGSC) sources. This new type of data source is becoming more practical due to its accessibility and locality coverage. As argued by Goodchild and Glennon (2010), during emergency situations, authoritative agencies sometime lack of resources and faced with imperfect communications to share and update the current situation, since data from them need to be verified before disseminate to the public. This is to avoid a potential panic situation and any unnecessary actions and evacuations; however the impact of a 'false negative' (when the data is true, but reported as false) when there is a delay in acquiring available data from official sources is high and could make a different between life and death situation; whereas in the case of data from crowd sources and VGI, which carries the risk of 'false positives' (when the data is false, but reported as correct), the cost of acting on this assumption is much lower compared to the cost of not acting in response to 'false negatives'. Although there is a trade-off in terms of quality and prone to errors, users tend to use and 'believe' this information. Users are conscious of the need to balance the rapid availability of VGI with the unverified nature of much of its content (Goodchild and Glennon, 2010, p.238).

One reason for many mainstream map providers disregarding user-generated sources of data is the perceived lack of quality control. The lack of control over lineage information which leads to unknown reliability and trustworthiness are the main challenge for the authorities to trust such data (Spinsanti and Ostermann, 2013). Notwithstanding, this type of data can be used to complement authoritative data as well as to support other data in a low accuracy application and a low risk situation. Such data have been used during a crisis when the official sources of data are inevitably delayed (Liu and Palen, 2010). As argued by Elwood et al. (2012), during a crisis, decision makers must make choices between acting immediately with questionable data and waiting for better data to arrive; these data have considerable potential to help initiate many actions during a crisis, despite the risks of false information from inaccurate data. Therefore, the issue of data quality and reliability has to be tackled thoroughly in a way that is not limited to the technical aspect of data accuracy, but examines the overall quality control and trust issues of data generated by this approach.

Several studies have highlighted the issues of credibility and examined approaches to support the use of user generated contents or VGI in supplementing authoritative data sources. For example, the credibility of VGI (Flanagin and Metzger, 2008) and their accuracy (Goodchild, 2008) have been discussed in the literature. Goodchild and Li (2012) proposed three general but practical approaches to validate VGI which include crowd source, social and geographic approaches; Spinsanti and Ostermann (2013) proposed one mechanism to gather and validate user generated data from social media such as Twitter and Flickr using machine learning and spatial clustering approach to complement authoritative data during crisis events. Mummidi and Krumm (2008) have proposed a method to address the quality problem which involves the development of a data mining algorithm to search through a large collection of VGI and assess the data/information in the form of user-supplied annotated pushpins that are consistently repeated. In this study the correct name of a place of interest (POI) is chosen on the basis of a simple cluster of pushpins that may represent a similar POI. Bishr (2007) has designed a spatial trust model to assess the value of data contributed collaboratively by examining the reputation of authors; in this study, the correctness of data/information presented depends on the reputation rating of the authors. Elwood et al. (2012) is in line with the previous study which highlighted the approach of checking feasibility and reasonableness of VGI against other information and the evaluation based on reviews and commentaries by varying numbers of peers.

Research related to credibility has become of interest in several domains, including communication, information science, marketing, psychology, interdisciplinary efforts in human computer interaction (HCI) and currently Web 2.0 applications. For example, several studies have investigated the credibility of Wikipedia (Luyt and Tan, 2010), blogs (Juffinger et al., 2009), and twitter (Al-Khalifa, 2011). Issues of credibility in user generated spatial contents, such as OpenStreetMap and the free web mapping service applications such as in Google Earth, have also been raised by several authors, including Flanagin and Metzger (2008) and Goodchild (2008). Credibility is synonymous with believability (Fogg and Tseng, 1999, Flanagin and Metzger, 2008). The main primary dimensions of credibility discussed in the literature which hold by this research are trustworthiness and expertise (Fogg and Tseng, 1999). As argued by Rieh and Danielson (2007, p.314), 'credibility refers to a perceived quality of a source (website), which may or may not result in associated trusting behaviour; trust frequently refers to a set of beliefs, dispositions, behaviours associated with the acceptance of risk and vulnerability'. The trust concept is defined more comprehensively than the credibility concept. Notwithstanding, assessing credibility is a pre-requisite before users may generate trust in the object. In traditional approach, credibility is granted by the perceived authority of the sources where there is one gatekeeper that responsible to control the quality of the supplied information. In contrast, credibility in VGI environment is generated through the perception of trustworthiness and believability of sources rather than in its accuracy. However, for fact, reference and scientific based information, accuracy of data is particularly critical where the inaccuracy will constitute errors and impact to the perceived credibility of information and sources (Flanagin and Metzger, 2008).

Research in other fields such as in communication, marketing, information science and HCI have established internal and external factors specific to their domains and contexts that influence users' perceived credibility assessment in offline and online media. The studies have added knowledge and understanding about individuals' (users) perceptions and the use of advance of communication and information technologies. Several characteristics have been identified to suggest developers to design application that able to tackle end users by focusing on how users assess the credibility of applications. Several potential danger in relation to the way of web users assess credibility of online information have been identified; for example only few users rigorously checking and verifying the information obtained from the Internet (Flanagin and Metzger, 2000); Studies by Scholz-Crane (1998) and Rieh and Hilligoss (2008) have raised greater concern of such behaviour among young adult; for example young adult demonstrates less interest in authority and currency of data but mainly consider to the relevant of that information to the topics and convenient aspect (Rieh and Hilligoss, 2008). Such studies contribute to the understanding that valuable for information sciences domain on how to educate and increase the awareness of such group on the good practices when obtains online information. These studies have led to a few schemes offering a 'seal of approval' to control and protect web users from obtaining misleading information. For example, the HONcode label exists to control the ethical issues of information presented on health information websites (Gaudinat et al., 2007a); stamped seal of approval in e-commerce applications (Cheskin, 1999).

These issues have been widely discussed in other domains but relatively new in geospatial domain. In the geospatial domain, a study by Skarlatidou et al., (2011) has highlighted the importance of designing a web based GIS application by considering not only the elements of usability, but including the elements that will make users believe the information and trust the analysis generated from the application. The issue of credibility is complex and challenging in the current web based mapping due to the emergence of user friendly and low cost of neogeography platforms (e.g. map mashups, Wikimapia, Google My Places) and the raise of neogeographers which contrary to the conventional way of mapping practices; the web-enable citizens are highly motivated to contribute and share location based data either by social media tagging, mashed-up various data sources into new web mapping application (i.e. map mashup) or joining crowdsourcing application (e.g. OpenStreetMap). As argued by Flanagin and Metzger (2008, p.144), among questions that need to be investigated in relation to credibility information of user generated contents or VGI includes whether end users and professionals will accept information largely supplied by volunteered driven mechanism, for what purposes and with what affect? What factors impact users' credibility perception? What technical and socio-technical tools that can help users and professionals navigate and assess the data appropriately?

According to Liu and Palen (2013, pp.75), there are several motivations that drive map mashups producers to create and share information through neogeography platforms (i.e. map mashup); for examples personal interest and gains, curiosity to the potential of that current technology, to expedite the communication of information and to make the information more accessible, usable and compelling than text based reports. Map mashup applications have been developed for several of purposes including entertainment such as mapping twitter usage (twitter.lab.idiap.ch/), commercial such as hotel locator by Travelodge (travelodge.co.uk), and community such as reporting broken streetlights (seeclixfix.com), reporting news and current events such as earthquake hazard program by the USGS (earthquake.usgs.gov). The data may draw from scientific community, commercial or user generated data. For examples, the data presented on West Nile Virus Disease Map published by the USGS (diseasemaps.usgs.gov) is supplied by scientific data by the Centre of Disease Control and Prevention; the listing of real estate properties published by the PropertyGuru (propertyguru.com.my) is supplied by the commercial data; the crime mapping map published by MalaysiaCrime (www.malaysiacrime.com) is supplied by the reports of web-enable citizens. According to Flanagin and Metzger (2000, p.531), the extent of how users assess the credibility of information varies and depends on the type of information sought. Reference (i.e. the information that user might want to refer to and look up) and news information commonly will be verified rigorously than either commercial or entertainment information. This could be implied to map mashup environment where users may critically judge the information if the information they sought associate with the community, news and current events since they want to refer to and use the fact or news obtain from the maps; but their assessments might less critical if the obtain information are for the purpose of entertainment and also due to either the risks or the impact of having misleading and inaccurate information are low in their contexts.

The question under consideration is why it is important to examine the issues of credibility in the map mashup context. Some reasons relate to the following:

- At present, the development of map mashup applications is dominated by non-geoliterate professional and amateur developers. This group, namely 'neogeographers', typically have a very little knowledge, skills or experience related to mapping and cartography. They typically do not follow the professional practice by, for example developing a map mashup that follows with established map design conventions. For example, results from interviews conducted by Liu and Palen (2013) with the producer of map mashups have demonstrated that the design of spatial and temporal data on the maps, the nature of data as well as the technology used to deploy map mashup, which in this case are the crisis maps, are depend on the creator of mashups.

Several studies have identified the high influence of visual design when users judge the credibility of online information. For example, Fogg et al. (2003) have identified the high influence of 'looking good' in users' credibility assessment of a website. A study by Albert and Van (2011) supports this by stating that website designers can increase perceived trustworthiness in a website by using an appropriate colour scheme. Users' perceived level of influence of visual cues is higher and contradictory to that of experts (Stanford et al., 2002b). If this happens in the map mashup context, users are easily exposed to purposely misleading information by a well-designed map.
As argued by Princeton (2002), although it is true that some less credible sites have flashy, distracting graphics it is not inversely true that clean, pro-looking sites are necessarily credible. Well-designed map mashups do not mean that the presented data are correct or accurate. The question is whether the visual cues highly influence map mashup users' judgement of information credibility.

- Other collaborative crowd source based user generated contents applications such Wikipedia and OpenStreetMap have their own moderator (gatekeeper) to deal with the issues of vandalism, copyright violation and disputes. These applications have mechanisms to validate and correct the errors by using the ability of the crowd to converge on the truth (Goodchild and Li, 2012, p.112). There is no gatekeeper, however to control the correctness of information presented through the map mashup medium. As argued by Cartwright (2008, p.24), the use of Web 2.0 as a means of providing geographic information presents different problems for assuring quality; a number of questions arise including who takes 'ownership/custodian of the product? Who guarantees the quality/integrity of the product? Who maintains the product? In the light to these arguments, map mashup applications are developed by mashed-up data and information from various sources. The data may come from authorised sources, supplied by national and commercial data providers and from non-authorised data sources, including individuals and communities. The power of map mashups is their ability to aggregate information coming from various sources (authoritative and non-authoritative) (Roche et al., 2013, p.33). The question is how to guarantee the correctness of the information presented on that medium?

This situation is in contrast to conventional mapping that uses authoritative data as the source of mapping applications. As stated by Elwood et al. (2012) data supplied by authoritative data sources has specific standards and policies resulting from rigorously defined procedures and periodic assessments of quality in collecting, processing and presenting the data. Moreover, the data usually comes with standard metadata that has been produced for users to assess the fitness of the data for their purposes. This is less likely to happen in the case of data and information presented in map mashup applications. The data are mashed up from various sources, where the metadata are recorded in informal and unstructured formats. Sometimes there is no basic metadata relating to the identification of sources such as the author (supplier of data), the last updated date or the level of accuracy or correctness of the presented data/information. As argued by Flanagin and Metzger (2008), in some cases the source information is unavailable, masked or entirely missing; in other cases if the source information is provided, it is hard to interpret such as whether the source is to represent the origin of data; or the source presented might only the producer who reproduced the data from other channel; or the source is actually the publisher who aggregate the data from several web data services; the multiplicity of sources, and the less clear of the origin of data, its quality and veracity result difficulties for the end users to accurately assess the fitness of data to serve their purposes. The authors pointed out that assessing credibility inaccurately can have serious impact to the scientific, social, personal, educational and even political (p.139).

The issue arises of whether users, typically those who are non-geoliterate, will judge the ‘believability’ (credibility) of information by using the metadata presented. Another issue is whether the users of a map mashup, which is typically created by a non-geoliterate developer, perceive the importance of source information and other metadata. The research question here is: **‘will map mashup users use the metadata provided on or with the map mashup to assist them assess the credibility of information presented?’**

Research in the GIS domain has proposed the use of graphic visualisation to increase users’ awareness of the quality and uncertainty of the data they use. Studies by Devillers et al. (2007) have proposed a tool that uses a colour coded traffic light scheme (CCTL) to present the accuracy of the features, layers and datasets in a GIS application. Yang (2007) has extended this approach to online and mobile GIS applications. The implementation of rating systems based on traffic light colour coded labelling on have been applied to certain food products, electrical appliance energy ratings and in the United Kingdom (UK) car carbon dioxide emissions. This colour coded scheme has been implemented voluntarily in certain supermarkets in the UK (BBC News, 2007) and become a food policy in Western Australian health and school services (Western Australian Department of Health, 2009). A study by Kelly et al. (2009) demonstrates that the probability of users identifying healthier foods from a traffic light labelling format was five times more likely than from a label using monochrome text information. As according to Fox (2006, p.12), users tend to pay attention to informative readily available label, such as labelling on food product more than checking the date and source of online information. Dissemination of data quality via graphic depiction, instead of using textual and numerical values, is definitely essential to developing users' awareness of the data they obtain from GIS applications. In fact, dissemination of the quality of data on a map mashup medium is more crucial due to the fact that development and use are not limited to professionals in the field. Wilson and Graham (2013) argue that it is crucial to provide information on uncertainties in data when disseminated through a neogeography tool to avoid misleading interpretation by citizen users. The research question here is **'will such graphic data quality visualisation, by means of a stamped credibility label on top of the map mashup, influence respondents' judgement of the credibility of information?'**

Therefore, this research deals with the issues to support the implementation of automated credibility assessment and labelling on map mashup applications. The research starts with identifying the elements users consider when assessing the credibility of map mashup information. The findings lead to an examination of the influence of stamped credibility rating and labelling as a solution to tackle the issues relating to the way in which users assess the credibility of online maps, specifically map mashup information. The next sections will specifically explain the objectives of this research and the research questions in order to fulfil the objectives.

1.3 Objectives

There are three objectives to achieve in this research, as detailed below:

Objective 1:

To examine the influence of metadata related to sources (specifically the identity of map producer and map supplier) on respondents' assessment of the credibility of map mashup information.

1.1 **Research question**: Are visual cues the dominant elements that influence respondents' judgement?
How significant do the influence of visual cues in users' credibility assessment?

Statements from literature to justify this research question:

A study by Fogg et al. (2003) has identified design look (visual design) including layout, typography, white space, images, colour scheme, professionalism etc. as one of the dominant elements assessed by web users when judging the credibility of websites. Other studies have also identified this element as one indicator to assess the quality of information (Barnes and Vidgen, 2003) and as a feature that induces trust in online websites (Wang and Emurian, 2005). A study by Skarlatidou et al. (2011) has indicated map and web design as one of the features that generate trust in the Web GIS applications. Notwithstanding, the aesthetic aspect of map design has been identified as having a low influence on users' level of trust in the map provider (Skarlatidou et al., 2010a).

Therefore, the **Hypothesis 1** (null) is**:**

Visual cues have no significant influence in respondents' credibility assessment on map mashup applications

1.2 **Research question**: Will map mashup users perceive the importance of metadata to help them assess the credibility and quality of information presented?
How significant does the influence of metadata variables related to sources in users' credibility assessment?

Statements from literature to justify this research question:
In Web 2.0 environment, content sharing is not limited to professional and scientific communities. Every web enabled citizen is able to contribute the data. This lead to a user-centric metadata, namely Metadata 2.0 that easy for contributor to create and easy for users to use and understand (Fu and Sun, 2011). The platforms typically incorporate very simple metadata into the map data structure (Poore and Wolf, 2010). The provision of metadata in map mashup applications is currently informal, unstructured and less comprehensive than that in geographic information systems. There is non-standardisation of layout to present a metadata element on a website. Fichter (2009, p.16) has arisen a concern by stating that casual mashup users often cannot easily discern who provided what

piece of data, hence making difficult to assess credibility and authority of the information. Moreover, a study by Metzger (2007) found that a tendency of web users was to be less than diligent in checking or verifying the sources of messages, for example, in verifying the author credentials. Additional assessments that seem time consuming and involve additional effort are less likely to be considered and tend to be neglected in a web user's browsing activity. A study by Scholz-Crane (1998) identified a tendency of users to conduct a less than thorough assessment when assessing the quality of web information. Activities that require additional tasks are unpopular among web users. Assessment of metadata may be considered to be an additional task to the main web browsing activity.

Hypothesis 2 is:

The metadata related to sources (i.e. map data supplier, map mashup producer) have significant influence in respondents' credibility assessment

1.3 Research question: For the above questions, what are the differences between the level of influences between geoliterate and non-geoliterate respondents?

Statements from literature to justify this research question:
The National Research Council (2006, p.4) argued that geoliterate, is a group of spatial literate users that has developed appropriate levels of spatial knowledge and skills in spatial ways of thinking and acting. One of their characteristics is able to evaluate the quality of data based on its sources and its likely accuracy and reliability. Geoliterate users are expected to understand the concept of ambiguity, uncertainty and data quality issues when dealing with geospatial data. In authoritative data sources, metadata have been compiled and embedded with the data to assist users, commonly professional users, to critically evaluate the fitness and suitability of data for specific purposes (Comber et al., 2006). In map mashup environment the metadata typically unstructured, and often informal. Nevertheless, it is postulate these geoliterate users will apply their spatial literacy skills and act as critical spatial thinker (Kim and Bednarz, 2013) regardless the context of map uses.

Hypothesis 3 is:

The metadata related to sources have significant influenced within geoliterate respondents

1.4 Research question: Do the users will notice the elements of metadata related to sources embedded on the experimental maps? Do they perceived that element(s) as important?

Statements from literature to justify this research question: Credibility not resides in the object or source itself. It is users who recognise dimension(s) of credibility based on the characteristic of object (Rieh and Danielson, 2007). Moreover, according to the Prominence-Interpretation (PI), a theory that has been specifically conceptualised to study the ways in which users assess the credibility of websites (Fogg et al., 2003); there are two stages of how users assess online information, which are; 1) users notice the credibility elements that looks prominent on the object of assessment; 2) judgement will then be based on that element(s). In other words, the influence of credibility element is dependent on whether users notice the element and the extent of the element looks prominent. Eysenbach and Kohler (2002) argued that there is substantially different between what users said they will do with what actually they did in practices. 'Consumers say that when assessing the credibility of a site, they primarily look for the source, a professional design and a variety of other criteria (i.e. interpretation is based on a statement). In practice however, online users do not check the 'about us' sections of websites, try to find out who are the authors or owners of the site or read' (p.576).
The question is 'do the users will notice (or check) the metadata related to sources embedded on the experimental maps? Do they perceived that element(s) as important?'

Therefore, **Hypothesis 4** (null) is:

There is no significant difference between the level of importance of the metadata related to sources between these two contexts - a prominence dependent (i.e. interpretation is based on what looks prominent) and a prominence independent setting (i.e. interpretation is based on a statement)

1.5 Research question: Do the perceived important of these elements differ if different level of map use contexts are applied?

Statements from literature to justify this research question: Olson (1978), Muehrcke (1979) and Board (1984), cited in Liebenberg (1998) have identified three levels of map use: level one comprises map reading tasks such as identification of individual symbols (lines, polygons, points) and the differences between these symbols' shapes and relative size; level two comprises of recognition of the spatial pattern where users visualise the data in the form of ‘space’, which is a term to describe the objects and events that occur in 2D and 3D views; whilst ‘place’ is used to describe the space according to its particular geographic location (Brown and Perry, 2001); level three comprises interpretation tasks where map users will use other information and also relate their knowledge and previous experience to answer the geographic ‘why?’ questions. Empirical study by Ferebee (2007) demonstrates that, in non-situational involvement (which means a low level of engagement with the context), users tend to notice more of the medium related elements: for example, the design look, design structure, functionality, security and technical capability. Meanwhile, more message related elements, such as information accuracy, usefulness, and clarity, are noticed when there is a high level of involvement.

Therefore, **Hypothesis 5** is:

There is significant difference between the levels of perceived importance of the metadata variables related to source between these two levels of engagement contexts – low level (level one) versus high level (level three)

Objective 2:

To examine the influence of colour coded traffic light (CCTL) labelling on respondents’ assessment of credibility

2.1 Research question: If there is credibility labelling stamped on a map mashup application, will this have a positive influence on users when making an informed judgement?

How significant does the influence of credibility labelling stamped on map mashup application?

Hypothesis 6 is:

Credibility label has a significant influence in respondents’ credibility assessment

2.2 Research question: Will the presence of such credibility labelling decrease the influence of other dominant elements in respondents’ judgement

Statements from literature to justify these research questions:

A few studies have proposed visualisation tools to communicate the quality of data to users. Devillers et al. (2005; 2007) proposes a model to support a development of visual spatial data quality assessment which can be integrated on a desktop GIS. Yang and Wang (2004) and Yang (2007) proposed a more comprehensive framework for visualising spatial data quality using an object-oriented approach specifically for a Web and mobile environment. Yang (2007, p.173) addressed a possibility to provide visual quality information on commercial maps such as Google Map and Microsoft Virtual Earth; this could be implemented by extending its Application Program Interfaces (APIs). A study by Mass et al. (2011) proposed the use of a geo-label on scientific maps. The label is based on information about data quality which will be extracted from metadata, the data itself and the validation process with in-situ sensors, provenance information, and user-feedback. This approach may educate and informed the general public on the uncertainty of the data and information they obtained from that medium. The same benefits might be generated if such a visual indicator is implemented on top of map mashup applications.

This research examines the influence of visual credibility indicator by the use of colour coded traffic light (CCTL) scheme. There are a few studies that became a basis of this research in proposing a visual credibility labelling using colour coded traffic light (CCTL) scheme CCTL). In a food industry domain, a study by Kelly et al. (2009) have identified a probability of user identifying healthier foods using a traffic light labelling scheme was five times more likely than from a label using monochrome text information. Fox (2006, p.12) argued users tend to pay attention to informative readily available label, such in labelling on food product, more than checking the date and source of online information. In geospatial domain, Devillers et al. (2005; 2007) proposes a visual spatial data quality indicator which can be integrated on a desktop GIS application using colour coded traffic light scheme indicator.

Hypothesis 7 is:

The presence of credibility label has a significant effect to the influence of visual cues in respondents' judgement

2.3 Research question: What are the differences between the level of influences between geoliterate and non-geoliterate respondents?

Statements from literature to justify these research questions:

Maps communicate in a language that a 'literate' person learns to understand, evaluate and use (Abilock, 2008, p.11). Such literate individual is known as geoliterate in this research. It is refer to individual who have information literacy skills not only into how to read, analyse and use the map but understand the concept of mapping, its limitations, and how to evaluate maps. Geoliterate users are expected to have certain level of skills to evaluate the map since measurement and information on map are rarely 100% accurate (Chang, 2014). They should understand the concepts of globe to 3 dimension (3D), 3D to 2 dimension (2D) mappings and the limitations caused from the transformations, how to read the coordinates system, analyse the visual cues including the symbol and code expressed in a map's legend and colour scheme. They are supposed able to evaluate the point of view defined by a map's title, orientation, scale, theme (Abilock, 2008) and assess its fitness of purposes by analysing the supplementary data including source, date, accuracy, projection. These elements are the common key elements of a map layout before print out for dissemination.

However, the rise of online maps in the era of neogeography particularly when users can simply create their own maps using mashup tools, at some extent disregard the cartographic principles in presenting map as an output. Nivala et al (2008) for example have studied the commercial web mapping sites such as Google Map, Yahoo Map and criticised so many issues of maps presentation in relation to cartographic principles and the usability aspects. The maps from these commercial providers have been used by many including the non-geoliterate users, which is to represent public users who do not have professional skills in mapping and cartographic concepts. For non-geoliterate users they might not have the skills on how to evaluate the map; they might even do not know that 'a map is not a mirror of the world, but rather a selective abstraction of some part of it' (Abilock, 2008, p.11). A map is a medium of persuasion (Muehlenhaus, 2012) and easily to be exposed with propaganda and misleading information; as stated by Monmonier (1991) 'it is easy to lie with maps'. The conventional keys elements on map layout such as the sources, date and accuracy (metadata) are not structurally presented on online medium. However, geoliterate users are speculated able to apply their skills to critically evaluate the data and information on maps without regard the medium of presentation. The question raised here is whether the influence of visual cues, metadata and credibility rating label among geoliterate users might be different with non-geoliterate users in map mashup medium.

Hypothesis 8:

There is significant difference between geoliterate and non-geoliterate respondents in the influence of visual cue variables when making judgement

Hypothesis 9:

There is significant difference between geoliterate and non-geoliterate respondents in the influence of credibility labelling when making judgement

Objective 3:

Through literature studies, to propose a conceptual framework to support the implementation of automated credibility labelling for map mashup applications.

3.1 Research question: What are the parameters that are practical to assess the credibility of map mashup applications in an automated manner?

3.2 Research question: In order to assess the parameters, how can a credibility rating index be developed?

3.3 Research question: What are the practical approaches and technologies that could be used to support the development of the automated tool?

3.4 Research question: What are the current issues that need to be tackled before this automated tool could be developed?

1.4 Scope of Work

This thesis is limited on the following areas;

1) The influence of credibility elements examined in this research were conducted within ‘prominence dependent’ research framework; this research concerned with the elements that look (notice) prominent by respondents; the results might be differences in those research that examine the influence of credibility by using method of ‘interpretation based on a statement’ (see section 2.5 and Chapter Discussion).
2) The influence of metadata related to the sources. The parameters tested in the experiments are related to sources - the map producer and foreground map data supplier only. Other parameters of metadata as stated in ISO19115 such as data accuracy are not examined. These parameters related to sources have been selected due to their relevance to the elements of credibility identified by Fogg et al. (2003). It is also bases on the argument by Flanagin and Metzger (2008, p.140) that ‘source information is often crucial to credibility because it is the primary basis upon which credibility judgements are thought to the rest’.
3) The experimental tasks to examine the influence of parameters tested in the experiments are conducted using a series of experimental map using an online map based questionnaire. A holistic functional of map use approach based on a real situation, such as in a ‘treasure hunt’ approach, is not conducted. This research also did not concentrated on the context of map use where the results were limited to the simulation of experimental based scenario; the implication of the results to other real scenario of map uses requires further investigation. The implication of this scope is discussed in details at the chapter Literature Review and Discussion (see Section 2.5.2).
4) The sample of this research did not filter according to their familiarity to the map area. This research also assumed the respondents understand the concept of foreground and background data sources applied on map mashups, although no assessment was made.
5) The proposed conceptual framework, including the proposed practical parameters to be evaluated in the rating and the formulation of an index rating is based on the literature review and proposed indicatively.

1.5 Contributions

The following is a summary of the original contributions made in this thesis.

1) The research has identified the **low influence of metadata** related to map producer and map data supplier and **high influence of visual cues** when users assess the credibility of a map mashup application. This contribution is demonstrated in detail in Chapter 4, Chapter 5 and Chapter 6.

2) The research has identified the **moderate positive influence of credibility stamped labelling** on map mashup applications. This finding is relevant to support the related research that is devoted to the development of a data quality visualisation tool in the GIS domain, including desktop, online and mobile applications. This contribution is demonstrated in detail in Chapter 7.

3) A proposed **conceptual framework** of the parameters (elements) that are practical to assess the credibility of map mashup applications is presented. The evaluated parameters include the elements of metadata, elements related to the data and elements related to a website that could be adapted to assess the credibility of other online maps, including web-based GIS and mobile GIS applications. This contribution is demonstrated in Chapter 9.

4) The research is leads to the **specification of practical approaches** and technologies that could be used **to support the implementation of automated assessment and credibility labelling** on map mashup applications. The technical issues that have to be tackled before this automated tool could be implemented have also been considered. This contribution is demonstrated in Chapter 9.

1.6 Thesis Outline

The rest of this thesis is organised in the following manner:

Chapter 2 provides a literature review presenting an overview of the credibility related studies and key definitions. It discusses the elements related to credibility in various domains as well as the relevance of the elements (parameters) tested in the experiments. Finally, this chapter highlights the need for automated credibility labelling on map mashup applications.

Chapter 3 describes the general methodology applied in this research.

Chapters 4, 5, 6 and 7 describe Experiments 1, 2, 3 and 4, respectively, conducted in this research. Each chapter describes in detail the hypothesis and methodology used, including the experimental design and the tested parameters. Then, the results and data analysis are presented. In the last section a discussion of the findings and the limitations of experiments are discussed.

Chapter 8 presents the overall discussions of this research that include the discussions of findings in the four experiments that have been conducted.

Chapter 9 presents the proposed conceptual framework to implement the automated credibility labelling on map mashup applications. The proposed practical parameters (elements) to be evaluated in the indexes drawn from literature studies are presented. This chapter also presents the proposed formulae to calculate the rating index. In the last section the technical issues to implement the automated credibility labelling that have been suggested for future research are discussed.

Chapter 10 summarises the thesis. The conclusions and research contributions of this research are presented, as well as a list of suggestions for future research.

2 LITERATURE REVIEW

2.1 Introduction

The history of GIS dates back to the era of mainframe computing. Applications emerged and have been widely developed by various agencies since the arrival of the desktop era. Today, we are in the era of distributed computing and many GIS applications have evolved and adapted these technologies. In the mainframe era, GIS programs were installed in the same mainframe computer in which users accessed the data and ran the analysis through dumb terminals via local-area-network (LAN). Access was limited to a single user at any one time. In the desktop era, which is still the basic architecture of the majority of today's GIS applications, GIS programs can be used either as a stand -alone application or accessed via the intranet. The architecture is based on a 2-tier client/server. In the present era, GIS is on distributed computers where GIS programs, data or functions can be accessed through both wired and wireless. GIS applications can be accessed via web or mobile technologies. And users can access them on-site, at home, or on the go.

According to Peng and Tsou (2003), Web GIS has evolved rapidly since 1993 from its origin in the Xerox PARC Map Viewer. Although it provided very basic functions, this application marked the first move to run GIS inside a web browser where users could access the application without having the program locally installed on their computers. Since then, the GIS community has adopted the concept of using GIS via web browsers, for example, the release of the first online version of the National Atlas of Canada in 1994 and MapQuest in 1996, the pioneer of commercial web mapping applications; thus led to the release of Google Map and Bing Map in early of 21st century.

Currently, the fusion of geospatial data (map) and the Web has entered into the Web 2.0 era. The applications generated under Web 2.0 share common characteristics, including the user-generated driven content, and the Web is used as a platform with the applications providing rich user experiences. The map mashup that has emerged in the present era is a type of interactive web application that integrates data retrieved from diverse web-based resources. Map mashup is an extension of Web GIS in the current neogeography era that widens the use to non-professional geoliterate map users, particularly for personal and community activities. This phenomenon encourages the availability of users-generated spatial contents (UGSC) and volunteered geographic information (VGI) and promotes the development of public participation GIS (PPGIS) (Fu and Sun, 2011).

2.2 Map Mashups

A mashup is a web application that uses content from more than one source to create a single new service displayed in a single graphical interface (Fichter, 2009). Liu and Palen (2010) have specifically define a map mashup as a web application where people aggregating two or more data feeds or functionalities from other websites using application programming interfaces (APIs). Most of mashups however do not follow this classic pattern of blending two or more sources of software (Batty et al., 2010). As mentioned by Haklay et al. (2008, p.2021) most of consumer mashups are the equivalent of push pins that have been located on a map with some multimedia information-mostly text but sometimes images of video clip-attached to the pin. Due to this, Li and Gong (2008) describe a mashup application as a web application that at least combines one map data source or service with added information, often geo-referenced to the map data to create a new map.

According to Fun and Sun (2011, p.89) the concept of mashup has existed since the late 1990s when web services originated, but map mashups have become popular since the release of free mapping APIs in 2005 by Google. Prior to 2005, map mashup applications were developed by professional programmers. During that period, mashups relied on what web services could offer and the processes were done on the server. The applications required professional programming tools and complex server-side programming. Most of current mashups, however are using browser-side architecture due to the rich contents and services available over the Web and the popularity of browser-side APIs, AJAX (asynchronous JavaScript and XML) and JavaScript. For example, commercial mapping providers including Google and Microsoft officially released their mapping capabilities through free APIs. Major software vendors are also adopted this easy and quick way to develop Web application such as ESRI ArcGIS APIs for Javascript, Flex and Silverlight. Mapping applications via browser-side APIs reduce the efforts to develop one application from scratch and simplify the user experience to develop and use the application. This APIs are built on top of web services; web services will only perform the server-side functions; all the client-side interaction will be performed at the web browser. Therefore, web application can be developed faster and easier and provide rich user experiences whilst using the applications. These characteristics have led to the increasing popularity to the development and the use of map mashups.

The sources of map mashups are not limited to citizen journalism and crowd sourced data, but drawn from publicly available scientific data, commercial and licensed data (Liu and Palen, 2010). The types of map mashups could be categorised into four which grouped according to the user and usage perspectives – informative, participatory, collaborative and enterprise mashups. The informative map mashup is the simple and lowest complexity in terms development process and the functionality offered whereas the enterprise mashups is more complex and requires high programming skills (Li and Gong, 2008).

In general, map mashups can be classified into three – consumer, data and business (enterprise) mashups (Li and Gong, 2008). At present, most of the developments of mashups are generally consumer based applications (Ogrinz, 2009, p.14). The deployment of more complex applications, such as for an enterprise mashup, is still not well developed. This is due to the ecosystem of map mashups that still not mature. A report by DuVander (2012) however, has shown the increasing number of enterprise APIs and this indicate a well progress of the use of map mashup technology into a more complex application. According to Fichter (2009), open content (data) is one of fundamental component in this ecosystem. To sustain the benefits from map mashup technology, the ecosystem has to be supported from a lot of open data and web services from government departments, academic research unit, and organisations that could offer open data related to social, economic and scientific data. This is due to the strength of mashup come from the Web services behind the API; web services are the programs that run on a web server and expose programming interfaces to the web. Example of web services are Web Map Service (WMS), WFS (Web Feature Service), WCS (Web Coverage Service), GeoRSS and KML.

The availability of web mapping applications, such as Bing Map, Ordnance Survey OpenSpace, OpenStreetMap and Google Maps have encouraged the creation of users' own maps by mashing up data from different sources. The availability of free basic portals such as MyMaps in Google Map, secondary software such as Google Map Map Maker and GMapCreator in MapTube, open data services and mapping providers' Application Programming Interfaces (APIs) contributes to the rapid deployment of map mashup applications. These APIs enable users to access aerial and space imagery, maps, postal codes and other features that have been created by other developers. The free tools and portals are easily configured without the need for high-level programming skills. The current state of map mashups is still primitive where deployment of map mashup applications still typically requires skills of developers who are able to code and run their own services (Batty et al., 2010)

Although no formal (often informal) data quality control is applied to this form of map publishing, web users tend to use the information. As Craglia (2009, p.33) has stated, volunteered information is sometimes more timely and accurate than official information. This can be seen when volunteer photos were uploaded and individuals' data were being mashed-up on Google Map during Hurricane Katrina; these data then were used to analyse the aftermath situation (Schutzberg and Francica, 2005). Map mashups have been used to mapping trivia information, current events, disease outbreaks and post disaster relief. Liu and Palen (2010) have examined the technical aspect of the development of map mashup applications, particularly during crisis; this study investigates the neogeography practices that were used including the purpose of application, the sources of data, the map design decisions and developmental issues that aroused.

The capability of neogeographers to generate user-generated spatial content (UGSC) has been seen as a potential approach to fill in the gaps in the top-down approach that is applied by the majority of map providers when producing geospatial related information. According to Bishr and Kuhn (2007), the top down approach has preserved the distinction between the data provider and the data consumer; the map provider acts as data provider whilst users act only as data consumers. In contrast, in a bottom-up approach, neogeographers have applied the reverse approach, whereby users act as a data consumers, as well as data providers. This approach has the potential to reduce the costs of data maintenance and meet the demand for timely data. It can also help tackle the issue of gathering geospatial information related to the following (Goodchild, 2007b):

1) Data that is not visible and cannot be extracted from satellite imagery or by any automation process

2) Information from local individuals that relates to specific and local activity

3) Descriptions on the use of land and buildings, particularly in suburban and urban areas

4) Environment related information, such as water, air and noise quality

However, one reason for many mainstream map providers disregarding user-generated sources of data is the perceived lack of quality control. The true value of UGSC data has to be evaluated; the methods used to assess, measure and control the quality of this data have to be examined and implemented so that this data can justifiably be perceived as credible and reliable. In fact, this type of data can be used to support other data in a low accuracy application and a low risk situation. Such data have been used during a crisis when the official sources of data are inevitably delayed (Liu and Palen, 2010). As argued by Elwood et al. (2012), during a crisis, decision makers must make choices between acting immediately with questionable data and waiting for better data to arrive; these data have considerable potential to help initiate many actions during a crisis, despite the risks of false information from inaccurate data. Therefore, the issue of data quality and reliability has to be tackled thoroughly in a way that is not limited to the technical aspect of data accuracy, but in a way that examines the overall quality control issues of data generated by this approach. The presence of an automated credibility tool which could check and assess the level of correctness of map mashup information can assist the decision processes during a post-disaster relief as well as protecting users from propaganda, incorrect, misleading and invalid sources of information.

A few scholars have suggested the potential of this user-generated content to support the implementation of spatial data infrastructure (SDI) (Goodchild, 2007b, Jackson et al., 2009). The adoption of this new data collecting approach may address challenges related to environmental and social needs via collaboration between government, industry and citizen. Haklay (2010) has examined the capability of user-generated content to complement data from authorised sources. However, several issues need to be tackled before this potential could be realised. One solution is to develop a mechanism to checking the credibility and reliability of the data. For a social aspect, a rating scheme and labelling might be useful to guide users making informed judgement of the data they obtained.

The credibility of information (Flanagin and Metzger, 2008) and data accuracy (Goodchild, 2008) generated from the UGSC approach have been discussed in the literature. Mummidi and Krumm (2008) have proposed a method to address the quality problem which involves the development of a data mining algorithm to search through a large collection of VGI and assess the data/information in the form of user-supplied annotated pushpins that are consistently repeated. In this study the correct name of a place of interest (POI) is chosen on the basis of a simple cluster of pushpins that may represent a similar POI. Bishr (2007) has designed a spatial trust model to assess the value of data contributed collaboratively by examining the reputation of authors; in this study, the correctness of data/information presented depends on the reputation rating of the authors. Elwood et al. (2012) is in line with that study which highlighted the approach of checking feasibility and reasonableness of VGI against other information and the evaluation based on reviews and commentaries by varying numbers of peers.

Apart from this research, only a few studies examine the elements that influence users and the factors that motivate users to ‘believe’ the map data. This issue, however, has been widely discussed in other domains. Such studies have examined the elements that make web users believe the information presented and have proposed a few schemes offering a ‘seal of approval’ to control and protect web users from obtaining misleading information. For example, the HONcode label exists to control the ethical issues of information presented on health information websites (Gaudinat et al., 2007a); stamped seal of approval in e-commerce applications (Cheskin, 1999). In the geospatial domain, a study by Skarlatidou et al., (2011) has highlighted the importance of designing a web based GIS application by considering not only the elements of usability, but including the elements that will make users believe the information and trust the analysis generated from the application. The next section discusses the concept of credibility and its elements.

2.3 Defining Credibility

Research related to credibility has become of interest in several domains, including communication, information science, marketing, psychology, interdisciplinary efforts in human computer interaction (HCI) and currently Web 2.0 applications. According to Rieh and Danielson (2007), early systematic research related to credibility may date back to the 1950s in the field of communication; a study by Hovland and Weiss (1951, p.1475) has been recognised as the first landmark in this research topic. Different domains, however, investigate credibility from different research perspectives For example, in the communication domain, one of the research issues focuses on sources of credibility by comparing perceived credibility from different media (Schweiger, 2000). In the information sciences domain, credibility has been examined from the aspect of how users assess documents and statements retrieved from information retrieval engines (for example see Rieh and Belkin 1998). In the human computer interaction (HCI) domain, credibility has been examined in terms of users’ perceptions and acceptance of information or output from computers relating to technology such as the Web, mobile devices and automated machines (Fogg, 2002). Moreover, in the new era of Web 2.0, research related to credibility has become one of the major interests in the applications generated from the technology of this new era. For example, several studies have investigated the credibility of Wikipedia (Luyt and Tan, 2010), blogs (Juffinger et al., 2009), and twitter (Al-Khalifa, 2011). Issues of credibility in user-generated spatial contents, such as OpenStreetMap and the free web mapping service applications such as in Google Earth, have also been raised by several authors, including Flanagin and Metzger (2008) and Goodchild (2008).

Credibility is synonymous with believability (Fogg and Tseng, 1999, Flanagin and Metzger, 2008). Credibility is an intangible concept and is related to a user's perception of an object of assessment. Credibility influences the viewer's perception of believability in the information conveyed by the object. The object of assessment may refer to the source, message, or the media itself. According to Fogg and Tseng (1999), there is no clear consensus on the underlying dimensions to describe the concept of credibility in the literature. A dimension provides a particular point of view for judging credibility; different domains and studies typically have varied dimensions for judging the perceived credibility of an object. For example, in a study by Hilligoss and Rieh (2007), five dimensions of credibility emerged from the results of their survey – truthfulness, believability, trustworthiness, objectivity and reliability. The main primary dimensions of credibility discussed in the literature, however, are trustworthiness and expertise (Fogg and Tseng, 1999). Hence, these two underlying dimensions were used in this present study when defining the concept of credibility. As described by Fogg and Tseng (1999, p.80), 'trustworthiness is defined by the terms well-intentioned, truthful, and unbiased; the trustworthiness dimension of credibility captures the perceived goodness or morality of the source. Expertise however is defined by the terms such as knowledgeable, experienced, and competent; this dimension captures the perceived knowledge and skill of the source'. Credibility research in general uses the term 'source' to describe media, an organisation or individual spokesperson; in an offline environment, the term 'source' however may be used to refer to a website (Hong, 2006).

Fogg (2002) suggests that the two dimensions of trustworthiness and expertise have a direct relationship; if an object of assessment lacks one of these dimensions, its credibility will decrease. In contrast, Hilligoss and Rieh (2007) argue that different people may define credibility differently in relation to the same object of assessment; in other words, different users might use different dimensions. However, they often hold multiple dimensions of credibility and apply certain constructs of credibility, depending on the situation or type of information encountered (Hilligoss and Rieh, 2007, p.1475). A user will recognise the dimension of credibility, based on the characteristics of the object of assessment, before making credibility judgement (Rieh and Danielson, 2007). Rieh (2002) claims that judgement related to credibility is concluded in an iterative process; Fogg and Tseng (1999) argue that multiple dimensions of credibility are assessed simultaneously before an overall judgement of credibility is made. It is still unclear whether these claims are relevant when users make assessments using a peripheral route, as suggested in the Elaboration Likelihood Model by Petty and Cacioppo (1986), a theory from the communication domain that suggested two distinct routes of focus when users process information, namely 1) a central route, which is when a user critically checks the contents and 2) a peripheral route when a user makes a judgement based on simple inferences.

2.4 Credibility Elements

The term 'credibility elements' refers to the cues or attributes of the object of assessment that influence users perceived credibility or believability in an object. In the literature, these credibility cues have been examined in various domains, although slightly different semantic terms are used. Some research into online trust examines the trust related cues of an object, such as a website or an online map. Several trust related cues that are similar to credibility cues have been identified in credibility related studies, particularly of website credibility. For example, in a study by Skalatidou et al. (2011), the static trust attributes identified are somewhat similar to the cues defined as credibility elements by Iding et al. (2009) and Fogg et al. (2003); in a study by Wang and Emurian, (2005), the trust inducing features identified are similar to the credibility elements discussed in the literature.

Fogg and Tseng (1999) have explored the issue of the different semantic terms used in studies to investigate credibility and trust in the psychology and HCI domains. The authors focus on the dissimilarity between the concepts of trust and credibility; these two terms are related but are not synonymous. In the HCI domain, the authors suggest that, in place of the word 'trust' the word 'dependability' could be used; the phrases in studies that use 'trust the information, 'accept the advice' or 'believe the output' are referring to perceived credibility (Fogg and Tseng, 1998, p.81). McKnight and Kacmar (2007, p.424) agree with this view and state that 'information credibility has the website information as its object, while trust typically has the website's vendor as it object. Credibility and trust have different etymological roots. While credibility and credible are from the Latin *credere* (to believe), trust is from the Old Norse word *traust*, meaning confidence; because their roots differ, their basic meanings differ fundamentally'. Notwithstanding, Wang and Emurian (2005, p.108) argued that 'trust is an abstract concept and is often used interchangeably with related concepts such as credibility, reliability or confidence in the literature; defining the term and to delineate the distinction between trust and its related concepts have proven challenging for researchers'.

Due to this semantic argument in the literature, this present research needs to define the term for the purpose of this research and it has adopted the views of Fogg and Tseng (1999) regarding the differences between credibility and trust. Moreover, the term credibility seems more appropriate to use in this present research as it focuses on the elements or cues that influence users' decisions when deciding between two online maps. The contexts, tasks and engagement level expected of respondents are not very high or complex and activities that involve risk and vulnerable situations are not assessed. Meanwhile, the proposed automated credibility index and labelling research uses the term 'credibility' instead of 'trust' because the automated calculated index and labelling that will be produced on the map mashup application may become a tool to generate a level of perceived credibility in the information provided. To what extent users will trust the information supplied by the labelling system, sources or map mashup application is unknown. As argued by Rieh and Danielson (2007, p.314), 'credibility refers to a perceived quality of a source (website), which may or may not result in associated trusting behaviour; trust frequently refers to a set of beliefs, dispositions, behaviours associated with the acceptance of risk and vulnerability'. The trust concept is defined more comprehensively than the credibility concept. Figure 2.1 indicates a model of online trust by Corritore et al. (2003). A discussion about trust related cues however could not be separated from a discussion about credibility elements because it is clear that they are related. For example, believing the information from a website may constitute a reason to trust the source of information. Notwithstanding, assessing credibility is a pre-requisite before users may generate trust in the object.

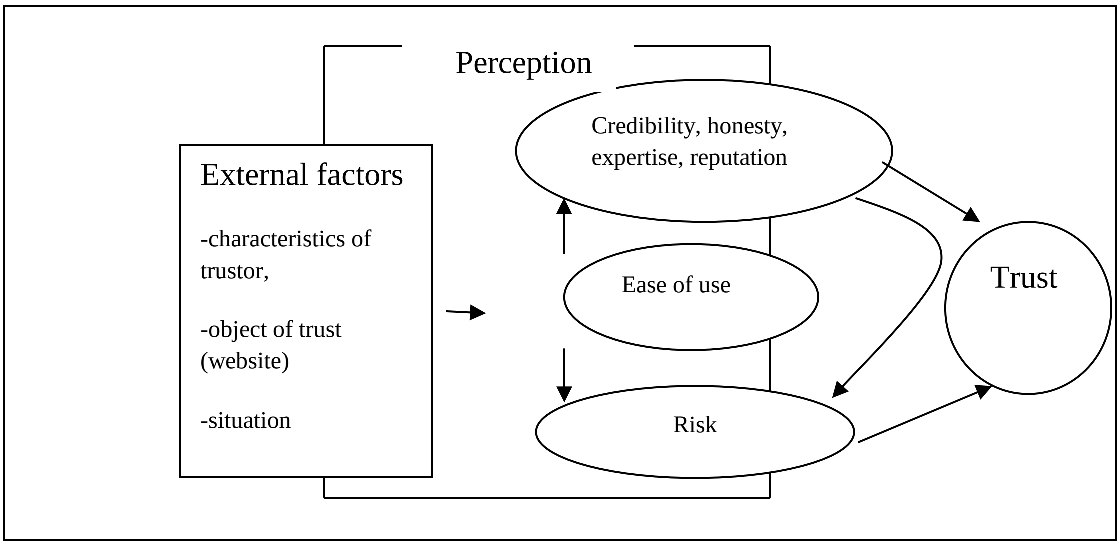


Figure 2-1 Model of Online trust (Source : Corritore et al. (2003))

Some studies of the information quality of websites also implicitly reveal a few indicators that are similar to credibility elements. Credibility plays an important role when judging objects which have several values of quality characteristics. For example, if two objects of assessment have a similar level of quality values, the object then will be assessed according to which object is perceived more credible (Rieh and Danielson, 2007). In the domain that studies the relevance of document during users' search, there are similarities between the criteria that become the basis of judgement. According to this authors, assessment of credibility has become a part of relevance judgement (Rieh and Danielson, 2007). Table 2-1 below summarises the studies in several domains that have demonstrated elements somewhat similar to the credibility elements identified in credibility research. The first four rows present credibility elements that have been identified in specific studies.

Table 2-1 Summary of a few studies that demonstrated several elements similar with credibility elements

Studies	Domains	Terms	Elements in the study	Similar elements related to credibility
Iding et al. (2009)	Information Science (IS)	Website credibility		Information focus, name recognition/affiliation, links, commercial interest/bias, reference, information design, currency, design look, expertise, bias, inaccuracy, information clarity, tone, corroboration
Fogg et al. (2003)	Human Computer Interaction (HCI)	Website credibility		Design look, information design/structure, information focus, company motive, usefulness of information, accuracy, name recognition/reputation, advertising, bias, writing tone, identity of site sponsor, functionality, customer service, information clarity, performance, readability, affiliation
Ulicny and Baclawski (2007)		Newsblog credibility		Commonality (more sources claim that a certain sentence is true), independent sources (two independent sources is more credible than information provided by two sources which depend on one another)
Nagura et al. (2006)		Credibility of News Documents on the Web		Commonality, numerical agreement (agreement on numerical statement in the articles), sources of news
Skarlatidou et al. (2011)	geospatial	Static trust	Links to online community, customer service (chat), feedback mechanism, forum, lineage (how the	Accuracy, map and website usability, map and website design, functionality, website provider reputation, logo, links to privacy policy, contact details, seal approval, testimonials, aesthetic

			map was constructed)	(background base map), visual design (colour scheme), readability, professionalism look
Bishr and Mantelas (2008)	Web GIS	Trust	Proximity to author's location to the event	Reputation (individual personality)
Cheskin (1999)	E-commerce	Trust	Previous experience, help services, use of technology	Navigation, brand, presentation
Corritore et al. (2005)	Online trust		Predictability, risk, safety, trust	Believable, expertise, reputation, perceived ease of use
Shek et al. (2003)	Online trust		Third party endorsement (individuals and affiliation)	
Wang and Emurian (2005)	Online trust	Trust-inducing features	Navigation reinforcement (i.e. guide, tutorials), disclosure company competence, security, privacy, financial, legal concerns, social cue design (face to face interaction)	Graphic design (screen size, colours, good-shot photo), structure design (navigation, broken links), page design techniques (i.e. white space, grouping, visual density), display brand-promoting information (i.e. logo, slogan), display seal of approval, third party certificates, comprehensive, correct, current information, relevant domain name
Nakamura et al. (2008)	Information Science (IS)	Trustworthiness	Positive and negative viewers' comments	
Charnock et al. (1999)	Health Information Science (DISCERN rating tool)	website quality rating tools	How treatment works, benefits and risks of treatment, overall quality of information, provide additional sources of information, reference to	Clear aims, relevance to users, sources of information, currency, bias and balance

			areas of uncertainty	
Fallis and Fricke (2002)	Health Information Science	Indicator of accuracy	Copyright acknowledged, exclamation points, citation, high number of in-links (hyperlinks from other websites to that site)	Website domain, up-to-date, HONcode logo, advertising, disclose author, disclose contact information, free of errors in spelling
Provost et al. (2006)	Health Information Science	Quality of health website	Content quality (peer-reviewed), disclose editorial process, , intended audience, disclose sponsor , disclose ownership and commercial interests, disclose acknowledgement, disclose funding, feedback mechanism, availability of support and documentation for users, security	Free of spelling errors, balance and neutral format of content, content comprehensiveness, relevance, disclose mission/purpose, currency (date posted, date revised), disclosure of authors and developers, author credentials, disclose affiliation, advertising issue (distinguish content, policy), design (navigation, layout, aesthetic), readability, accessibility, links availability, contact address, privacy policy, referencing
Rieh and Belkin (2000)	Information Science (IS)	Information quality and cognitive authority	Characteristics of information object (type of object, title, content topic), type of source, domain knowledge, system knowledge, first-hand experience, second hand knowledge, situation, ranking in search output, general assumptions	Organisation/structure, presentation, free spelling errors, graphics, functionality, URL organisation type (domain), reputation of source, one-collective source, author/creator's credentials

Barnes and Vidgen (2001; 2003; 2004)	Website Quality	Web Qual 1.0, Web Qual 4.0, E-Qual	Use of technology, reliability (dependable information, guaranteed services), security, empathy (customer services, customisable, sense of community, personalisation, easy to communicate), responsiveness	Aesthetics, design look, accurateness, up-to-date, navigation, professionalism, reputation, readability, scope (details), believable information, timely, relevant, format, Website overall view
Parker et al. (2012)	geospatial	Relevance of search document		Accuracy, clarity, currency, depth and scope, quality of sources

2.4.1 Credibility elements tested in the experiments

Table 2-2: Summary of credibility elements examined in the experiments

Experiments	Credibility elements
Experiment 1	Sources - Map producer (author)
Experiment 2	Sources - Supplier of foreground data
Experiment 3	Source - Map producer (author)
Experiment 4	Source - Map Producer, seal of approval (i.e. credibility label)

Of the many credibility elements that have been identified in the literature, the present research is examining only a few elements. The major elements being examined in the series of experiments are mostly concerned with the influence of the authority of sources on users' perceived credibility. Further research to examine other influences on users' perceived credibility of map mashups, such as data accuracy, information correctness or third party endorsement (i.e. comments, reviews, testimonials) could be conducted by future research. Table 2.2 presents the variables that tested in the experiments.

The identity of the author or creator of a website has been identified as one of the elements by which web users assess the perceived credibility of information presented on the Web. There are two aspects typically being assessed; the first is whether the website has disclosed the identity of the creator or author behind the website; the second concerns the background of the authors (creator), which may indicate their credentials for providing information, to web users. In an offline map product, such information is typically indicated in the form of metadata, where a separate document gives information about the data presented on a map. In an online website, as well as Web GIS applications, this type of information is typically provided in a separate link, for example in an 'About Us' section in order to present the details of the creator behind the website. Cambridge Advanced Learner's Dictionary (Woodford and Jackson, 2003) defines credentials as the abilities and experience which make someone suitable for a particular job or activity. The credentials of an author (creator), such as educational or professional background which can have an affiliation (i.e. membership, attachment) to a respected organisation, awards and certificates gained and conferred titles, including Dr. and Prof., may influence users' perceptions of credibility (Fogg and Tseng, 1999). Nevertheless, a study by Rieh (2002, p.154) has identified that the author/creator credentials are of importance to those web users looking for research and medical information; the influence of this element in users' judgement of map mashup credibility is still unclear.

Web users are more concerned about source reputation and type of source when judging the quality and authority of information (Rieh, 2002, p.154). Reputation is the estimation of the consistency over time of an attribute of an entity. The concept of reputation depends on a user's initial beliefs and observations of a firm's past behaviour (Herbig and Milewicz, 1995, p.5). In a study by Metzger et al. (2010) source or site reputation has been identified as one of the most important heuristics used for evaluating credibility. This reputation heuristics signals a reliance on the reputation or name recognition of website or sources of Web-based information as a credibility cue, rather than a close inspection of site content or source credentials (Metzger et al, 2010, p.426). Rieh (2002) has indicated the significant influence of the institutional level of a source (i.e. source reputation, URL domain type) rather than an individual level (i.e. author/creator credentials) when they making judgements of quality and authority. This might be because of the lack of availability of information about an author's credential on many websites.

The supplier of map data or information is used in this present research experiment to examine the influence of disclosed sources of reputation on users' perceived credibility. The sources of map information are divided into two - the supplier of the base (background) map and of the foreground map. This is due to the nature of map mashup technology that uses various sources of maps and data to present the information. On a website, the source of information is typically denoted in the form of a citation or a reference to other studies or links to other web pages. On an offline map, this type of information may be provided in the form of metadata in a separate document or embedded in the map legend. In an online map, this information may be identified at the bottom of the map screen or in a separate section of a web page.

There is an issue, however, about where to find this type information in a map mashup medium since there are no set guidelines about where to locate this element. The supplier of the background data of a map mashup may be easily identified at the bottom of the map layout, but it is unclear how to identify the source of the foreground data. Since some mashup developers are ordinary citizens who lack knowledge of mapping practices, this element is often not presented appropriately on a map mashup. Some mashups present this element implicitly, with the result that users have to scrutinise each feature to establish where the information has been sourced; sometimes it is neglected and not clearly presented on the map mashup. A study by Skarlatidou et al. (2010a) has identified the significant influence of a website provider's reputation in a Web GIS application on users' level of trust. The influence of information disclosed about the supplier of the map data (i.e. background and foreground data), as well as its reputation, on users' judgement of the credibility of a map mashup environment is, however, still unclear.

There is also an issue when the supplier of data uses an anonymous identity, for example, when the data are supplied in collective mode and informal IDs are used to supply the information. This might occur in a map mashup application that has been developed for participatory involvement, such as reporting broken utilities in a community or reporting lost and found items. Users' perceived credibility of the reputation of sources information supplied by such methods needs further investigation. A study by Bishr and Mantelas (2008) has proposed a reputation model in the domain of collaboratively contributed geographic information, which may be used to tackle this issue. The author has proposed a model to calculate a rating, based on the number of people contributing the information, the reputation of the contributors (supplier) and the distance between the contributor and the event.

Cambridge Advanced Learner's Dictionary References (Woodford and Jackson, 2003) defines affiliation as a connection with a larger organisation. In this present research, such a connection is expressed in terms of an embedded brand logo on top of the web page. Other styles of presenting an affiliation on a web page include the provision of a hyperlink exchange to the affiliated organisation, such as a text link, button or banner advertising (Park et al., 2002). Affiliation differs from individual credibility factors, such as the author/creator credentials, as affiliation involves the transfer of perceived credibility between networks of organisations. Judging credibility on the basis of a website's affiliation may be categorised as making a judgement using seeking strategies that minimise users' cognitive effort and time, namely cognitive heuristics. Metzger et al. (2010, p.417) indicate that 'cognitive heuristics as information-processing strategies that consist of useful mental short cuts, rules-of-thumb, or guidelines that reduce cognitive load during information processing and decision making'. Affiliation has been identified as one of the elements that influence the perceived credibility of an object (Fogg et al., 2003, Provost et al., 2006, Iding et al., 2009). An empirical study by Fogg et al. (2003) has identified affiliation as being generally among the elements least mentioned by respondents, but it is one of the dominant elements identified by health-expert reviewers (Stanford et al., 2002a). The influence of the affiliation of map producer who publish the map element on the top of a map mashup on users' perceived credibility is still unclear. The parameter of affiliation tested in this research was to represent the metadata related to sources; the term affiliation was referred to the connection of the producer of online map with other organisation.

Currency has been identified as one of the elements users take into account when judging the credibility (Iding et al., 2009) and quality of information (Charnock et al., 1999, Provost et al., 2006) and it has also been identified as a feature that induces trust in the online object (Wang and Emurian, 2005). The date of publication and the date of revision are among the indicators that assist web consumers in identifying how up-to-date the information presented on the medium is. In a mapping context, currency is refers to whether data are up-to-date. It is also regarded as temporal validity in the ISO spatial data quality standard (ISO19157), which is the validity of data with respect to time (Harding, 2005). Meanwhile, in a map mashup environment, the state of currency may be indicated in three ways:

1) the date of the map and information being published on the Web
2) the date of the map/information being updated
3) the date of geospatial data being captured from the ground (i.e. age of data)

The status of geospatial data currency depends on the update policy of the mapping agency that provides the data. For example, the Ordnance Survey of Great Britain update policies dictate that significant business and national interest areas, such as housing, commerce or industry, are updated within a six month period; less significant features, such as vegetation and small building extensions are updated within every five years; mountain and moorland areas are updated within ten years (Harding, 2005). The ages of images layered on Google Map are updated according to the coverage area. The currency of data varies, depending on the area. The age of the data, however, is not implicitly indicated on the Google Map. In order to identify the currency of images of certain areas, check-ups have to be conducted using a tool in Google Earth.

To what extent online map users, particularly citizens, are aware of such update policies is unknown. In map mashup applications using Google Map as their base map, the copyrighted date is always shown as the current year on the bottom of the map, which might be used by non-geo-literate map users as an indicator that the sources of map data are up-to-date. They might believe that the data presented are current because of the stamped date of copyright. Such misconceptions may affect the level of correctness of the data they obtain from the map. The currency of the foreground data and attribute information presented on map mashups may vary according to the sources of data. It may get complicated, particularly when the attribute information is derived from collaborative sources from anonymous and non-anonymous map users. A formal, structured way to describe the currency of such data from various sources, particularly in a map mashup environment, is required to guide developers in presenting the currency descriptions of the data. It can also be valuable in assisting map users to evaluate the credibility and quality of map information from that medium and, consequently, might prevent them from obtaining misleading information. Although the currency of data and information has been identified as one of the critical elements in judging credibility and quality of information in several domains, to what extent this element is used by map consumers when judging the credibility of map mashups has not yet been identified.

Displaying a seal of approval has been identified as one of the elements to indicate the accuracy of information (Fallis and Fricke, 2002) and is valuable in assessing the credibility of online information as well as generating trust in the web providers (Skarlatidou et al., 2011, Cheskin, 1999, Wang and Emurian, 2005). For example, the HONcode label acts as a seal of approval on many websites that disseminate health related information to consumers; another example is a VeriSign Trust Seal that is used within websites which do not require Secure Socket Layers (SSL) certificates when online transactions are made by consumers (Knight, 2010). According to Cheskin (1999, p.3), a Web based 'security brand' seal of approval, may generate perceived trustworthiness in websites when it is recognised by users. In the e-commerce domain, such seals of approval may act as a 'security brand' that has been designed to reassure consumers that security has been established when they connect with the websites. The influence of a seal of approval for online map mashup applications is as yet unidentified. In an offline environment, the colour coded traffic light (CCTL) seal of approval on food products has been implemented voluntarily in certain supermarkets in the UK (BBC News, 2007) and has become part of food policy in Western Australian health and school services. Kelly et al.(2009), which demonstrates that the probability of users identifying healthier foods from a traffic light labelling format was five times more likely than from a label using monochrome text information. As according to Fox (2006, p.12), users tend to pay attention to informative readily available label, such as labelling on food product more than checking the date and source of online information. The diminish diligent to check the sources and date might due to the presence of sources and currency of information on the web pages not clearly presented and not disclosed. This approach may educate and informed the general public on the uncertainty of the data and information they obtained from that medium. The same benefits will be generated if such a visual quality indicator is implemented on top of map mashup applications. Hence, one of the experiments in this present research will examine the influence of a Colour Coded Traffic Light (CCTL) label, which acts as a seal of approval on top of a map mashup, with regard to users' perceived credibility in the information presented.

Other parameters examined in this present research relate to the design look of the website, which includes colour scheme and symbol design. A study by Fogg et al.(2003) has identified design look as one of the dominant elements assessed by web users when judging the credibility of websites. In that study, the term is used to describe the keywords used by respondents to describe visual design aspects of the websites, such as layout, typography, white space, images, colour schemes etc. Other studies have also identified this element as one indicator to assess the quality of information (Barnes and Vidgen,2003) and as a feature that induces trust in online websites (Wang and Emurian, 2005). A study by Skarlatidou et al.(2011) has indicated map and web design as one of the features that generate trust in the Web GIS applications. Notwithstanding, the aesthetic aspect of map design has been identified as having a low influence on users' level of trust in the map provider (Skarlatidou et al., 2010a). The influence of design look on users' perceived credibility in the information presented on map mashup applications needs further investigation.

2.4.2 The issues of metadata

In geospatial domain, metadata is typically presented in a document that contains information about the data. The identification of data, spatial extent, quality, spatial and temporal schema, spatial references and distribution of data are all key elements of information that should be presented in metadata (ISO 19115). Generally, one purpose of metadata is to assist prospective users to determine the fitness of the data to suit their particular requirements. Metadata also uses as a basis to assess the credibility of the data and information. Various standards for metadata have been developed at international, national and local levels. Nevertheless, the effectiveness of metadata in guiding end user assessments has been questioned, particularly on the grounds that end users have difficulty understanding the contents.

- Documentation of metadata is producer-centric

It is claimed that documentation is more producer-centric when the report is not based on the expectations and requirements of end users. Current metadata standards and specifications are considered as static and grounded in data production (Comber et al., 2007; 2008). Documentation is based on the perceptions of the data producer. An analogy that could be used to explain this is that it is similar to producing a manual for a kitchen gadget, which should include relevant information to aid users to utilize that tool rather than describing how the tool was produced.

- A particular skill is required to document and to interpret metadata

The creator of metadata has to take a step forward when documenting metadata by including relevant information that might be expected by data consumers. This might be a reason why metadata is frequently not used by some end users, particularly non-professional users. They may not know how to use the documented information to define the suitability of the data for their purposes. A particular skill is required to interpret the contents described in metadata before they can determine the fitness of the data for their purposes.

- The need for user- focused metadata

If a similar form of metadata were to be implemented in map mashup applications, those users drawn from non-professional and non-geoliterate backgrounds might face difficulties when using it. Comber et al. (2007, p.2) emphasise the need to document metadata that is user-focused and which emphasises the operational use of the data in order to allow users to determine the usefulness for the task in hand. User-focussed metadata is ‘Information that helps the user assesses the usefulness of a dataset relative to their problem’ (Comber et al., 2007, p.6). These authors suggests that there is a need for a new form of metadata that would incorporate third party reviews, such as expert opinions and end users’ feedback, stating their knowledge and experience about the data; an automated mining tool is required to analyse such reviews before they can help users determine the fitness of data for the tasks in hand.

- The unstructured metadata embedded on map mashup

The provision of metadata in map mashup applications is currently informal, unstructured and less comprehensive than that in geographic information systems; typically mashup developers provide a very basic identification of the data, such as the author or sources of data and the currency of data. Meanwhile, in crowd sourced based applications such as those using OpenStreetMap data, the platforms typically incorporate very simple metadata into the map data structure (Poore and Wolf, 2010). The non-standardisation of layouts to present a metadata element on a website may be a further reason for the failure by users to consider metadata during assessment. Metadata may be located at any frame in a web page or any page of a website. Such non-uniformity may cause web users to spend time looking for that element which they are not prepared or motivated to do. It might be even more problematic if the design or placement of that metadata is unclear or not highlighted. Fichter (2009, p.16) has arisen this concern by stating that casual mashup users often cannot easily discern who provided what piece of data, hence making difficult to assess credibility and authority of the information. The assessment of metadata may be considered to be an additional task to the main web browsing activity. As argued by Scholz-Crane (1998), activities that require additional tasks are unpopular among web users.

Generally, there is no obligation for websites to display their metadata as part of the data and the information disseminated. As argued by Warnick (2004), a website is a medium comprising an authorless environment where information about the data or messages presented, particularly in terms of the identification of sources, is not always available. It is typical to have information and data on map mashup applications that state neither the contributor nor the currency of information. The presence of sources and currency of information on the web pages that do not disclosed and not clearly presented lead to the diminishing of web users' diligence to check the sources and the date of information (Fox, 2006, p.12). An example is the case of a map mashup that showed the locations of reported cases of swine flu during the pandemic in the United States. The first version of this application did not clearly state the author of mashups or the currency of data; the information was presented in a chaotic form to visitors to the site due to the increasing number of cases being reported.

Another issue is related to the unavailability of at the point of use of metadata associated with the data. Although various standards for metadata have been developed, it is frequently either not captured or linked to the core data. As claimed by Sidda (2009), it is typical to have a dataset with no metadata at all. This might be because of the tedious process of documenting comprehensive metadata, which requires the developer to allocate a certain amount of time to the task and the fact that it is labour intensive. The process of re-documenting metadata becomes more difficult if the data, without metadata, have been obtained from secondary sources. Producing metadata can be more expensive than the cost of producing a dataset (Devillers et al., 2007). As metadata is typically kept in a separate file from the data, there is a possibility that metadata is not transferred when data are exchanged between the data producer and users; the case of metadata being unavailable is quite common. This issue also occurs in an online environment when using a website as a communication medium, including map mashup applications.

By having an credibility assessment tool as proposed in this study (see Chapter 9), more elements in the metadata, including credibility elements, quality information as well as third party reviews about the data, can be extracted from the online maps itself or relevant resources on the Internet. An analogy to this notion would be the independent portals such as 'booking.com' and 'Tripadvisor.com' that allow customers to give reviews about a specific hotel; a tool to compile relevant information about an identical hotel on the World Wide Web might be useful for users to find the hotel that would best suit their requirements, without the need to visit several portals to obtain the information.

In map mashup environment, information in metadata is valuable to help users determine the credibility of the information presented in that application. The metadata might not limited to the elements proposed by the geospatial data standards, but it could include the aspects that other studies used to assess credibility as have been reviewed in the previous sections. The proposed elements to assess credibility of online map mashups (see Chapter 9) have considered this issue; hence proposed credibility elements that covered the elements from the geospatial data standards and elements of users' perceived credibility from several domains. Metadata is commonly used to assess the fitness of purposes of geospatial data and mapping applications; Due to several issues related to current metadata and the nature of current map mashups, credibility could be evaluated through metadata that not limited to the rigid aspects of quality and accuracy as defined by the standards; but more flexible and adapts relevant supporting credibility and trust elements that could be used to assess map mashups from misleading, false, and inaccurate information.

This section describes the elements that users used to assess credibility of online information in several domains. Then, the credibility elements tested in the experiments are discussed in the middle section. The latter section briefly discusses the issues of metadata in general context and relate the nature of metadata in current map mashups. Next section will describe how web users assess credibility of online information

2.5 The way web users assess the credibility of online information

In general, several studies have examined how web users make a credibility assessment of an online environment. For example, a study by Metzger (2007) found that the tendency of web users was to be less than diligent in checking or verifying the sources of messages, for example, in verifying the author credentials. Additional assessments that seem time consuming and involve additional effort are less likely to be considered and tend to be neglected in a web user's browsing activity. Moreover, a study by Scholz-Crane (1998) identified the tendency to conduct a less than thorough assessment when web users are assessing the quality of web information; the findings showed many cases of the use of single criteria when judging the quality of sites. These findings support the views from the Elaboration Likelihood Model (ELM) (Petty and Cacioppo, 1986) about the two routes of users processing information. Central processing occurs when users have the motivation and ability to scrutinise the message; in this route, users will make the effort to process the content of the message, checking the truthfulness of the message that comes across to them. With the peripheral route, however, users are not diligent in critically checking and validating the message; they make a judgement based on simple inference, for example looking at the source of the message, at aesthetic aspects or at the structure or number of arguments. Judgement based on the peripheral route typically occurs when users lack the motivation to process the message or do not have the ability to process the message, possibly due to lack of knowledge or experience to process information on that topic.

Prominence-Interpretation theory has been specifically conceptualised to study the ways in which users assess the credibility of websites. There are two stages of this theory, whereby users make judgements on the basis of the elements they notice on the object of assessment; judgement will then be based on that element(s). According to Fogg (2003, p.722), ‘prominence’ is the likelihood that a website element will notice or perceived. The author suggests five factors that affect this prominence component:

1) The involvement of users who have the motivation to inspect the truthfulness of the message and the ability, including the relevant skills, to conduct the assessment and knowledge on the topic to support the assessment.
2) Experience of a user. The level of user experience - either novice or expert - in the subject or in the usage of Web technology
3) The background differences between individuals. For example, the literacy level and the learning style.
4) The topic of the website. For example, users may be less critical when assessing the credibility of an entertainment website, compared to a news website.
5) The task conduct; the influence of the level of the task, either for surfing websites, seeking information or making transactions.

Meanwhile, ‘interpretation’ is a person’s judgement about an element under examination (Fogg, 2003, p.723). The author suggests that three factors affect this component:

1) User assumptions to interpret the element under examination, which may be influenced by culture and past experience
2) Skill and knowledge of users concerning the topic of the message and in conducting appropriate judgement
3) User context; for example, the level of constraint when assessing the message

This theory is drawn particularly from a series of qualitative studies by Fogg et al. (2003). In that research, the most dominant credibility element given by respondents is based on the design look of the object under assessment. The majority of respondents in that study most likely did not conduct the central route processing suggested by Petty and Cacioppo (1986) in Elaboration Likelihood Model (ELM). The percentage of respondents (n = 2684) who mentioned the importance of information accuracy and authorship when judging the perceived credibility of a website was below 14% compared to respondents who mentioned design look in their criteria for selection (46%). These findings contradict other studies that demonstrate the importance of critical elements, such as the expertise related to sources (Fogg, 2002, Hovland and Weiss, 1951) in determining the credibility of object assessment. The authors argue that there are a few reasons why less central route processing occurred among respondents. This might be due to (Fogg et al., 2003):

1) The judgements being made on the basis of the elements that respondents noticed

2) Lack of motivation to check or validate the message
3) Lack of ability, knowledge or experience to judge the topic or conduct appropriate credibility assessment
4) The low level of engagement to the given task in that study, where the tasks given did not explicitly request respondents to scrutinise the information on websites, which might have affected the responses given by respondents.

An empirical study by Ferebee (2007) supports this view by demonstrating that, in non-situational involvement (which means a low level of engagement with the context), users tend to notice more of the medium related elements, p. for example, the design look, design structure, functionality, security and technical capability. Meanwhile, more message related elements, such as information accuracy, usefulness, and clarity, are noticed when there is a high level of involvement. This occurs when users have deep engagement with the context; for this study, respondents were requested to search for specific information from the experimental websites. Moreover, this study indicates the lesser influence of users' endurance or experience level on the topic to a change in the focus of processing (i.e. either central or peripheral) and the elements being noticed. In this study, there was no significant difference in the categories of elements being noticed between low enduring (experience) and high enduring groups, whereas there was a significant shift in the elements being noticed when respondents were requested to deeply engage with the tasks. Therefore, Ferebee (2007) argues that deep engagement with the task appears to be the primary driver for the shift of focus in information processing (i.e. either central or peripheral) as well as the elements being noticed by respondents.

Some research in the literature discusses how users assess the credibility or trustworthiness of online websites. There are limited studies that examine how web users assess the credibility of online maps, including map mashup applications. There is an empirical study by Skarlatidou et al. (2010a) that assesses the perceived trust of non- expert users in web GIS applications. The findings show the low influence of aesthetic issues on the perceived level of trust due to the reputation of source, a finding that might indicate the tendency of web GIS users to assess the map critically by considering the sources of information. Models of how users assess online information from the Web have been developed by a few researchers (for example see Hilligoss and Rieh, 2007; Metzger, 2007; Wathen and Burkell, 2002). It is still unclear however, whether there is a similarity in the way users assess and judge the credibility of messages from a website and the information from a map mashup medium. Table 2.3 below presents the comparison between the studies that might be related to this present research.

Table 2-3 The comparison of related studies that examine how web users assess online websites, including online maps

Studies	Sample	Methods	Stimuli	Purposes	Finding
Skarlatidou et al., (2011)	Eight respondents (unidentified)	Evaluate one Web GIS application and rate it using rating scales	One Web GIS application	To examine the influence of users' trust to the trust inducing features on Web GIS	1)The high influence of a source's reputation to the level of users' trust on the application 2) The low influence of any aesthetic element on users level of trust
Fogg et al. (2003)	More than 2500 respondents (citizen)	Compares two websites and rates their credibility using an open-ended questionnaire	Two websites	To examine the influence of credibility elements on web pages to users judgement on their perceived credibility	1) The high influence of design to users' judgement of the credibility of a website 2) The low influence of authority elements
Metzger (2007)	274 to 718 respondents (students and general public adults)	Compares two web documents and rates the frequency using 5-point scales	Two web documents	To examine the frequency of checking credibility elements	1) Respondents check the currency, comprehensiveness and objectivity most often. 2) Checking of the authority related elements was carried-out least often.

2.5.1. Credibility assessment and map use

Map is a standardise description of geographical features but is read in specific places. Map has been used in various contexts including navigation, communication, planning, ownership, general references, to understand spatial phenomena as well as forecasting. To understand how and why people use a map, six dimensions of map uses, which were initiated by the International Cartographic Association (ICA) Map Use Commission, have been identified. According to Carter (2005), the dimensions relate to:

1) Users of maps – consumer versus producer
2) Uses of map – generic uses, functionality, levels of map uses, tasks during use
3) Environment of map– paper based map versus digital map, currency, interactivity, access
4) The nature of map – languages, readability, misuse, data classification interpretation issues
5) Map user communities
6) Societal impact – the use and abuse of maps

The studies examining the issues related to map use are divided into two broad categories (Elzakker and Wealands, 2007)

1) Those focused on cognitive and perceptual map use
2) Those applied to more holistic functional map use

The first group focused on the cognitive and perceptual aspects of how maps are used. According to Slocum et al. (2001, p.68), such research has a long history in cartography domain where experiments were developed for studying static map use; the focus has been on comparing relatively narrow alternatives for a narrow tasks. The latter group consider the context of how a map is used (operated) in a real situation. These studies are based on clear notions that a map is made for a particular purpose; hence, it is important to conduct a study of map use according to its context. This type of research is more comprehensive and emphasises more than the cognitive process investigating the aspect of map use as a whole (Elzakker and Wealands, 2007).

There are several studies that applied more holistic approach in investigating the map use. For example, Yarnal and Coulson (1982) distributed maps and questionnaires to the users at recreational park to examine the influence of map use and map design upon trail use in the Yoho National Park, Canada;

Parker et al. (2013) using focus groups and observations in real Kayak trip to identify the elements that user used to assess the sources.

Several studies have examined the issues related to map use, but have focused on the perceptual and cognitive approach. For example, Linden and Sheehy (2004) examined the methods that were appropriate to elicit environmental related perceptions using maps and verbal questionnaires; the context of map use was general and volunteer respondents were requested to rate the cleanliness of each counties in Ireland. Hurst and Clough (2013) examined how people used and viewed online and paper maps. In this study, respondents were given a list of tasks to simulate a situation of a map user choosing a map to plan their long, medium and short-distance travel.

Among the studies that examine the aspect of credibility and trust, a few were conducted by considering the context of map use in real situations. For example, Scholz-Crane (1998) and Metzger (2003) examined how undergraduates assessed the credibility of online information by using the right samples and appropriate contexts. Notwithstanding, there are a few studies that have focused on the cognitive aspect of how web users assessed online information such as below:

1) Fogg et al., (2003) investigated the credibility elements using volunteer subject sampling where respondents had to evaluate the credibility of a list of websites.
2) Muehlenhaus (2012) used simulated contexts (USA radiation maps) and distributed questionnaires to online volunteered respondents to examine the influence of rhetoric map design in map users' trust.
3) A study by Skarlatidou et al. (2010b), examined the static trust (credibility) elements and usability of 'What is in your back yard' websites by using map based questionnaires. In this study, respondents drawn from university students' volunteers had to browse the websites and give their perception related to their usability and trust of the website via a survey. This study also ran a simulated scenario to stimulate respondents.

According to Bazire and Brézillon (2005, p.38) context plays an important role in domains where reasoning occurs, particularly during the understanding, interpretation and diagnosis processes; context is a collection of relevant conditions and surrounding influences that make a situation unique and comprehensible; the context acts like a set of constraints that influence the behaviour of a system (a user or a computer) embedded in a given task. A context is influence by the entity concerned, by the context, its focus of attention, its activity, its situation, its environment and its observer. In other words, to understand how and why a situation occurs, it needs to be interpreted according to its context.

User's perceived credibility is influenced by the context of user and the context of assessment. From prominence-interpretation theory by Fogg (2003), a user assesses credibility based on the elements that they noticed on the map. What appears prominent on the map will be assessed and interpreted and will be used in their judgement. The author summaries there are a few factors affecting the prominence, including involvement of the user, topic of the Website, user task, user experience and literacy level. A few factors affect the interpretation of credibility elements including culture, past experiences, skill/knowledge of a user and context (e.g. the user's environment, user expectations). Specifically investigates at these users contexts could aid in understanding how and why the findings are as such.

Princeton (2002) examined credibility elements between two different contexts (i.e. health and e-commerce) from the perception of experts and demonstrated slight differences of dominant elements. For example on health websites, the dominant elements indicated reputation and source elements greatly influenced the health-site's credibility ranking. In the financial context, however the experts' perceptions were more heavily influenced by the motivations that drove the website authors to share the information. Table 2.4 presents the elements, in the view of the experts, influenced credibility in two different contexts.

Table 2-4 Credibility elements and its ranking in two different contexts

Rank	**Health context**	**Finance context**
1	Name/reputation/affiliation (43.9%)	Information focus (40.3%)
2	Source (25.8%)	Company motive (35.8%)
3	Company motive (22.7%)	Information bias (29.9%)
4	Information focus (19.7%)	Design look (16.4%)
5	Advertising (13.6%)	Information design (13.4%)
6	Design look (7.6%)	Name/reputation/affiliation (13.4%)
7	Information bias (4.6%)	Writing tone (10.5%)
8	Information design (3%)	Advertising (1.5%)
9	Writing tone (3%)	Accuracy (0%)
10	Accuracy (1.5%)	Source (0%)

This indicates that the elements perceived as prominent were slightly different and subject to external and internal influencing factors as well as the contexts. O'Donovan (2012) argued that in order to determine which features (elements) affect credibility, studies must concentrate on specific circumstances. Not all the elements that have been identified to determine credibility are embedded in the medium in each instance. This author has identified that the perceived credibility for each context scenario may vary and is subject to population and culturally dependent. According to Flanagin and Metzger (2000, p.531), the extent of how users assess the credibility of information varies and depends on the type of information sought. Reference (i.e. the information that user might want to refer to and look up) and news information commonly will be verified rigorously than either commercial or entertainment information. This could be implied to map mashup environment where users may critically judge the information if the information they sought associate with the community, news and current events since they want to refer to and use the fact or news obtain from the maps; but their assessments might less critical if the obtain information are for the purpose of entertainment and also due to either the risks or the impact of having misleading and inaccurate information are low in their contexts.

According to Phillips (1984), studies on map use must be conducted with particular maps and within a particular group of map reader. It is regrettable if map reading tasks are not conducted in regard to the nature of how the map is being used (Board, 1978). The way a map is viewed by a user is influenced by the map viewing environment, the knowledge to interpret the information, the motivation of the user, and the design and presentation of the graphic image (Carter, 1988). As argued by Perkins (2008, p.151), the studies that applied a scientific approach by conducting experiments, commonly oversimplified the contexts, overplayed the functional explanations and marginalised the irrational and the feelings of map users; the map use processes are hard to model and are not concerned with the wider social contexts of map use. Scientific approaches do not deal well with the diverse concerns which reflect the current real world, everyday uses of mapping in society and cultural concerns. On the other hand, according to Elzakker and Wealands (2007), the results generated from a qualitative approach are lacking in terms of data validation due to very few respondents being sampled in the studies; hence the results could not be used to generalise into the real population.

From this review, there is no doubt that to appropriately conduct an examination of the aspects of map uses, one must consider the nature of how the map is actually operated in real situations. As argued by O'Donovan (2012), studies in understanding credibility in general (without specific context) might not be able to provide a mechanism to customise the general elements to specific scenarios. The elements identified in general contexts might be useful to understand the common credibility elements and general views of users' perceptions to those elements. However, these general elements might not be all available or relevant to assess credibility within certain contexts or scenarios. This present research did not concentrated on the context of map use where the experimental settings were limited to the simulation of experimental based scenario; the implication of the results to other real scenario of map uses requires further investigation. Nevertheless, it is not a fatal flaw for those studies that are not holistically conducted and do not consider the contexts including how the maps are actually being used in real operations; if the results yielded similar findings, a fairly clear picture could be generated by analysing the relative merits of the results from similar studies (Phillips, 1984).

This section discussed the theory and concepts drawn from a few domains of how web users assess credibility of online information. A few studies that examine this issue are highlighted. This section also discusses the issue of map use and relates it within credibility assessment. Next section will discuss the issues that relate to the needs of an automated credibility rating assessment in map mashups.

2.6 The need for automated credibility assessment and labelling

As discussed in the Section 2.3 certain elements of an online medium influence the perceived credibility of users. This issue has been widely discussed in several other domains, particularly those using websites as their communication medium. At present, websites have been used in the geospatial domain to disseminate geospatially related information by integrating a web based GIS, a map mashup and crowd sourcing applications within that medium. Previous sections have discussed the issues that led to several reasons why such a tool has become a necessity in geospatial applications, particularly those implemented in an online environment. These reasons relate to:

1) The way web users assess messages and information they obtain from an online medium

2) The efficiency of metadata

2.6.1 The influences from other domains and research area

The research from other domains also leads to the need for an automated credibility labelling on map mashup applications. This section describes the research focuses from other domains in tackling the issue of credibility and quality of information presented on website medium. The latter part of this section describes the research in geospatial domain that focuses on disseminating the quality information of geospatial data via visualisation approach.

- **Research towards the implementation of automated credibility labelling in a few domains**

Transparency of information disseminated on websites is important in reducing the risks of web consumers getting misleading information or being trapped by fraud or scam activities. Assessment tools designed for an online environment have been well implemented in other domains to help consumers evaluate the credibility and quality of messages, services and products in their fields. For example, in the medical informatics domain, a Health on the Net Code of Conduct (www.healthonnet.org.) labelling system has been well implemented on websites that disseminate information related to health.

The system is managed by a non-profit making organisation, namely the Health On The Net Foundation (HON). In order to obtain the certificate label, an application has to be made to this foundation. Specified websites able to be endorsed by this organisation are eligible to obtain the stamped HONcode label on their websites. The role of the HONcode is to set rules of transparency among those websites that have managed to get the stamped label. There are eight criteria evaluated under this labelling system - authority, complementarity, privacy, reference, justifiability, transparency, sponsorship and advertising. These criteria have been selected on the basis of their suitability for the code of ethics in the medical informatics domain. Griffiths et al. (2005) however, have argued the lack of HONcode principles that too focused into accountability criteria where it should consider as well the quality of content disseminated ; therefore the authors emphasise the need to conduct an evidence based assessment before giving accreditation.

Nevertheless, most of the process of granting these labels is currently conducted as a semi-automated process whereby human involvement is still required. The process of detecting certified HONcode stamped websites on the Web however is currently conducted in an automated manner by the means of an embedded HONcode tool bar. A study by Mayer et al. (2006) has developed a scheme and tools to support and improve the current medical quality labelling implementation. This study proposed a labelling system that could locate medical websites and monitor already labelled medical websites in automated manner. The architecture utilises semantic web resources including a web crawling system, a website spidering system and an information extraction system. Quality experts are required for this system, particularly at the final stage to make informed decision of adding or withdrawing a site from the directory.

A few portals have been proposed as gateways to a collection of trustworthy and reliable websites; For example, Oprhanet (2013), a portal for rare diseases and orphan drugs, which provides access to reliable and trustworthy websites within those topics; Provisu (2013), a portal that provides access to a collection of high quality websites which relate to vision disorders and eye diseases and WRAPIN (2011), a portal that indexes trustworthy medical databases, including HONcode accredited websites as well as medical scientific articles. Nevertheless, most of portals rely on a typical search engine concept, namely, using search tools to search for the relevant documents. Several studies have demonstrated more advanced search where users can customise the content to filter through information based on individual interests and quality requirements. Gaudinat et al. (2007a) have proposed a method of filtering the collection of websites that have been accredited onto the HONcode database through an automated statement detection system. In that portal, users can use a filter tool to find websites that match their requirements based on HONcode principles. Eysenbach et al. (2001) have proposed an automated tool that can detect several elements of metadata which can then be used for automatic comparison with users' own sets of preferences.

Meanwhile, in the e-commerce domain, a few labels endorsed by third parties have been used to generate the credibility and trustworthiness of the messages and services provided by companies conducting business in an online environment. Third party endorsement has been found to be associated with credibility perception (Hong, 2006). For example, labels such as VeriSign or TRUSTe have been stamped on most online banking websites to generate a sense of safety and confidence among consumers when conducting transactions via websites. As demonstrated by Cheskin (1999), such stamped labelling is one of the criteria to generate users' trust in e-commerce applications.

The implementation of rating systems based on traffic light colour coded labelling on have been applied to certain food products, electrical appliance energy ratings and in the United Kingdom (UK) car carbon dioxide emissions. This colour coded scheme has been implemented voluntarily in certain supermarkets in the UK (BBC News, 2007) and become a food policy in Western Australian health and school services (Western Australian Department of Health, 2009). A study by Kelly et al. (2009) demonstrates that the probability of users identifying healthier foods from a traffic light labelling format was five times more likely than from a label using monochrome text information. As according to Fox (2006, p.12), users tend to pay attention to informative readily available label, such as labelling on food product more than checking the date and source of online information. The diminish diligent to check the sources and date might due to the presence of sources and currency of information on the web pages not clearly presented and not disclosed.

- **Research related to visualisation of geospatial data quality**

Presenting quality information via visualisation approaches could increase users' awareness to the quality and uncertainty of the data they used. In the geospatial domain, research into quality assessment tools has been conducted by several studies. This is partly on account of problems relating to data quality presentation (Goodchild, 2007c):

1) Lack of the means to present hierarchical quality information that typically describes data at a global level
2) Lack of capacity to link quality information with spatial data, error modelling or visualisation packages
3) Dynamic updating being disallowed due to the static statement of data quality. There is no updating mechanism to produce quality information for derived secondary data. There is difficulty in automatically updating unstructured text-based descriptions

A few studies have proposed visualisation tools to communicate the quality of data to users. Devillers et al. (2005, 2007) proposes a model and a prototype tool to support a development of visual spatial data quality assessment which can be integrated on a desktop GIS application. This study highlighted the need to structure heterogeneous geospatial data types (i.e. global, layers and features) and data quality indicators using a multidimensional database approach.

It is typical used in the field of Business Intelligence, particularly to support data warehousing and online analytical processing (OLAP) applications. The advantage of using this database is due to its capability to rapidly process users queries related to summarising business operations and trends. Data in this approach are structured using a multidimensional data model. For example, in a business perspective, dimensions (D) are Product, Store and Time. Measure (M) is Sales. Each D stores details about the dimension, while M is stored in a fact table

(F). Each tuple of F contains the measure plus pointer (i.e. foreign keys) to the dimension tables (see Papadias et al., 2001).

In the GIS domain, dimension (D) refers not only to numeric spatial (x,y,z) and temporal (t), but includes semantic, temporal and spatial hierarchical concepts (e.g. continent→ country→region/state or object class→object instance→geometric primitive); each fact (F) contains measures resulting from the intersection of all dimensions at a given level in their respective hierarchies (Devillers et al., 2007, p.269). This study integrated Spatial On-Line Analytical Processing (SOLAP) into a GIS application to enable the rapid and dynamic communication of quality information at various levels of detail. The proposed Quality Information Management Model (QIMM) allowing the management of spatial data information within a data cube.

The prototype system is designed to enable navigation assessment of data quality information in two different ways - by quality indicator dimension or analysed data dimension; in quality indicator dimension, user can explore quality information along a quality indicator hierarchy. Table 2.5 presents the example of an indicator hierarchy at four levels of granularity. In analysing the data dimension, assessment could be conducted by navigating from different levels of detail; for example from the quality of the whole dataset down into the quality of a feature (object).

Table 2-5 Examples of Indicators hierarchy at four different levels (source: Devillers et al., 2005)

Global	**First level**	**Second level**	**Third level**
	ISO indicators	Completeness	Commission
			Omission
		Logical consistency	Conceptual consistency
			Domain consistency
			Format consistency
			Topological consistency
		Positional consistency	Absolute accuracy
			Relative accuracy
		Temporal accuracy	Accuracy of time measurement
			Temporal consistency
			Temporal validity

Quality		Thematic accuracy	Classification correctness
			Non-quantitative attribute correctness
			Quantitative correctness
	Other quality indicators	Data producer reputation	
		scale	
		Spatial extent	
		Temporal extent	
		Timeliness	
		Legal constraints	

Yang and Wang (2004) and Yang (2007) addressed the issues of spatial data quality visualisations in online environments. The authors proposed a more comprehensive framework for visualising spatial data quality using an object-oriented approach specifically for a Web and mobile environment. The proposed framework organises information around three basic components:

1) visualisation contexts;

2) visualisation contents;

3) visualisation techniques.

If a visualisation context is specified by a user, the corresponding visualisation content will be defined; then visualisation techniques can be determined according to these two combinations. Map use and information stage form the visualisation contexts component; hierarchical level and error models supplement data quality information in the visualisation contents component; the visualisation techniques component provides approaches to visualised spatial data together with quality information; three possible visualisation methods have been demonstrated including animation based on confidence level settings, feature filtration according to threshold values and statistics graphs. The object-oriented model has been chosen in their study in order to (Yang, 2007, p.167);

- maintain the connection between data and quality information;
- maintain the close ties between quality information at multiple levels. It is recognised that spatial data are structured in hierarchical levels from the global dataset level down to the primitive object level. The inheritance concept in object-oriented modelling can maintain the relationship associated between the upper level and lower level;

- it could minimise update efforts and data redundancy problems

In that study, Geographical Markup Language (GML) and Scalable Vector Graphics (SVG) were used to develop the prototype visualisation tools; GML is chosen due its ability to store and transport quality information with spatial data; GML is an Open Geospatial Consortium standard using a XML schema where the structure definition of document and data integrity of the information transmitted between client/server can be automatically validated against the specifications (Yang, 2007, p.114).

A study by Mass et al. (2011) proposed the use of a geo-label on scientific maps. The label is based on information about data quality which will be extracted from metadata, the data itself and the validation process with in-situ sensors, provenance information, and user-feedback. In this study, data quality information may be acquired in two conditions:

1) a static state using qualitative and quantitative testing, and
2) an operational state when tracking the processing state.

This is in line with the approaches suggested by Devillers et al. (2007) that suggested a bottom-up and a top-down approaches to feed the quality information database. In bottom-up approach, metadata and any other data quality information that can be assessed, are valuable to supply a quality database; data aggregation could be used to document metadata at a detailed level as well; for example, horizontal spatial accuracy of a specific road segment can be aggregated into higher level information to generate average spatial accuracy of the road layer of a selected area.

Meanwhile in a top-down approach, quality information is collected from third party sources such as an expert who has very good knowledge and experience elating tot the relevancy of the data to suit the application in hand. Experts may provide insights on the spatial heterogeneity of the quality of certain datasets as well as assessing data with respect to the period of measurements and other informal criteria. The top-down approach is a good complement to the bottom-up approach, particularly due to the incomplete, too general or non-existent metadata in many datasets (Devillers et al., 2007, p.270).

A common stance among several studies of spatial data quality visualisation is on the agreement of the unsuitability of storing data quality information using graphic depiction solely at the overall level. The fact that geospatial data as well as its data quality are heterogeneous in one dataset is undeniable. As indicated by Devillers et al. (2005), such heterogeneity is not adequately recorded in most current metadata to properly assess data quality for the subset of data being used. Instead, metadata is often documented as the average quality of the entire map sheet which may hide significant variations at more detailed levels (Devillers et al., 2007, p.276). Quality information has to be managed in hierarchical format that stores quality data at global level, down to layer level, object level and primitive level. This is in line with Nuth et al. (2007) who argue that a single overall quality statement may not suffice to help end users make an assessment to determine the fitness of purpose of the data. Therefore, the authors suggest splitting the visualisation of quality information into several parameters in order to provide a more detailed quality overview.

Communicating data quality information down to the finest details of an object however may only be practical for professionals and expert users. By considering this, the tool proposed by Devillers et al. (2007) is designed for such target users; the authors believe however, the ideal system should provide clear output regarding the fitness of data for a given use in an automated manner. Nevertheless, due to the constraints of knowledge and technological advancement at present and in the near future, the informed decision has to be made by experts. A question is raised, however as to the appropriateness and cost efficiency of providing such detail of quality information if the end user is drawn from the general public without the knowhow or tools to use it. Yang (2007) also in line with the notion of an ideal solution proposes providing quality information at each level, including at the lowest level, in order to respond different users' need. However, due to the issue of storage efficiency, data quality information has to be in aggregated models. The author suggested the aggregation level has to be as high as a value can be accepted and at the same time as representative and as deep as necessary. At this highest level of aggregation, it can minimise the data quality update efforts and data redundancy issues (Yang, 2007, p.99). Another issue that has to be considered is the basis to aggregate the spatial extent. A prototype system developed by (Devillers et al., 2007) presented visual quality indicator values that are aggregated based on the current view of the map being displayed. Other possible solutions are to base the aggregation on the whole dataset, layers or collection of features.

Yang (2007, p.173) addressed a possibility to provide visual quality information on commercial maps such as Google Map and Microsoft Virtual Earth. This is due to the current trend where web citizens tend to use such maps to browse, locate and query spatial-reference information; this could be implemented by extending its Application Program Interfaces (APIs). This approach may educate and informed the general public on the uncertainty of the data and information they obtained from that medium. The same benefits will be generated if such a visual quality indicator is implemented on top of map mashup applications.

2.7 Summary

This chapter begins with the introduction to map mashups where apparently inherits from Web GIS domain. Then brief reviews of the definitions of mashup and the mashup use contexts are discussed. The current issues related to neogeography products based, particularly map mashups are highlighted. The correctness, credibility (believability), trust and validity of data and information from such medium are highlighted. These issues lead to the reviews of the concept of credibility assessment and the elements of credibility that have been identified in the literatures, including the justifications of credibility parameters tested in this research. Brief discussions of the relation between credibility elements and metadata and the issues of current metadata, particularly in mapping domain are given. These lead to the reviews of the approaches of web users in assessing credibility in offline and online mediums. The relations of map use during credibility assessment are discussed.

The first part on this chapter has discussed the issues that lead to the need to have credibility labelling on online map, particularly on map mashup applications. Therefore the latter parts strengthen this need by providing reviews of other research from other domains and research area that have demonstrated similar visions.

As can be seen from the review undertaken, the ways web users assess credibility of information in online medium and the inadequate informal methods in documenting and presenting metadata in map mashup reinforce the need:

1) to examine the influence of metadata when users assessing credibility of map mashup information and

2) to test a hypothesis that presentation of detailed metadata does not influence a user's assessment of credibility (trust) to map mashup information

Several studies from other domains and in geospatial domain that devoted to increase awareness of the credibility, quality, accuracy and uncertainty of the data and information by the means of graphic depiction strengthen the basis of this present research, which was to examine the influence of a credibility rating labelling on map mashup applications.

3 METHODOLOGY

3.1 Introduction

This chapter discusses the experimental approaches used to clarify the central questions relating to the tested elements that influence respondents' perceived credibility of information presented on map mashups. These approaches enable the exploration of the elements by the means of online map-based questionnaires. These questionnaires examine the influence of metadata related parameters specifically map data producer and data supplier and the influence of colour-coded traffic light labels on online map users who are making informed judgements about the perceived credibility of information presented on a map mashup application.

The flowchart below provides an overview of the steps taken to achieve the research objectives. The methodology is firstly described in general terms and followed by a description of the experimental design used in each experiment. The samples, dependent experimental variables, experimental design and questionnaires used specifically for each experiment are explained in each of the experiment chapters.

3.2 Research Methodology

The research methodology consists of the strategies used in the study to conduct the research and to achieve its objectives. Creswell (2009) refers to this approach as 'strategies of inquiry'. In general, several strategies for research may be possible, including quantitative, qualitative and mixed method approaches. The research framework in this present research applied mixed methods, a combination of qualitative and quantitative approach.

This research implied mixed method design in collecting and analysing the responses. According to Creswell and Clark (2007, p.5), mixed methods research is a research design that involves the mixture of qualitative and quantitative approaches in many phases in the research process. The methods focus on collecting, analysing, and mixing both quantitative and qualitative data in a single study or series of studies. The main idea is to provide a better understanding of research problem by the combination of quantitative and qualitative approaches.

The main approach however was quantitative methods. In this approach, there are two main strategies of inquiry - survey research and experimental research. According to Creswell (2009, p.12), survey research provides a quantitative or numeric description of trends, attitude, or opinions of a population by studying a sample of that population; it includes cross-sectional and longitudinal studies using questionnaires or structured interviews for data collection. Meanwhile, experimental research seeks to determine if a specific treatment influences an outcome. This present research combined these two strategies- survey (questionnaires) and experiments. A survey is used to collect responses from respondents, while the experimental component is based on a map mashup, to create the context and tasks, tests the research variables and acts as a stimulus of research.

Qualitative methods namely think-aloud protocol and structured interviews were conducted after running the survey based experiments. Think-aloud protocol is a common method in usability test domain. In this method, respondents are requested to think-aloud while completing the tasks. During usability test, respondents are able to vocalise the difficulties with the system and in some cases recommended changes to interface design and functionalities (Boland et al., 2013). This method is also applied in psychology and social science domains. In this present study, think-aloud protocol enables respondents to vocalise their perception whilst completing the experimental tasks, what they thought and why they chose such decisions. Structured interviews were also conducted after each think-aloud sessions. Think-aloud protocol also has been applied in the studies that examine the aspect of map uses. For example, Elzakker (2004) applied a combination of think-aloud protocols, observations, video recording and screen transaction recording to analyse the way users used maps in regional context.

In this research, these two types of data (quantitative and qualitative) were mixed by using embedded design. According to Creswell and Clark (2007, p.5), there are three types of mixed method procedures – triangulation, embedded, explanatory and exploratory design. The embedded design is when one dataset provides a supportive, secondary role in a study based primarily on the other data types. The basis of this design is that a single data set is not sufficient to answer each type of question. The uses of mixed methods in this research were to support the results generated from quantitative approaches.

The purpose of this research was to examine the influence of metadata when judging credibility of map mashup information. Several researches related to credibility and trusts in the literature were conducted by collecting quantitative data. For example see Fogg et al., 2003, Metzger (2003), Scholz-Crane (1998). Muehlenhaus (2012) used simulated contexts and distributed map based questionnaires to online volunteered respondents. Yarnal and Coulson (1982) distributed maps and questionnaires to the users at recreational park.

However, a few other studies collect qualitative data. For example Parker et al. (2012) using focus groups and observations in real situation approach to identify the elements that user used to assess the sources. This study investigates the aspect of the relevance of professional and volunteered sources in a real expedition where the criteria to determine the ‘information relevance’ are quite similar with the element of credibility. As argued by Suchan and Brewer (2000), qualitative methods could bring research closer to understand the real world problem in solving the issues related to mapmaking and map use.

Nevertheless, there are other relevant studies that implement mixed methods. A study by Skarlatidou et al. (2010a, 2010b) investigates the elements that influence trust on web based GIS application by using mixed methods; they combined heuristic evaluation (HE) which is a common method in evaluating usability of application that rated the respond in a rating scale with cognitive walkthrough (CW); cognitive walkthrough is an approach that collects users' respond in the aspect of cognitive and affective dimensions. A study by Elzakker (2004) applied mixed methods in the study of map uses; the study combine the questionnaires with think-aloud protocols, observations, video recording and screen transaction recording during the experiments.

There are particular reasons for using mixed methods (quantitative and qualitative approaches) in this research.

1) The use of open-ended questions (Experiment 1) (qualitative approach) was to allow the respondents to provide the information from their perspectives without giving any clue or hints from checklists. The drawback was that the answers led to in-depth explanation but were limited to a single variable response.

2) The use of closed-ended questions (quantitative approach), including check boxes and rating scale instruments, in Experiment 2, Experiment 3 and Experiment 4 were;
    a. to focus on the variables from the literature
    b. to test the research questions either to support or refute the relationships statements of the theories
    c. to test the theories to see how they applied to a number of people
3) The use of a think-aloud approach (Experiment 3 and Experiment 4) (qualitative approach) was to allow the respondents to provide the information from their thoughts and opinions whilst conducting the tasks. This would develop an understanding as to what they actually thought and why they did as in the experiments.

The main dataset of this research was quantitative data. Qualitative data were used to support the findings. The methods were implemented in a series of experiments. Figure 3.1 presents the phases involved in this research. The approaches in the four experiments in this research adopted mixed methods for the following purposes;

1) To collect data from more than 30 respondents to enable the data to fit in a statistical model and confirm the significance of the results (quantitative approach)
1) In Experiment 1, open-ended questions were used to collect the responses of 'why respondent accepts or reject the map'. This type of question was used to allow respondents to answer freely without being given any hint or clue. This was important for not limiting the view (responses) of respondents in giving their perceptions or answers (qualitative approach)
2) In Experiment 2, structured interviews were conducted to gain insight as to why they chose the map and the influence of the map elements in their judgements. The purpose was to validate the responses collected from the quantitative approach (qualitative approach). Spot the differences activity was conducted before respondent involved with the experimental tasks. Experiment 2a was conducted to confirm the findings demonstrated in previous experiment.
3) In Experiments 3 and 4, observation and think-aloud approaches were conducted to observe how users assessed the map and to understand the cognitive processes that influenced their decisions during the assessment, which were prompted by a series of questions (qualitative approach).

3.3 Research framework

This section describes the research framework including the flows and justifications of each experiment. Altogether, this research involved six phases as shown in Figure 3.1. The first stage consists of gathering and analysing the literature, then developing the issues of research based on problem identification. The results from phase 1 are addressed in detail in Chapters 1 and 2. Phases 2, 3, 4 and 5 are described in separate chapters that discuss the details of the conducted experiments, including the findings. The results from phase 6 are addressed in detail in Chapter 9 in which the proposed framework, including practical parameters, approaches and an algorithm to implement automated credibility assessment and labelling, is demonstrated.

Figure 3-2 depicts the experimental model or framework that applied in this research, namely embedded design. As have been discussed in previous section, there are three types of mixed method procedures – triangulation, embedded, explanatory and exploratory design. Embedded design is one of the approaches in mixed methods. This design is commonly used to support the dataset or findings between qualitative and quantitative methods. The experiments conducted in this research could be divided into two sets – pre-CCTL and post CCTL. The pre-CCTL is a term used to represent the experiments before proposing the colour coded traffic light (CCTL) credibility rating; whereas the post CCTL is a term to describe the experiments after proposing the CCTL rating on the map mashups. Three series of experiments were conducted in the pre-evaluation stage (pre-CCTL) and one experiment in the post evaluation stage (post CCTL).

Experiment 1 used open questions to collect qualitative data. Experiment 2 used closed-ended questionnaire to collect quantitative data. Experiment 3 and 4 used questionnaires to collect data. Think-aloud protocol and structured interviews were conducted to support the responses from questionnaires. The acronyms of ‘QUAN’ and ‘Qual’ applied in the Figure 3-2 are used to indicate that the major approach is quantitative (QUAN) where qualitative approach (Qual) was used to support the quantitative data. Figure 3-3 presents the details including the flow, procedures and output of each experiment. Table 3-1 presents the justifications of the progress of each experiment.

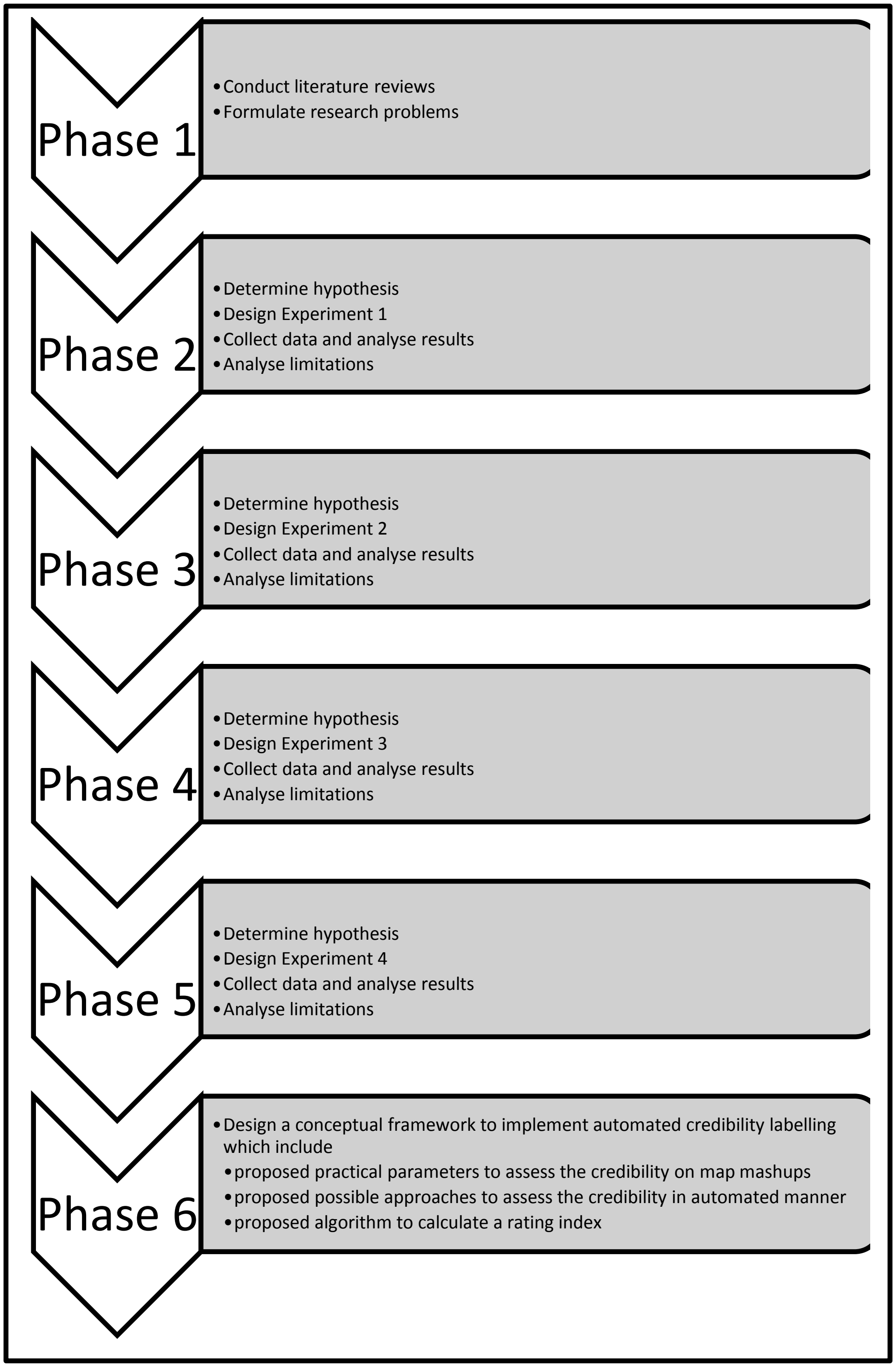


Figure 3-1Phases in the research framework

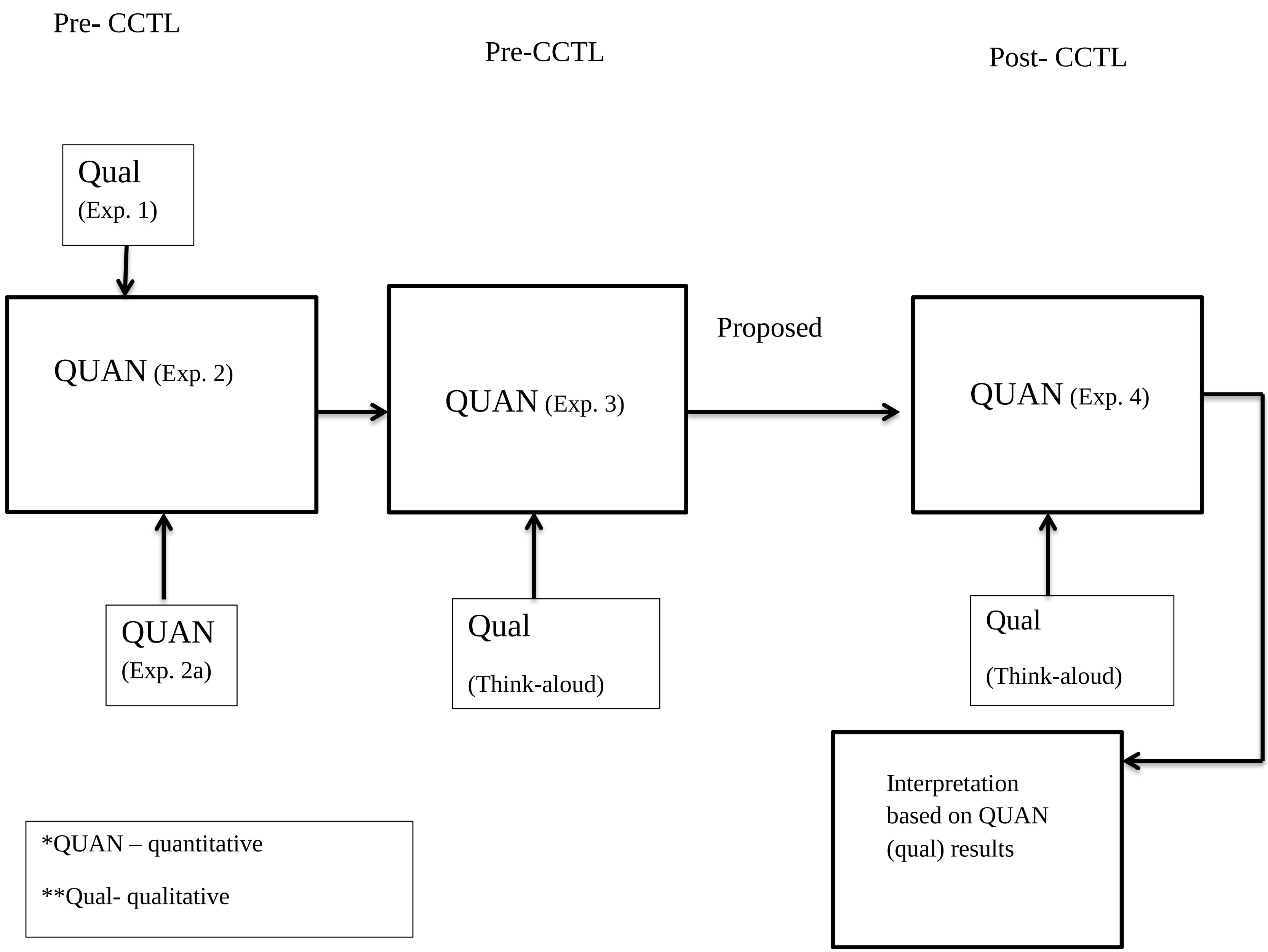
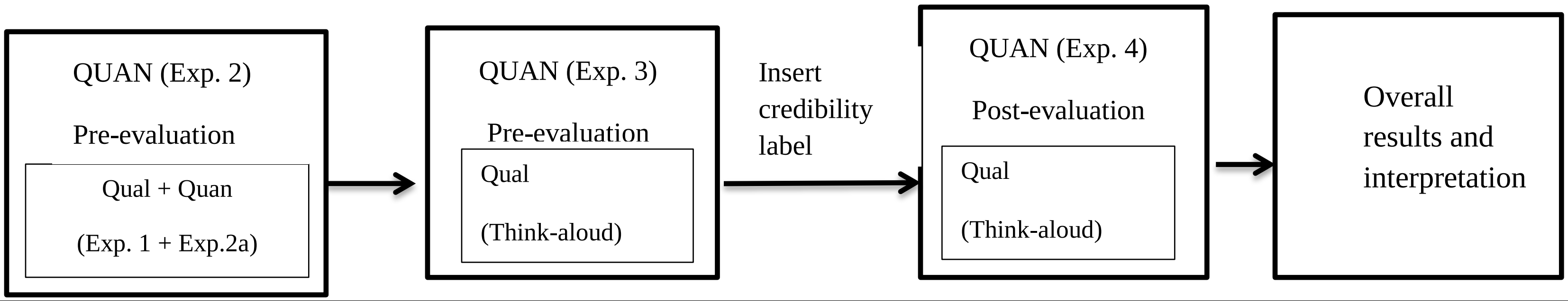


Figure 3-2 Embedded Design: Embedded Experimental Model



Procedures	**Procedures**	**Procedures**	**Procedures**
<u>Data collection</u> Structured questionnaires - Open ended question Section B (Qual) (Exp 1) - Closed ended checklist Section A (Exp 1) and Section A and B (Exp 2) - Spot the differences activity (Exp 2) - closed ended checklist + open ended question (Exp. 2a) <u>Data analysis</u> Score responses Rank order Categorical data Content (thematic) analyses (Qual)	<u>Data collection</u> Structured questionnaires - closed ended questionnaires (Exp 3) Think-aloud (observation, transcript (sound recording), semi-structured interviews) <u>Data analysis</u> Score responses Rank order Content (thematic) analyses (Qual) Comparison analysis (direct vs. indirect experimental context Content (thematic) analyses (think-aloud, interviews)	<u>Data collection</u> Structured questionnaires -closed ended questionnaires (Exp 4) Think-aloud (observation, transcript (sound recording), semi-structured interviews) <u>Data analysis</u> Score responses Rank order Content (thematic) analysis (Qual) Comparison analysis (with CCTL label vs. without label) Content (thematic) analyses (think-aloud, interviews)	Discuss the prominence of map elements that influencing respondents' judgement Discuss the prominence of metadata elements that influencing respondents' judgement Discuss the influence of CCTL labelling in respondents' judgement
Outputs	**Outputs**	**Outputs**	**Outputs**
-Numerical item scores -Themes and quotes -Group classification (geoliterate vs. non-geoliterate) -The basic elements influencing respondents' judgement -The influence of metadata elements in influencing respondents' judgement	-Numerical item scores -Themes and quotes -Group classification (geoliterate vs. non-geoliterate) -The influence of metadata elements in two different assessments (context vs. non-context based)	Numerical item scores Themes and quotes Group classification (geoliterate vs. non-geoliterate) - The influence of credibility label vs. without label	 Discussions

Figure 3-3The progress of experiments

Table 3-1 The structure and the progression of the series of experiments

Experiment No.	**Item**	**Justifications**	**Comments**
Experiment 1 (pilot)	Section 1: The map they chose (QUAN*)	To prompt respondents to analyse the set of maps	
	Section 2: The basis of their decision in selecting and rejecting the maps (in open ended question) (Qual**)	To collect responses of the map elements that influenced respondents' judgement (Qualitative (Qual) method)	Single item responses given by most of the respondents limited the coverage of answers. Maybe there were other factors that influenced them but were not written down.
	Tested parameter: identity of map producer The map design used unconventional symbology and colour scheme, which were contrary to professional mapping practices	To test the influence of source (i.e. map producer) The experimental maps were designed using the pre-set map styles available in CloudMade$_{TM.}$ This application, used to offer free services, but had just discontinued its services since May 1st 2014 to allow other application, namely MapBox$_{TM}$ to continue the services; this application used to offer custom maps (i.e. maps created by other players who then upload to use by other users) and had more than dozen	**Action taken (the next experiment):** include closed-ended questions with a list of possible answers include 'spot the differences' activity

		collection of pre-set maps that powered by OpenStreetMap data that could be used by citizens to create their online maps. The intention of not designed the experimental maps by following the conventional cartographic practices was to replicate the map design styles by neogeographers (i.e. via personal observations on the collection of map mashups reported by programmablewebapi portal between 2008 to 2010) which was still immature and often relied to available default and pre-set styles setting; the rendering of map mashups are dependent on the personal preferences of the map creator (Liu and Palen, 2013).	
Experiment 2	Section 1: The map they chose (QUAN*)	To prompt respondents to analyse the set of maps	
	Section 2: The basis of their decision in selecting and rejecting the maps (in order of ranking) (QUAN*) Tested	To collect responses about map elements that influenced respondents' judgement; the responses were weighted by order of ranking. The first main reason was to represent higher weight of influence. To test if there is any	The reputation level of data sources between maps were not easy to distinguish; hence may not strong enough to influence respondents to use that element as a basis in rejecting map.

	parameters: -foreground data supplier	influence of sources in other parameter (i.e. foreground data supplier), due to low influence of sources (i.e. map producer) in Experiment 1.	**Action taken (the next experiment):** Conduct another experiment to compare the influence of foreground data supplier if the comparison maps were supplied by different level of reputations.
	Experimental design: -Data on both maps are identical (i.e. no contradictory) -visual cues on both maps are different (symbols, colour scheme, labels) -data sources (same level of reputation)	To test the influence of sources if the data on both maps were no contradictory except on the visual cues (i.e. the design of symbols, visual scheme used and the style of labels); the reputation level of sources were indistinguishable (e.g. the University of Nottingham vs. the Ordnance Survey; city council vs. student union)	
	The map design used unconventional symbology and colour scheme, which were contrary to professional mapping practices	The experimental maps were designed using the Google My Places$_{TM}$ and the colour schemes selections were based on the pre-set map styles available in CloudMade$_{TM.}$ The selection of symbology was based on the default symbols available by Google Map. The intention was to replicate the map design styles developed by neogeographers which often relied to the available default setting; the rendering of map mashups are dependent on the personal	

		preferences of the map creator (Liu and Palen, 2013).	
Experiment 2a	The structure of questionnaire was similar with the main Experiment 2 Experimental design: -Data on both maps are identical -visual cues on both maps are slightly identical (differences only on the symbol design) -foreground data sources (were significant different in terms of reputation level) Tested parameters: -foreground data supplier	This is additional experiment to confirm the results of the previous experiment. To test the influence of sources if the data and visual cues on both maps were no contradictory except on the sources of data; the reputation level of sources were distinguishable (the University of Nottingham vs. the Starbuck café) To confirm the influence of sources and focus on one variable (i.e. foreground supplier)	The low level of the map use task may implicitly affect the findings. Is there any influence of source if the data between two maps appear contradict each other? **Action taken (the next experiment):** Increase the map use task level to include map analysis and interpretation. Comparing maps that present contradictory data
Experiment 3	Section 1: The map they chose (QUAN*)	To prompt respondents to analyse the set of maps	
	Section 2: To test the influence of map elements in a given case (QUAN*)	To collect responses of map elements that influenced respondents' judgements; the responses were weighted by using order of ranking. The first main reason was to represent higher weight of influence. The tested variable (i.e.	The results confirmed the previous experiments findings (Experiment 1,2, and 2a) **Action taken (the**

		the map producer) was maintained to compare the results and confirm finding of previous experiment (Exp. 1)	**next step) :** 1) to conduct think-aloud protocol and structured interviews.
	Experimental design: -Data on both maps are slight different (the locations of roadblock and the number of reported landslides) -map producer (were significant different in terms of reputation level)	To test the influence of sources if the data on both maps were contradict in showing the location of roadblock (work construction) and the number of reported landslides. Visual cues on both maps were presented with minimal different (i.e. symbol design, colour scheme) The comparison maps were produced by different producers that hold different level of reputations (the University of Nottingham vs. Jane Smith)	2) Conduct Experiment 4 to test the hypotheses to support the suggestions
	Section 3: To test the influence of map elements which were not restricted by the experiment task (QUAN*)	To compare the responses in two difference approaches of eliciting answer. To test the Prominence Interpretation Theory	The results indicated low influence of metadata in respondents' judgement **Action taken (the next experiment):** to test the influence of CCTL to assist respondents in

			making judgement
	Think-aloud protocol and structured interviews (Qual**)	To support the results of quantitative method	
Experiments 4	Section 1: The map they chose (QUAN*)	To prompt respondents to analyse the set of maps to compare the responses to the map without CCTL (Experiment 3) and with CCTL (Experiment 4)	The results yielded significant findings
	Section 2: Questions related to the influence of CCTL (QUAN*)	To test the influence of CCTL in respondents' credibility assessment	The results confirmed the previous findings
	Think-aloud protocol and structured interviews (Qual**)	To support the results of quantitative method	
*QUAN = quantitative *Qual = qualitative			

From Table 3-1, there are three experiments to serve several hypotheses of Objective 1 and one experiment for the hypotheses in Objective 2. The Experiment 1 was served as a pilot study of this research where it then became the basis in designing Experiment 2. Experiment 2a was conducted as an additional test to support Experiment 2. Experiment 3 was then conducted to support the findings of Experiment 2 and 2a.

In Experiment 1 and 2 the information differences between two set of maps were related to the map design styles (i.e. visual cues), including the colour scheme, symbol design and the value of metadata elements; for example the map data supplier of Students' Union versus The Nottingham City Council. Whereas the information differences between the two set of comparison maps (Map A and Map B) within Experiment 3 and 4 were related to the number and locations of landslides and roadblocks, map design styles and the value (parameter) of who produced the maps. The parameters differences are describe in details in each experimental Chapter 4, 5, 6 and 7 in section Experimental design and Material. The information between Map A and Map B in each experiment were designed to have slight differences; the intended purposes were to stimulate respondents to make informed judgement which map they will believe or perceived credible if there were two maps of a similar context but the data or information were different and conflicted between each other. Table 3.2 below presents the setting differences between comparison maps in each experiment; the last column presents the added activity of each experiment.

Table 3-2 The experimental setting differences between maps and experiments

Differences between comparison maps (Map A vs. Map B)	**Data**	**Visual cues**	**Sources**	**Added activities**
Experiment 1 (pilot)	Almost identical	Different - Symbols design, colour scheme, labels	Within same reputation level (i.e. map producer)	
Experiment 2	Almost identical	Different - Symbols design, colour scheme, labels	Within same reputation level (i.e. foreground data supplier)	Spot the differences activity
Experiment 2a	Almost identical	Very minimal difference (i.e. symbol design). No different on colour scheme	Significant different of reputation levels (i.e. foreground data supplier)	
Experiment 3	Contradictory data	Different - Symbol design, colour scheme, labels	Significant different reputation levels (i.e. map producer)	Interactive online map experimental task level increased think-aloud protocols
Experiment 4	Contradictory data	Different - Symbol design, colour scheme, labels	Significant different reputation levels (i.e. map producer)	Added CCTL rating label Think-aloud protocols

Although Experiment 2 and 3 served the same objective, the experimental tasks were slightly different. In Experiment 2, the task was to evaluate the maps in order to only choose one map that they perceived as more credibility to assist in self-guided campus tour. In Experiment 3, the task level was slightly increased; respondents had to evaluate and choose the map they perceived more credible to suggest the safest route for an ambulance to rescue the trapped victims during landslide disaster in the campus. In this task, respondents had to act as if they were the responsible officer to give the advice. The task level in Experiment 4 was slightly similar; the context was changed to analyse the safest route for them to cycle from one building to another building during landslide events.

In cartographic communication domain, Board (1978) in Nyerges (1991) have identified three major tasks for map use related experiments – navigation, measurements and visualisations. The common tasks in navigation involve movement from one place to another and incorporate subtasks such as search, identify and locate position on map, search for the optimum route on map, search for and recognise landmarks, search for and identify destinations and verify data. The measurement tasks include subtasks such as search, identify, count, compare, contrast, estimate, interpolate and measure. The visualisation task involve picturing mentally a terrain scene that incorporate subtasks such as search, identify, describe, compare, contrast, discriminate, generalise, prefer and like the pattern(s).

According to Elzakker (2004), the map uses tasks could be categorised into four groups – elementary, intermediate, temporal and overall. The elementary level generally deals with simple questions to the object itself; whereas the intermediate level commonly deals with another level of cognitive processes where map users tend to analyses the map, for example by looking at the relation of one class of object with other classes and its spatial distribution and measurement; temporal level is deals with analysing the map or series of map in terms of the changes within certain periods of time; overall level general deals with more complex analyses that involve interpretation (see Kimberling et al., 2012) and higher order map using tasks (see Board, 1984). Figure 3.4 presents the tasks that commonly conducted under these levels.

Geographic questions	Tasks
Elementary	
What is there?	1. to **recognize** objects (external identification)
At a given place, what is there?	2. to **identify** objects (internal identification)
At a given place, ho much is there?	3. to **estimate** amounts
Where is that geographic object?	4. to **locate** an object
Intermediate	
What is near that geographic object?	5. to **position** with respect to other objects
What is the distance to similar/other objects?	6. to **define** relative / absolute distance
Is that geographic object linked to other objects?	7. to **encounter** spatial linkages
Why is a geographic object there?	8. to **explain** a location
What is the spatial distribution of that object?	9. to **find** order, patterns or spatial anomalies
Where is the most / least?	10. to **quantify** spatial anomalies
Where are the limits of a spatial distribution?	11. to **delimit** a distribution
What comes in / what goes out?	12. to **connect** a region to the outside world
Temporal	
Has that geographic object always been there?	13. to **determine** changes
Have the spatial distribution patterns changed?	14. to **establish** trends
Which spatial processes are taking place?	15. to **detect** processes
Overall	
What are the influences from outside the region?	16. to **contemplate** spatial context
What relevant patterns are there?	17. to **recapitulate** the found patterns
Are there relationships between spatial patterns?	18. to **discover** correlations / dependencies / conflicts
Which factors cause the regional structure?	19. to **structure** the geographic information
Can different (sub-)regions be identified?	20. to **regionalize**
What are the region's geographic characteristics?	21. to **obtain** insight into and overview of the region

Figure 3-4 Map use tasks level (source: Elzakker, 2004)

Muehrcke (1979) in Kimberling et al. (2012) identified three strategies of map uses – map reading, map analysis and map interpretation. Liebenberg (1998) has summarised the levels of map use from the studies debated by Olson (1978), Muehrcke (1979) and Board (1984); level one comprises of map reading tasks such as identification of individual symbols (lines, polygons, points) and the differences of these symbols’ shapes, relative size, level two tasks comprise of recognition of the spatial pattern where at this level, users are still visualise the data in a form of ‘space’. According to Brown and Perry (2001), space is a term to describes the object and event that occurs in 2D and 3D views but still not relate to its actual event in the real world; place is a term to describe the space according to its particular geographic location. Level three comprises of interpretation tasks where map users will use other information and also tend to relate their knowledge and previous experiences to answer the ‘why’ geographic questions.

Timpf et al. (1992) have identified three levels of tasks that commonly applied in way finding (navigation) context. In this context, map users apply their knowledge and previous experiences with geographic space and place to find their ways. The tasks include planning, instructional level and driver (user) levels. In planning level, a common task include users plan the trip including estimates journey duration. In instructional level, users give and receive instruction; for example set the start (from) and stop (to) destinations. In driver level, users make a final decision which route (lane) to go.

Clarke (2003) is in line with Liebenberg (1998) that proposed skill levels to determine map literacy by using the level of map use as a basis. This author proposed three skill levels as following;

> Entry level: Get the main idea from a single or simple symbol (search, locate, identify, and compare a single symbol). Simple estimation (measure, calculate, relative size) on familiar symbols.
>
> Level 1: Recognising properties of symbol groups on the map as a whole and analysing spatial patterns (more complex recognition, reorganisation, decoding, detection, compare, discriminate, contrast) More complex estimation.
>
> Level 2: Complex tasks leading to understanding the meaning of spatial phenomena for knowledge enhancement. At this level inferential reasoning is used from the spatial relationships, patterns and map phenomena of one or more referents or source.

From these reviews, this research conducted the experiments that applied experimental tasks related to navigation context. There were increment in terms of the tasks in Experiment 3 and 4 compared to the earlier experiments. Table 3-3 present the tasks level incorporated in each experiment.

Table 3-3 The levels of map use tasks in each experiment

<table>
<tr><th>Experiment</th><th>map use contexts (simulated)</th><th>Map use task (Board, 1978 in Nyerges, 1991)</th><th>Map use strategies Muehrcke (1979) in Kimberling et al. (2012)</th><th>Task levels (Liebenberg, 1998)</th><th>Experiment tasks</th></tr>
<tr><td>1</td><td rowspan="2">campus tour</td><td rowspan="2">1) search, identify, and locate position on map
2) search for and recognise landmarks
3) compare, contrast, discriminate, prefer, like</td><td rowspan="2">Map reading</td><td rowspan="2">Level one</td><td rowspan="2">1) choose the map that more credible to assist in self-guided campus tour</td></tr>
<tr><td>2</td></tr>
<tr><td>3</td><td rowspan="2">route planning during disaster crisis (navigation)</td><td rowspan="2">Consists of the tasks 1, 2, 3 as above.

4) search for the optimum (safest) route on map
5) search for and identify from and to destination</td><td rowspan="2">Map analysis, map interpretation</td><td rowspan="2">Level three</td><td>1) choose the map that more credible
2) to suggest the safest route for the ambulance to rescue the trapped victims</td></tr>
<tr><td>4</td><td>1) choose the map that more credible
2) to plan the safest route to cycle from building A to building Z</td></tr>
</table>

3.4 Experimental Procedures

The procedures used for the experiments in this research are depicted in general form in the Figure 3.5 below.

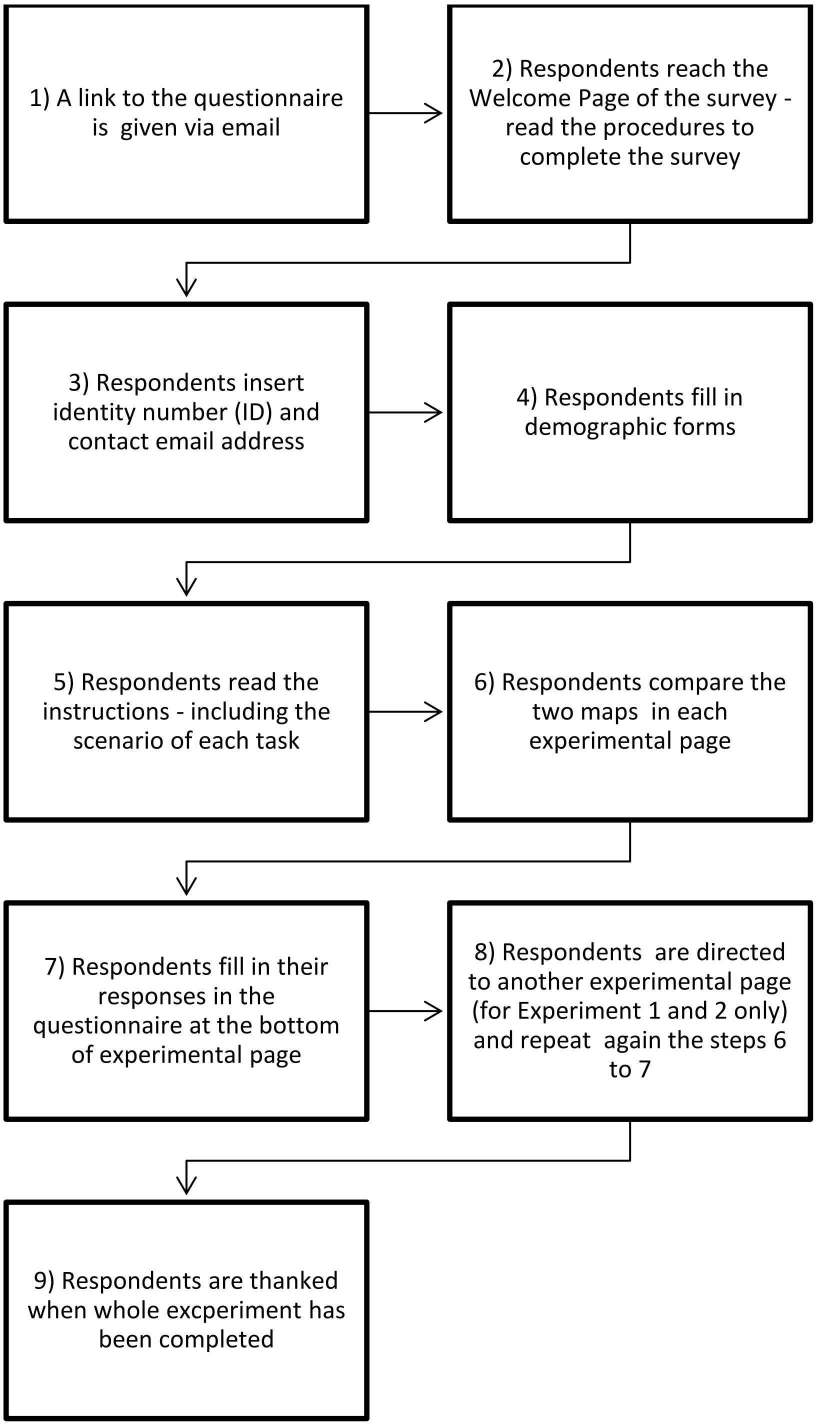


Figure 3-5 The procedures to complete the experimental map questionnaires

A hyperlink to the online survey was given to respondents by email. The hyperlink directed respondents to the study's welcome page. On that page, simple explanations were given describing the procedures to complete the survey. The aims and benefits of joining the survey were provided to give respondents an overview of the importance of their responses to the study and how their responses could help achieve the study's objectives. The chance to win a prize draw was provided to increase response rates to the survey. The prize for the winner is a £100 cash incentive. Via this page, respondents were directed to a form providing a contact email address and IDs to use for each experimental task's page. Apart from the prize draw, the IDs were required to match the page for each task to a respondent. This study was conducted using Google Spreadsheets to create the survey's forms. Session management, whereby the survey can remember the forms submitted by one user, such as in the SurveyMonkey$_{TM}$ application, is not provided by default; advance customisation is required. Therefore each respondent was required to fill in their ID before submitting each experimental task. The identification of each page submitted by one respondent is valuable for analysing the demographic data. Nevertheless, the data were analysed on an anonymous basis, where analysis of specific individuals was not conducted.

A respondent was directed to an experimental task page. The instructions for the task were provided at the top of the page. Each experiment has a different level of tasks. Next, a respondent would assess the given set of maps and give their responses to the questionnaire at the bottom of the page. The survey was completed with a thank you message in a pop-up window.

3.5 Think-aloud protocols

Think-aloud protocol refers to a method known as 'thinking aloud' or 'concurrent verbalisation'. This method asks respondents to perform a task and to verbalise whatever crosses their mind during the task. The written transcripts of the verbalisation are called think-aloud protocols (TAPs) (Jaaskelainen, 2010, p.371). This method enables respondents to vocalise their perception whilst completing the experimental tasks, saying what they thought and why they chose such decisions. In this study, the think-aloud method was conducted after quantitative data were collected from the experimental based survey. Elzakker and Wealands (2007) defined this method as introspection which is questioning and prompting during the problem solving tests; in this method, either the investigator will ask questions during the process or prompt the participants to convey what they are doing or the participants may chose the moments when they will report on their cognitive process during the test. This method was conducted to confirm the results of Experiment 3 and 4. In this present research, structured interviews were also conducted after think-aloud sessions. Elzakker and Wealands (2007, p.34) defined this approach as retrospection where the investigator interviews the participant either in a structured or unstructured way after the task completion. The authors also highlighted the drawbacks of these methods; for example, the data may become less valid due to disruption during the cognitive process (introspection); the data may also be invalid and incomplete due to memory errors during retrospection; the danger of participants feeling disposed to interpret and rationalises their problem solving behaviour; participants may be steered too much into directions anticipated by the investigators. The authors also suggested solutions to tackle these issues by conducting observations during the tasks and incorporating video recording and other advance technique such as recording eye movements.

In this method, a few respondents were solicited to be involved in think-aloud sessions. These sessions were conducted in combination with structured interviews. During the session, respondents were requested to read the same scenarios as given in the Experiment 3 and Experiment 4. The same tasks and contexts as in these previous experiments were used to trigger respondents' actions in the think-aloud sessions. The task was – 'please browse the two (2) map mashups and evaluate which of the two maps you perceived as having more credibility (more believable information) to assist you in the experimental task'. The actions of respondents to complete the tasks (e.g. zooming, panning, identify, scrolling, and clicking) were recorded using an audio and video recorder. The verbalised thoughts of their perception of the maps and the elements that influenced their decisions were also recorded. Table 3-4 below presents the protocols during the think-aloud session.

Table 3-4 Think-aloud protocols and questions in think-aloud session

Main task (question)	1) please browse the two (2) map mashups and evaluate which of the two maps you perceived as having more credibility (more believable information) to assist you in the experimental task'
Respondent actions:	zooming, panning, identify, scrolling, clicking, switching between sites
	Respondents respond to the question
Trigger questions	2) Which map you perceived believable in this task? 3) Are there any elements on the map that influence your decision? 4) Has the colour scheme of the map influenced your decision? 5) Does the design of symbol have any influence on your decision? *if a respondent do not notice (mentioned) of the map author* 6) Does the map producer (author) (at the side bar) have any influence on your decision? 7) If you had noticed this element earlier, would it have influenced your decision? *In Experiment: 4 if a respondent did not noticed the rating label* 8) Does the credibility rating label have any influence

	on your decision? 9) If you had noticed this element earlier, would it have influenced your decision? *To conclude the session* You are allowed to change your earlier decision (respond) which is whether to choose Map A or Map at question 2. So, you chose this map because of what bases?

3.6 The Research Material

The source of data for this research is an online map based questionnaire; respondents were given specific tasks using the supplied map before giving responses to the questionnaire. This research used online questionnaires, an approach made practical due to the increasing use of the internet at work. In contrast to paper based questionnaires, online questionnaires are relatively easy to administer and the results can be collected regardless of the location of the respondents. Participants only need to receive the online link or website address of the questionnaires. This approach allows respondents to give their responses from their office using their own computers in their own time, without the concern of returning a paper copy (Rossukorn, 2011, p.102).

The use of a written questionnaire to elicit environmental related perceptions was thought to be of limited use if the locations are not portrayed in visual form (i.e. on a map) due to people's poor geographical knowledge (Linden and Sheehy, 2004, p.34). The use of a map in a survey is not uncommon. For example, a study by Linden and Sheehy (2004) examines the reliability of a study that used a map and questionnaire to elicit environmental perceptions (level of cleanliness) of counties in Ireland. This study found that the addition of a spatial component (i.e. the map) may give participants more to think about and increases the variability of their responses. The comparison of the levels of reliability, using test-retest, between a map and verbal (written) questionnaire in this study has shown a significant correlation, thus indicating that both methods produce reliable results. The use of a map, however, produces a more sensitive measure due to increased variation in the results. A study by Hurst and Clough (2013) that examines how people use and view online and paper maps, used an online map based questionnaire to conduct the study and collect the responses. In this study respondents were given a list of task, including finding, investigating, planning and identifying a series of locations/routes using both map formats (i.e. a digital and a paper based map).

The use of a map embedded with questionnaires could also be of relevance to other research domains, which examine the usability of the application. For example, a study by Tiits (2003) examined the usability of a web based GIS application that provides a journey planning service; in this study respondents were given a series of tasks to manipulate the map from the applications. Questionnaires were distributed online in this study in order to collect responses from respondents drawn from countries that might use the tested services. A study by Skarlatidou et al., (2010b) also examined the static trust (credibility) elements and usability of 'What is in your back yard' websites by using map based questionnaires. In this study, respondents have to browse the websites and give their perception of the website through a survey.

In this present research, the maps given to respondents are in the form of online map mashups embedded in a website. In the first two experiments, the maps presented take the form of static maps, whereas in the last two experiments, the maps presented are dynamic. In the case of static maps, respondents are not able to interact with the map; in dynamic setting maps respondents can interact by using the zoom, pan or identify functions. In the static map layout, the two maps are presented side-by-side in order to enable respondents to make comparisons more easily. Meanwhile, in the dynamic map setting, respondents are given separate hyperlinks to the two comparison maps that are stored on the university server. Maps are given a static setting in-order to control the experiment so that each respondent will examine the same map view. This design was also to setup a low level involvement setting to test the influence of the parameter in a low level of users' engagement. In the other experiments, the maps are given a dynamic setting to create an interactive environment that simulates the way in which map users are likely to interact with a map mashup application in real life. This approach was used for Experiment 3, to test the influence of a parameter where there is a high level of users' engagement.

3.7 Questionnaire Design

Issues considered when designing the questionnaire include:

1) **The wording of questions so as to avoid biased questions and ambiguous answers**

According to Weisberg and Brown (1977, p.45), biased questions are questions that make one response more likely than another, regardless of the respondents' opinions. In other words, this type of question leads a participant to a specific response. For example, some people conduct surveys not to learn people's opinions on an issue but to show that people agree with them. A subtle form of bias typically occurs in social science research where insufficient consideration is given to how a person with limited knowledge will respond to a question. The author suggests that one solution is to explain the scenario before presenting the questions to respondents. The bias issue in the question design might not be relevant to this research that asked questions about a scenario or experimental tasks and as also, in this present research, the researcher could not explain to respondents the problem that she wanted to identify: because if respondent became aware of the problem, they might respond on the basis of the solution to this problem.

This present research has carefully phrased questions to avoid answers that could be interpreted ambiguously. The wording of questions was checked for double-barrelled questions; questions were designed to address one single issue only. Another issue in question design to have been considered is the possibility of no response being given; to tackle this issue, the option of a 'neutral' response was included. A further pitfall in survey design is the inclusion of a question that assumes familiarity with a given topic. In this research the fact that questions are based on the scenario from the experimental task has tackled this issue. This is because the task provides respondents with the experience of using a map for a particular set of circumstances; hence they can draw upon that experience to answer the questions on the basis of their opinions.

2) **The format of the questions and answers**

The format of the questions and answers in a questionnaire is important for structuring the collected responses. The chosen format also has a relationship to the analysis to be conducted from the survey. The format for questions used in this research has mainly involved the use of closed-ended questions, except in Experiment 1. In this experiment the question that asked 'the reason the respondents chose and rejected the map' was an open-ended question. The purpose of using this format was to allow respondents to answer in a way that would avoid giving hints and directing them to the expected answers. At first, it was assumed that, if using a closed-ended question, bias might occur as a result of providing them with a series of answers. After conducting the data analysis, however, it was found that respondents tend to give their answers on the basis of a single issue; they do not include other criteria that influence their judgement.

Another issue concerns the fact that it is difficult to identify whether the answers given are the main reasons or just answers that might come to mind at the point at which they filled in the form; this present study postulates that there might be other factors that influence their judgement. Therefore, in the later experiments, this research applied a closed-ended question format in the questionnaires.

According to Weisberg and Bowen (1977, p.49), the main advantage of using a closed ended question format is the uniform frame of reference for respondents to use in determining their answers. In this format, respondents are assisted with a list of potential answers. Several question formats are used in this research, as shown in Table 3-5.

Table 3-5 Question format used in this research

Question formats	Question themes
Multiple choice questions	Demographic related questions
Rank-order scale questions	To identify the rank of importance of the tested parameters on their judgement
Rating scale questions : Likert Scale	To identify the weight of agreement of the tested parameters on their judgement
Rating scale questions: Semantic Differential scale	To identify the level of perceived credibility of respondents in the maps they chose and rejected.

3) Constructing the questionnaires

Some issues related to the order of the questions were also considered in this research, including the flow of questions. It is important that questions on similar topics be grouped together, with an introductory sentence. In this research, similar questions were grouped together in various sections of the questionnaire. The issue of the flow of one question to another question was also taken into consideration.

A pre-test was conducted before distributing the map based questionnaire to the mailing lists and student portal in the case of each experiment. The pre-tests were distributed to five respondents to check the functionality of the map mashup, the clarity of sentences, the flow of questions and the length taken to complete the whole questionnaire. The experiments in this research were designed to be completed by respondents within the range of 5 to 10 minutes.

According to Litwin (1995) a successful data collection survey is not simply a set of well-designed questions that are written down and administered to a sample population; bad surveys produce bad data, which is unreliable, irreproducible, or invalid. It is therefore important to measure the reliability and validity of the survey instrument (i.e. the questionnaires). This research developed its own questionnaire due to a lack of research into the psychometric issues relating to the instruments that measure credibility, particularly the perceived credibility of online maps, including map mashups. Other domains that already have well established survey instruments do not face this issue. This is because the reliability and validity of items used in the questionnaires have been identified, which means that any related research can use these items in questionnaires. To tackle this issue, this present study has needed to check the reliability and validity of the items (questions that are being asked). Reliability is a statistical measure of how reproducible the survey instrument's data area (Litwin, 1995). There are a few ways to test the reliability, including test – re-test reliability, alternate-form reliability, internal consistency reliability, inter-observer reliability and intra-observer reliability. In this present research, internal consistency reliability was used. This involves assessing survey instruments and scales that are not limited to single items, but to a group of items that are thought to measure different aspects of the same concept. Internal consistency is an indicator of how well the different items measure the same issue (Litwin, 1995, p.13).

3.8 Sampling Procedures

According to Weisberg and Bowen (1977), an early decision in conducting a survey is to define the relevant population: whose attitudes do we want to describe or make generalizations about. The findings resulting from this research may be generalized to the population of young adult web users (age between 18 to 35 years old) who use online map mashup applications. To date, there are no statistics recorded of the number of online map users as well as map mashup application users in general. Vaibhav (2012), however, indicates that Google Maps has become the main default mapping service used by digital map users around the world and has more than 150 million mobile users; from these statistics, the current research postulates that the population of online map users, including other mapping services such as Bing Map and OpenStreetMap, is well in excess of 100 million.

The assumption made in this research is that online web users might be more likely to use online maps, including map mashups, than offline web users. The approach to sampling in this research, however, applied non-scientific sampling methods (Weisberg and Bowen, 1977), namely nonprobability sampling methods (Weisberg et al., 1996). There are several methods in this type of sampling – typical people, haphazard sampling, purposive samples and volunteer subjects. In this research, an inexpensive sampling approach, namely a volunteer subjects sampling method is used.

The sample of respondents is drawn from those who are willing to contribute to the survey distributed via the School of Geography, University of Nottingham (UoN), and the alumni of the Faculty of Geoinformation, Universiti Teknologi Malaysia (UTM) mailing lists and the University of Nottingham intranet student portal. The think-aloud protocols respondents were also selected based on volunteer participants based in Malaysia. The specific demographic information and numbers of respondents for each experiment were addressed in detail in each experiment chapter. The problem of this method, however, is that respondents who volunteer may not be typical: the volunteers are generally more interested in the topic of the study than other people, thus they are unrepresentative of the larger population (Weisberg et al., 1996, p.40). This sampling method may be biased against respondents who do not have access to the Internet, those who are not alumni of Universiti Teknologi Malaysia (UTM) or students (undergraduate and postgraduate) with access to the School of Geography mailing lists and those without access to the University of Nottingham (UoN) student portal.

A few other studies that assess users' perception on map uses applied volunteer subjects sampling. For example Muehlenhaus (2012) and Lautenschutz (2012) used undergraduate students and volunteered respondents solicited via social networking and emails as their samples of experiments. Sudman (1976, p.2) highlighted the fact that there is no uniform standard of quality that must always be reached by every sample. The quality of the sample depends entirely on the stage of the research and how the information will be used; whether or not a sample design is appropriate depends on how it is to be used and the resources available. In some cases, it may be fair to say the sample design is appropriate for the available resources. For example, in the case of student who is doing unfunded research, it would be inappropriate for that student to attempt or be advised to attempt a large national study; the resources available are not adequate for the task. From this argument, the sample used in this present research may be indicated as low-quality sampling. However, it is worth pointing out that this type of sampling is appropriate at the earliest stages of research design, namely exploratory research when the researcher is attempting to develop hypotheses and procedures for measuring them. Sudman (1976, p.9) argues that any sort of sample may be useful when very little is known; for example, only a few interviews can reveal major problems and dimensions of topics that researchers have ignored.

3.9 Chapter Summary

This chapter discussed the research methodology that has been applied to achieve the objectives of this research. The research materials and the design of experimental map based questionnaires are explained in general. This chapter also highlighted the sampling procedure and procedure for the experiments conducted in this research. The next chapter will describe the series of experiments, including the specific experimental design, results and discussions of findings that have been conducted to achieve the objectives of research.

4 EXPERIMENT ONE

4.1 Introduction

This section describes the hypothesis and methodology used for Experiment One. Several studies have examined how web users evaluate information credibility on websites and which factors influence users' perceived credibility in the medium before they believe the information. However, there is little understanding of the ways in which web users evaluate the credibility of a map, particularly a map published on a website. Experiment 1 consists of an experiment conducted to examine how map users evaluate online map information, particularly when using a map mashup medium. In order to conduct this investigation a series of self-completed online map based questionnaires was designed. This experiment was conducted to achieve the objective 1, which is

'to examine the influence of metadata related to sources (specifically the identity of map producer and map supplier) on respondents' assessment of the credibility of map mashup information'

4.2 Hypotheses

The hypotheses of this experiment were based on the research questions highlighted at Section 1.3 (Chapter 1).

Hypothesis 1 is**:**

Visual cues have no significant influence in respondents' credibility assessment on map mashup applications

Hypothesis 2 is:

The metadata related to sources have significant influence in respondents' credibility assessment

Hypothesis 3 is:

The metadata related to sources have significant influenced within geoliterate respondents

4.3 Experimental Design

Three experimental conditions were developed to test the hypotheses. The conditions are summarised in Table 4.1. The conditions were designed according to the experimental suppositions of credibility levels – low credibility, intermediate credibility and high credibility.

Table 4-1 Experimental Conditions

Experiment Labels	**T5**	**T2**	**T6**
	Condition 1	**Condition 2**	**Condition 3**
Experimental Assumptions	High level of credibility	Intermediate level of credibility	Low level of credibility
Variable 1 : Identity of Map Mashup creator (producer)	University researcher versus mapping agency's researcher	An undergraduate student versus a member of the off campus community	No identity for mashup's creator versus identity of mashup creator provided
Full descriptions (as in the experimental task sheets) Map A	John Cullen is a researcher at one of the research centres at the University of Nottingham. He joined the centre in 2005, after graduating from	Philip Albert is currently an undergraduate student at the University of Nottingham. He is actively involved in the students' society	No background information provided

	Loughborough University	programme	
Map B	Fred Smith is a researcher at a map production agency operating in London	Karen Ryan is a member of the Nottinghamshire Community Foundation, which is a community network committed to improving the quality of life of people in Nottingham. She has been a community member since 2005	David Crossley is an administrator in a communication company in Birmingham. He has a level 2 National Vocational Qualification (NVQ) in Business and Administration from Newcastle City Learning

4.3.1 Dependent Variables

Core dependent variables used in this experiment are shown in Table 4-2. Full details of these dependent variables can be found in the Section 2.3.

Table 4-2 Dependent variables used in Experiment 1

Dependent variable	Measurement
Identity of map mashup creator	Qualitative data – open question
Perceived credibility of the selected map	5-point Likert scale
Perceived credibility of the rejected map	5-point Likert scale

4.4 Sampling Procedure

Sixty-seven to eighty-one respondents completed the different questionnaire of the study. The number of respondents varied in each of the six experimental tasks. This was due, as described below, to the ability of the server database used to record simultaneously the submitted number of self-completed questionnaires by respondents. This study used the Google spread sheet database server to record users' responses. According to the Google spreadsheet forum, there is a limit to the number of simultaneous responses that can be recorded by the spread sheet server (Strickland, nd) The sample was selected using a volunteer subjects sampling method, whereby the self-completed survey was distributed to respondents via email. The respondents were selected on the basis of their availability to give their responses to the survey. The sample was drawn from members of the University of Nottingham (staff and students) and non-university members who came from outside that community, who resided in the United Kingdom or Malaysia.

Data from seven respondents were excluded from the analysis pertaining to geoliterate and non-geoliterate respondents since these respondents did not mention their academic or professional background on the demographic form. Data from five respondents were excluded because the ages were greater that the targeted sample age, 18 to 35 years old. Table 4-3 presents the full details of respondents' demographic backgrounds.

Table 4-3 Respondents' Demographic Information

Task	Geoliterate	Non-geoliterate	Unknown	Total respondents
2	23	45	7	75
5	29	41	7	68
6	19	41	7	67
Total number of respondents (average)	71.83			
Age (mean average)	29.02			
No. of University members (average)	21.5			
No. of Non-University members (average)	50.3			
No. of respondents residing in the UK (average)	42.33			
No. of respondents residing in Malaysia (average)	29.5			
Male (average)	33.5			
Female (average)	38.33			

4.5 Materials

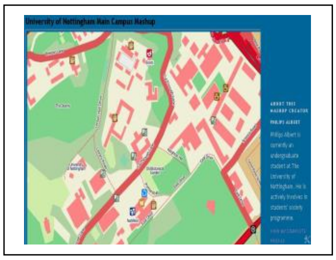
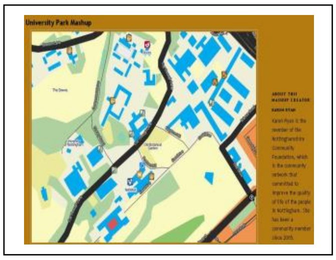
Two static map mashups were distributed for each experimental condition. The same maps were presented in three different scales – small scale, medium scale and large scale- in order to provide respondents with several views of the data. OpenStreetMap (OSM) map was used as a base map whilst the point of interests (POI), such as the location of the bank, café and shops, provided additional data that were mashed up from other sources. This experiment used the OSM map as the need to use a base map that did not provide a stamped date of copyright; such a stamped date is usually provided on top of Open Web

Mapping Service (OWMS) maps, such as Google Map and Bing Maps, but not OpenStreetMap map.

Figure 4.1 shows a structured example of a map based questionnaire used in each experimental task. Figure 4.2, Figure 4.3 and Figure 4.4 below, show the maps used to test the variable ‘map producer’. Tables 1 to 3 (See Appendix A) present the style of map features. The map styles were designed and not based on any convention pertaining to cartographic design. Justification for this approach was based on:

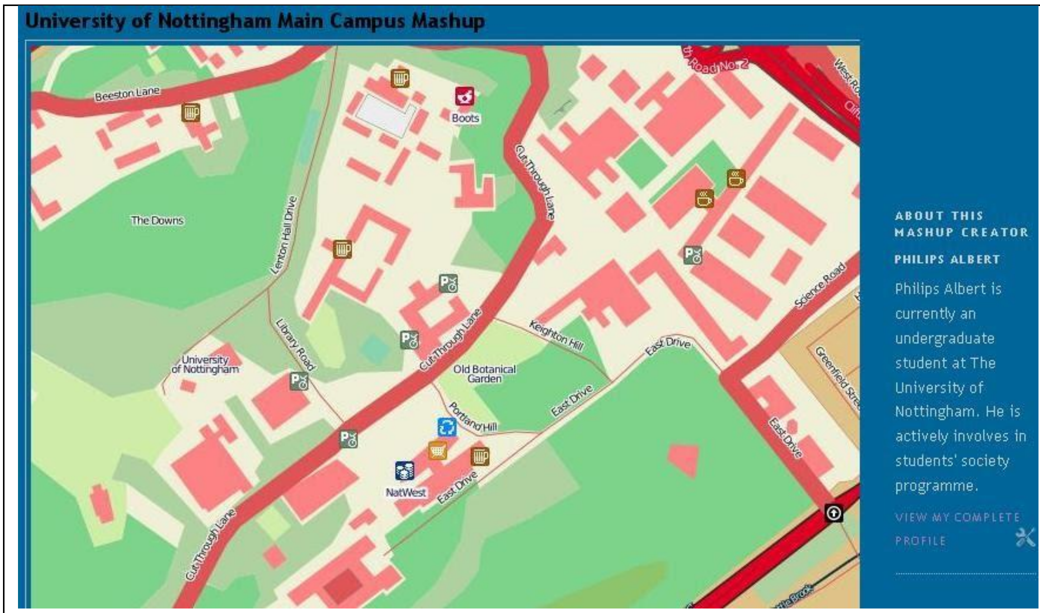
1. The purpose of the first experiment was to test the influence of metadata related variables on users’ judgement pertaining to map information credibility, not the map design.
2. The goal was to implicitly direct user attention to make judgements based on the metadata elements embedded, not on design appearance
3. There is no dominantly accepted source that provides cartographic guidelines, particularly where the focus is to help citizen map mash up developers to design a good map. Most citizen developers have little knowledge of cartographic mapping. Hence, map design is based on their intuitions and preferences.

Task 2/6:
Compare the set of Map A with the set of Map B and their attached information. Then choose only 1 map mashup you will use in your self-guided campus tour

MAP A (see Figure 4.2)	MAP B (see Figure 4.2)
Scale 1:4000 (Main Map)	Scale 1:4000 (Main Map)
	
Scale 1:9000	Scale 1:9000
	
Scale 1:2000	Scale 1:2000
	

Question 2 of 6:
Compare the set of Map A with the set of Map B and their attached information. Then choose only one (1) set of map mashups you will use in your self-guided campus tour.
Then answer the questions below.

Figure 4-1 A layout of map based questionnaire for each experimental task



Map Mashup A (see Table 4.1 for a clearer text version)

Variable: Identity of Mashup producer; **Parameter:** Undergraduate student

Map Mashup B (see Table 4.1 for clearer a text version)

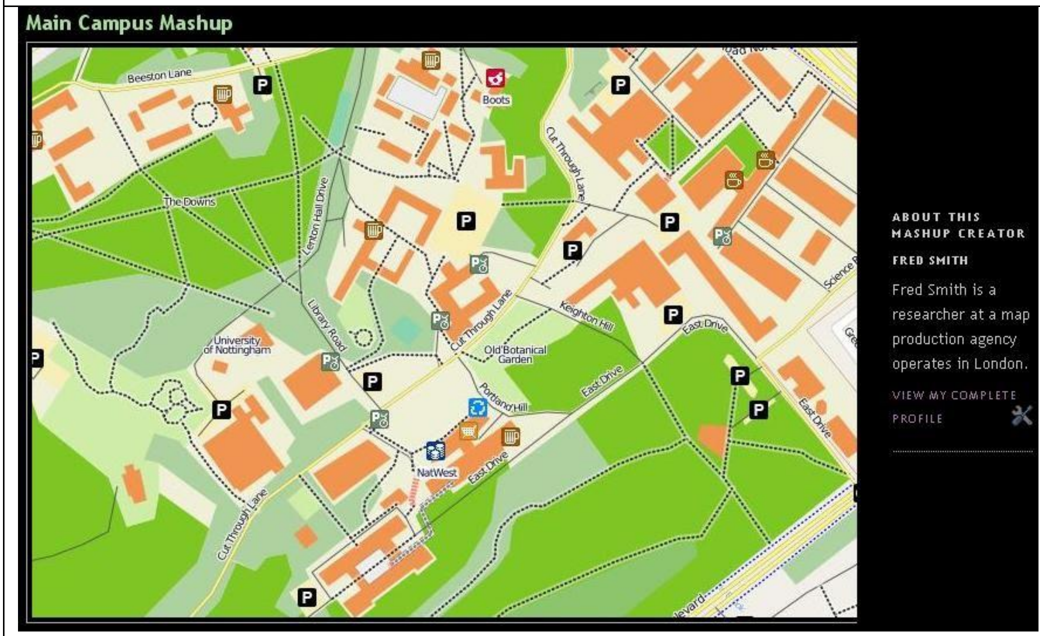
Variable: Identity of Mashup producer;

Parameter: Member of outside community

Figure 4-2 A Snapshot of Experimental Task 2 (Condition 1) Map A (left) and Map B (right)

Map Mashup A (see Table 4.1 for a clearer text version)

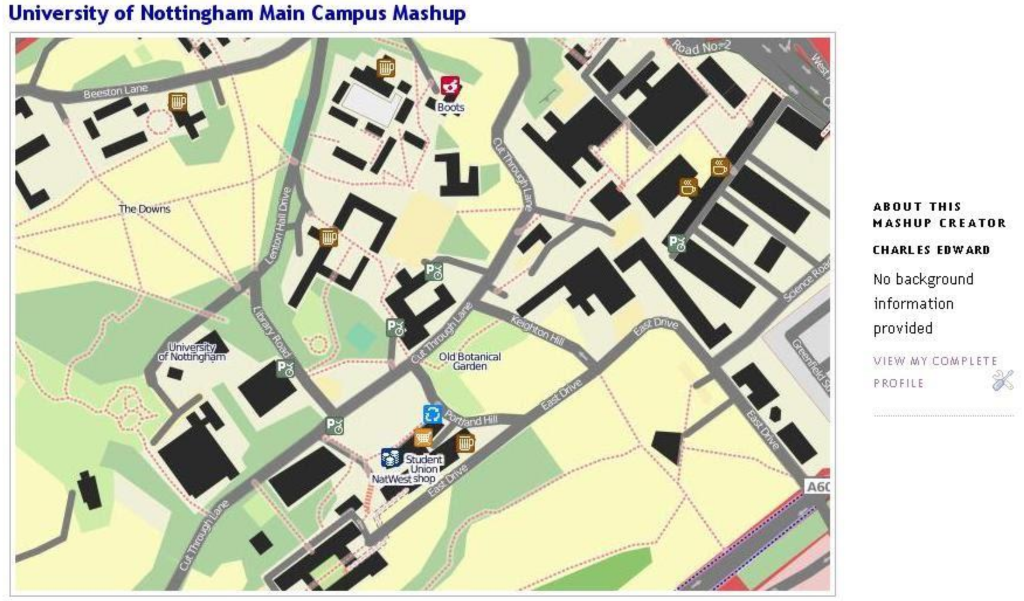
Variable: Identity of Mashup producer; **Parameter:** University researcher



Map Mashup B (see Table 4.1 for a clearer text version)

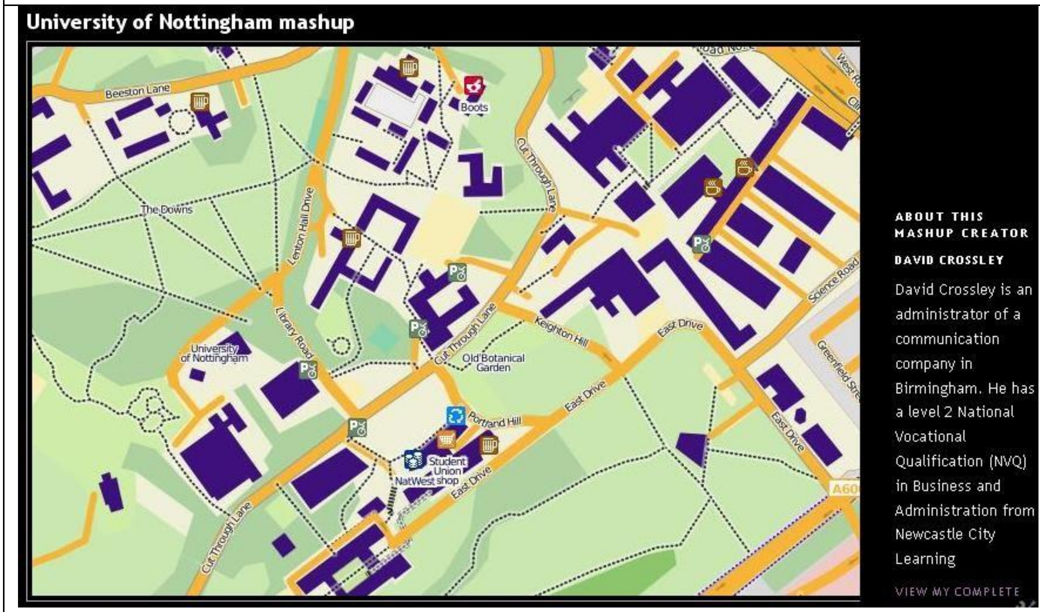
Variable: Identity of Mashup producer; **Parameter:** Mapping Agency

Figure 4-3 A Snapshot of Experimental Task 5(Condition 2) Map A and Map B



Map Mashup A (see Table 4.1 for a clearer text version)

Variable: Identity of Mashup producer; **Parameter:** No information provided



Map Mashup B (see Table 4.1 for a clearer text version)

Variable: Identity of Mashup producer;

Parameter: Administrator, Level 2 NVQ

Figure 4-4 A Snapshot of Experimental Task 6 (Condition 3) Map A and Map B

4.6 Pre-test study

This section will summarise the pre-test that has been conducted. A pre-test study with four respondents was conducted in order to test the designs of the map based questionnaire. One respondent came from a geoliterate background, whilst the other three respondents were from engineering backgrounds. After analysing the data and feedback from the pre-test study, changes to the questionnaire design were made.

The initial layout of the map based questionnaire was designed to provide only one map, instead of two maps, in each experimental task. Each map displayed each tested variable with different values or parameters. Respondents had to rate their perception of credibility on one map, and fill in the form on the basis of the ratings they had given to that map. Comparison between two maps was not required. However, there were several drawbacks to this initial design. Since there was only one map and no indications given suggesting what to look at, respondents had difficulty finding a basis on which to rate the map. They seemed to have no idea of which aspects to analyse or what basis to use before they could rate the map. In addition, the 12 experimental tasks, consisting of 12 parameters for the two variables, took more than 30 minutes to complete and were reported as very time consuming by respondents.

Therefore, changes were made to use a comparative approach in each experimental task. This design assisted respondents by providing a guide directing them where to look in order to form the basis of a perceived credibility rating perception, responding to the element(s) on the map that became the basis of their judgement. Moreover, designing two maps for each task to test the 6 parameters reduced the time required to complete the whole experimental task to less than 20 minutes.

4.7 Data Analysis

Qualitative text analyses were conducted on responses to open-ended questions - the basis of respondents' selected and rejected map. The software package, NVivo, was used to code and conduct the analysis. There are three approaches to content analysis: thematic, semantic and network text analysis (Roberts, 2000, Popping, 2000). Thematic text analysis analyses the word count in a text block before inferences about the predominance of themes in the text are made (Roberts, 2000). This analysis is also known as contingency analysis, whereby the frequency of occurrences and co-occurrence of themes are encoded. Occurrences indicate the prominence of themes (or concepts) and co-occurrences examine associations among themes (or concepts) in text blocks. Thematic text analysis could not determine a cause and effect question (Roberts, 2000) or preserve the narrative flavour of the original text (Popping, 2000, p.27).

Semantic text analysis examines the relationships among encoded concepts or themes in sentences or clauses. This analysis is designed to overcome the limitation in the thematic text analysis, not by encoding the occurrence of themes but by encoding a concept or theme from what was written in a sentence or clause. The relationship between concepts is examined using a sequence of subject-action-object triplets and encoded using a predetermined template of semantic grammar (Roberts, 2000, p.264).

Network text analysis examines the locations of interrelated themes by depicting them as a network; concepts and linkages are analysed according to their position within a network. This analysis originates with the view that one can construct networks of semantically linked concepts after conducting semantic text analysis, which has the encoded semantic links among the concepts (Popping, 2000).

In this study, thematic text analyses were conducted on the responses. The data generated in this analysis are presented in a data matrix, where one row represents each theme (or concept) and one column represents each sampled text block. Cells in the data matrix refer to the number of occurrences of particular themes or concepts within a specified block. Themes usually refer to broad classes of concepts (Popping, 2000, p.26). There are two approaches to the interpretation of the texts: instrumental and representational interpretation approaches. In the instrumental approach, themes are constructed and texts are interpreted by reflecting the researcher's perspective, which is built upon a theory. In the representational approach, the themes/concepts are constructed and texts are interpreted on the basis of the representation of the responses; in this approach, a researcher tries to identify the sources' intended meaning (Roberts, 2000, p.262).

This present study used both representational and instrumental perspectives as the words and phrases in the texts were coded representationally from the text, but interpreted instrumentally on the basis of the themes/concepts that were constructed from the informed literature. The themes and concepts constructed in this study are informed by the literature, in particular by Fogg et al (2003).

Each text block was analysed and coded on the basis of the frequency of the concepts that emerged. In this process, a dictionary containing specific keywords for each concept was developed to maintain the reliability of the coded text analysis. Concepts that emerged in each text block at a certain frequency were coded representationally in specific classes. Next, the keywords (sub-concepts) were instrumentally reinterpreted and redefined according to the literature. Then the processes of merging and eliminating certain concepts were conducted before they were classified into specific themes. Table 4.4 presents the finalised themes and concepts that emerged from the responses.

Table 4-4 Themes, concepts and keywords that emerged in the open-ended responses

Themes	Concepts	keywords
Visual cues	Overall presentation (Design Look)	e.g. Professional, amateur, complexity, aesthetic
	Information clarity	e.g. Contrasts, identify, read, clear
	Colour Scheme	e.g. make sense, convention, harmonize, dull, colour
	Individual preferences	e.g. like, keen, prefer
	Information details	e.g. highlight, details
	Map design	e.g. Cartographic, format, design, layout, generalisation, scale, label
metadata	Identity of site operator	e.g. author, creator

4.8 Analysis and Findings

4.8.1 Data Analysis

The main aims of this chapter are:

1) to explore the themes/concepts that emerged in respondents' responses to 'what was the basis of choosing and rejecting the map?'

2) to examine the occurrence and co-occurrences of the most dominant concepts that emerged in respondents' open-ended responses;

3) to examine the occurrence of metadata element (i.e. identity of site operator) in respondents' responses.

Contingency analyses were conducted on the responses. Contingency analysis involves counting the occurrence of themes/concepts within a sampled block of text. A data matrix is produced with distinct themes heading the column, unique text block heading the rows, or vice versa and a count of occurrences in the cells. The analysis proceeded by computing a matrix of association between pairs of themes/concepts. A researcher could develop explanations of why some themes/concepts occurred and why other themes/concepts are disassociated (Roberts, 2000, p.260). More advanced analyses could be conducted, such as path analysis, cluster analysis and factor analysis, which would depend on the research questions (Popping, 2000).

4.9 Results: The map respondents chose to assist in their self-guided campus tour

Frequencies analysis was conducted to analyse the question of 'for the purpose of self-guided campus tour, which map you will use?' This question yielded two categorical variables, which were either Map A or Map B. Table 4.5 presents the results of this question.

Table 4-5 Frequencies analysis of the map that respondents chose or rejected in each experimental task

	Most voted	%	Count	Least voted	%	Count	Chi-square test
Task 2	B	66.67	50	A	33.33	25	X^2 (1, n = 75) = 8.33, p <0.05)
Task 5	B	52.94	36	A	47.06	32	X^2 (1, n = 68) = 0.23, p <0.63)
Task 6	B	71.64	48	A	28.36	19	X^2 (1, n = 67) = 12.55, p <0.001)

Chi-square tests for independence were conducted on the results in Table 4.5. This test compares the observed frequencies that occur in each of categories, the most voted map versus the least voted map, with the values that would be expected if there were no association between the two categories being measured. The shaded cells in Table 4.5 indicate the experimental tasks which have high proportions of voted map and show a statistically significant difference in comparison with the alternative maps. From the above table, there were two tasks in which the differences of votes were statistically significant - Task 2 (Condition 1) and Task 6 (Condition 3).

4.10 The occurrences of concepts in the responses

Hypothesis 1 (null) is**:**

> Visual cues have no significant influence in respondents' credibility assessment on map mashup applications

Contingency analysis was used to count the frequencies of concepts that emerged in the open-ended responses collected in this study. Table 4.6 presents the occurrences (in frequencies) of concepts in each experimental task. This table indicates the patterns of the concepts that emerged in respondents' responses to 'what was the basis of choosing the map and rejecting the alternative map'. From the table, the colour was the most dominant concept that respondents used as a basis to choose a map and reject the alternative. 'Colour scheme' was the most dominant concept in the experimental tasks. Then, the second dominant concept was 'information clarity', followed by the concept of the 'design look'.

Table 4-6 The occurrence of concepts in the responses to ‘what was the basis of you choosing the map and rejecting the other map?’

Rank	Concepts	T2	T5	T6	Average
1	Colour scheme	47	29	29	35
2	Information Clarity	31	20	15	22
3	Overall (Design look)	15	18	17	17
5	Information Details	4	1	3	3
6	Map design	4	3	4	4
7	Author (map producer)	4	6	9	6

In task 6 (condition 3), the frequency of comments from respondents on the ‘colour scheme’ seems more significant than the comments of concept of ‘information clarity’. This might be because the ‘colour schemes’ applied on both maps were significantly contrasted. For example in Task 6 (see Figure 4.5 below) both maps were using a good contrast of ‘colour scheme’.

Figure 4-5 Both comparison maps in Task 6 use a good contrast of colour scheme (Refer Figure 4.4 for a clearer image)

These initial findings **not supported the Hypothesis 1** of this experiment that postulate visual cues have no significant influence in users’ credibility assessment. Further analysis is conducted in the next section, to examine the extent of the ‘colour scheme’ concept in respondents’ responses, particularly on the pattern connection of ‘colour scheme’ to other concepts that emerged.

4.11 The co-occurrences of concepts emerged with 'colour scheme'

From the results in Table 4.6, colour was the most dominant concept that emerged in respondents' responses. Thematic text analysis cannot detect the relationship between concepts but could describe the co-occurrence of other concepts that emerged with a certain concept. From the results, there were certain concepts that seem to be associated with the concept of colour in the respondents' responses. A matrix of association between these pairs of concepts is presented in Table 4.7. The table presents the co-occurrence (in frequencies) of concepts that emerged with the concept of colour in the responses. It can conclude that the 'colour scheme' that dominant in the responses was not independent (not solely because of colour). The concept was used with other concepts including 'information clarity', 'overall (design look)', 'information details', 'individual preferences' and 'map design' (which were the concepts that emerged from respondents responses).

Table 4-7 A matrix of association between the concepts of "colour scheme" with other concepts

	Clarity	Design look	Combination (scheme)	Details	Individual preferences	Map design
T2 Colour	22	12	14	2	5	3
T5 Colour	11	8	15	1	3	2
T6 Colour	7	6	18	2	4	2

4.12 Visual cues as a dominant variable

Visual cues emerged as the most dominant concept in respondents' assessment of the credibility of map mashup information. Analysis of the aspect of gender was then conducted. Table 4.8 presents the academic background of respondents who indicated the concept of colour in the responses.

Table 4-8 Analysis of academic background of those who indicated concept of colour in the responses

	Geography, GIS, Land Survey, Cartography, Remote Sensing	Engineering	Sciences	Social Sciences, Law Education	Medical health sciences	others	Not respond
T2	15	12	6	5	2	3	6
T5	9	5	6	1	2	3	3
T6	9	4	6	2	2	2	4

From this analysis, visual cues were the dominant influence that became the basis of respondents' judgement selecting and rejecting a map. Surprisingly, the geoliterate group of respondents in this experiment tend to use visual cues, when making judgement, particularly to assist them in the experimental tasks, rather than using the metadata element as their basis of judgement.

4.13 Metadata as the least dominant variable in users' perceived credibility

Hypothesis 2 is:

> The metadata related to sources have significant influence in respondents' credibility assessment

Hypothesis 3 is:

> The metadata related to sources have significant influenced within geoliterate respondents

The element of metadata related to sources (i.e. 'identity of map author' (map producer)) was the concept least indicated in the responses. Therefore, **Hypothesis 2 is not supported**. Table 4.9 presents the academic background of respondents who indicated the concept of metadata in their responses.

Table 4-9 Analysis of academic background of those who indicated the concept of metadata in the responses

	Geography, GIS, Land Survey, Cartography, Remote Sensing	Engineering	Sciences	Social Sciences, Law Education	Medical health sciences	others	Not respond
T2	0	1	2	1	0	0	0
T5	1	1	4	0	0	0	0
T6	2	3	4	0	0	0	0

The number of geoliterate respondents that use the critical metadata element (i.e. identity of author) was very low. Hence, **Hypothesis 3 is not supported.** Example responses from the questionnaires are given below;

'the author is working in this university, which may make him more familiar with the uni [university]' (F, geoliterate)

'though it is done by professional, that person might not be as familiar as the person creating map A' (F, geoliterate)

4.14 Discussions

According to the findings in Section 4.9, the number of differences between respondents that chose either Map A or Map B in each task was significant, with the exception of Task 6. In this task, the number of respondents that chose either Map A or Map B were not statistically different. Analysis on the perceived credibility level of the map they chose or rejected, as in Table 12 (see Appendix A) demonstrates respondents perceived the rejected map as having lower credibility than the selected map.

Textual analysis on the basis of respondents' judgements' in selecting the maps that they will use in the experimental task, indicates the high influence of 'colour scheme' and 'information clarity' in their decisions (see Table 4.6). This finding **not supported the Hypothesis 1 and 2** of this experiment that expects the low influence of visual cues and significant influence of the critical metadata related to sources (i.e. identity of map producer) in respondents' judgements. Although the results show the 'colour scheme' concept was the most dominant keyword found in respondents' answers, the frequency of this concept with the second dominant concept, which is 'information clarity', was statistically insignificant. Table 4.10 below shows a sample of respondents' comments on the basis of their decision when selecting and rejecting the map that related to 'colour scheme' and 'information clarity' concepts.

Table 4-10 Sample of respondents' comments on the basis of their decision when selecting and rejecting the map (in Task 4)

Map selection	Select	Reject	Concept
Map A	Like the colours more (F, geoliterate)	Don't like the colour	Colour
	Light green more attractive than purple (M, non-geoliterate)	Striking purple not really suitable.	colour
	because the map is simple and easy to study. (F, non-geoliterate)	because there are too many contrasting colours and it is confusing.	Information clarity
	Eye catching and easy to read (F, non-geoliterate)	Hard to read because of that purple color	Information clarity
Map B	Acceptable use of colour and symbols. (F, non-geoliterate)	Colours are too soft (among them)	colour
	Better colour contrast (M, non-geoliterate)	Taking extra time to identify a building or road	
Map B	the buildings are more easily defined, despite the bad colour clash (M, geoliterate)	it is very difficult to make out the buildings and roads	Information clarity
	easier to distinguish the landmark (M, non-geoliterate)	difficult to see the landmark	

In this experiment, the ‘colour scheme’ concept refers to keywords relating to colour selection and colour combination that emerged in respondents’ answers. Meanwhile, the ‘information clarity’ concept refers to keywords that relate to discrimination of the text, features, patterns, colours and the understanding of the meanings of the signs on the maps, such as ‘easy to read’, ‘able to define’ and ‘distinguishable’. As stated by MacEachren, (2004, p.213), maps are imbued with meaning by virtue of semiotic relationships. Semiotics is the science of signs, with a sign considered to be a relationship between expression (the sign-vehicles) and its referent (content). Colours, symbols and patterns are the sign-vehicles that represent objects in the real world that are subject to the interpretation of cartographers and map users. The meanings in maps can be interpreted either by reference to a map legend or assumed to be part of the normal readers’ general map schema (e.g. blue = water) (MacEachren, 2004, p.311). Colour schemes used in maps have explicit meanings that represent spatial features. The colour scheme of a map is not like a regular colour used on a textual based medium, but it represents special functions to deliver messages to map readers.

Although ‘colour’ is the dominant keyword found in respondents’ answers, this keyword tends to emerge with other keywords to describe the relation of ‘colour’ to the ‘clarity of information’, ‘combination of schemes’ used, the influence of ‘overall presentation’, ‘information details’, ‘the design of the map’ and ‘individual preferences’, as shown in Table 4.7. This indicates that the keyword ‘colour’, found in respondents’ answers, is not a single keyword but has emerged to relate with other concepts. From Table 4.7, ‘clarity’ and ‘combination’ tend to be the dominant keywords used to relate to the colour keyword.

In the aspect of demographics, the ‘colour’ keyword tends to be used dominantly by both genders. Table 13 and Table 14 (see Appendix A) support this by indicating no significant difference between females and males in using colour as the basis of their judgement. The results shown in Table 4.8 also indicate no difference in responses between the group of respondents who have an academic background that is geospatially related (i.e. geoliterate) and the group of respondents drawn from other domains (i.e. non-geoliterate). These surprising results indicate that the majority of geoliterate respondents tend to use the concept of colour when making judgements to select the preferred maps in the tasks. This is in line with the non-geoliterate group responses, which show the dominant use of the concept of colour when making judgements in the tasks. Hence, these findings **reject the Hypothesis 3** of this research that postulates respondents drawn from geoliterate group would be more aware of the critical metadata elements and will use those as the basis of judgements when selecting and rejecting a map.

In a study by Williams (1967) that examines how a person searches for a visual object, colour became the main basis during a search. In that study, a person had to search for a visual object in a cluttered visual field. It was found that subjects were much better at discriminating and identifying an object in a cluttered visual field using the colour characteristic than the object's size or shape. This may support the findings in this experiment, where the colour scheme concept is dominant in respondents' answers, whether drawn from geoliterate or non-geoliterate respondents. Colour scheme seems to become the basis of their search for flaws in the comparison maps. The identified flaws then become the basis of their judgement to decide which map is perceived as credible to assist them in their tasks.

In this experiment respondent made judgements by using metadata element, the 'identity of the map author/creator' (map producer) was less in numbers compared visual cues when making judgements. Moreover, in this group of respondents, there was no significant difference between geoliterate and non-geoliterate participants' use of metadata related elements as the basis of their judgement. Table 4.11 presents the summary of the hypotheses in this experiment.

Table 4-11 Hypotheses summary

No		The hypotheses statement in this experiment	Result
1.	Hypothesis 1	Visual cues have no significant influence in respondents' credibility assessment on map mashup applications	Not supported
2.	Hypothesis 2	The metadata related to sources (i.e. map mashup producer) have significant influence in respondents' credibility assessment	Not supported
3.	Hypothesis 3	The metadata related to sources have significant influenced within geoliterate respondents	Not supported

4.15 Conclusion

This experiment provided some useful insights into the elements that become the basis of map readers' selection and rejection of a map. The open-ended questions in this experiment allowed respondents to state the basis of their judgement freely without restriction. The responses drawn from this type of survey may reveal the 'true feeling' of respondents when making judgements in selecting and rejecting a map. From the results, visual cues formed the basis of elements that respondents used when making judgements.

‘Colour scheme’ and ‘information clarity’ emerged as the dominant concepts in responses when selecting and rejecting a map. It can conclude that visual cues that include the ‘colour scheme’, ‘information clarity’, ‘overall (design look)’, ‘information details’, ‘individual preferences’ and ‘map design’ (which were the concepts that emerged from respondents responses) were dominant factors compared to the tested metadata variables.

This surprising result rejected the hypothesis of this experiment, which had anticipated that the critical element in metadata, namely sources (i.e. identity of map producer) would be the dominant basis of judgement. Further experiments have to be conducted to examine the extent to which visual cues constitute the dominant element and whether metadata related elements are the least important when map users make judgements relating to the perceived credibility of map mashups. This first experiment became a pilot study for the design of the next series of experiments.

5 EXPERIMENT TWO

5.1 Introduction

This section describes the aim, hypothesis, methodology, results and analyses used in Experiment Two. The objective of this experiment is similar with the previous experiment (Experiment 1), which is;

'to examine the influence of metadata related to map producer and map supplier on respondents' assessment of credibility when selecting and rejecting a map mashup'

The purpose of this experiment was to confirm the findings and to enhance the methodology applied in Experiment One; in which open questions were used to collect respondents' responses about the factors they measured to determine the credibility of a map mashup. In Experiment Two, open questions were replaced by closed questions whereby a predetermined set of answers was provided with the questions. This is to tackle the limitations in Experiment 1 that explained in Chapter 10: Section 10.3, specifically on the issue of single item respond given by respondents. Experiment Two consists of a series of experiments that were conducted to examine the influence of metadata related to 'map data supplier'. One parameter was used to test this element in this Experiment 2, which was the 'identity of map data supplier'.

5.2 Hypotheses

The hypotheses of this experiment were based on the research questions highlighted at Section 1.3 (Chapter 1).

Hypothesis 1 is**:**

Visual cues have no significant influence in respondents' credibility assessment on map mashup applications

Hypothesis 2 is:

The metadata related to sources (i.e. map data supplier) have significant influence in respondents' credibility assessment

Hypothesis 3 is:

The metadata related to sources have significant influenced within geoliterate respondents

5.3 Experimental Design

Three conditions were designed to test the experimental hypotheses. The conditions are summarised in Table 5-1. The conditions were designed according to the experimental assumption of credibility levels – low credibility, intermediate credibility and high credibility level.

Table 5-1 Experimental Conditions

	Condition 1	**Condition 2**	**Condition 3**
Experimental Assumptions	High level of credibility	Intermediate level of credibility	Low level of credibility
Variable 1 : Supplier of foreground data	Nottingham City Council versus Student Union	BBC versus Mix Bistro Cafe (café in the university)	Anonymous supplier versus Jane Smith
Full descriptions (as in the experimental task sheets)	The top data layer (i.e. the placement of map symbols and information) are supplied by the Nottingham City Council	The top data layer (i.e. the placement of map symbols and information) are supplied by the British Broadcasting Corporation (BBC) Nottingham	The top data layer (i.e. the placement of map symbols and information) are supplied by anonymous individuals
	The top data layer (i.e. the placement of map symbols and information) are supplied by the Students' Union	The top data layer (i.e. the placement of map symbols and information) are supplied by the Mix Bistro-café. The Mix Bistro company runs several	The features on the top data layer (i.e. the placement of map symbols and information) are supplied by Jane Smith

		cafes at student halls of residence at the University Park campus.	

5.3.1 Dependent Variables

Core dependent variables used in this experiment are shown in Table 5-2. Full details of these dependent variables are discussed in experimental methodology (Section 2.3).

Table 5-2 Dependent variables used in Experiment Two

Question ID	**Dependent variable**	**Measurement**
Q1	Selected map to serve the experimental task	Binary
Q2	Data Supplier of foreground data	Ordinal (ranking question)
Q3	Perceived credibility in the selected map	7-point Likert scale
Q4	Perceived credibility in the rejected map	7-point Likert scale
Q5	Perceived credibility in the supplier of data	5-point Likert scale
Q6	Perceived credibility of the data source	5-point Likert scale

5.3.2 Participants

There were 114 respondents who completed the study. The numbers of respondents were varied in each of six experimental tasks. This was due to the ability of the server database to record simultaneously the self-completed questionnaires submitted by respondents. As discussed in Chapter 3: Section 3.8, the sample in this research was selected based on volunteer subjects sampling. The sample was drawn from the members of University of Nottingham (staff and students) and non-university members from outside community who resided in the United Kingdom or Malaysia. The average age in the sample was 25 where the mode values fell in the groups of 22 to 24 years and 25 to 30 years. There were more female (58.9%) than male (41.1%) respondents in the sample. The majority of respondents had experience of using maps supplied by a national mapping provider (e.g. Ordnance Survey and JUPEM (Department Survey and Mapping Malaysia) and maps from commercial providers (e.g. Google Map, Bing Map, and Yahoo Map). 46.10% of respondents had experience of visiting online maps administered by individuals, groups or maps that were community based.

Data from seven respondents were excluded from the analysis pertaining to geoliterate and non-geoliterate respondents since these respondents did not identify their academic or professional background on the demographic form. Table 5.3 shows full details of respondents' demographic backgrounds.

Table 5-3 Respondents' Demographic Information

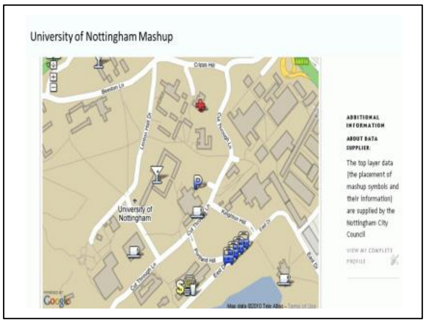
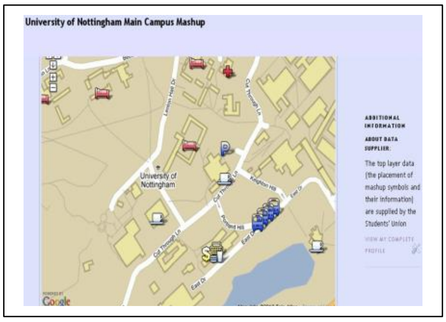
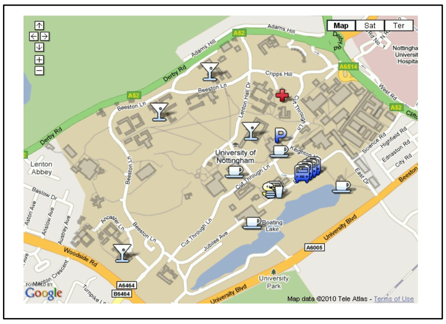
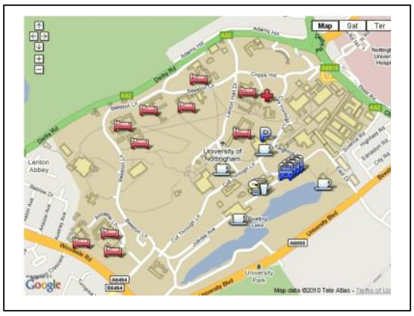
Task	**Geoliterate**	**Non geoliterate**	**Unknown**	**Total respondents**
1	57	78	6	141
3	45	61	5	111
5	43	59	5	107
Average total number of respondents (average)	114			
Age (average)	25.35			
University members	60%			
Non-members	40%			
No. of respondents residing in the UK	70%			
No. of respondents residing in Malaysia	29%			
Male	41%			
Female	60%			
Highest level of formal education attended	High school (33 %), Bachelor's degree (31 %), Master's degree (18%), others (17 %)			
Current primary occupation	Student (51 %), researcher (22%), others (26 %)			
Experience of using paper maps	99%			
Experience of using maps from a national map provider	79 %			
Experience of using maps from a commercial provider (Google Map, Yahoo Map etc.)	96 %			
Experience of visiting online maps administered by an individual, group, or community based website	46 %			

5.3.3 Materials

Two simulated static map mashups were distributed for each experimental condition. The same maps were presented in three different scales – small scale, middle scale and large scale- to provide several views of the maps. They were labelled as simulated mashups because the maps displayed on the questionnaire were print-screen maps; they were not interactive and did not allow the function of zooming and identifying features. Although the maps were not interactive, they were embedded in three different scales that displayed information if a feature was clicked by an identify function. The reason for using non- interactive maps was the need to control the views that respondents would see on the maps; respondents could make judgements based on the same view. The maps were created using a simple tool to develop mashups.

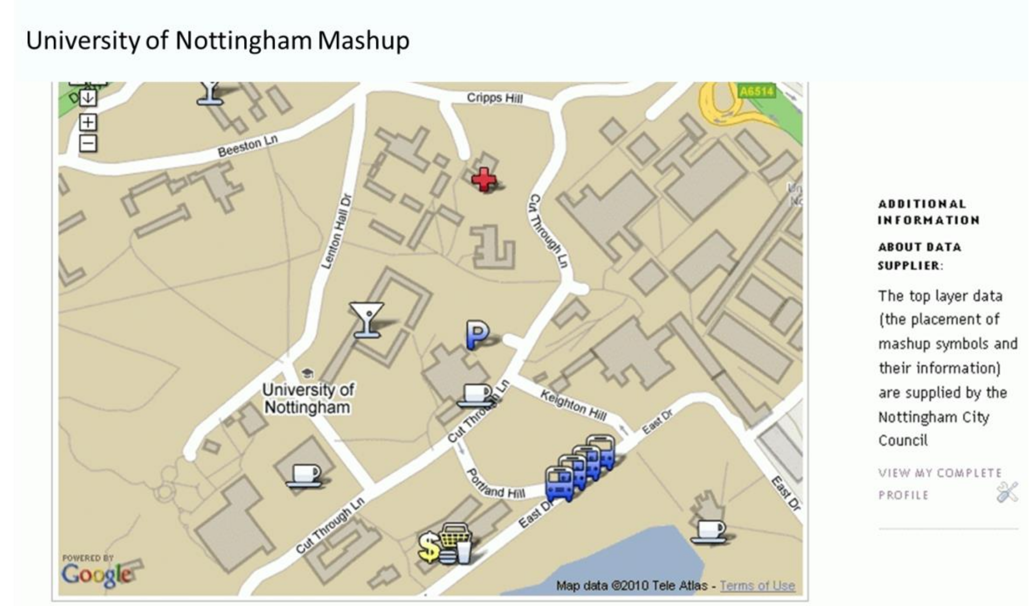
Figure 5.1 shows an example of a map based questionnaire used in each experimental task. Figure 5.2 to Figure 5.3 below and Table 18 to Table 20 (see Appendix B) show the maps and the style features used to examine the variables of the foreground data supplier. The previous Experiment One indicated ‘colour scheme’ to be the major element used by respondents to determine the map they wanted to use for the self-guided tour. In that experiment, the map styles were randomly designed and not based on any convention related to cartographic design; in Experiment Two the maps were well designed by considering the hues in the ‘colour scheme’ so that its maintain similar contrast with the background (Brewer, 2005).

Task 3/6:
Compare the set of Map A with the set of Map B and their attached information.
Then choose only one (1) map mash up you will use in your self-guided campus

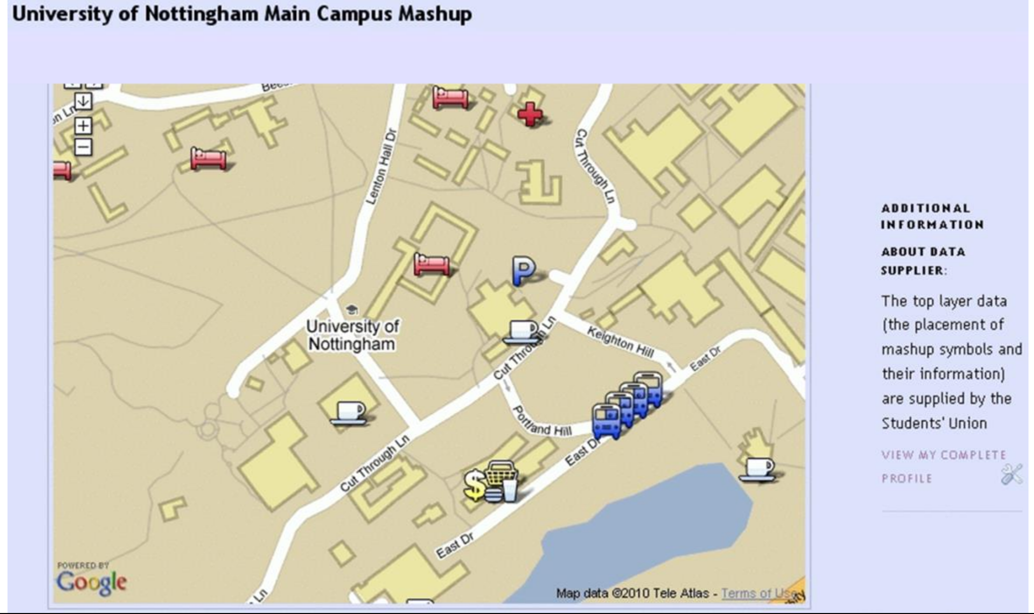
A Set of MAP A **(Main Map)** **(See Figure 5.2)**	A Set of MAP B **(Main Map)** **(See Figure 5.2)**
	
Map A (Second Map) -The view if you zoom in the map at a larger scale and click at the bus symbol	Map B (Second Map) – The view if you zoom in at a larger scale and click at the basket symbol
	
Map A (third map) – The view if you zoom out the map at a smaller scale	Map B (third map) – the view if you zoom out at a smaller scale
	

Question 3/6:
Compare the set of Map A with the set of Map B and their attached information.
Then choose only one (1) set of map mashups you will use in your self-guided campus tour
Then answer the questions below.

Figure 5-1 The layout of questionnaire in Experiment Two

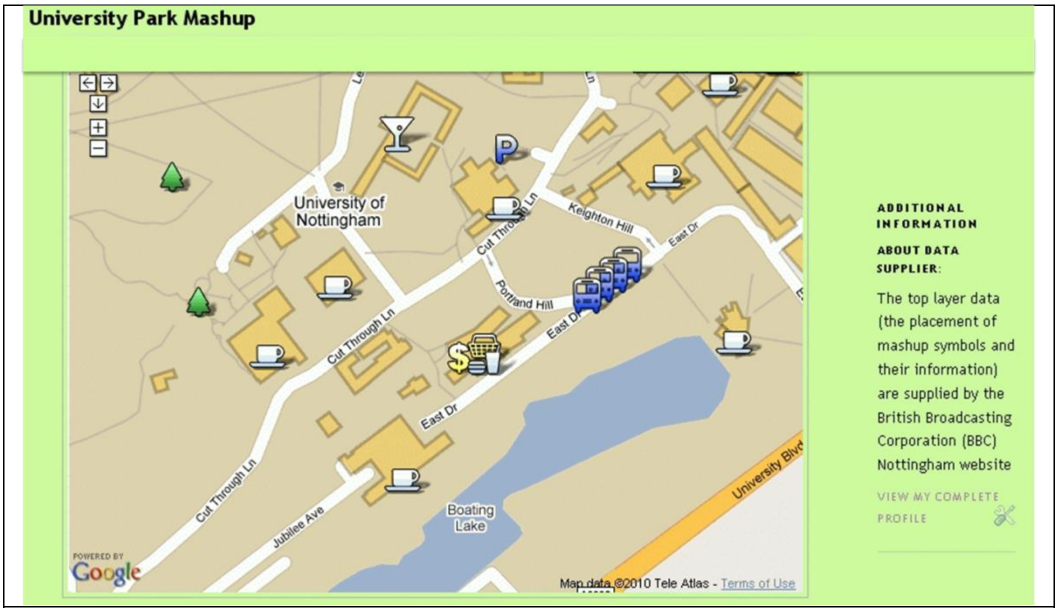


Map Mashup A (see Table 5.1 for a clearer text version)
Variable: Supplier of foreground data
Parameter: Nottingham city council

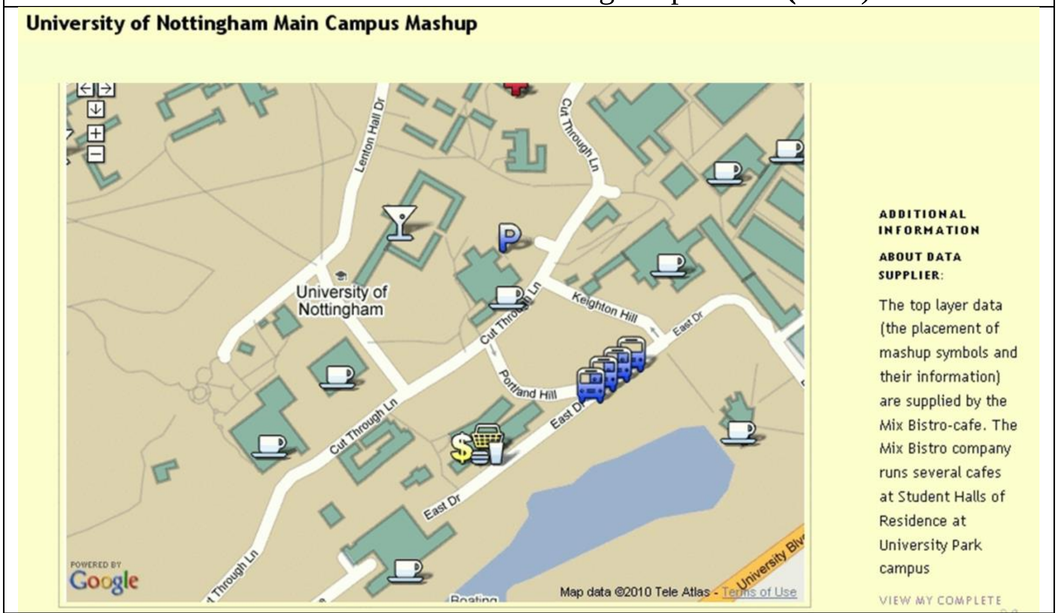


Map Mashup B (see Table 5.1 for a clearer text version)
Variable: Supplier of foreground data
Parameter: Students' Union

Figure 5-2 A Snapshot of Condition 1 (Experimental Task 3) Map A and Map B

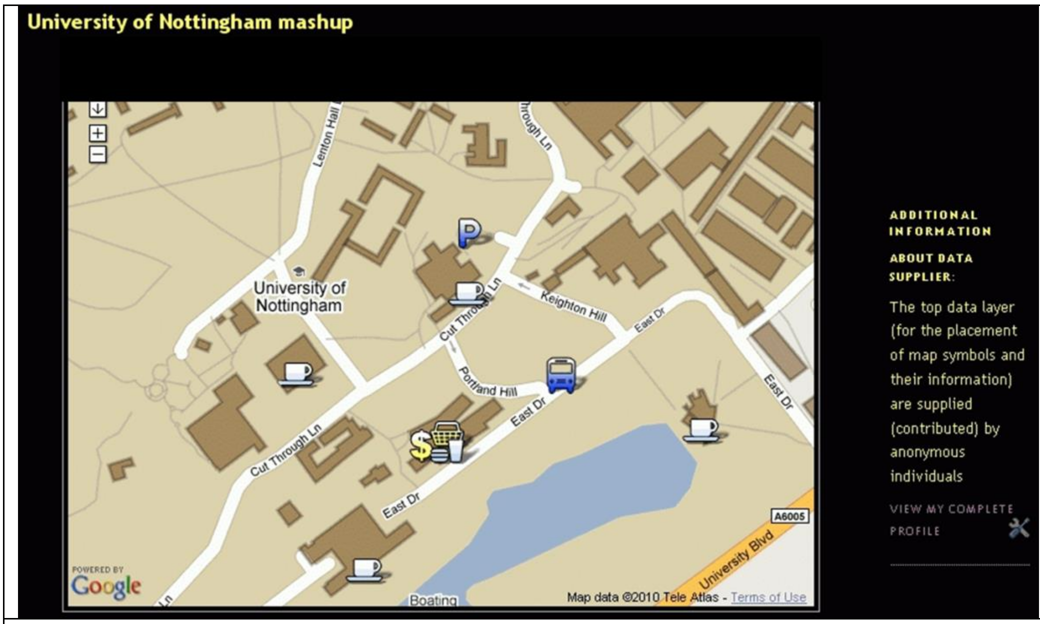


Map Mashup A (see Table 5.1 for a clearer text version)
Variable: Supplier of foreground data
Parameter: British Broadcasting Corporation **(**BBC)

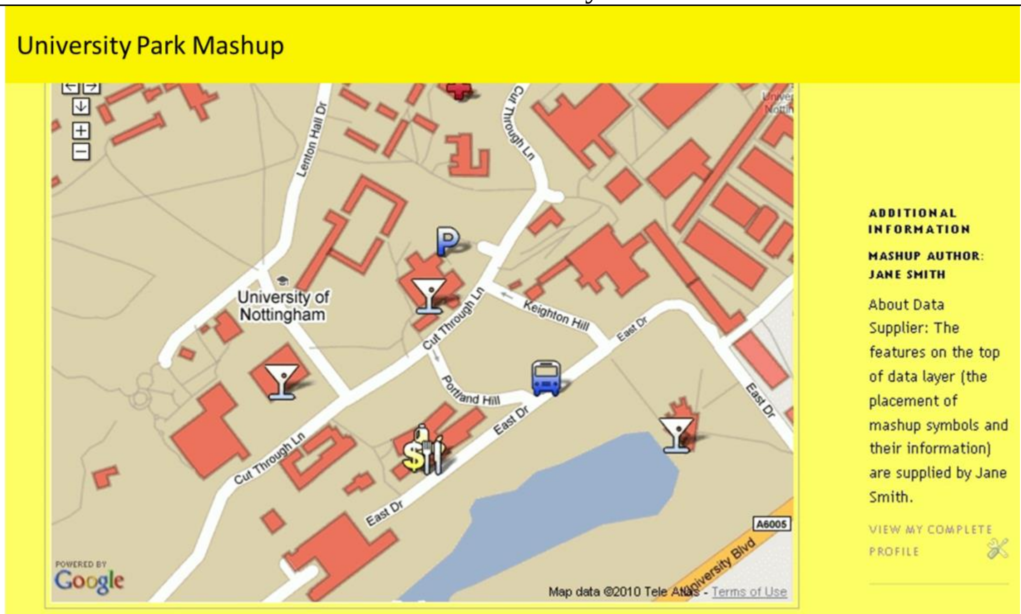


Map Mashup B (see Table 5.1 for a clearer text version)
Variable: Supplier of foreground data
Parameter: Mix Bistro

Figure 5-3 A Snapshot of Condition 2 (Experimental Task 5) Map A and Map B



Map Mashup A (see Table 5.1 for a clearer text version)
Variable: Supplier of foreground data
Parameter: Anonymous



Map Mashup B (see Table 5.1 for a clearer text version)
Variable: Supplier of foreground data
Parameter: Individual

Figure 5-4 Snapshot of Condition 3 (Experimental Task 2) Map A and Map B

5.4 Questionnaire Design

The questionnaire in this experiment measures two constructs:

1) The influence of metadata element related to 'data supplier' and 'map producer' of the map mashups (items of Q2f, Q2g, Q5,Q6,Q7)
2) the influence of visual cues of map mashups (items of Q2a, Q2b, Q6,Q7)

The questionnaires conducted in this experiment were experiment based, whereby responses were stimulated by main questions under three different experimental conditions. The main question was to analyse and make a judgement between two sets of maps that a respondent would use in a context of self-guided campus tour. Based on the responses to this main question, a respondent was given a series of questions to measure the two items. Table 5.4 presents the series of questions used in the questionnaire and the specific measures that were applied.

Table 5-4 The questions in the questionnaire and the measures

ID	Question	Measurement
Qa	Section: Spot the differences Spot three differences between the FIRST MAP in set A and the FIRST MAP in set B. Then choose your answer(s) from the drop down lists given below First difference? Second difference? Third difference?	Nominal
Q1	For the purpose of a self-guided tour, which map mashup you will use?	Binary
Q2	What was the basis for your decision in Q1 in selecting this set of maps and rejecting the other set of maps? Please sort your reason(s) from the list by ranking them according to your order of priority a) Colour scheme on map b) The map mashup design look (amateur/professional c) Readability of text (feature's label) d) The underlying motive of the webpage to be published e) Usefulness of information	Ordinal (ranking question)

	f) The data supplier for the top data layer g) The web producer's affiliation h) Coverage (scope) of information i) Unbiased information	
Q4	How much will you perceived the credibility of the set of map you selected?	7-point Likert scale
Q5	How much will you perceived the credibility of the set of map you rejected?	7-point Likert scale
Q6	Q6: The map mashup design looks professional	5-point Likert scale
Q7	Q7: I like the colour(s) applied on the mashup feature(s)	5-point Likert scale
Q8(a)	Q8: I am familiar with the data supplier(s) for the top data layer	5-point Likert scale
Q9(a)	Q9: I perceived the information supplied by the data supplier(s) of the top data layer is credible	5-point Likert scale
Q10(a)	Q10: The data source(s) of the top layer for this set of maps is more credible than the other set of maps	5-point Likert scale

To understand the samples in the experiments and the differences between responses, a series of demographic questions were asked before respondents conducted the experimental tasks. Table 5.5 presents the series of demographic questions in the questionnaire.

Table 5-5 A series of demographic questions used in the questionnaire

	Items
Q1	Age
Q2	Gender
Q3	Location of residence
Q4	Highest level formal education
Q5	Current primary occupation
Q6	Professional and academic qualifications
Q7	Experience of using paper based maps
Q8	Experience of using maps supplied by a national map provider
Q9	Experience of using online maps from commercial providers
Q10	Experience of using online maps administered by, an individual, group or community based
Q11	Member or non-member of the University of Nottingham, UK

5.5 Experimental Procedure

The procedure used in this experiment is depicted in the figure below.

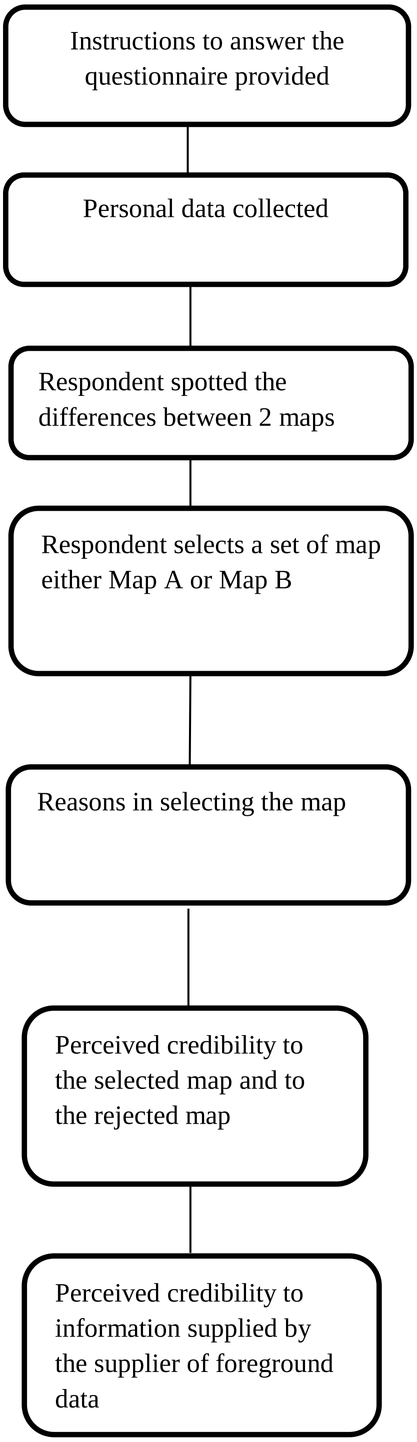


Figure 5-5 Procedures used in Experiment Two

5.6 A Pre-test study

A pre-test study consisting of four respondents was conducted in order to test the designs of the map based questionnaire. One respondent came from a geoliterate background, while the other three were from engineering backgrounds. After analysing the data and feedback from the pre-test, slight changes were made particularly on the questions wording of the questionnaire.

5.7 Results and Analysis

This section describes the analysis of the data and the results. The first part describes the characteristics of the data, and the statistical tests to test the hypotheses. This section also describes the reliability of the data. The latter section describes the results in relation to the research hypotheses.

5.7.1 Data Analysis

Statistical analysis using SPSS v.16 was used to study the response data. The data collected from the questionnaire were measured as binary variables and ordinal variables, respectively. The responses data were not in continuous variables and have bee transform from discrete variables using Terrell's Transformation technique. However, the data have been tested do not fit with normal distribution model. Hence, data analysis focused on selecting appropriate non-parametric statistical tests. The Chi-square test and Mann Whitney U test were used to analyse the data. Non parametric tests are free from the restrictive assumption that the data needs to be normally distributed; The Chi-square test has a specific assumption of the data before it can be modelled correctly. The assumptions in the Chi-square test are that the data is independent and each entity contributes to only one cell in the contingency table and the expected frequencies should be greater than 5. These assumptions were not violated with regard to the experimental data modelled by this test.

As mentioned in Chapter 3 (Methodology) in Section 3.4, this research used internal consistency test to check the reliability of the items (questions) in the questionnaires. This test is to check how well the items measure the same issue (construct). Table 5.6 to Table 5.7 below presents the items in the questionnaire that measure the same construct with the value of Cronbach's alpha if the item deleted. Table 5.8 presents a summary of the value of Cronbach's alpha for each measured construct in this experiment.

Table 5-6 The items that measure the construct of the influence of 'data supplier'

items	**Construct: The influence of data supplier**	Cronbach's alpha if item deleted
QS8	I am familiar with the data supplier(s) for the top data layer	0.726
QS9	I perceived the information supplied by the data supplier(s) for the top data layer is credible	0.699
QS10	The data source(s) of the top layer for this set of maps is more credible than the other set of maps	0.825

Table 5-7 The items that measure the construct of the influence of visual cues

Items	**Construct: the influence of visual cues**	Corrected Item-Total correlation
QV6	The map mashup design looks professional	0.741
QV7	I like the colour(s) applied on the mashup feature(s)	

Table 5-8 A summary of the Cronbach's alpha for each construct

Constructs	**Cronbach's alpha**
The influence of Data supplier	0.820
The influence of Visual cues	0.850

From the results of reliability analysis in Table 5.8, the results of Cronbach alpha coefficient of each construct was more than 0.7. A value of 0.7 to 0.8 is an acceptable value for Cronbach's alpha (Field, 2009, p.675). This indicates the scales used in the questionnaire of this experiment are reliable to measure the constructs.

5.7.2 Results: Exploration of the map that respondents chose

Frequency analysis was conducted to analyse the main question: 'For the purpose of a self-guided campus tour, which map you will use?' This question yields two binary variables - Map A and Map B. Table 5.9 presents the results.

Tests of Significance using the Chi-square tests for independence were conducted on the results in Table 5.10. This test compares the observed frequencies in the categories of the map with most selected versus the map with the least selected with values that would be expected if there were no association between the two categories being measured. The shaded cells in Table 5.9 indicate the experimental conditions under which the differences between the map with most selected and the map with least selected were statistically significant. From the table below, there was one experimental task in which the difference in responses was statistically significant – Condition 2 (Task 5) and Condition 3 (Task 1).

Table 5-9 Frequency analysis of the maps that respondents chose in each experimental task

	Most selected	%	Count	Least selected	%	Count	Chi-square test
Condition 1 (T3)	Map A city council	57.66	64	Map B Student union	42.34	47	X^2 (1, n = 100) = 2.56, $p < 0.110$)
Condition 2 (T5)	Map A BBC	70.09	75	Map B Mix bistro	29.91	32	X^2 (1, n = 100) = 16.0, $p < 0.001$)
Condition 3 (T1)	Map A anonymous	55.96	93	Map B Jane Smith	34.04	48	X^2 (1, n = 90) = 5.378, $p < 0.026$, $p < 0.05$)

5.7.3 Results: The pattern of factors, according to respondents' priority ranking order, that became the basis of respondents' decision in selecting and rejecting the map mashup

Hypothesis 1 (null) is**:**

Visual cues have no significant influence in respondents' credibility assessment on map mashup applications

Hypothesis 2 is:

The metadata related to sources (i.e. map data supplier, map mashup producer) have significant influence in respondents' credibility assessment

Frequency analysis was conducted for the response of 'what was the basis of your decision in Q2- selecting this set of maps and rejecting the other set of maps-; please choose up to five reasons from the list by ranking the reasons according to your order of priority'. The responses are presented in Table 5.10. Only factors that collected a high number of responses in each level of importance (i.e. first priority to fifth priority) were shown in this table. From the results, the dominant reasons to choose and reject the maps were related to visual cues; hence **Hypothesis 1 is not supported**. There were only two experimental conditions, as below, rated 'data supplier' as their reasons.

- Condition 1 (City council vs. student union),
- Condition 3 (anonymous vs. Jane Smith) and

However, this high number of responses was rated at the fourth and fifth ranks, not at the first rank. Shaded cells in Table 5.10 indicate these results. Therefore, **Hypothesis 2 is not supported**.

Table 5-10 The pattern of factors according to respondents' priority ranking order for selecting the map

Rank	Condition 1 city council vs. student union	Condition 2 BBC vs. Mix Bistro	Condition 3 Anonymous vs. Jane Smith
1	Useful (48.6%)	Useful (36.4%)	Colour scheme (46.10%)
2	Coverage (21.6%)	Coverage (24.3%)	Design look (26.95%)
3	Readability (16.2%)	Readability (16.8%)	Useful (20.57%)
4	Supplier (9.9%) Design look (9.0%)	Design look (9.3%)	Coverage (12.06%)
5	Colour scheme (7.2%)	Colour scheme (6.5%)	Supplier (7.80%)

5.7.4 Results: The proportion that measured the variable of foreground data supplier

Hypothesis 3 is:

The metadata related to sources have significant influenced within geoliterate respondents

Frequency analysis was conducted to the response of ‘what was the basis of your decision in selecting this set of maps and rejecting the other set of maps; please choose up to five reasons from the list ranking the reasons according to your order of priority’.

Table 5.11 shows the results, based on the ranking order from the responses. Of the results, the number of responses was higher in Condition 2 where the variables of BBC and Mix Bistro Café were manipulated. However, the difference in responses compared to Condition 1 was 6.4% and was not significant. From the results:

1) in Condition 3, a high proportion of respondents did not measure the factor of ‘data supplier’ in their judgement. In Condition 3, the variable of anonymous and an individual named ‘Jane Smith’ was manipulated on the maps; 64.54% of respondents did not measure the variable of ‘data supplier’ in this condition. The lowest number of responses (35.46%) measured the metadata in Condition 3 compared to the other five conditions; the tested variables were between ‘Anonymous’ and ‘Jane Smith’ as the data supplier of the map.

2) the highest numbers of responses (54.1%) that measured the metadata was in Condition 1; the manipulated variables were between ‘City Council’ and ‘Student Union’. A low number of respondents in Condition 1 compared to the other conditions did not measure the factor of ‘data supplier’. In this condition, the variables of ‘city council’ and ‘student union’ were manipulated on the maps; 45.9% of respondent did not measure these factors in their judgement.

Table 5-11 The number of respondents (as a percentage) who measured foreground data supplier factor according to ranking order

Ranked factor	Condition 1 (%)	Condition 2 (%)	Condition 3 (%)
Experimental assumption	High level credibility	Intermediate level credibility	Low level of credibility
Variables	City council vs. Student Union	BBC vs. Mix Bistro cafe	Anonymous vs. Jane Smith
First	12.60	17.80	6.38
Second	14.40	13.10	6.38
Third	12.60	10.30	7.80

Fourth	9.90	1.90	7.09
Fifth	4.50	4.70	7.80
Respondents not measured	45.9 (51) n=111	52.30 (56) n=107	64.54 (91) n=141
Total respondents measured	54.1 (60)	47.7 (51)	35.46 (50)
Average ratio	1:1		

Table 5-12 The proportion that had spotted the differences of data supplier

Conditions	Spotted differences of data supplier	Measured (ranked) data supplier	Not measured data supplier
Condition 1 (T3)	Yes	20	8
Condition 2 (T5)	Yes	16	13
Condition 3 (T1)	Yes	11	19
Average		15.7 ~ 16	13.3 ~ 13

Table 5.12 presents the results of ‘spot the differences’ activities that conducted before respondents analyse and choose the map that they perceived credible for the given task. From the results, on average, there was no significant different within the sample that spotted the differences of data supplier parameters with the proportion that measured (ranked) and not measured data supplier. In other words, respondents noticed the data supplier of both maps were different, however this did not necessarily influence them to assess credibility of information by using that element (i.e. data supplier).

Comparisons of ranking scores between the geoliterate and non-geoliterate groups in the sample that measured the data supplier factor in the assessments were conducted. From the Mann Whitney U test, which compares the responses between groups by looking at the differences in the ranked positions of scores, indicates that there were no significant differences in the scores rated by the geoliterate and non-geoliterate groups. There were no differences in the priority levels rated between the two groups. Hence, the **hypothesis 3 is not supported**.

Charts 5.1, 5.2 and 5.3 present the distribution of scores between the groups. In each condition, a high number of responses between groups rated at point 0 (not measured the factor). The ranked positions of factors according to respondents’ priorities for selecting or rejecting the map varied from rank 1 to rank 5.

From the charts, both groups rated a high number of responses for not measuring the variable of ‘data supplier’ in their assessment. However, the total responses that rated the factor on positions 1 to 5 shows there were no significant differences between the number of responses that measured (sum of position 1 to 5) or did not measure (position 0) the factor.

There was an exception in Condition 3 (see Table 5.13) within the geoliterate group, where there was a significant difference between the number of responses that measured the factor and did not measure the factor, X^2 (1, n = 57) = 12.79, p < 0.001). In this case, the proportion of respondents within the geoliterate group that measured the factor was considerably lower than the proportion that did not measure the factor.

Specific analysis was conducted to the sample that spotted the differences of parameters of data supplier in 'spot the differences section'. From Table 5.13, there were no significant of the proportion that assessed and not assessed 'foreground data supplier' in their judgements between geoliterate vs. non-geoliterate.

Table 5-13 Results comparison on the differences of ranked scores between groups based on Mann Whitney U test in Conditions 1 to 3

	Geoliterate	**Non-geoliterate**		
Condition 1	Mean rank = 56.3 Mdn = 1 , n = 45	Mean rank = 51.43 Mdn = 1, n = 61	U = 1246.5, z = -0.852, p = 0.394, r = 0.08	Not significant at p > 0.05, r= small effect
Measured the factor	26	31		
Factor not measured	19	30		
Spotted 'foreground data supplier'				
Measured the factor	8	10		
not measured	3	5		
Condition 2	Mean rank = 50.78 Mdn = 0, n = 43	Mean rank = 52.03 Mdn = 0, n = 59	U = 1237.5, z = -0.229, p = 0.819, r = 0.02	Not significant at p > 0.05, r= small effect
Measured the factor	19	29		
not measured	24	30		
Spotted 'foreground data supplier'				
Measured the factor	5	9		
not measured	7	6		
Condition 3	Mean rank =	Mean rank =	U = 1869.5,	Not

	61.8 Mdn = 0, n = 57	72.53 Mdn = 0, n = 78	z = -1.843, p = 0.065, r = 0.16	significant at p > 0.05, r= small effect
Measured the factor	15	33		
Spotted ‘foreground data supplier’				
Measured the factor	4	6		
not measured	8	9		

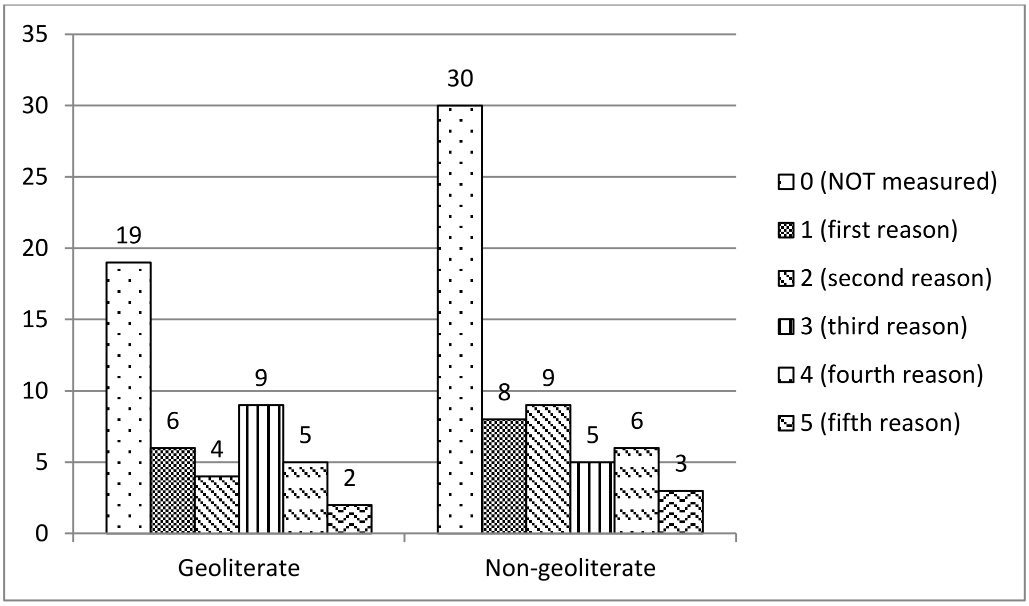


Chart 5-1 Score distribution between geoliterate and non-geoliterate groups in experimental Condition 1

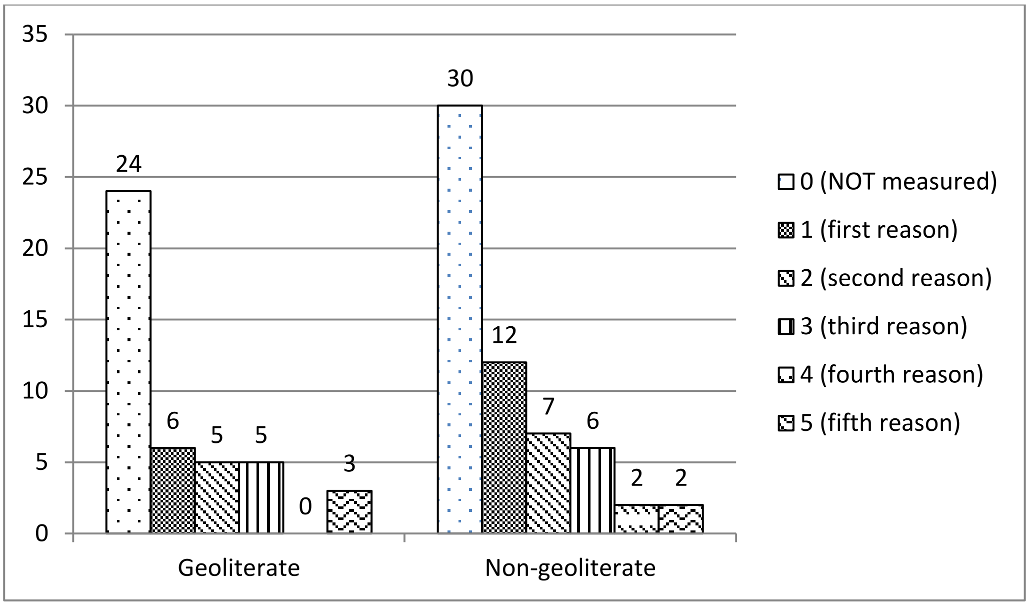


Chart 5-2 Score distribution between geoliterate and non-geoliterate groups in experimental Condition 2

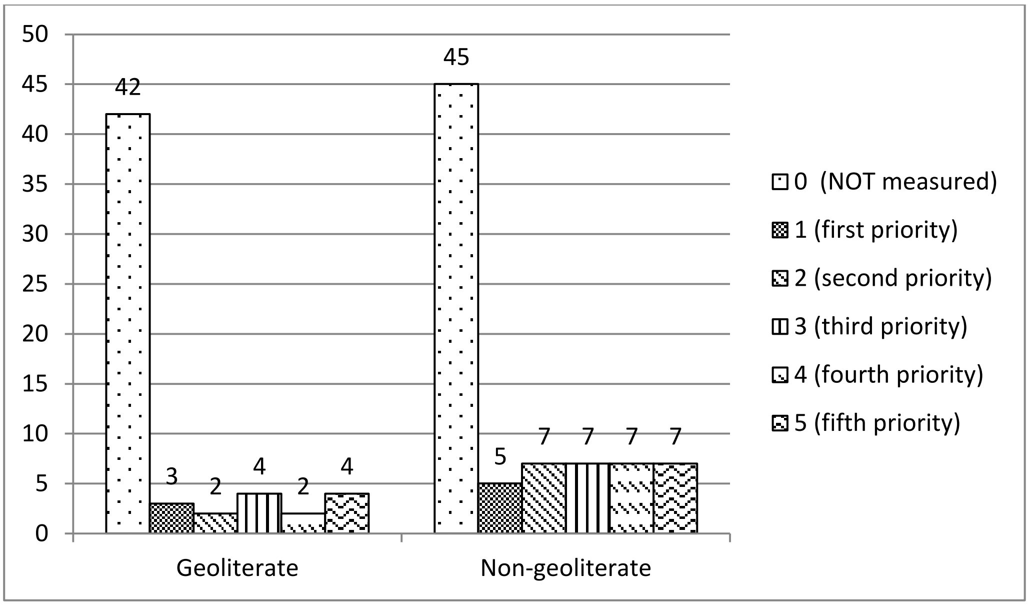


Chart 5-3 Score distribution between geoliterate and non-geoliterate groups in experimental Condition 3

5.7.5 Results: The influence of visual cues vs. the influence of metadata

Hypothesis 1 (null) is:

Visual cues have no significant influence in respondents' credibility assessment on map mashup applications

A list of items in three constructs (see Table 5.7) was used to assess the influence of visual cues and metadata in respondents' assessment. Terrel's transformation technique was used to convert ordinal data from the scale into indices (as in Table 5. 14) in order to compare the mean and check the significance differences between scores.

The formula used as below; (Equation 5.1)

$$Transformed\ score\ (T) = \frac{[\text{Actual raw score} - \text{lowest possible raw score}]}{\text{possible raw score range}}\ x\ 100$$

Table 5-14 Formulae for scoring and transforming scale

Construct	Items to be summed for actual raw score values	Lowest possible score range	Possible score range
Visual cues	V6 + V7	2 x 1 = 2	2 x (5-1) = 8
Metadata (Data supplier)	S8 + S9+ S10	3 x 1 = 3	3 x (5-1) = 12

After transforming the ordinal to the interval value, the total scores for each construct from each respondent were calculated. The scores then tested into Kolmogorov-Smirnov test to check the normality of the distribution. From the test, the scores from the construct that measure the influence visual cues, $D(107) = 0.23$ $p < 0.001$ and the influence of data supplier $D(107) = 0.12$, $p < 0.001$, were significantly non-normal. This finding indicates that the distribution of scores significantly differs from a normal distribution and do not met the assumption to use parametric tests, namely dependent t-test to compare the mean of scores. Therefore, the non-parametric counterpart of the dependent t-test that called Wilcoxon-signed rank Test was used.

The Wilcoxon signed ranked Test works based on the differences between scores in the two conditions. Once the different calculated, they are ranked but the sign of difference (positive and negative) is assigned to the rank. Table 5.15 below presents the results from Wilcoxon test that comparing visual cues construct with data supplier construct respectively. From this test:

'the scores of the influence of visual cues was significantly higher than the influence of data supplier'

Hence, **hypothesis 1 is not supported**. There was exception to Condition 2 where the different between the group scores was not significant different. The effect sizes (r) that indicate the differences between constructs were large at Condition Page 5 ($r = 0.6$); in the others conditions, the differences were small (below 0.3) and medium size (below 0.5 to 0.3).

Table 5-15 The results from Wilcoxon Signed Rank Test

Conditions	Construct Visual Cues (median)	Construct Data supplier (median)	Wilcoxon test
1	75.0	50.0	$T = 884.0$, $p < 0.001$, $r = 0.4$
5	75.0	66.7	$T = 1071.5$, $p < 0.05$, $r = 0.6$.
3	75.0	66.7	$T = 1095.0$, $p < 0.05$, $r = 0.18$

5.8 Additional experiment (Experiment 2a)

Further online map based questionnaires was conducted to confirm the findings of previous Experiment 2. In this additional experiment, several changes were made to the experimental design; the intention was to control and minimise all the design factors that might affect respondents' judgement, but concentrated on the differences of reputation level of foreground data suppliers. In this experiment, the data and colour scheme between maps were controlled to be identical. However, the symbols of points of interest (i.e. café and bus stop) were designed to be slightly a little bit different; the intention was to prompt respondents to engage with the maps exercises (i.e. to choose the map that they perceived credible) but in very minimal distraction. The data suppliers of maps were manipulated to be easily distinguishable in terms of their reputations in supplying data related to the campus. Table 5.16 below present the map setting differences between these comparison maps in this experiment.

Table 5-16 Map setting differences in Experiment 2a

Experiment 2a	**Data**	**Visual cues**	**Reputation level of data supplier**
Map A	Identical	Identical except on design of symbols	The Starbucks Coffee
Map B			The University of Nottingham

The online map based questionnaires were distributed through the intranet student portal of the University of Nottingham and via mailing lists between 4 until 20 June 2010. There were 137 respondents had joined the questionnaires. Of this sample, 11 respondents were from Malaysia and not familiar with the map area. 17 respondents were drawn from the geoliterate group which had professional or academic background in geography, land surveying, cartography or Geographic Information Sciences (GIS); whereas others were non-geoliterate respondents who had professional or academic background in other fields. Table 5.17 below presents the questions that were collected in this experiment.

Table 5-17 A list of questions collected in Experiment 2a

ID	Question	Measurement
Task	Evaluate the Site A and the Site B and their attached information. Then decide which one has the most believable information	
Q1	Please evaluate the Site A and Site B published above, and rank them from most to least credible.	Nominal 1 = most credible 2 = less credible
Q2	Please tick the reason(s) for your answer in Q1 - the coverage area of map - the design of the map - readability of labels **- the source (supplier) of foreground data** - the accuracy of information - clarity of map symbols - advertising -identity of site operator (creator) -colour scheme	Nominal
Q3	Please share your comments (optional)	Open-ended question (qualitative data)

Figure 5.6 presents the snapshot of the comparison maps. The maps presented on the experiment were in static mode where respondents could only view and assess the screenshot maps; the intention was to control respondents' assessment so that they would make judgement based on what given in the test.

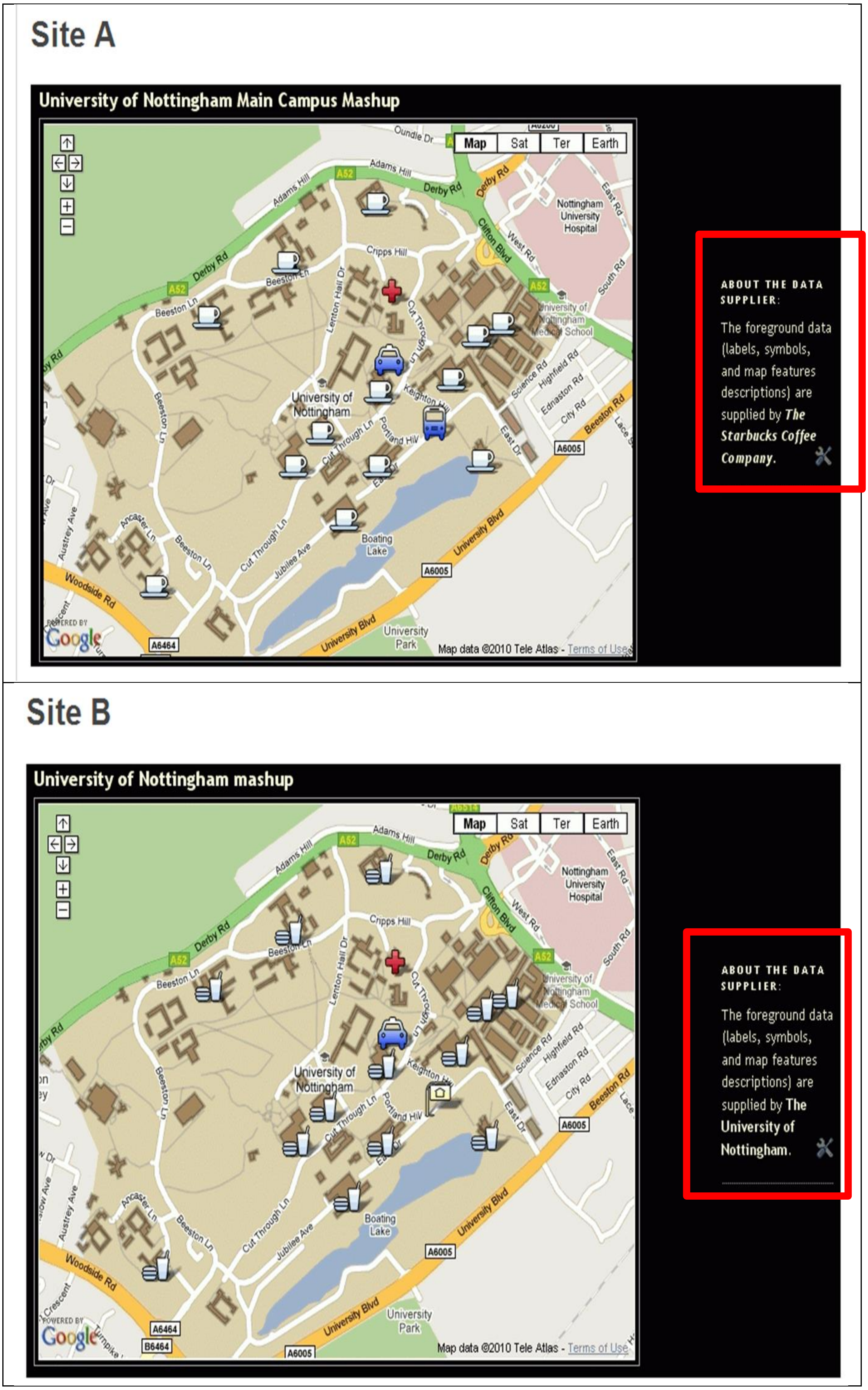


Figure 5-6 Snapshot of comparison maps in Experiment 2a

Table 5.18 below presents the results of the questionnaires. Analysis on the results of Q1 which was ‘please evaluate the Site A and Site B published above, and rank them from most to least credible’ demonstrated the number of respondents that chose either Map A and Map were not significant different. Of the sample that chose Map A as more believable (i.e. the foreground map data supplied by the Starbucks Coffee), 70 respondents stated their reason to choose that Map due to the influence of visual cues, including the ‘design of map’, the ‘readability’ of labels, the ‘clarity of map symbols’. Hence, **Hypothesis 1 is not supported**. Meanwhile, 68 respondents chose Map B (i.e. the map data supplied by the University of Nottingham). Of this sample, 49 respondents were influenced by the foreground data supplier; whereas other 19 respondents chose this map not due to the data supplier, but had been influenced by the visual cues of the map. This might indicate **Hypothesis 2 is not supported**. There was also no significant different within the geoliterate group that chose either Map A and Map; the results indicated seven geoliterate perceived Map A as more credible and ten geoliterate perceived Map B as more credible. Therefore, **Hypothesis 3 is not supported**.

Table 5-18 Results of Q1 and Q2 in Experiment 2a

	Frequencies	
Map A as more credible (believable) (The Starbucks Coffee)	70	
Influence by the foreground data supplier	2	
Influence by the visual cues	68 (geoliterate = 7; non-geoliterate = 61)	
Map B as more credible (believable) (The University of Nottingham)	68	
Influence by the ‘foreground data supplier’	Yes	49 (geoliterate = 8; non-geoliterate = 41)
	No	19 (geoliterate = 2; non-geoliterate = 17)
Influence by visual cues	19	

5.9 Discussion

This experiment was to test three hypotheses of this research;

Hypothesis 1 is**:**

Visual cues have no significant influence in respondents' credibility assessment on map mashup applications

Hypothesis 2 is:

The metadata related to sources (i.e. map data supplier, map mashup producer) have significant influence in respondents' credibility assessment

Hypothesis 3 is:

The metadata related to sources have significant influenced within geoliterate respondents

Results indicate that most respondents used visual cues (design) to assess, judge and select the map to serve the experimental tasks. The selection and combination of 'colours scheme' and 'design look (amateur/professional)' on map features influenced respondents' judgement in the assessment. The survey indicates this by showing that about 84% of responses judged the map by reliance on the differences in visual cues in each experimental condition except in Condition 1, where both comparable maps used dull colour schemes. The visual cues appeared to be the main factor influencing respondents' judgement in Experiment 2. It makes a great impact on respondents' understanding of the displayed information. The findings in the previous experiment, Experiment One, confirmed the important role of visual cues (specifically colour scheme and design look) to respondents. Therefore, this finding **supports the hypothesis 1**. Although controlled were made in Experiment 2a to the colour scheme applied on both maps and a very slight dissimilarity was introduced on the symbols between maps, respondents made judgement based on their preferences and the clarity of the symbols design. Excerpt from open-ended questions (in Question 3) that mentioned the influence of 'symbol design' in their judgement as below;

'The only downside to Site A was the source of foreground data. The Site B although good lacked the bus sign and also had the coffee house sign board which was less impressive' (M, non-geoliterate)

'I much prefer the bus symbol on Site A, although I prefer the food/drink symbol on Site B' (F, non-geoliterate)

'Don't like the little burger and soft drink symbols. Very theme park'

(M , non-geoliterate)

‘It's fairer to display the food and drinks provision with a tea cup than with a burger and soft drink, as not all facilities have fast food (of that sort) available. Moreover, it is a more recognisable symbol for refreshments’

(M, non-geoliterate)

David and Jason (2008) argue that judgements based on visual cues would operate within the first few seconds when a respondent makes first time contact with the online medium; in this phase users tend to make judgements based at an intuitive level. Later on, users tend to rely on cognitive judgement when they proceed to scrutinize the contents in depth. At a cognitive level of judgement, users tend to measure more critical elements than visual cues such as the information details, accuracy and the authority of the information. This argument might support the findings of the present study where the factor of ‘usefulness’ tends to emerge as a factor that influences respondents’ judgement. The influence of this factor was as dominant as the influence of visual cues on judgement and indicates that respondents made critical judgements at an intuitive and a cognitive level in the experimental tasks. As argued by Metzger and Flanagin (2011, p.45) people will engage with information that they can access easily, and that they perceive as relevant (in this case ‘useful’) to their information seeking goals and credible.

Nevertheless, the metadata related factors of ‘data supplier’ was still not perceived as the critical metadata elements that were supposed to be the measured in respondents’ judgement when assessing map information. These critical elements were measured by half of the respondents, whilst the other half did not measure these elements in the assessment. From the results in the six experimental conditions, these elements were not ranked as the most important influence. The results of ‘spot the differences’ activities that conducted before respondents analyse and choose the map that they perceived credible for the given task supported this; there was no significant different within the respondents that spotted the differences of metadata (i.e. foreground data supplier) parameters with the proportion that measured (ranked) and not measured these metadata. In other words, respondents noticed the data supplier of both maps were different; however this did not necessarily influence them to assess credibility of information by using that metadata (i.e. data supplier and map producer).

Influence of visual cues were controlled to be minimal in Experiment 2a; the intention was to implicitly highlights the ‘foreground data supplier’ at the sidebar; the ratio of respondents that perceived either Map A (i.e. the map supplied by the Starbucks Coffee) and Map B (i.e. the map supplied by the University of Nottingham) as more credible was 50:50; this result supports the finding of previous Experiment 2; in this experiment, 72% of the total respondents that chose Map B had influenced by the ‘foreground data supplier’ that hold high reputation to supplied campus map compared to Map A. Therefore, **hypothesis 2 is not supported**. There were variations however in the order of priority and the differences in response were not significant between the first to fifth orders of importance. Additional analysis of the respondents' perceived credibility between the elements' variables in each experimental condition is discussed in Appendix B

The findings from both experiments indicate there were no significant differences between the scores from the geoliterate and the non-geoliterate groups. In Experiment 2, both groups respond more to visual cue elements, which were 'colour scheme' and 'overall design' but they gave a low response to the evaluated variables of metadata, which was 'data supplier' on the experimental tasks. Moreover, there were no significant differences between the ranking rated between the two groups on the evaluated metadata (specifically data supplier) where about half of responses within the group varied from the first ranked to the fifth ranked, whilst another half did not measure the variables. In Experiment 2a, half respondents in geoliterate group chose Map A as more credible due to the influence of visual cues; another half of geoliterate group chose Map B as more credible due to the foreground data supplier; there was no significant difference between the results. Hence, **hypothesis 3 is not supported.** Table 5.19 below presents the results of hypotheses of this experiment.

Table 5-19 The results of experiment hypotheses

No		The hypotheses statement in this experiment	Result
1.	Hypothesis 1	Visual cues have no significant influence in respondents' credibility assessment on map mashup applications	Not supported
2.	Hypothesis 2	The metadata related to sources (i.e. map data supplier, map mashup producer) have significant influence in respondents' credibility assessment	Not supported
3.	Hypothesis 3	The metadata related to sources have significant influenced within geoliterate respondents	Not supported

5.10 Conclusion

The contribution of this experiment to the overall thesis is that metadata related to data supplier is not the dominant factors being measured in respondents' assessment of the credibility of map information. Visual cues are the dominant factors that influence respondents' judgement, which operates on first time contact with the map until the stage of scrutinising the map. This is due to the important role of 'colour schemes' and 'design look' (amateur/professional) in respondents' comprehension of the information displayed on map. Assessment at a cognitive level might occur in the experimental tasks due to a high response to the factor of 'usefulness'. The metadata related to sources (specifically foreground data supplier) however were not sufficient to produce a high impact in respondents' judgement in analysing map information during the cognitive stage. Half of the respondents did not measure these critical factors when assessing map information although; 1) they noticed (spotted) these metadata and 2) the influence of visual cues (i.e. colour scheme, symbol design) was controlled to be minimal.

6 EXPERIMENT THREE

6.1 Introduction

This section describes the aim, hypothesis, methodology, results and analyses used in Experiment Three. This Experiment 3 was still to achieve the Objective 1:

‘To examine the influence of metadata related to sources (specifically the identity of map producer and map supplier) on respondents’ assessment of the credibility of map mashup information’

The aim of this experiment was to enhance the experimental approach used in Experiment Two and to confirm the results in different contexts. In Experiment Two the tasks given did not incorporate an interactive approach; the stimuli applied did not offer deep involvement context for respondents. Experiment Three therefore involves an interactive task whereby respondents are required to attempt a task that requires a degree of cognitive judgement. Experiment Three consists of one experimental task to examine the influence of metadata related to map data producer. Respondents are invited to make judgements on the basis of the author (map producer) of the mashup, as identified on the maps.

6.2 Hypotheses

Hypothesis 1 is**:**

Visual cues have no significant influence in respondents’ credibility assessment on map mashup applications

Hypothesis 2 is:

The metadata related to sources (i.e. map data supplier, map mashup producer) have significant influence in respondents’ credibility assessment

Hypothesis 3 is:

The metadata related to sources have significant influenced within geoliterate respondents

Hypothesis 4 (null) is:

There is no significant difference between the level of importance of the metadata related to sources between these two contexts - a prominence dependent (i.e. interpretation is based on what looks prominent) and a prominence independent setting (i.e. interpretation is based on a statement)

6.3 Design and Materials

This Experiment 3 was designed to evaluate the influence of metadata related to map data producer in active involvement with the task where respondents were required to use cognitive judgement before giving responses to the questionnaire. Table 6.1 presents the details of the experimental task. Two conditions were designed to test the experimental hypothesis. The two comparison maps, as shown in Figure 6.1 and 6.2 differ in terms of the value of controlled variables and visual appearance, including the colours of building features, the designs of road works and landslide event symbols. The number and the locations of road work points and landslide events are also different. Table 6.2 presents full descriptions of the condition differences between the two comparison maps. Figure 6.3 shows snapshots of possible routes selections (route 1, route 2 or route 3) that have to be decided by respondents in the task.

In the task, respondents were given a situation in which they had to pretend to be an Ambulance Incident Officer and had to find the safest route for the ambulance to evacuate trapped victims from inside a red-coloured building. In this setting, a respondent has to apply cognitive judgement before suggesting the safest route. They have to judge the two maps critically because the information presented on the two comparison maps is different.

The colour of building features as well as symbols for road work points and landslides differ. Map A was simulated to be designed by a high credibility organisation, ‘the University of Nottingham’ and was designed to look unappealing and not well designed by showing the colour scheme of building features and point symbol in black and not easily distinguishable. As stated by Brewer (2005, p.122) a well-designed scheme is usually represent data in a different unique hues. This was to test the influence of a poorly designed on the map produced by a high credibility organisation. Meanwhile Map B was simulated to be designed by low credibility individual, ‘Sarah Smith’. This map was designed to look well designed by the use of similar contrast of colours; in order to test the influence of ‘well design’ on a map produced by a low credibility individual.

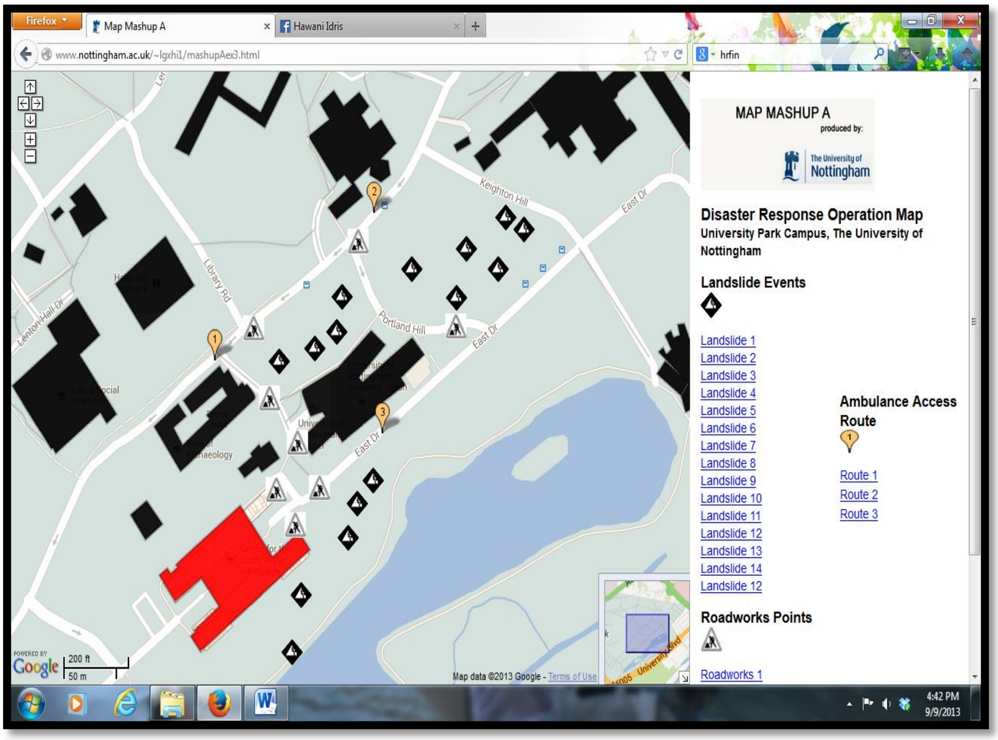


Figure 6-1 Snapshot of Map Mashup A

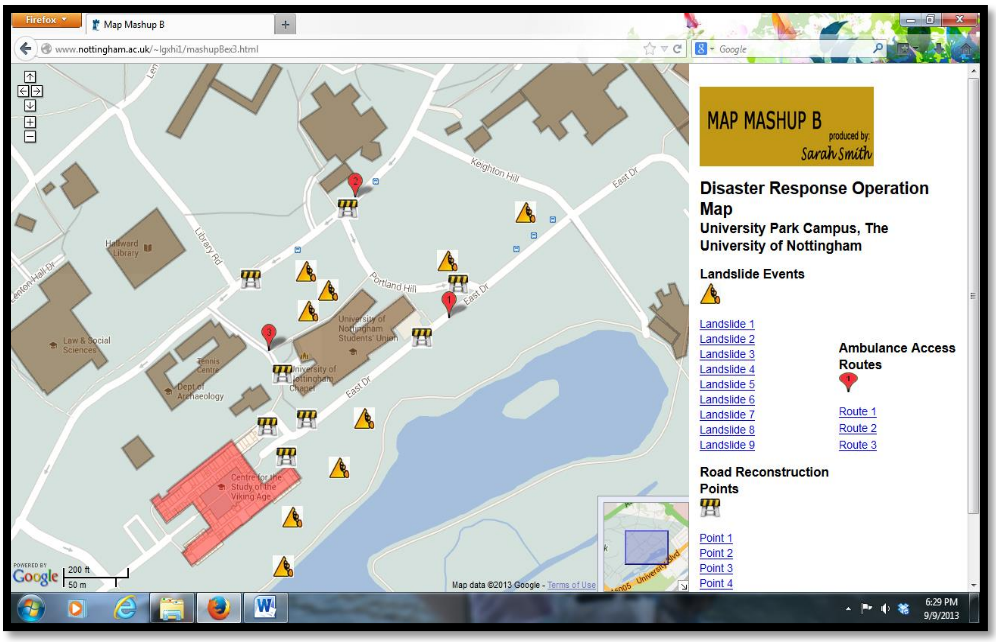
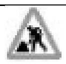


Figure 6-2 Snapshot of Map Mashup B

Table 6-1 Experimental Design

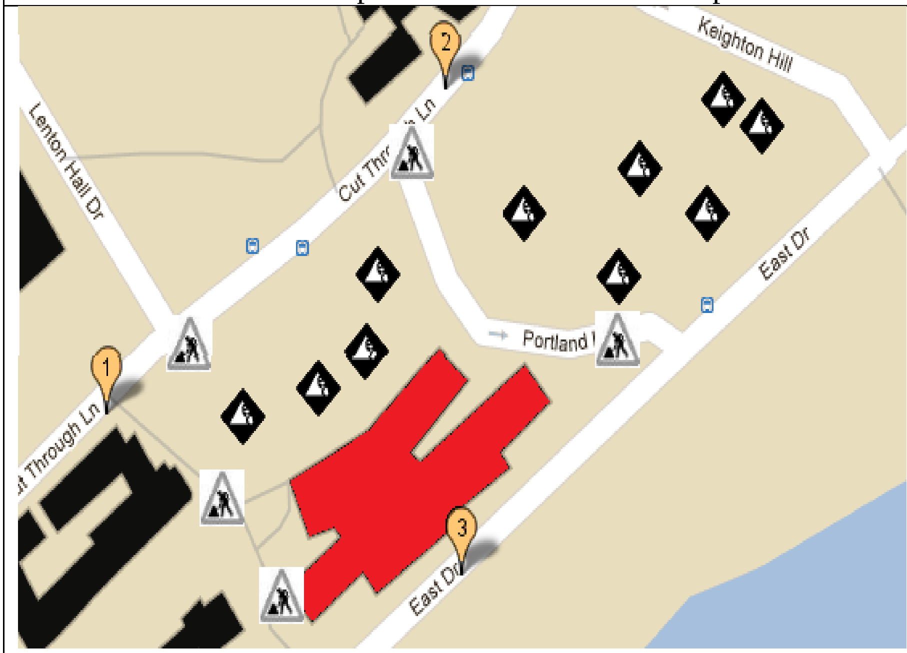
	Map A	Map B
Experimental assumption	High level of credibility	Low level of credibility
Manipulated variable: The producer of map mashup	The University of Nottingham (UoN)	Sarah Smith
Full descriptions (as in the task sheet)	Map Mashup A: Disaster Response Operation Map produced by the University of Nottingham	Map Mashup B: Disaster Response Operation Map produced by Sarah Smith
Experimental task	Situation 1: There are a few landslide events around the University Park Campus. Unfortunately a group of students are trapped in the Portland Building (this building is marked in red on the map). You are the Ambulance Incident Officer (AIO). You are required to assess the scene to determine an ambulance entry and exit route to evacuate the victims. The task: 1) Browse and evaluate the two interactive maps 2) Decide which map mash up you will use (either Map Mash up A or Map Mash up B) to determine the ambulance access route (either Route 1, Route 2 or Route 3) that is safe to be used for the entry to and exit from the Portland Building 3) Submit your answers using a questionnaire.	

Table 6-2 Condition differences between Map Mashup A and Map Mashup B

Experimental conditions	Map A	Map B
Building colours	Black	Brown
Roadwork points	8 points	8 points
Roadwork symbols		
Landslide points	12 points	9 points
Landslide symbols		
Route 1	Stop at Cut Through Lane, in front of Archeology building	Stop at East Drive, a few metres before juction to Portland Hill

Route 2	Stop at Cut Through Lane, in front of the Sir Clive Granger building	Stop at Cut Through Lane, in front of the Sir Clive Granger building
Route 3	Stop at East Drive, in front of the Portland building	Stop at the end route between Archeology and Admission building

Locations of possible safest routes in Map A



Locations of possible safest routes in Map B

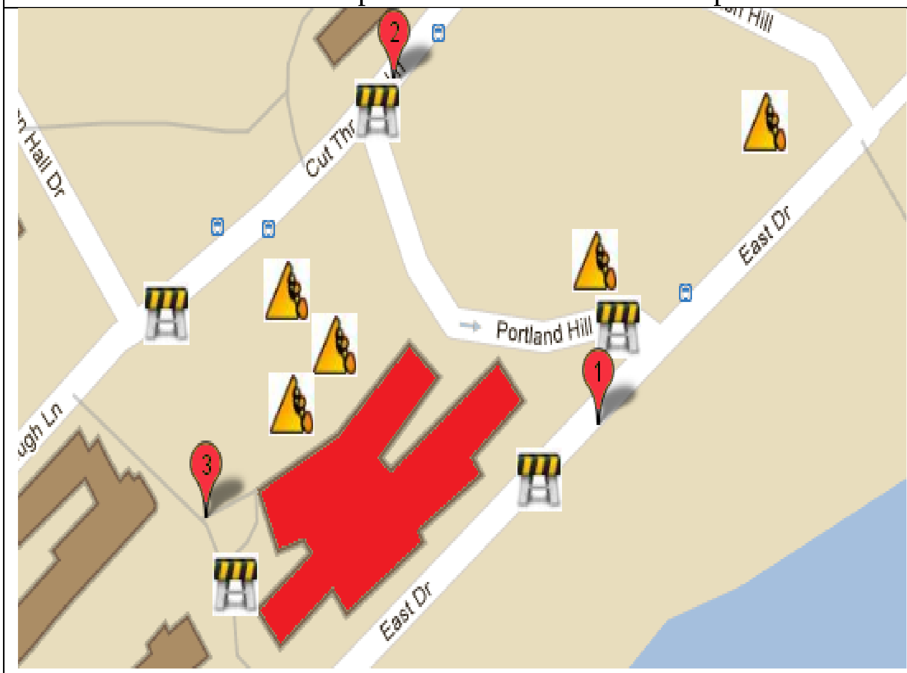


Figure 6-3 The possible of safest routes between the two maps

In previous experiments, Experiments One and Two, the data and information presented were similar on the two comparison maps; in those contexts, respondents tended to rely on visual cues and 'usefulness' when judging the map. Metadata variables related to 'map supplier' were least measured in these assessments. In Experiment Three, information on the location of the road works and landslide events were designed differently. In order to test the influence of the metadata, the 'producer' of a map mash up, was the assessed variable. This was to test whether the metadata factor would be a major influence on the judgement of the two maps when the information was contradictory or whether visual cues were still the major influence on respondents' judgement.

The maps were created using a simple tool to develop mashups: Google Earth tool was used to create the 3 layers in *.kml format- the layers for road works, landslide events and safest routes. Map Makers tools provided on the Google Map interface were used to create the building layers. Then Google Map APIs were used to mash up the 4 layers into one functional map. The links distributed to respondents used the university's domain. For example this link (http://www.nottingham.ac.uk/~lgxhi1/mashupAex3.html) displayed the mash up A and this link (http://www.nottingham.ac.uk/~lgxhi1/mashupAex3.html) displayed mash up B. The layout of the experimental based questionnaire was designed using Google Sites interface and the questionnaire was designed using Google Docs.

6.3.1 Participants

There were 133 respondents who completed the study. The samples selection is explained in Section 3.6. The respondents were drawn from the members of University of Nottingham (staff and students). The ages in the samples ranged from the category of 19 and below, to the category of 21 to 35 years old. Responses from three respondents at the age between 35 to 50 were excluded in this study in order to generalise the results to young adult web users' population. The average age in the sample was 22.3. There were more female (53.4%) than male (46.6%) respondents in the sample. A majority in the samples use websites every day (95.5%), on average 1 to 4 hours a day. The majority of respondents (96.2%) had experience of using interactive online maps supplied by commercial providers (e.g. Google Map, Bing Map, and Yahoo Map) and crowd source (e.g. OpenStreetMap) applications. There were 23.3% (average age 20 years old) of the respondents in the sample (31 respondents) drawn from the geoliterate group who have a background in attending geography, cartography, remote sensing, land surveying or geographic information science courses, whilst 76.7% in the sample (102 respondents) (average age 22 years old) were drawn from the non-geoliterate group who have background in attending other non-geospatial related courses, such as engineering, sciences, and social sciences.

The respondents involved in think-aloud session was selected by voluntary basis. There are six respondents where four of them were drawn from non-geoliterate group whereas the other two had background in cartographic and mapping. There are 3 males and 3 females in the samples. The average age of all respondents was 27 years old. The respondents were Malaysian nationality and alumni of the Universiti Teknologi Malaysia (UTM). Of the six respondents, three work as a researcher and another two work as a lecturer at UTM. One respondent was the undergraduate student. Of the sample within geo-literate group, one respondent work as lecturer and another respondent work as a researcher.

6.3.2 Questionnaire Design

The questionnaire was divided into two sections. Single item measures were used to measure the responses. The responses to the questions in the first section were dependent on the experimental context which, in turn, was dependent on responses to the main question. The main question was

'to analyse and make a judgement between two maps that a respondent would use (either Map Mash up A or Map Mash up B) to determine an ambulance access route (either Route 1, Route 2 or Route 3) that would be safe to use for entry to and exit from the Portland Building'

Based on the responses to this main question, a respondent was given a series of questions to assess the importance of metadata related to map data producer and visual cues in their assessments of map mashup information credibility in context and non-context dependent situations.

In the first section, respondents have to rate their responses based on the experimental task. The first section measures factors that have been identified as part of credibility elements in the literature – visual cues (i.e. colour scheme and symbol design) and 'identity of the map mashup producer'. These three elements were controlled in the experimental dataset where the influences to these elements were measured. The questionnaire was designed to measure the perceived importance of the metadata variable, specifically producer of the mashup as well as the perceived importance of the visual cues measured in in terms of 'colour scheme' and 'symbol design'. In this section, respondents have to rate the extent to which these factors influence their judgement in the experimental task.

The responses to the questions in the second section are independent of the experimental context. In this section, respondents rated their responses on the basis of their previous experience or knowledge. Respondents have to rate to what extent the importance of the factors on the list influence their assessment of the credibility of any online map. The parameters on the list consists metadata related to ‘data supplier’, ‘map producer’, ‘affiliation’ and ‘currency’. Table 6-3 presents the list of questions used in the questionnaires.

Table 6-3 The list of questions in the questionnaire and types of measurement

Q-ID	Items	Question	Measurement
Section One – context dependent			
Question 1	Main question	Please browse the two map mashups above and evaluate which of the two maps you perceived as having more credibility (more believable information) to assist you in determining an ambulance safe access route to evacuate the trapped victims in the Portland building	Binary
Question 2	Level of confidence on to the response in Q1	Please indicate your level of confidence on your answer in Q1	Interval 0% to 100%
Question 3 (a)	The importance of visual cues in respondents' perceived credibility to map mash up information	Please indicate how important the following criteria are in influencing you to choose the map in Q1 and reject the other map i. Visual attractiveness ii. Colour scheme iii. Symbol design	3-point rating scale 0 = do not know, 1 = not important, 2 = important
Question 3 (b)	The importance of metadata related to map producer in respondents' perceived credibility to map mashup information	Please indicate how important the following criterion are in influencing you to choose the map in Q1 and reject the other map i. Map producer (map author)	3-point rating scale 0 = do not know, 1 = not important, 2 = important
Question 4		I decided to suggest this route for the ambulance entry and exit route	Categorical (Route 1 or Route 2 or Route 3)
Section Two – context independent			
Question 5 (a)	The importance of metadata in respondents' perceived credibility to	Please indicate how important the following criteria are in influencing you in assessing the credibility of online maps	3-point rating scale 0 = do not know, 1 = not important, 2 = important

	map mash up information	i. The supplier of base (background) map. E.g. Google Map, OS OpenSpace map ii. The supplier (contributor) of foreground map data iii. Reputation of map producer iv. Website affiliation (e.g. the site that published a map has a connection/ relationship with the University of Nottingham) v. Map data last updated	

To understand the samples in the experiment and the differences between responses, a series of demographic questions were asked before respondents conducted the experimental tasks. Table 6-4 presents the series of demographic questions in the questionnaire.

Table 6-4 The demographic questions used in the questionnaire

	Items
Q1	Age
Q2	Gender
Q3	Professional or academic qualifications
Q4	Experience of using online maps from commercial providers and crowdsourcing map applications
Q5	Experience of using websites and average time spent
Q6	Member or non-member of the University of Nottingham, UK

6.3.3 Procedures

The procedure used in this experiment is depicted in the figure below. Further explanation has been described in Section 3.7.

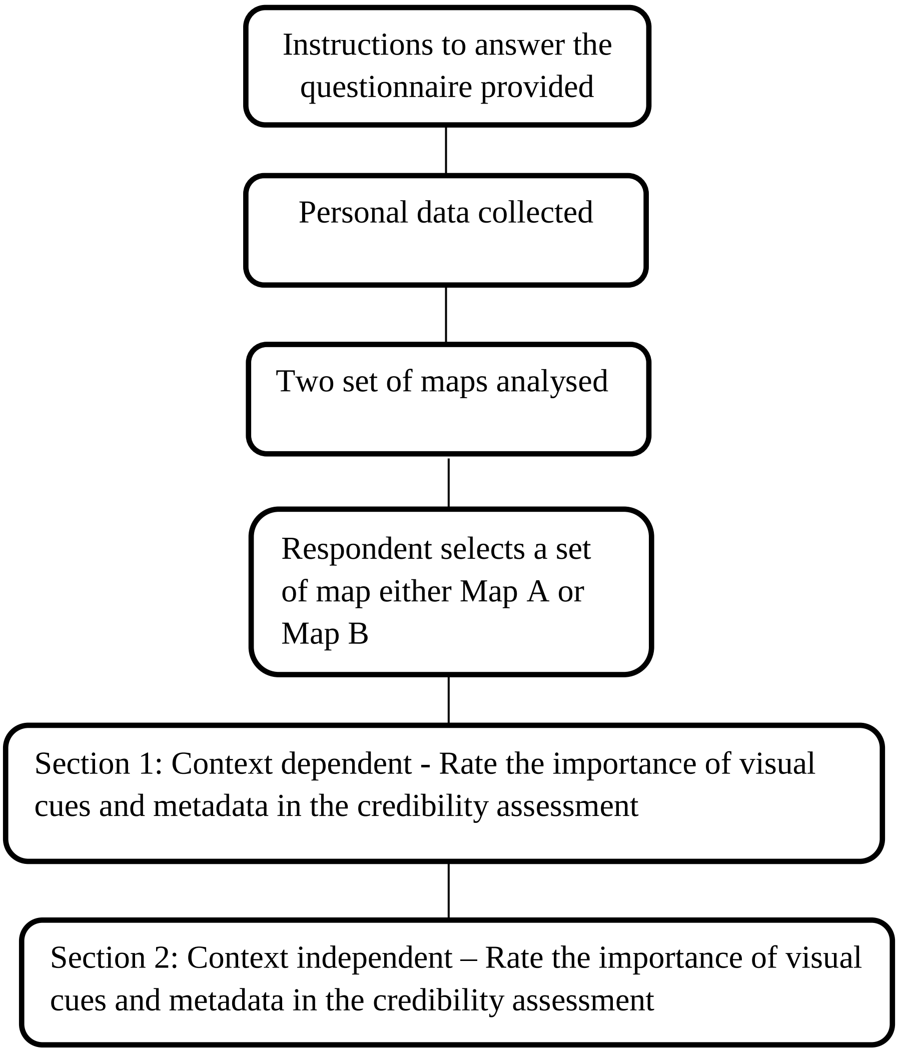


Figure 6-4 Procedures used in Experiment Three

6.3.4 Pre-test

This section will summarise data collection during this study. A pre-test study consisting of ten respondents was conducted in order to test the design of the map based questionnaire. Five respondents came from a geoliterate background, while the remaining five were from other backgrounds. The purpose of this pre-test was to check if there were any issues in the wording of the experimental task and the questionnaire. Since the map based questionnaire used single item measures to measure the responses, it was important to eliminate ambiguity in a question that could have different meanings for respondents. This study used concrete direct questions so that they would be understood unequivocally by each respondent. From the feedback, changes were made to the some part of the questionnaire wording. No comment was found related to the design of questionnaire; hence no changes related to design were made.

6.4 Results and Analysis

6.4.1 Data Analysis

Statistical analysis using SPSS v.16 was used to study the response data. The data collected from the questionnaire were measured as binary variables and ordinal variables, respectively. The responses data were not in continuous variables and transform into discrete variables using Terrel's Transformation technique. However, the data have been tested do not fit with normal distribution model. Hence, data analysis focused on selecting appropriate non-parametric statistical tests. The Chi-square test and Mann Whitney U test were used to analyse the data. Non parametric tests are free from the restrictive assumption that the data needs to be normally distributed; The Chi-square test has a specific assumption of the data before it can be modelled correctly. The assumptions in the Chi-square test are that the data is independent and each entity contributes to only one cell in the contingency table and the expected frequencies should be greater than 5. These assumptions were not violated with regard to the experimental data modelled by this test.

As mentioned in Section 3.4, reliability is a statistical measure of how reproducible the survey instrument's data area (Litwin, 1995). Table 6-5 present the results from internal consistency test to check the reliability of items in the instrument that measure the influence of visual cues.

Table 6-5 Corrected total item correlation and Cronbach's alpha for the items that measure the influence of visual cues

	Corrected item total correlation	Cronbach's alpha if item deleted
Colour coding	0.383	0.259
Symbol design convention	0.387	0.247
Clarity of symbol	0.230	0.536

From Table 6-5, the Cronbach's alpha for these three items was 0.49 which indicate a poor reliability instrument (scale) to measure the same issue (the influence of visual cues). The value of Cronbach's alpha reflects the homogeneity of the items in a scale, complementing each other in their measurement of different aspects of the same variable or quality (Litwin, 1995:24). However, from intraobserver test, which is another approach to test reliability, that reported a correlation coefficient between respondents, indicated intraclass correlation as r = 0.493 p < 0.001. According to Field (2009, p.173), r = +- 0.1 represent small effect, r = +-0.3 is a medium effect and r = +-0.5 is a large effect. Therefore, this result indicates a large effect of correlations between the responses of items from different respondents. Although the intraclass correlations within the construct was large, but the results of Cronbach's alpha that reflects the homogeneity of the scales to measure the visual cues construct was poor. Therefore, other research who want to replicate the instrument (i.e. questions) used in this experiment should cite it with caution due to this limitation.

Descriptive analyses were conducted on the four questionnaire items for which the responses were dependent on the experimental context. Of the three items were the items that measure the construct of the influence of visual cues when respondents making judgement related to credibility. Table 6.6 presents the correlation between these items.

Table 6-6 The correlation between items that measure the influence of visual cues

Construct : the correlation between items that measure the influence of visual cues			
	Item 1:The coding of colours on map	Item 2:The clarity of map symbol	Item 3:The map symbol design convention
The coding of colours on map	r = 1.0	r = 0.259 p < 0.003	r = 0.335 p < 0.001
The map symbol design convention	r = 0.276 p < 0.001	r = 0.274 P < 0.002	r = 1.0

From the table above, the correlation coefficient (*r* value) which resulted from Spearman's correlation coefficient demonstrated medium effect (+- 0.3). As stated by (Field, 2009:173), values of +- 0.1 (small effect), +-0.3 (medium effect) and +-0.5 (large effect) are the commonly used measure of the size of an effect of correlation.

Table 6-7 presents the results of internal consistency test to check the reliability of the items in the scale to measure the influence of metadata. From the table below, the Cronbach's alpha for the five items was 0.561 that indicates a poor reliability of scale. Nevertheless the results from intra class correlation indicated $r = 0.561$, $p < 0.001$. This indicates a large effect of correlation between the responses from different respondents.

Table 6-7 Corrected total item correlation and Cronbach's alpha for the items that measure the influence of metadata

	Corrected item total correlation	Cronbach's alpha if item deleted
Hosted website affiliation	0.376	0.471
Foreground data supplier	0.430	0.441
Basemap data supplier	0.353	0.490
Map currency	0.123	0.590
Map data producer	0.332	0.508

6.4.2 Results: Exploration of the map that respondents perceived/ chose as having the most credibility to assist in determining an ambulance safe access route

Frequency analysis was conducted to analyse the main question: 'please evaluate which of the two maps (either Map A or Map B) you perceived as having more credibility (more believable information) to assist you in determining an ambulance safe access route to evacuate the trapped victims in the Portland Building'. This question yields two binary variables, which are either Map A or Map B. Chart 6.1 presents the results of this question.

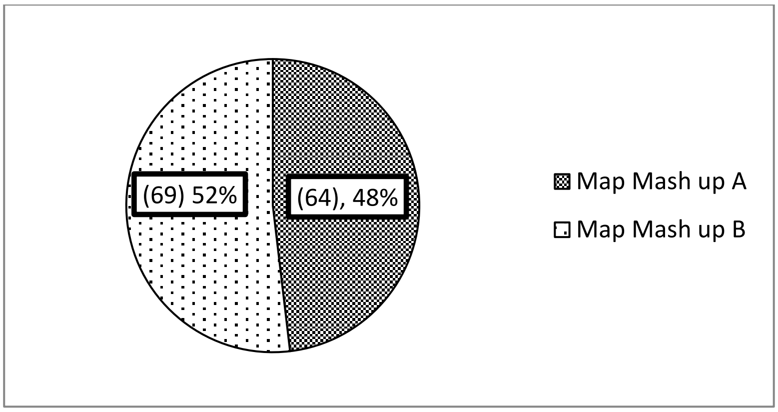


Chart 6-1 The frequencies of responses to the main question (Q1)

From the results, 48% of respondents perceived Mashup A as having more credibility. In contrast, 52% of respondents perceived Mashup B as having more credibility. Mashup A was the map simulated as being produced by the 'University of Nottingham', meanwhile Mashup B was simulated as being produced by 'Sarah Smith'. The responses between the two mash ups, however, were not significantly different in the Chi-squared test X^2 (1, n = 133) = 0.188, p > 0.729).

Analysis of the routes evaluated by respondents for ambulance access showed that most of them suggested Route 3 in mashup A and mashup B as the safest route to choose in the experimental task. Chart 6.2 presents the routes suggested by respondents in percentages. Chart 6.3 presents the distribution of respondents' level of confidence in their responses to the main question (Q1). From the chart below, more than 50% of respondents determined Route 3 as the safest for ambulance access. This was in line with the answer that had been pre-determined when designing the task, where the Route 3 was calculated to be the correct route suggested by respondents if they critically analysed the mashup.

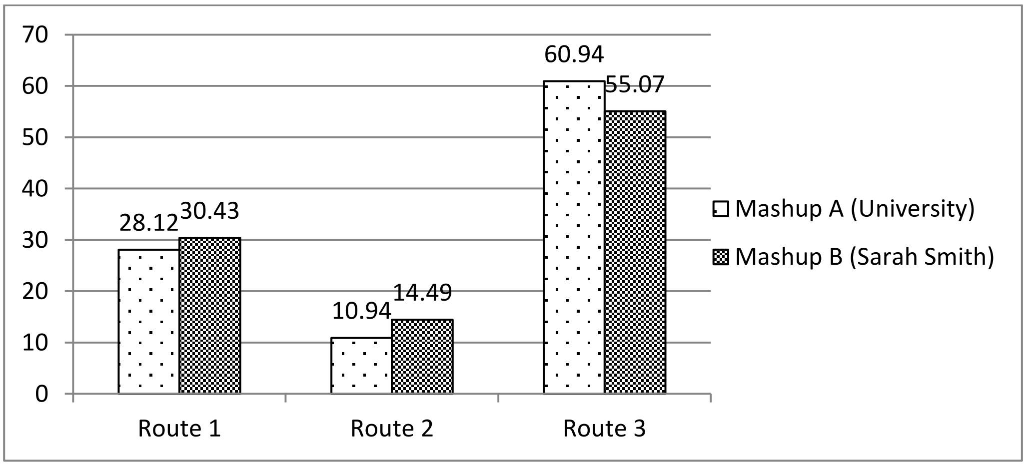


Chart 6-2 Routes suggested by respondents (in percentage)

Descriptive analysis of respondents' level of confidence in their responses to the main question (Q1) indicate the mean (68.8) and standard deviation (SD = 16.74); the mode and median at 70 per cent, respectively. This indicates that, on average, respondents have moderately high confidence in their responses to determine the most believable mashup and suggest the safest route for ambulance access.

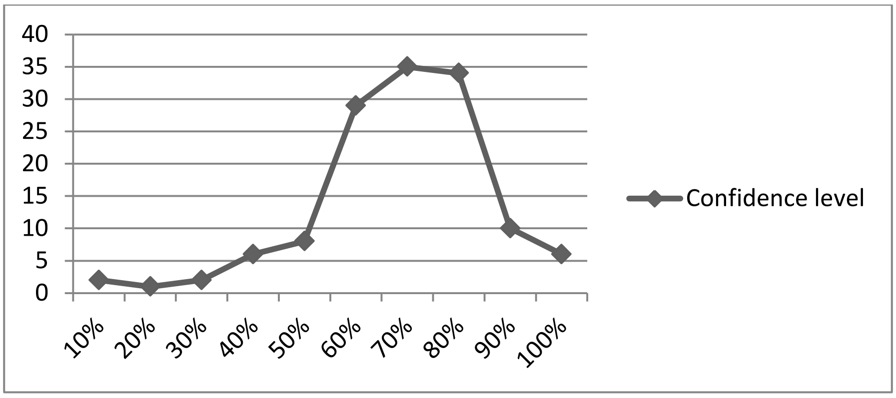


Chart 6-3 Distribution of respondents' level of confidence in their responses to the main question

The results from the think-aloud observation and assessment indicated of six respondents participated in this test, four of them chose the Map A. Other two respondents chose the Map B.

6.4.3 Results: The perceived importance of visual cues and metadata (prominence dependent setting)

Hypothesis 1 is**:**

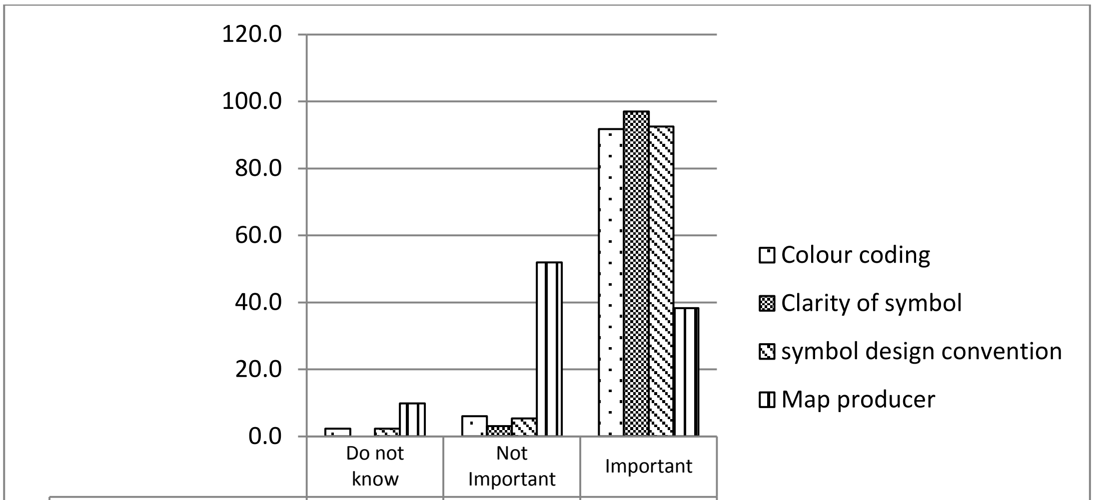
Visual cues have no significant influence in respondents' credibility assessment on map mashup applications

Hypothesis 2 is:

The metadata related to sources (i.e. map data supplier, map mashup producer) have significant influence in respondents' credibility assessment

From the Chart 6.4, below, about 90% of respondents perceived the importance of the items that measure the influence of visual cues, specifically 'colour coding', 'clarity of symbol' and 'symbol design convention' as influencing their judgement in the experimental task. On other hand, 38.3% respondents perceived the element of map producer to be important.

In contrast, 51.9% respondents perceived the ‘map producer’ element to be unimportant in influencing their judgement pertaining to map information credibility. Likewise, a high number of responses rated the importance of the ‘map producer’ element at the undecided point (9.8%) compared to other manipulated elements. The response differences were statistically significant in the Chi-squared test X^2 (1, n = 16) = 6.25, $p < 0.05$).



	Do not know	Not Important	Important
Colour coding	2.3	6.0	91.7
Clarity of symbol	0.0	3.0	97.0
symbol design convention	2.3	5.3	92.5
Map producer	9.8	51.9	38.3

Chart 6-4 The number of responses (as a percentage) to the items dependent on experimental context

Hypothesis 1 is not supported. The results from the think-aloud observation test confirmed this. The basis of respondents’ judgements on choosing the map was due to its visual cues, specifically the ‘colour scheme’ and the symbols used. Excerpt from the audio transcripts (translated from Malay to English and the labels (ids) after each excerpt was not represent the real identities of respondents) are indicates as below;

‘I chose this map because the presentation was not too crowded and the symbols used were easy to understand’ (Am, geoliterate)

‘I chose this map because the colour [red, black] and presention looks serious and suit for this critical situation (context)… the other map (Map B) is more suitable for the use of public user for navigation’ (Fad, geoliterate)

‘I chose this map because it looks attractive. The other map used black. It blocks the building label’ (Fik, non geoliterate)

‘I chose this map because it used the symbol road block instead of road construction to represent the no access route’ (A, non geoliterate)

However, the basis of judgement of the other two respondents was the detail of the information. They perceived a map would be credible if it presented more data compared to another one. They perceived more data as more detail. The transcript excerpt as below;

'This map is more credible; it has the most important element which is the detail of information. This map displayed more landslide points compared to another one' (R, non-geoliterate)

'This map is more credible because it displayed more information (i.e. presentation more landslide points); this information is important to indicate the best access route for an ambulance' (Fa, geoliterate)

In this experimental context, however **hypothesis 2 is not supported**. Of the sample, the respondents that perceived the important of 'map producer' element were only 40 percent (38.3%). In contrast, 52 percent (51.9%) of respondents perceived this element as not important.

From the think-aloud experiment, of the sample, two respondents mentioned 'map producer' element as not important in their decision; this element was not considered when they made the judgement to choose the credible map. For example;

'Map producer element has no influence in my judgement. I just focused on the events [task] and on the centred of map [map producer element was displayed at the sidebar, the top corner of map]' (Am, geoliterate)

'I do not look at the map producer [which displayed at sidebar, at the top corner of map]. The important is about the clarity of information. I just look the powered by Google which displayed at the bottom of the map. Both maps displayed this. It means both maps produced by Google' (Fik, non-geoliterate)

From the think-aloud experiment, four of respondents mentioned the element of 'map producer' has a certain extent of influence to their judgement. For examples;

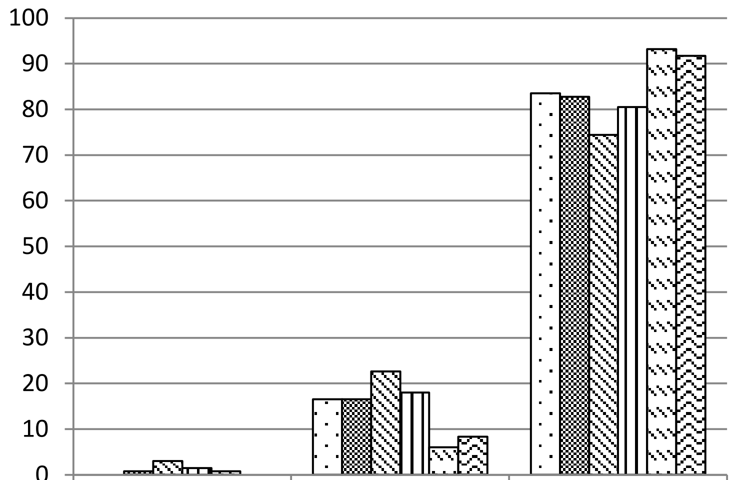
'the map A is more believable because it is produced by the university ..and the map B is less believable because it produced by nobody, we do not know the background of the author. She might make up the data. So it is less believable. Map A was produced by an authorised source; we know its [reputation] and we can believe them' (Fa, non-geoliterate)

'the map from the university is I believe more reliable. The map B is from private individual, isn't? it is better to have a map produced by an organisation, rather than an individual' (A, non-geoliterate)

6.4.4 Results: The perceived importance of metadata variable (prominence independent setting)

Hypothesis 2 is:

The metadata related to sources (i.e. map data supplier, map mashup producer) have significant influence in respondents' credibility assessment



	Do not know	Not important	Important
base map data supplier	0	16.5	83.5
foreground data supplier	0.8	16.5	82.7
map producer	3	22.6	74.4
affiliation	1.5	18	80.5
currency	0.8	6	93.2
design look	0	8.3	91.7

Chart 6-5 The number of responses (as a percentage) to the questionnaire items independent of experimental context

Descriptive analysis was conducted on the six items that were independent of the experimental context. From the results in Chart 6-5, more than 80% of responses to four items indicate the perceived importance of the metadata elements in their credibility assessment; responses indicate the metadata related elements, which were as below;

1) The 'supplier of base map data' (83.5%),
2) The 'supplier of foreground data' (82.7%)
3) The 'website's affiliation' (80.5%),
4) 'Currency' (93.2%)
5) The 'design look' (professional/amateur) (91.7%)
6) The 'map producer' (74.4%)

A high number of respondents (93%) perceived the importance of assessing the' currency' of map; this high rate of responses was consistent with the high responses to the importance of 'design look' (92%), which is a professional or amateurish look to the map. The perceived importance of the 'map producer', however, was rated lower at 74.4% of responses.

In this general context, which interpretation was not based on what looks prominence however, **hypothesis 2 is supported**. This finding indicates that in a context independent setting where respondents made interpretation (judgment) based on a statement, not basis on what looks prominent, a high number of responses rated the importance of critical elements of metadata in their assessment. These results did not seem to be consistent with the results found in this study, in the setting that is dependent on what looks prominence in the experimental context.

6.4.5 Results: The differences between the context of prominence dependent and prominence independent settings

Hypothesis 4: There is no significant difference between the level of importance of the metadata related to sources between these two contexts - a prominence dependent (i.e. interpretation is based on what looks prominent) and a prominence independent setting (i.e. interpretation is based on a statement)

Analysis of the results presented in the previous Chart 6-4 and Chart 6-5 yield inconsistent findings. In Chart 6-4, the degree of importance of the metadata related element, which in this context was the 'map producer', was inconsistent with the results for the degree of importance of the same element in Chart 6-5. **The difference in the ways of user interpreted the contexts was the main cause of this inconsistency.** In Chart 6-4 respondents have to rate their responses based on what looks prominent in the experimental context. This setting was in contrast with the latter context in which respondents have to rate their responses either 1) on basis of their experience or knowledge with other online maps or 2) respond in accordance to social norms, or giving answers that they believed would be socially acceptable answer, rather than what they actually did. The differences in response in the category of importance was 36.1% (74.4% – 38.3%) and was statistically significant in the Chi-squared test X^2 (1, n = 112) = 11.571, $p < 0.01$).

Other metadata related elements, such as 'foreground data supplier' and the 'website's affiliation' indicate a high number of responses in the category of importance. This was inconsistent with the findings in the context dependent experiments in the previous experiment shown in Tables 5.11 and Table 30 (see Appendix B). In the context dependent setting (i.e. interpretation was based on what looks prominent), the numbers of responses that measured and did not measure (assessed) these two elements in the credibility assessment were in the ratio 50:50. About 50% of respondents measured this element whilst the other 50% of respondents did not measure the element. Meanwhile in the context independent setting ((i.e. interpretation was based on a statement), the number of responses that measured these elements in terms of perceived importance for map information credibility was higher and response rates were about 80%.

In the experimental context (i.e. interpretation was based on what looks prominent)**, hypothesis 2 is not supported**. The way users interpreted the importance of the credibility elements, which was either basis on what looks prominent or interpreted based on a statement was the main cause of this inconsistency. From the results on Chart 6.5, more than 80% of responses perceived the importance of these metadata elements in credibility assessment.

However, when these elements were tested in experimental contexts (i.e. map producer and the foreground data supplier), they were perceived important by only half of the respondents. The results from Table 5.11, Table 5.12 and Chart 6.4 indicate these elements had less influence on their judgement in these experimental contexts.

One reason for these findings was that they might be aware of the importance of these metadata when evaluating the believability of maps. But they did not notice these elements when assessing credibility during the experiments. Findings from the think-aloud experiment confirmed this. Excerpts are given below;

Observer: How about the influence of the map producer?

Respondent: mmmm.. [thinking] map producer?

Observer: At the top sidebar

Respondent: Ok…but it did not influence me

Observer: Either Sarah Smith or the University of Nottingham [parameters of map producer], which one is more credible?

Respondent: Oooh if had looked at this element…it might have influenced me…the University is more credible….more believable

(Fad, geoliterate)

Observer: Was the map producer important and did it influence your decision?

Respondent: Err…actually I did not look at who produced the map

Observer: [observation] respondent then quickly searched to find where The map producer element was on the map

Respondent: Yes, Map A is more believable because it is produced by the university. Map B is less believable because it produced by nobody, we do not know the background of the author. She might make up the data. So it is less believable.

Map A was produced by an authorised source, we know its [reputation] and we can believe [the source]

(Fa, non-geoliterate)

Observer: Was there any influence from the map producer?

Respondent: 'Oooh, I just realised this element. Yes indeed. Map A was produced by the university. We consider it's from an authorised source compared to Map B that was produced by an unknown individual'.

(Ra, non-geoliterate)

Observer: So how you want to evaluate this map, correct or incorrect?

Respondent: mmm…since it was from Google. So I trust the map because it was created by Google.

Observer: Map A was produced by the university and Map B was produced by Sarah Smith. Did these elements of map producer at the top side bar influence your decision?

[observation] respondent did not notice the element of map producer at the top at the sidebar.

Respondent: I think the map produced by the university is more credible. Map B was produced by a private individual isn't? I think it is better to have a map produced by one organisation compared to an individual.

Observer: So, did it influence you?

Respondent: Yes, I chose Map A because the data are more detailed, the symbols used, and because of the producer of the map. In Map B, they did not mention who Sarah Smith was.

(A, non-geoliterate)

Therefore, **Hypothesis 4 is not supported.**

6.4.6 Results: Analysis of non-geoliterate and geoliterate groups -the perceived importance of the 'map producer' variable in a prominence dependent setting

Hypothesis 3 is:

The metadata related to sources have significant influenced within geoliterate respondents

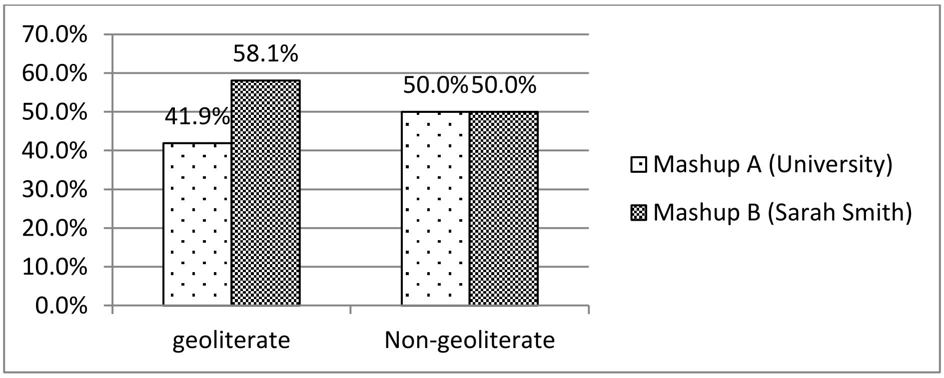


Chart 6-6 The perceived importance of map producer between groups

From the Chart 6.6 above, there were more respondents (58.1%) from the geoliterate group perceived Mashup B, the map produced by ‘Sarah Smith’, as having more credibility than the Mashup A (41.9%), the map produced by the ‘University of Nottingham’. The response differences however were not statistically significant (X^2 (1, n = 100) = 2.56, p < 0.5). Meanwhile, within non-geoliterate group, the number of responses that chose either Mashup A or Mashup B as more credible was 50:50. There were no significant differences in the scores between the geoliterate and non-geoliterate group in selecting the map they perceived as having more credibility (X^2 (1, n = 92) = 0.696, p < 0.5).

Chart 6.7 and Chart 6.8 below present the number of responses (as a percentage) between the geoliterate and non-geoliterate groups that perceived the assessed elements on the map as important and not important, respectively. From the chart below, more than 70% of responses perceived the visual cue elements (i.e. colour scheme, symbol design and visual attractiveness) as important in influencing their judgement in the experimental task. The number of responses that perceived the element of ‘map producer’ as important in their judgement however was less than 50% in the two groups (37% and 42%). Meanwhile, the responses in the two groups that perceived this element as not important in influencing their judgement was comparable, with both groups showing the number of responses to be around 52%.

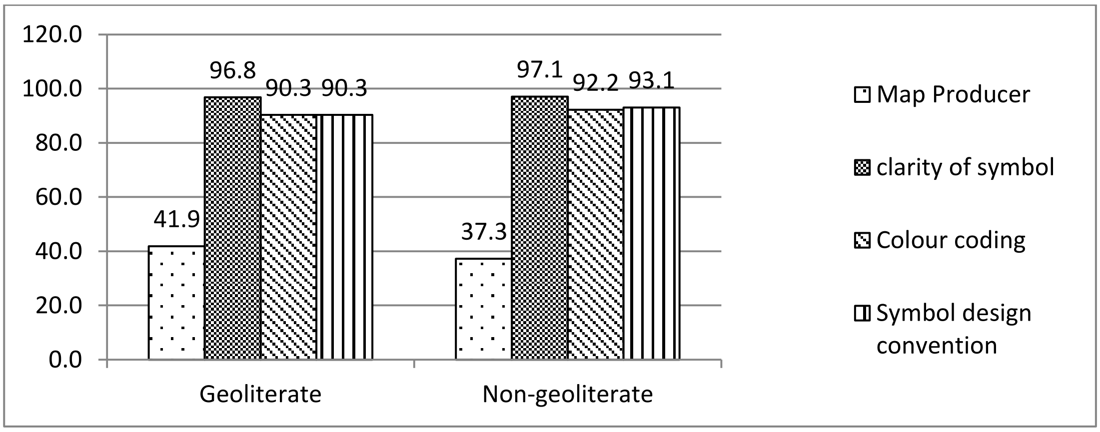


Chart 6-7 The number of responses (as a percentage) in the category of perceived importance for each group

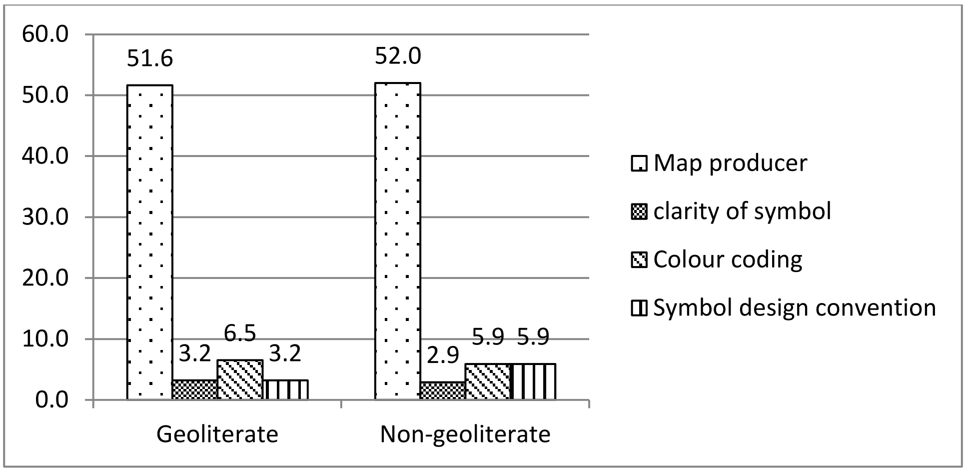


Chart 6-8 The number of responses (as a percentage) in the category of perceived unimportance for each group

The findings from the think-aloud experiment, confirmed this. Of the two geoliterate respondents, only one of them perceived the importance of the 'map producer' element. The excerpt from the transcripts as below;

'Ooh, if I had looked at this element [Sarah Smith versus University of Nottinham], it might have influenced me. The map came from [produced by] the University…I believe…it looks more believable' (Fad, geoliterate)

Whereas another geoliterate respondent respond as below;

'I was not influenced by that element [the producer of map]. I just focused on the events and the symbols used' (Am, geoliterate)

Hypothesis 3 is not supported. Some of the geoliterate respondents were influenced by the element of 'map producer' and used this element to support their decisions; some of them did not have felt any influence from this element.

6.4.7 Results: Analysis of non-geoliterate and geoliterate groups -the perceived importance of metadata in a prominence independent setting

Hypothesis 3 is:

The metadata related to sources have significant influenced within geoliterate respondents

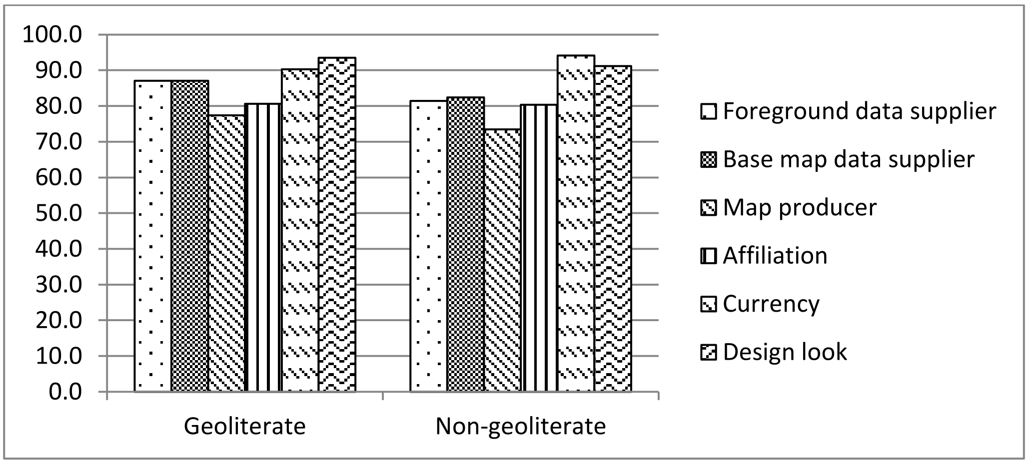


Chart 6-9 The number of responses for each group (as a percentage) in the category of perceived importance of metadata variables in a non-context, independent setting

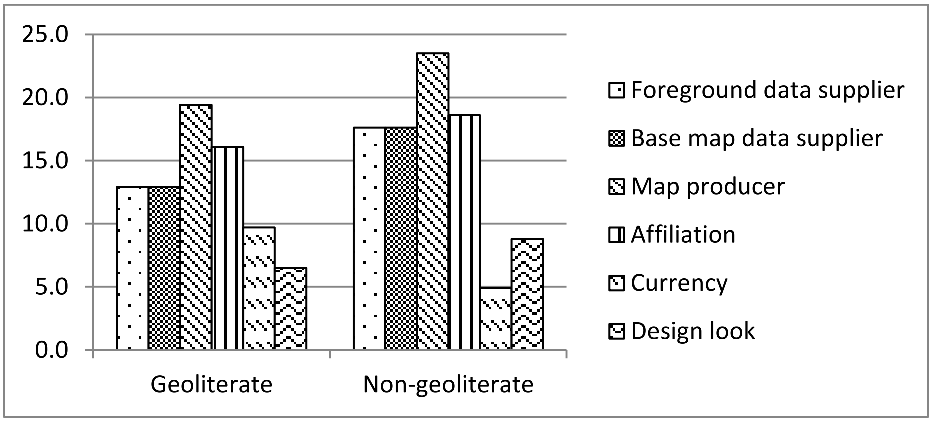


Chart 6-10 The number of responses in each group (as a percentage) in the category of perceived unimportance of metadata variables in a non-context, independent setting

From the Chart 6.9, both groups perceived the assessed metadata variables (specifically foreground and base map data supplier, website's affiliation and currency) as important when judging the credibility of an online map. The responses to these elements were more than 80%; except for the responses for the perceived importance of the 'map producer' element, which was about 77% and 74% in the geoliterate and non-geoliterate groups, respectively. The response to the 'currency' element was higher than other metadata elements, with more than 90% perceiving this element as important. This number of responses was high and in line with the number of responses to the importance of visual cues in judging the credibility of online map information.

Likewise, from the Chart 6.10, a low number of responses in both groups perceived these metadata variables as unimportant (approximately below than 25%). There were no significant differences between the responses from either group- geoliterate and non-geoliterate - on the perceived importance of the metadata variables when judging the credibility of online map information.

Hypothesis 3 is supported. This might indicates, in general context, geoliterate users perceived the 'foreground' and 'background data supplier', 'website's affiliation', 'currency' and 'map data producer' elements would have influenced in their credibility assessments. However in certain contexts (for example in the experimental context), some of them did not influenced by these elements because they did not notice. For example, from the think-aloud study, two geoliterate respondents mentioned this;

'In the context of navigation, I do not care who produced the map…..some users might have influence by this [the producer of map was the University of Nottingham], but for me, I have to use the application first. Just like the TomTom application. At first, the brand TomTom was nothing. Then, after try the application, people slowly acknowledge it. The name of producer is not important' (Fad, geoliterate)

'I have not influenced by the map producer. If the map is produced by a mapping authority such as NASA or survey mapping department, then it might have influenced me. If it is produced by a university, it is only an academia. Not an authoritative source to produce a map' (Am, geoliterate)

6.4.8 Analysis of structured interviews (qualitative data)

Three respondents from the sample were selected to participate in a short interview to obtain further explanation of their decisions after conducted the online surveys in the University of Nottingham in October 2010. Analysis from the interviews indicates that visual cues applied to the mashup, comprising 'colour scheme' and 'symbol design', became the basis of respondents' judgement for selecting or rejecting the map. Excerpts from the responses are shown below:

Interviewer: Why did you choose the map?

Respondent: 'The map was more colourful. There was a small route that showed clearly the way to access the building'

(B, non-geoliterate, age 25)

Respondent: 'Symbol colour was more meaningful because the black indicates danger. I do not like the colour to be similar to that of the building. The style of the symbols was more meaningful; the image on the road works symbol indicates clearly there was an obstruction on the road. The rejected map was too complex, too colourful and so many colours

(A, geoliterate, age 28)

Respondent: 'The roadblock symbols were clearer and more meaningful. I was confused by the landslide symbol used in mashup A. My decision was based on the symbols

(B, non-geoliterate, age 30)

In addition, interviewees were asked about the influence of the 'map producer' element in their judgement. The two interviewees who came from a non-geoliterate background and chose Map B indicated that this element had no influence on their decisions. One interviewee said she did not notice the element on the sidebar, whilst another interviewee said this element was not important to his judgement. Meanwhile, the interviewee who came from a geoliterate background chose Map A due to the influence of map producer on the map. An extract from the responses is shown below;

Interviewer: How important is the influence of map producer in your judgement?

Respondent: 'I **did not look** at the producer on the sidebar'

(B, non-geoliterate, age 25)

Respondent: 'The two mashups showed different information so I considered the 'supplier' of the map, 80% because of trust in the brand. At the beginning I did not notice that element. Users will focus their attention on the centre of the map, not on the side bar; users will refer to the suppliers of the base map shown on the top of the map, which were Google and Tele Atlas. I suggest changing the placement of that element to the top of the map, not on the sidebar. I will be more confident if the data is updated'

(A, geoliterate, age 28)

Respondent: 'Who produced the map **is not important**. What is more important is that the map is clear and easy to understand'

(B, non-geoliterate, age 30)

6.5 Discussion

Results of the survey indicate no significant differences between the responses who perceived 'Mashup A (produced by the University of Nottingham)' or 'Mashup B (produced by Sarah Smith)' as more credible or having more believable information. The number of responses choosing mashup A and mashup B were in the ratio of 50:50. Analysis of the correctness of routes suggested by respondents for the ambulance access shows that about half of respondents suggest Route 3; Route 3 had been predetermined as the correct route to suggest if they analyse the mashup critically. On other hand, respondents were quite confident in their responses to the mashup that had more credibility where, on average, their level of confidence was about 70%.

In the context dependent task (i.e. interpretation based on what looks prominent), the assessed visual cue elements of 'colour coding', 'clarity of symbol' and 'symbol design convention' had a high number of responses (about 90%) for perceived importance in influencing respondents' judgement of map information credibility. Hence, the **hypothesis 1 is not supported**. From the think-aloud protocols, the basis of judgement of two respondents was the details of information. They perceived a map will be credible if its present more data compared to the other map. They perceived the more data as more details. The transcript extraction as below;

'This map is more credible because the more important thing is because of the details of information. This map displayed more landslide points compared to the other one' (R, non-geoliterate)

'This map is more credible because it displayed more information (i.e. presentation more landslide points)…this information are important to support the best access way for ambulance' (Fa, geoliterate)

The question that arises here is how to actually determine which map is the more credible, if the data presented contradict each other; additional data does not always indicates that the map is more current; there might be a possibility of a map author/creator to have hidden motive as a map could be used as a communication medium for propaganda and rhetoric intentions (Muehlenhaus, 2012; Monmonier, 1999). But this was not the case; a few excerpts from the experiment indicate that when a map presented more data than the other map, they perceived the map was showing more current data; they did not perceived the map as showing contradictory data.

Nevertheless, the metadata related variable of map producer, assessed in the survey, was perceived to be a less important element, with only 38% of respondents perceiving this element as important and 52% of respondents perceiving this element as unimportant to their judgement. See Appendix C for distributions of responses. Therefore, **hypothesis 2 is not supported**.

This finding however was inconsistent with the results found in the non-context, independent experiment where respondents made judgement (interpretation) based on a statement. Metadata related elements, which in this case were the importance of the 'website's affiliation' and 'the foreground' and 'background map data supplier', were perceived as important in 80% of responses. This is inconsistent with the findings in the previous experiment, Experiment 2, where the ratio that perceived the importance of these two metadata elements in their assessment of credibility was about 50%. The importance of the 'currency' of map data was perceived higher than that, at about 93% and in line with the responses for the importance of the 'design look (amateur/professional)' of the map (92%). Nevertheless, the importance of 'map producer' was lower than other critical elements, at about 75%. However, this response was still higher than the responses in the context dependent experiment (where interpretation was based on what looks prominent), where the importance of the 'map producer' element was only rated by 38%. Hence, the **hypothesis 4 is not supported.**

One explanation for this inconsistency might be the differences in the ways respondents made interpretation (judgement). In the context dependent experiment, the applied approach involved judgements about importance based on prominent elements that a respondent notices. Meanwhile the non-context dependent experiment involved judgements by interpreting statements given in the questionnaire. Credibility assessment via a prominence based approach is supported by the theory of Prominence-Interpretation (Fogg, 2003). This theory posits users' assessment of information credibility occurs in two stages;

1) users would notice element(s) on the map that look prominent;
2) users will make judgement about information credibility by interpreting the element.

During assessment of online information credibility, users will judge on the basis of what they notice. If they do not notice an element, no judgement will be made. In the context dependent setting, respondents might judge on the basis of what they see and notice when assessing the map. In a non-context, independent setting, respondents might judge by interpreting given statements, with their responses likely to be influenced by previous experiences, culture, skills, knowledge (Fogg et al., 2002, p.85) when dealing with a map or their responses might be based on what they consider to be socially acceptable answers.

This issue is raised in a study by Fogg et al. (2003); the findings from that survey, which applied the prominence approach, contradicted the findings from other research (for example see Princeton (2002)) that applied an interpretation based on a statement approach in the assessment. In the first study, with regard to the authority related element, 8.8% of respondents measured the identity of site operator element (8.8% responses) when judging the credibility of online information (Fogg et al., 2002, p.23). Meanwhile, in the latter study, the identity of site operator was rated by 67% of respondents as one of important factors when choosing websites.

Analysis between the geoliterate and non-geoliterate groups, however, did not show any significant difference between the responses in both groups. Half of each group perceived either Mashup A or Mashup B as more credible. The **hypothesis 3** posited before the experiment, was **not supported** because a low number of responses from the geoliterate group identified the importance of the 'map producer' element (as metadata related variables) in judging map information credibility. In contrast, a high number of responses perceived the metadata related elements as important, but in the non-context, prominence independent experiment (i.e. interpretation based on a statement).

Visual cues of 'colour scheme', 'clarity of symbol' and 'symbol design convention' were rated as the most important factors in this study. This might be because of the prominence of these elements on the map mashup. Although the design of the 'map producer' was enhanced from textual information (see Experiment 1) to image based information (depicted as a logo), this still seems insufficient to increase the impact of that element. Respondents might not notice the element of 'map producer' embedded on the sidebar due to their attention being focused on the centre of the map; the additional information attached to the sidebar may not be perceived as a crucial element to be scrutinised. This is in line with the argument by Metzger et al. (2003) that the need for extra checking of information was the cause of critical elements rarely being assessed by web users. Generally, metadata elements for judging credibility, such as date, authority, policy and privacy are presented as additional information in the form of a separate layout or in new site pages. The presentations and design of these critical elements, however, does not have special prominence in a web page. Users sometimes do not want to spend extra time checking the validity before disseminating and relying on the information. There is therefore a need for specific research to design these critical elements to appear prominently on an online medium. Nevertheless, from these findings there is a group of map users who do not have knowledge to the importance of analysing critically the information disseminate on the WWW as well as the importance to assess the metadata related to authority of information when judging the credibility of map information. A public education campaign to the importance of critically judging information on the website as well as online map mashups is needed to educate web users from the bias, propaganda, threats and false information from the Internet. Table 6.8 summaries the results from this experiment.

Table 6-8 The results of experiment hypotheses

No		The hypotheses	Results
1.	Hypothesis 1	Visual cues have no significant influence in respondents' credibility assessment on map mashup applications	Not supported
2.	Hypothesis 2	The metadata related to sources (i.e. map data supplier, map mashup producer) have significant influence in respondents' credibility assessment	Not supported (in prominence dependent setting); Supported (in prominence independent setting)
3.	Hypothesis 3	The metadata related to sources have significant influenced within geoliterate respondents	Not supported (in prominence dependent setting); Supported (in prominence independent setting)
4.	Hypothesis 4	There is no significant difference between the level of importance of the metadata related to sources between these two contexts - a prominence dependent (i.e. interpretation is based on what looks prominent) and a prominence independent setting (i.e. interpretation is based on a statement)	Not supported

6.6 Chapter Summary

The contribution of this experiment to the overall thesis is to highlight the dominance of visual cues in users' judgement of online map information credibility. Even in a context that involves a deep level of involvement, where the stimulus used in the experimental task was quite critical (specifically in search and rescue disaster operation), visual cues are still the dominant factors that became the basis of respondents' judgements. The visual cues elements of 'colour scheme', clarity of symbol and the 'symbol design convention' were strong influences when judging the credibility of online map information. On the other hand, it concluded that clarity of information; understanding of information and utilising a map by the means of 'colour scheme' and 'symbol design' were considered to be the most important aspects, according to the majority of respondents. Moreover, these elements are prominent on any map and hence attract users' attention when assessing information.

7 EXPERIMENT FOUR

7.1 Introduction

This section describes the aims, hypothesis, methodology, results and analyses used in Experiment Four. The aim of this experiment was to achieve objective 3:

‘to examine the influence of colour coded traffic light (CCTL) labelling on respondents’ assessment of credibility’

Previous experiments have been conducted to examine the influence of metadata on users’ credibility assessments. The findings of these earlier experiments have indicated that the provision of metadata has little influence on users’ assessments of map information credibility. Moreover, findings have demonstrated the major influence of visual cues to users’ judgements. Therefore, a further experiment was conducted to examine the influence of credibility labelling in the form of visual ‘Colour Coded Traffic Lights’ (CCTL) on users’ assessments of credibility.

In this experiment, respondents were given a slightly similar task that shared a few common properties and dataset as in Experiment Three. They were provided with an interactive task that involved a degree of cognitive judgement. A major difference between these two experiments was the presence of the visual credibility indicator that was presented as a CCTL rating label. Respondents were invited to analyse both maps, which had been labelled with two categories of credibility ratings – high and low credibility - and were asked to make a judgement about which map they would choose to help solve the given task.

7.2 Hypotheses

Hypothesis 6:

Credibility label has a significant influence in respondents' credibility assessment

Hypothesis 7:

The presence of credibility label has a significant effect to the influence of visual cues in respondents' judgement

Hypothesis 8:

There is a significant difference in the influence of visual cue variables between the geoliterate and non-geoliterate respondents when making judgement

Hypothesis 9:

There is a significant difference in the influence of credibility label between the geoliterate and non-geoliterate respondents when making judgement

Hypothesis 2 (to confirm the previous hypothesis) is:

The metadata related to source (i.e. map mashup producer) has a significant influence in respondents' credibility assessment

7.3 Experimental Design and Materials

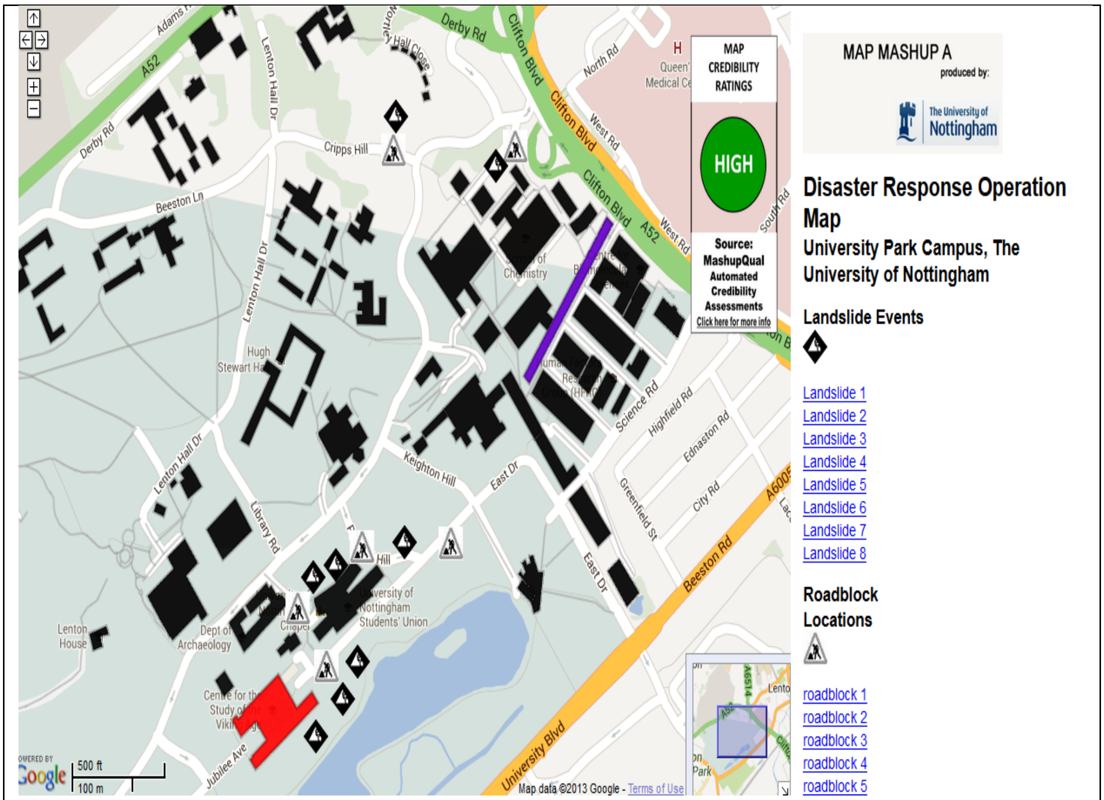
Experiment 4 was designed to stimulate a degree of involvement with the task with respondents being required to use cognitive judgement before giving responses to the questionnaire. Table 7.1 presents the details of the experimental task. Two conditions were designed to test the experimental hypotheses. The two comparison maps, as shown in Figure 7.1, differ in terms of sources of map producer and ratings category as well as visual appearance, including the colours of building features, the designs of road works and landslide event symbols. The number and the locations of roadwork points and landslide events were also different. Table 7.2 presents full descriptions of the condition differences between the two comparison maps.

In the task, respondents were given a situation in which they had to analyse a map to determine the shortest cycle route from one area to another on a university campus. Respondents were given two map mashups that presented the locations of landslide events and roads that had been blocked by the authorities. The planning of the shortest route has to be made on the map mashup respondents perceive to have most credibility. They have to judge the two maps critically because the representation of information presented on them is different: the colour of building features, as well as symbols for roadwork points and landslides, differ.

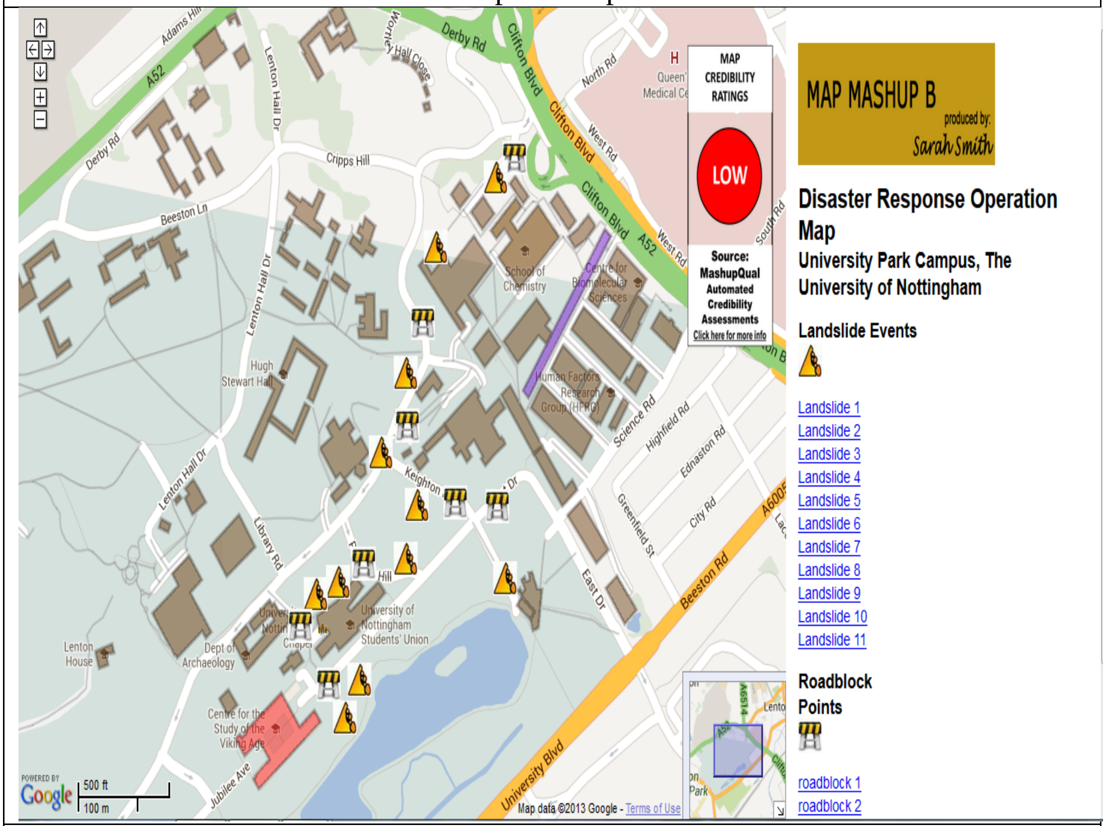
Map A was simulated to be designed by a high credibility organisation, ‘the University of Nottingham’ so a ‘high credibility’ label, in a form of visual CCTL, was stamped on top of the map mashup. This map was designed to look as not well designed by showing the colour scheme of building features and point symbol were in black and not easily distinguishable. As stated by Brewer (2005, p.122) a well-designed scheme is usually represent data in a different unique hues. The aim was to test the influence of a poorly designed map produced by a high credibility organisation. Meanwhile, Map B was simulated to be designed by low credibility individual, ‘Sarah Smith’ so a ‘low credibility’ label, in a form of a visual CCTL, was stamped on top of the map mashup. This map was designed to look well design by the use of similar contrast of colours; in order to test the influence of ‘well design’ on a map produced by a low credibility individual.

In the two previous experiments, One and Two, the data and information presented on the two comparison maps had been similar; in those contexts, respondents had tended to rely on visual cues, such as ‘colour scheme’ and ‘symbol design’ when judging a map. Metadata elements, such as ‘map data producer’ and ‘map data supplier’, were least measured in these assessments. In Experiment Three, information on the location of road-works and landslide events were designed differently in order to test the influence of the authority (metadata-source) related element, in this case the ‘producer of a map mashup’, on respondents' credibility judgement. The findings showed that the influence of the metadata related to authority was divided 50:50 between –unimportant and important. Experiment Four was in contrast conducted to examine the influence of stamped credibility labelling on the top of the map mashups. This was designed to act as a visual element that would harvest the credibility elements underlying the map mashup application, which are critical to users' assessments.

The maps were created using a simple tool to develop mashups: Google Earth tool was used to create the two layers in *.kml format- the layers for road works, and landslide events. MyMap Makers tools provided on the Google Map interface were used to create the building layers in *.kml format. Then Google Map APIs were used to mashup the three layers into one functional map. The maps were stored in the university server under the researcher's student account. The links distributed to respondents used the university's domain. For example this link (http://www.nottingham.ac.uk/~lgxhi1/mashupAex4.html) displayed the mashup A and this link (http://www.nottingham.ac.uk/~lgxhi1/mashupBex4.html) displayed mashup B. The layout of the experimentally based questionnaire was designed using Google Sites interface and the questionnaire was designed using Google Docs.



Map Mashup A



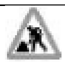
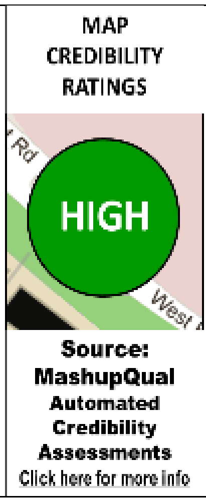
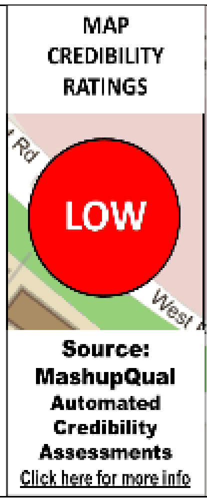
Map Mashup B

Figure 7-1 Comparison of the two maps

Table 7-1 Experimental Design

	Map A	Map B
Experimental assumption	**High level of credibility**	**Low level of credibility**
Manipulated variable: The producer of map mashup	The University of Nottingham (UoN)	Fictional and therefore unknown producer -Sarah Smith
Full descriptions (as in the task sheet)	Map Mashup A: Disaster Response Operation Map produced by the University of Nottingham	Map Mashup B: Disaster Response Operation Map produced by Sarah Smith
Experimental task	Situation 1: You are a student of the University of Nottingham. You have to attend a course at the Trent building. Unfortunately, there are a few landslide events and roadblocks reported around the campus. You want to access a map to plan your shortest route by bike from the Coates Building (this building is marked in purple on the map) to the Trent Building (the building marked in red). After browsing through a search engine, you find two interactive map mashups. Both maps disseminate the information about the disaster. But you must choose only one map mashup that you perceive as having the most credibility to meet the given situation. Both maps represent a similar area, but came from different sources and present slightly different information on the number of slides and the roadblock locations. You must evaluate the two maps and decide which map you will choose to assist you in your route planning. The maps provide a function to zoom in, zoom out, pan and identify features (by clicking on an object on the map) The task: 1. Browse and evaluate the two interactive maps 2. Decide which map mashup you will use (either Map Mashup A or Map Mashup B) to assist you in planning your shortest cycle route from the Coates Building (the building marked in purple to the Trent Building (the building marked in red) 3. Submit your answers using a questionnaire.	

Table 7-2 Condition differences between Map Mashup A and Map Mashup B

Experimental conditions	Map A	Map B
Building colours	Black	Brown
Roadblock points	6 points	8 points
Roadwork symbols		
Landslide points	8 points	11 points
Landslide symbols		
Rating visual indicator		

7.3.1 Dependent Variables

Dependent variables used in this experiment are shown in Table 7-3 . Full details of these variables are discussed in Section 2.3.

Table 7-3 Dependent variables

Question Id	Dependent variables	Measurement
Q1	The map that perceived as having more credibility	Binary
Q2	The perceived level of credibility	7-point scale
Q3	The influence of colour scheme	7-point scale
Q4	The influence of symbol design	7-point scale
Q5	The influence of overall presentation	7-point scale
Q6	The influence of credibility label	7-point scale
Q7	The influence of credibility analysis	7-point scale
Q8	The influence of identity of map producer	7-point scale

7.3.2 Participants

There were 75 respondents who completed the study. The sampling selection was discussed in Section 3.6. The sample was drawn from the members of University of Nottingham (students). The selection of respondents in the sample was based on a pragmatic approach whereby the map based questionnaire was distributed via online in the University of Nottingham intranet portal and student group mailing lists. The intranet portal can be accessed by any student of the university; the link of the map based questionnaire was published as an advertisement in the Opportunity Section and open for all students who were interested to join in the online survey. The questionnaire was distributed in the student groups' mailing lists which comprised of undergraduate, master and postgraduate students that registered under the School of Geography, the University of Nottingham. The link of the questionnaire was distributed via email where every student who has accessed to the mailing lists was a potential respondent. Respondents in the sample were drawn from those who were interested to join in the survey and who have accessed to the application services mentioned above.

However, a sample selection was made to a group of respondents that in a range of 18 to 35 years old. This was to represent a population of young adult map users. The sample does not represent individuals who did not have access to the Internet, to the university application services (i.e. student portal and mailing lists), university staff and non-university members. The ages in the samples ranged from the age categories of 18-19 to 25-35 years old. One respondent aged over than 36 was excluded from the sample. The average age in the sample was 22 where the mode values fell in the groups of 20 to 21 (26%).

The sample overrepresented the respondents aged between 18 to 24 years old (90.7%). The sample may represent a population of young adult map users who study and do research in the university community in the UK. There were more female (53.3%) than male (46.7%) respondents in the sample. A majority in the samples use websites every day (94.7%) where 61.3% spend 1 to 4 hours a day. The majority of respondents (94.7%) had experience of using interactive online maps supplied by commercial providers (e.g. Google Map, Bing Map, and Yahoo Map) or crowd source (e.g. OpenStreetMap) applications.

There were 37.3% in the sample (28 respondents) drawn from the geoliterate group who have a background in attending geography, cartography, remote sensing, land surveying or geographic information science courses, whilst 62.7% in the sample (47 respondents) were drawn from the non-geoliterate group who have background in attending other non-geospatial related courses, such as engineering, sciences, and social sciences.

Of this non-geoliterate group, 17% have background in social sciences, law or education, 23% were from sciences background and 9% from engineering. The average age between geoliterate and non-geliterate groups was 21 years old. The sample of think-aloud protocols were similar with the sample applied in Experiment 3. See section 6.3.1 for further descriptions.

7.3.3 Questionnaire Design

The responses to the questions in the questionnaire were dependent on the experimental context which, in turn, depended on responses to the main question. The main question was to analyse and make a judgement between two maps that a respondent would use (either Map Mashup A or Map Mashup B) to plan the safest cycle route from building A to building B in a post-disaster situation. Based on the responses to this main question, a respondent was given a series of questions to measure the importance of metadata related to 'map data producer' and visual cue elements in their assessments of map mashup information credibility.

The questionnaire was divided into two sections; the first section was to measure the extent of respondents' perceived credibility in the map mashup they had chosen; multi-item measures were used in this section to assess the extent of users' perceived credibility in their map. In the second section respondents' level of agreement with the pre-determined elements that may influence perceived credibility were examined. Both sections required respondents to rate their responses on the basis of the experimental task.

The first section measures respondents' perceived credibility with regard to the map mashup that they have chosen. Multi-item measurement is referred to several questions which are quite similar and one-dimensional, measuring to the underlying principle of credibility: 'credible', 'believable', 'trustworthy', 'source competency' and 'expertise level' were other terms connected with the notion of credibility. The second section measured respondents' agreement with the elements that might influence their perceived credibility in the maps they chose and rejected. Several elements have been identified in the literature as credibility elements: visual cues (i.e. colour scheme, overall presentation, and symbol design), source authority (i.e. identity of the map mashup producer) and a visual credibility indicator in a form of a CCTL label. Respondents have to rate the extent to which these elements influence their judgement in the experimental task.

The main purpose of the questionnaire, however, was to measure the perceived importance of the visual credibility indicator in respondents' credibility assessment of map mashup information. Table 6-3 presents the list of questions used in the questionnaires and specific levels of measurement.

Table 7-4 The list of questions in the questionnaire and specific levels of measurement

Q-ID	Items	Question	Measurement
Section One			
Q1	Main question	Please browse the two map mashups above and evaluate which of the two maps you perceived as having more credibility (more believable information) to assist you in planning your shortest cycle route from the Coates Building to the Trent Building in a post disaster situation	Binary
Q2	The extent of respondents' perceived credibility to the map they chose in Q1	On a scale of 1-7, (1 = low, 7 = high) indicate how much you perceived the map mashup you chose in Q1 as; a) believable b) trustworthy c) credible On a scale of 1-7, indicate d) the expertise level e) the competency level of the source(s) of information on the map you chose in Q1.	7-point scale – in rating scale
Section Two			
Q3 (i)	The extent of respondents agreement to the visual cue elements in influencing their judgement of credibility	a) I chose the map because of the colour scheme used on the map b) I chose the map because of the symbol design used on the map c) I chose the map because I have been influenced by the overall presentation of the map	7-point scale Strongly disagree to strongly agree
Q3 (ii)	The extent of respondents agreement to the metadata element in influencing their judgement of credibility	d) I chose the map because I have been influenced by the identity of the map mashup producer (author)	7-point scale Strongly disagree to strongly agree

Q3 (iii)	The extent of respondents agreement to the visual credibility indicator in a form of CCTL label in influencing their judgement of credibility	e) I chose the map because I have been influenced by the label (result) of credibility ratings provided with the map f) I chose the map because I have been influenced by the additional information of credibility rating assessment (analysis) provided with the map	7-point scale Strongly disagree to strongly agree)

To understand the samples in the experiment and the differences between responses, a series of demographic questions was asked before respondents conducted the experimental tasks. Table 7.5 presents the series of demographic questions in the questionnaire.

Table 7-5 A series of demographic questions used in the questionnaire

	Items
Q1	Age
Q2	Gender
Q3	Professional or academic qualifications
Q4	Experience of using online maps from commercial providers and crowdsourcing map applications
Q5	Experience of using websites and average time spent
Q6	Member or non-member of the University of Nottingham, UK

7.3.4 Procedure

The procedure used in this experiment is depicted in the figure below. Further explanation to this diagram was discussed in Section 3.7.

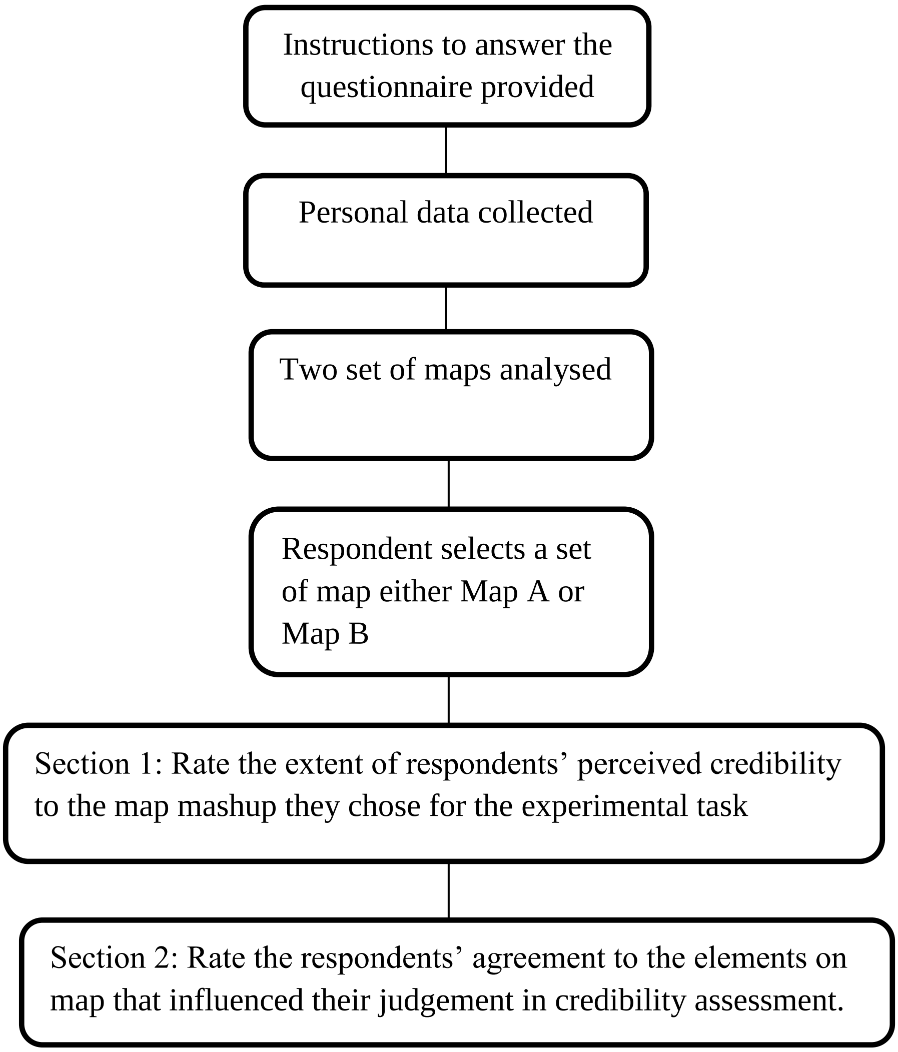


Figure 7-2 Procedures used in Experiment Four

7.4 Results and Analysis

7.4.1 Data Analysis

Statistical analysis using SPSS v.16 was used to study the response data. The data collected from the questionnaire were measured as binary variables and ordinal variables, respectively. The responses data were not in continuous variables and transform from discrete variables using Terrell's Transformation technique. However, the data have been tested which do not fit with normal distribution model. Hence, data analysis focused on selecting appropriate non-parametric statistical tests. The Chi-square test and Mann Whitney U test were used to analyse the data. Non parametric tests are free from the restrictive assumption that the data needs to be normally distributed; The Chi-square test has a specific assumption of the data before it can be modelled correctly. The assumptions in the Chi-square test are that the data is independent and each entity contributes to only one cell in the contingency table and the expected frequencies should be greater than 5. These assumptions were not violated with regard to the experimental data modelled by this test.

Questions were used to measure different facets of credibility. According to (Fogg and Tseng, 1999, p.128), credibility can be measured in several facets, with the main facet being trustworthiness and expertise. Some questions in the questionnaire were therefore based on these two dimensions. Such types of multi-item measures have been widely used to measure psychological or sociological constructs, enabling assessment of the validity of responses to questionnaires in which several similar questions measure the same construct.

Consistency of scores from one respondent to those questions might assess the validity of the responses given. Validity of responses was measured by the reliability of scales (multiple items/questions) used in measuring a construct; individual items or sets of items should produce responses consistent with questions of same construct. Split-half reliability was used to measure the consistency of the responses from respondents, which randomly splits the response data into two. A score for each participant is then calculated, based on each half of the scale; if a scale is very reliable, a person's score on one half of the scale should be similar and have a high correlation with their score on the other half (Field, 2009, p.674). Cronbach's alpha, a common measure of scale reliability was used. Values of 0.7 to 0.8 are acceptable for Cronbach's α.

From the analysis (see Table 7-6), each item in the scale has a value of Cronbach's α 0.8, which is within an acceptable value that enables the scale used to produce high correlation results and to have enough reliability to measure the credibility construct. The values of item-total correlation, shown in the table below, that indicate the validity of the scale, were within 0.65 to 0.8.

This is within the acceptable value, which is more than 0.3. An item should have at least 0.3 item-total correlations to provide evidence for the unidimensionality of a scale; items that do not correlate well with other items probably do not belong to the scale, since they are tapping a different concept (Vaus, 2008, p.128).

Table 7-6 Reliability and Validity analysis on the multi-items scale

variables	Scale mean if item deleted	Corrected item-total correlation	Cronbach's alpha if item deleted
Q2(a) On a scale of 1-7 indicate how much you perceived the map mashup you chose in AQ as believable	19.97	0.651	0.878
Q2(b) On a scale of 1-7 indicate how much you perceived the map mashup you chose in AQ as trustworthy	20.16	0.785	0.847
Q2(c) On a scale of 1-7 indicate how much you perceived the map mashup you chose in AQ as credible	20.35	0.776	0.847
Q2(d) On a scale of 1-7 indicate the competency level of the source(s) of information on the map you chose in Q1	20.09	0.768	0.849
Q2(e) On a scale of 1-7 indicate the expertise level of the source(s) of information on the map you chose in Q1	20.49	0.668	0.876

Two items were used to measure respondents' agreement with the influence of the credibility labelling element in influencing their assessment of map mashup information. The table below presents the items/questions given in the questionnaire. These two items, which measure the two related agreements with the influence of ratings indicator in credibility assessment, produced a value of Cronbach's α 0.77. This alpha value was within the acceptable value, which was greater than 0.7. This indicates a good reliability of the overall scale to measure the same construct, which in this case was the influence of a ratings indicator in credibility assessment. The values of item-total correlation were more than 0.3 as in Table 7.7 and may indicate the unidimensionality of the responses.

Table 7-7 Reliability and validity analysis on the multi-items scale to measure the influence of credibility rating

Items	**Scale mean if item deleted**	**Corrected item-total correlation**	**Cronbach's alpha**
Q3d) I chose the map because I have been influenced by the result of credibility ratings provided with the map	3.96	0.626	0.77
Q3e) I chose the map because I have been influenced by the additional information of credibility rating assessment provided with the map	4.41	0.626	

Three items were used to measure respondents' agreement with the influence of visual cue elements on their assessment of map mashup information credibility. These items, however, were selected and included in the questionnaire because of the context experiment dataset that manipulates these elements. In this study, the element of 'colour scheme', 'symbol design' and 'overall presentation' were classified under visual cue elements. The table below presents the items/questions given in the questionnaire. These three-items, which measure agreement with the influence of visual cues in credibility assessment, produced a value of Cronbach's α of 0.79. These values were within the acceptable value to indicate a good reliability of the overall scale to measure the same construct, which, in this section, was the influence of visual cues on credibility assessment. The item-total correlation for each item demonstrates the correlation at 0.6, which is within the acceptable value to indicate the validity of the item to measure the same construct- in this case the influence of visual cues. Table 7.8 presents the values of item-total correlation and Cronbach's alpha for each item.

Table 7-8 Reliability and Validity analysis on the multi-item scales to measure the influence of visual cues

variables	Scale mean if item deleted	Corrected item-total correlation	Cronbach's alpha if item deleted
Q3a) I chose the map because of the colour scheme used on the map	9.88	0.655	0.694
Q3b) I chose the map because of the symbol design used on the map	8.95	0.630	0.717

Q3c) I chose the map because of I have been influenced by the overall presentation of the map	9.09	0.618	0.730

7.4.2 Results of the map that respondents chose as having perceived credibility

Hypothesis 6 is:

Credibility label has a significant influence in respondents' credibility assessment

Frequency analysis was conducted to analyse the main question: 'Please evaluate which of the two maps (either Map A or Map B) you perceive as having more credibility (more believable information) to assist you in planning the safest cycle route from Coates Building to Trent Building in a post-disaster situation. This question yields two binary variables, which are either Map A or Map B. Chart 7-1 presents the results of this question.

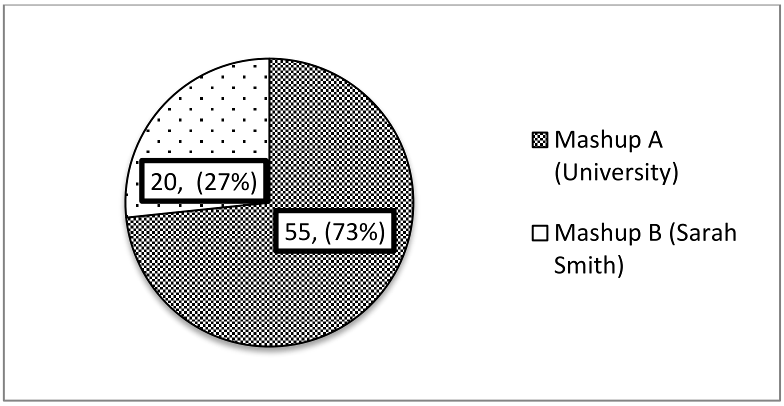


Chart 7-1 The percentage of responses that chose Mashup A and Mashup B

From the results, 73% of respondents perceived Mashup A as having more credibility. In contrast, 27% of respondents perceived Mashup B as having more credibility. A 'high credibility' rating label was stamped on top of the Map Mashup A, whereas a 'low credibility' rating label was stamped on top of Mashup B. The response differences between the two mashups were statistically significant in the Chi-squared test X^2 (1, n = 75) = 16.333, p <0.001).

Table 33 (see Appendix C) presents the results of Question 2a to 2e which is how much the respondent's perceived credibility in the map they chose. Comparison analysis was carried out between Experiment 4, which stamped a visual credibility indicator label on the mashup, with the previous

Experiment 3, which did not stamp any credibility label. The analysis demonstrates significant changes and increases in the responses that chose mashup A as a more credible map than mashup B. Chart 7-2 below depicts the response differences between both experiments and Table7.9 presents the number of response differences (as a percentage).

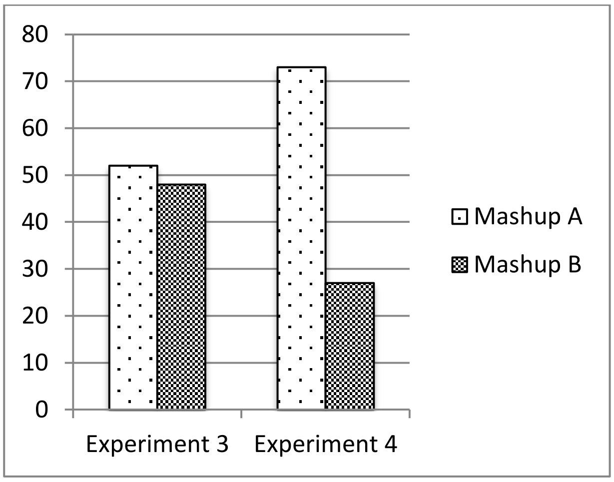


Chart 7-2 Responses differences (as a percentage) between the results of the main question in both experiments

The Table 7.9 below indicates significant differences between the responses that chose mashup A in Experiment 4 with responses that chose mashup A in the previous Experiment 3. The 21% increase may indicate the positive impact in the usage of credibility labelling on users' credibility assessment of map mashup information. The 46% increase in the responses that chose Mashup A rather than Mashup B in Experiment 4 compared to only 4% increase in Experiment 3; this strongly indicate the positive influence of credibility labelling on map mashup applications.

Table 7-9 Differences in responses that chose either mashup A or mashup B between Experiment 3 and Experiment 4

Experiment	**No. of responses (%) to Mashup A**	**No. of responses (%) to Mashup B**	**Differences (%) on the responses that chose Mashup A and Mashup B within experiment**
Experiment 3	64 (52%)	69 (48%)	5 (4%)
Experiment 4	55 (73%)	20 (27%)	35 (46%)
Differences (as a %) in responses that chose mashup A and mashup B between experiments	9 (21 %)	49 (21%)	

Table 7-10 Cross-tabulation of the number of cases falling into each combination of categories

			Ratings	
			High credibility map	Low credibility map
Influence	No label	Count	64	69
		% within influence	48.1	51.9
		% within category	53.8	77.5
	With label	Count	55	20
		% within influence	73.3	26.7
		% within category	46.2	22.5
Total		Count	119	89
		% within influence	57.2	42.8
		% within category	100	100
		% of total	57.2	42.8

Pearson's chi-square test was conducted to examine whether there was any association between the two categorical variables (which in this case were the influence of the rating label and whether respondents chose the high credibility map (Mashup A) or the low credibility map (mashup B)). The number of responses to the main question were coded into another variable, namely frequency by SPSS tool and weighted by the frequencies of responses that fell into each combination of categories. From the cross tabulation in Table 7.10 in total 57.2% chose the high credibility mashup (Mashup A); of the responses that chose high credibility map, 41.6% were influenced to choose the map in the context that presented no rating label and 53.8% were influenced to choose the map in the context that presented a credibility rating label.

In contrast, in total 42.8% chose the low credibility mashup (Mashup B) and of these responses 77.5% were influenced to choose the map in the context that presented no rating label and 22.5% were influenced to choose the map in the context that presented a credibility rating label. Regarding the percentages within the influence category – the presence of rating label and without the presence of rating label – of those responses in the context without the rating label 48.1% of responses chose the high credibility map and 51.9% of responses chose the low credibility map. In contrast, of those responses in the context with the presence of a rating label, 73.3% of responses chose the high credibility map compared to 26.7% that did not chose the map. Therefore, this finding may indicate the positive influence of credibility rating labels on map mashups with regard to respondents' credibility assessment.

From the result of the Pearson chi-square test X^2 (1, n = 208) = 12.453, p <0.001), the value is statistically significant (p<0.001). This value is small enough to reject the hypothesis that the categorical variables are independent, increasing confidence in the hypothesis that the two variables are related to some extent. This highly significant result indicates that there is an association between the presence of a credibility rating label and whether respondents choose a high credibility map or low credibility map. This indicates that the presence of a stamped rating label on the mashup significantly assists respondents to choose the high credibility map mashup rather than low credibility mashup. An odds ratio was used to calculate the effect size of this significant difference. The formula below presents the calculation of the odds ratio which indicates that the likelihood of respondents choosing the high credibility map, **if they were given a map with a visual rating indicator (CCTL label), was three times higher than when given a map without the visual rating indicator (CCTL label)**.

Odds $_{\text{respondents chose high credibility mashup in labelling context}}$ (Equation 7.1)

$$= \frac{number\ \ that\ chose\ high\ credibility\ map\ in\ \ labelling\ context}{number\ that\ chose\ low\ credibility\ map\ in\ labelling\ context}$$

$$= \frac{55}{20} = 2.75 \qquad \text{---------- (1)}$$

Odds $_{\text{respondents chose high credibility mashup in no labelling context}}$

$$= \frac{number\ \ that\ chose\ high\ credibility\ map\ in\ no\ labelling\ context}{number\ that\ chose\ low\ credibility\ map\ in\ no\ labelling\ context}$$

$$= \frac{64}{69} = 0.93 \qquad \text{---------- (2)}$$

$$\text{Odds ratio} = \frac{\text{Odds respondents chose high credibility mashup in labelling context}}{\text{Odds respondents chose high credibility mashup in no labelling context}} \qquad \text{------ (3)}$$

$$= \frac{(1)}{(2)} = \frac{2.75}{0.93} = 2.96 \sim 3.0$$

Therefore, **hypothesis 6 is supported**.

7.4.3 Results of the influence of the credibility labelling on their judgement of credibility

Hypothesis 6 is:

Credibility label has a significant influence in respondents' credibility assessment

Table 7.11 presents the mean and standard deviation (SD) of the individual items that measure the influence of the credibility rating element in respondents' credibility assessment (see Chart 7 and Chart 9 in Appendix C for the responses distribution of these items).

Table 7-11 The Mean and Standard Deviation (SD) of the responses measure the influence of a credibility indicator on the experimental dataset

Items	Of the responses that chose Map A (n =55)		Of the responses that chose Map B (n=20)		Total responses	
	Mean	SD	Mean	SD	Mean	SD
Q3d) I chose the map because I have been influenced by the label of credibility ratings provided with the map	5.09	1.76	2.55	1.61	4.41	2.05
Q3e) I chose the map because I have been influenced by the additional information of credibility rating assessment provided with the map	4.33	2.17	2.95	1.638	3.96	2.12

The total scores of these two items were then classified into three equal-sized groups. The rating value of point 4 (i.e. undecided and neutral) was recoded into value 0 before the classification was made. Table 7.12 and Table 7.13 present the classification of scores according to the categories. A high category indicates a high rating agreement with the influence of a credibility rating label in respondents' judgement; an intermediate category indicates a middle rating agreement; a low category indicates a low rating agreement with the influence of a rating label, including the undecided and neutral ratings. The mean of total scores was 7.3, with a standard deviation 4.4, which fall within the intermediate level of influence (see Table 7.11). Table 7.13 presents the total scores of the specific responses that chose Map A (University of Nottingham) for their task. The mean was 8.47, with a standard deviation 4.3

which fall within the intermediate level of influence. This result indicates that there is moderate (intermediate) influence of credibility labelling on users' credibility judgement.

Table 7-12 Classification of total scores according to low, intermediate, high influence categories

Total scores	**Frequency (n=75)**	**Percentage**	**Categories**
0	4	5.3	Undecided/neutral
1-5	24	32.0	Low
6-10	25	33.3	Intermediate
11-15	22	29.3	High

Table 7-13 Classification of total scores of the responses that chose Map A (labelled with a high credibility rating)

Total scores	Frequency	Percentage	Categories
	n = 55		
0	2	3.6	Undecided/neutral
2-6	20	36.4	Low
7-11	15	27.3	Intermediate
12-15	18	32.7	High

In other words, hypothesis **6 is supported**. The credibility label had moderate (intermediate) influence in users' judgement. Findings from the think-aloud assessment confirmed this.

From the think-aloud experiment, of the six respondents involved five of them chose Map A (the map that labelled with high credibility label). They were influenced by the credibility rating label stamped on the map. However, only three respondents noticed the credibility rating label at the early stage of assessment; hence they used this label as a basis of their judgement.

Excerpts from the respondents as below;

'I chose this map [Map A] because its rating is high compared to the Map B'

(A, non-geoliterate)

'Because this map showed low and this map has high rating; I will chose Map A because it showed high rating' (R, non-geoliterate)

'I chose Map A because there is one organisation who reviewed this map…because it had high credibility rating (Fik, non-geoliterate)

Meanwhile, the other two respondents did not notice the credibility rating label. At the early stage of assessment, they chose Map B. When the interviewer queried them about any influence of the credibility rating label, then they had second thoughts and eventually chose Map A due to the stamped high rating label. Excerpts from the respondents' transcripts are given as below;

Observer: Which map did you perceived as more credible?

Respondent: I chose Map B because it looks suitable for navigation purposes; it looks creative, easy to use. Map A looks more professional, but Map B looks more suitable for public users

Observer: Does the credibility rating have any influence on your decision?

Respondent: Ok. If there is a rating, there might be some influence. This was rated by whom? If it was rated from one organisation, I will be influenced. I chose Map A because there was one organisation which rated the map.

(Fad, geoliterate)

However there was one respondent that did not choose the Map A on the basis of the presence of a high credibility rating label. Extract as below;

Observer: Which map did you perceive more credible for this task?

Respondent: I chose Map A because the data presented was more suitable. It does not have [many] landslides and constructions. I would suggest using this map because there is one route at the back of this building [Trent Building] that is not affected by the landslides and roadblock. If Map B, it is difficult…..there is no suitable route. The data presented so many roadblocks, landslides etc.

Observer: Did the rating label have any influence on your decision?

Respondent: I just focused on how to get from A to B. Before you mentioned about the credibility rating, I had not noticed. I might have been influenced if I had noticed that…my argument [decision] was based on the details of the information. But my second choice…another reason [to choose] was due to the high credibility rating….at first, I did not notice this rating; although if I had noticed, I would not have been influenced. I would look [focus] on the main purpose. This element [rating label] is an additional element to strengthen my decision. I will believe the map more if there is a credibility rating.

(Am, geoliterate)

From these responses, the presence of a credibility rating label had a certain level of influence on some of the respondents. Some respondents used that element as the main basis in their judgement; whilst the others used it not as a main element but only to support their decisions. One respondent was not influenced by the rating label. Excerpt from respondent transcript as below;

'I chose Map B as a reference map to decide the shortest route [in this task]. This map is more believable because it has more details… it showed more landslides and roadblocks. The colour used influenced me because more building names can be seen [clear to read the label of building]. If we ride a bike, we have to know more landmarks. The name of the building is important so that we know which junction to go…. [When observer queried about any influence of the credibility rating label] The credibility rating label did not help me much. I did not look at it. I just focused on the map and symbols. I was not aware of the importance of this element' (Fa, non-geoliterate)

Therefore, **Hypothesis 6 is supported**.

7.4.4 Results of the influence of visual cue elements on their judgement of credibility

Hypothesis 7 is:

The presence of credibility label has a significant effect to the influence of visual cues in respondents' judgement

Individual item analyses were conducted as in Table 7.14. From the table below, respondents' agreement with the influence of 'colour scheme' used were, on average, at point 4 (neutral/undecided). Respondents' agreement with the 'symbol design' used on the map and the 'overall presentation' were, on average, at point 5 (slightly agree).

Table 7-14 The Mean and Standard Deviation (SD) of the responses measure the influence of visual cue elements on the experimental dataset

Items	Of the responses that chose Map A (n =55)		Of the responses that chose Map B (n=20)		In total responses	
	Mean	SD	Mean	SD	Mean	SD
Q3a) I chose the map because of the colour scheme used on the map	3.93	1.70	4.50	1.78	4.08	1.73
Q3b) I chose the map because of the symbol design used on the map	5.20	1.42	4.50	1.70	5.01	1.52
Q3c) I chose the map	4.84	1.44	4.95	1.61	4.87	1.47

because of I have been influenced by the overall presentation of the map						

The total scores of the three items were then classified into three equal-sized groups. The rating value of point-4 (i.e. undecided and neutral) was recoded into value 0 before the classification was made. Table 7.15 presents the classification of scores according to the categories. A high category indicates the high rating agreement with the influence of visual cues in respondents' judgement; an intermediate category indicates middle rating agreement; a low category indicates low rating agreement with the influence of visual cues, including the undecided and neutral ratings. The mean of the total scores was 12.5 with a standard deviation of 4.7; **on average, the influence of visual cues in respondents' judgement of map information credibility was likely to fall within the intermediate level of influence.**

Table 7-15 Classification of scores into low, intermediate, high influence categories

Total scores	Frequency (n=75)	Percentage	Categories
2-10	26	34.6	Low
11-16	27	36	Intermediate
17-20	22	29.3	High

Further analysis was conducted on the ratings between Experiment 4, with previous experiments. The response rating data for two items – influence of 'colour scheme' and 'symbol design'- were recoded into two values, '1 = no influence' and '2 = influence'. A cross tabulation table based on Table 7.16 was generated to compare the two individual items – the influence of 'colour scheme' and 'symbol design' in the two experiments. This table contains the 4 categories – the presence or absence of credibility rating (Exp4) and the level of importance – important or not important. Pearson's Chi-square test was then conducted.

From the test, there was a significant association between the presence of credibility rating and whether or not the 'colour scheme' element influenced respondents' judgement, $\chi2(1) = 33.65$, $p < 0.001$. This seems to represent the fact that based on the odds ratio, **the odds of the influence of 'colour scheme' is 8 times higher if the map was not labelled with a credibility rating than when labelled. In other words, the influence of 'colour scheme' on users' judgement will be decrease if a map is labelled with credibility rating.**

Meanwhile, the result of chi-square test on the influence of ‘symbol design’ demonstrates a significant association between the presence of credibility rating and whether or not the ‘symbol design’ element influenced respondents’ judgement, $\chi2(1) = 6.11$, $p <0.05$. This seems to represent the fact that based on the odds ratio, **the odds of the influence of ‘symbol design’ is 3.3 times higher if the map was not labelled with credibility rating than when labelled. In other words, the influence of ‘symbol design’ on users’ judgement will be decrease if a map is labelled with credibility rating.**

Table 7-16 The frequency of individual items – ‘colour scheme’ and ‘symbol design’ – between experiments

	Influence		
	Yes (rating = 2)	No (rating = 1)	Neutral/undecided (rating = 0)
Variable: colour scheme			
Experiment 4 (n = 75)	38 (50.6%)	25 (33.33%)	12 (16.0%)
Experiment 3 (n = 133)	122 (91.7%)	8 (6.0%)	3 (2.3%)
Variable : symbol design			
Experiment 4 (n = 75)	58 (77.3%)	11 (14.7%)	6 (8.0%)
Experiment 3 (n = 133)	123 (92.5%)	7 (5.3%)	3 (2.3%)

Responses from the think-aloud experiment confirmed this. There were two respondents who changed their decisions and had second thoughts due to the influence of the high credibility rating stamp. The transcript excerpts are below;

‘I chose Map B because it looks suitable for navigation purpose…..Map A looks more professional, but Map B looks more suitable to use for public users……..Ok. If there is a rating, there might be some influence. This was rated by whom? If it was rated from one organisation, I will be influenced. I chose Map A because there was one organisation which rated the map’

(Fad, geoliterate)

‘I chose Map A because the data presented is more suitable. It does not have [many] landslide and constructions…. I just focus on how to get from A to B. Before you mentioned about credibility rating, I did not noticed. I might be influenced if I noticed that’ (Am, geoliterate)

From these results, **hypothesis 7 is not supported**.

7.4.5 Results of the analysis of non-geoliterate and geoliterate groups on the influence of visual cue elements in their credibility assessment

Hypothesis 8:

There is a significant difference in the influence of visual cue variables between geoliterate and non-geoliterate respondents when making judgement

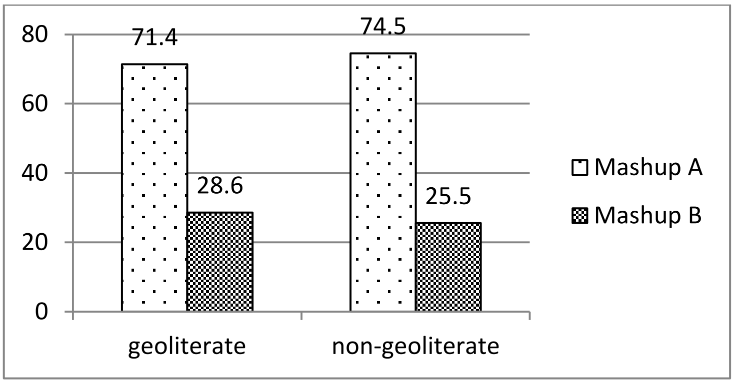


Chart 7-3 Results comparison of the responses that chose Map A or Map B between groups

From the chart above, both respondents from the geoliterate (71%) and non-geoliterate groups (75%) perceived Mashup A, the map labelled 'high credibility map' as having more credibility than Mashup B, the map labelled 'low credibility map'. A low number of responses, 29% and 26%, in the geoliterate and non-geoliterate groups respectively, perceived Mashup B to have more credibility.

Exploratory analysis was conducted to check the normality and variances of the sampling distribution. The results from the Kolmogorov-Smirnov test indicated the response data distribution for the geoliterate, D (28) = 0.13, $p > 0.2$ and for the non-geoliterate, D (47) = 0.16, $p < 0.05$ where the significance was less than 0.05; therefore the distribution in each group was non normal. The Mann-Whitney non parametric test was then selected to test the hypothesis. From Mann-Whitney, the influence of visual cue elements from geoliterate respondents (Mdn = 61.1) did not differ significantly from those of non-geoliterate respondents (Mdn = 72.2), U = 521.5, z=-1.5, p>0.12, r = 0.2 (small effect size). Therefore, the **Hypothesis 8 is not supported**. Chart 7.4 below presents the median value of each response to individual items.

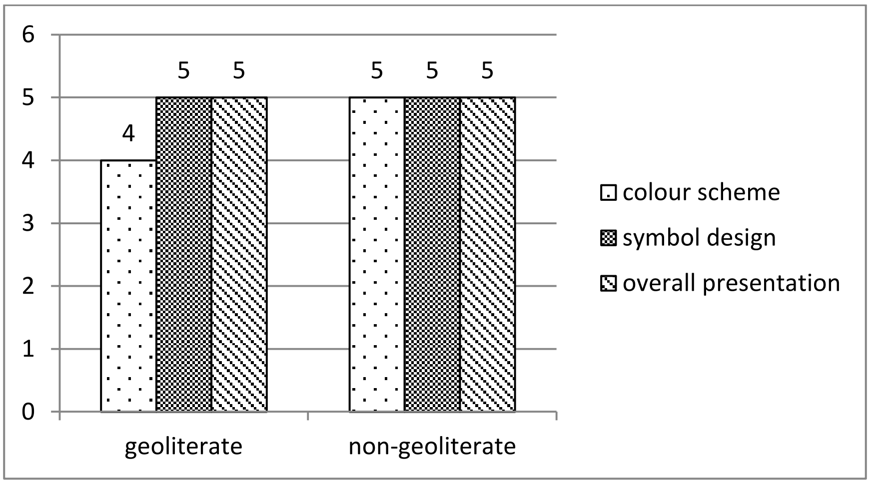


Chart 7-4 The median value for the responses to the influence of visual cue element between groups

The chart above presents the median responses for each element, according to geoliterate and non-geoliterate groups (see Chart 10 to Chart 12 in Appendix C for the patterns of response distributions).

Transcript excerpts from think-aloud experiment are below;

'I did not look at the map producer. It was not important. What was more important was the clarity and ease of use' (Fik, non-geoliterate)

'I chose Map B, route 1 because in terms of design, it was more attractive, neat and was not too crowded…I did not influenced by the map producer. I just focused on the events and all the symbols' (Am, geoliterate)

These results might indicate both groups used visual cues (for example the labels, colour scheme, symbol design) to form the basis of their judgement in assessed the credibility of map information. The levels of influence of these elements in their credibility assessment might be different. Some of them might use visual cues as main element to form their judgement, whereas the tested parameters (i.e. map producer and credibility rating) as supporting elements. Excerpt from think-aloud experiments as below;

'….although if I noticed, I would not be influenced. I will look [focus] at the main purpose. This element [rating label] is an additional element to strengthen my decision. I will believe the map more if there is a credibility rating'

(Am, geoliterate)

'I have been influenced by the visual attractiveness, symbol design, clarity of symbol, symbol convention……the importance of map producer was just ‘important’ [if using a scale of important]… the ‘very importance’ was the symbol…the symbols [number] represent the details of information’

(Fa, non-geoliterate)

7.4.6 Results: Analysis of the non-geoliterate and geoliterate groups - the influence of the credibility labelling in credibility assessment

Hypothesis 9:

There is a significant difference in the influence of credibility label between geoliterate and non-geoliterate respondents when making judgement

The total scores of the two items that measure the influence of credibility ratings were then analysed according to the knowledge background of respondents –geoliterate and non-geoliterate. Exploratory analysis was conducted on the sampling distribution of the the total scores between the two groups to check for normality. From the Shapiro-Wilk significance test, the data from the geoliterate and non-geoliterate groups appeared to be normal, $D(28) = 0.14$, $p > 0.05$ and $D(47) = 0.99$, $p > 0.05$, respectively. The significant values (p) more than 0.05 indicate no deviation from normal distribution.

To test the hypothesis, the mean difference was compared using a parametric test, namely the independent t-test. From this test,on average the influence of the credibility labelling from geoliterate respondents (mean = 48.8, SE = 6.6) was lower than from non-geoliterate respondents (mean = 55.67, SE = 4.2). This difference was not significant $t(73) = -0.92$, $p > 0.05$: however, it did represent a small sized effect $r = 0.1$. Therefore, the **Hypothesis 9 is not supported**. Chart 7.5 below compares the mean value of each response to the two constructs – the influence of credibility label and the influence of metadata related to map data producer.

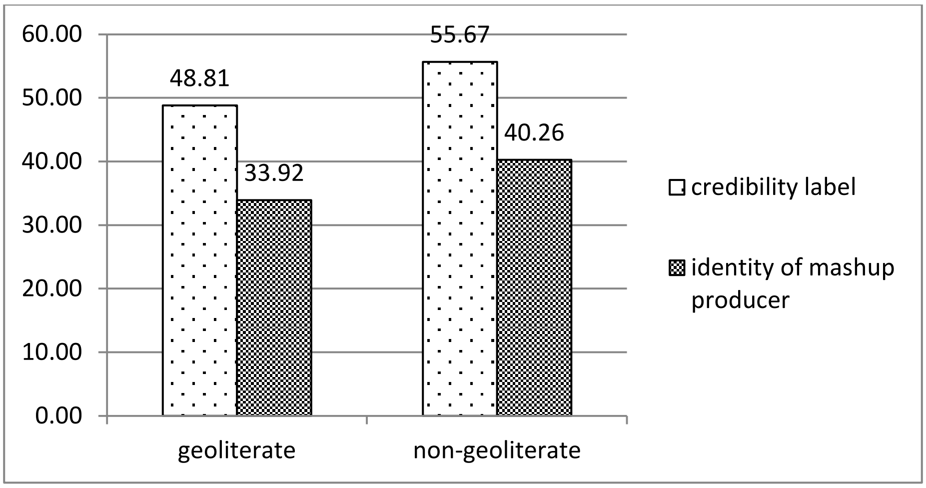


Chart 7-5 The mean value for the responses of the influence of visual credibility indicator element and the identity of map producer between groups

From the chart above, both groups tend to have similar pattern of agreement, to the influence of a credibility labelling element and the influence of 'the identity of map producer' in their credibility assessment (see Chart 15 and Chart 16 in Appendix C for the responses distribution of the credibility label element between the geoliterate and non-geoliterate group).

Table 7.17 presents the score classification that equally divided the categories of influence into three levels. According to this table, both group perceived intermediate influence to the credibility labelling on map. Non-geoliterate group perceived positive intermediate influence to the 'identitiy of map producer' (mean = 45.74; SE = 5.7). Geoliterate group however perceived low influence to this element (mean = 33.92; SE = 7.5).

From the Shapiro-Wilk significance test, the data from the geoliterate and non-geoliterate groups appeared to be non-normal, $D(28) = 0.31$, $p < 0.001$ and $D(47) = 0.21$, $p < 0.05$, respectively. The significant value of less than 0.001 indicates deviation from normality and resulted in rejection of the assumption to apply the independent t-test.

The Mann-Whitney non parametric test was then selected to test the hypothesis. From Mann-Whitney, the influence of 'the identity of map producer' from geoliterate respondents (Mdn = 16.66) did not differ significantly from non-geoliterate respondents (Mdn = 50.0), $U = 552.0$, $z = -1.2$, $p > 0.5$, $r = 0.1$ (small effect size).

Table 7-17 Classification of scores

Total scores (in percentage)	**Categories**
0 – 35	Low
36- 70	intermediate
71 - 100	High

Transcript excerpts from think-aloud experiment as below confirmed this. Although this study speculated that all the geoliterate respondents will be influenced by the element of 'map producer' when making judgement, however some of them were not. For examples;

'I chose Map B, route 1 because in terms of design, it was more attractive, neat and was not too crowded…I did not influenced by the map producer. I just focused on the events and all the symbols' (Am, geoliterate)

Therefore, the **Hypothesis 9 is not supported.**

7.4.7 Results: the influence of map mashup producer (author)

Hypothesis 2 is:

The metadata related to source (i.e map mashup producer) has a significant influence in respondents' credibility assessment

Descriptive analysis of the influence of 'the identity of the map mashup producer (author)' was conducted. This was to examine the influence of single-item metadata in respondents' assessment when judging map information credibility. Of the responses that chose 'high credibility map', descriptive analysis of the results yielded mean responses M=3.91, standard deviation (SD) = 2.44, and median =4 (undecided). Of the responses that chose 'low credibility map', descriptive analysis of the results yielded mean responses M=2.30, standard deviation (SD) = 1.87, and median =1.5 (undecided) (see Chart 15 in Appendix C for the response distribution of the influence of 'map producer' between the geoliterate and non-geoliterate groups).

Further analysis was conducted to compare the responses between these experiments, Experiment 4 with previous experiments. The response rating data for 'the identity of map producer'- were recoded into one binary values, which were 1 = no influence and 2 = influence. A cross tabulation table, based on Table 7.18, was generated to compare this in two different experimental contexts. This table contains the 4 categories – the presence or absence of a credibility rating (Exp4) and the category of influence – (Yes) influence or (No) no influence. Pearson's Chi-square test was then conducted. From the test, there was no significant association between the presence of a 'credibility rating' and whether or not the identity of the 'map producer' influenced respondents' judgement, $\chi 2(1) = 0.021$, $p > 0.8$. **There are no significant different between the influence of 'map producer' in these two experimental context.**

Table 7-18 Frequency of responses to the influence of 'identity of map producer' between Experiment 4 and Experiment 3

	Influence		
	Yes (rating = 2)	No (rating = 1)	Neutral/undecided/do not know (rating = 0)
Variable: identity of map producer			
Experiment 4 (n = 75)	29 (38.7%)	41 (54.7%)	5 (6.7%)
Experiment 3 (n = 133)	51 (38.3%)	69 (51.9%)	13 (9.8%)

Further analysis was conducted to compare the ratings between these experiments, Experiment 4 with previous experiments. The response rating data for 'the identity of map producer'- were recoded into one binary values, which were 1 = no influence and 2 = influence. A cross tabulation table, based on Table 7.18, was generated to compare this in two different experimental contexts. This table contains the 4 categories – the presence or absence of a credibility rating (Experiment 3 vs. Experiment 4) and the category of influence – (Yes) influence or (No) no influence. Pearson's Chi-square test was then conducted. From the test, there was no significant association between the presence of a 'credibility rating' and whether or not 'the identity of the map producer' influenced respondents' judgement, $\chi 2(1) = 0.021$, $p > 0.8$. From the odds ratio, the odds of the no influence of 'the identity of the map producer' in respondents' judgement was **1.1 times higher if the map were not labelled with a credibility rating than if labelled.**

This might indicate the 'map producer' element is not the dominant element that influences users in credibility assessment. Some users will be influenced by this element; whereas others do not perceive this element as important. With or without stamped credibility rating label on a map mashup do not have significant influence to the perceived important on this element. Excerpts from think-aloud experiments as below;

'who produced the map doesn't help much. We have to try the map first'

(Fad, geoliterate)

'I did not look at the map producer. It was not important. The more important was on the clarity and easy to use' (Fik, non-geoliterate)

'I did not care about the map author [producer]. On Google Map, they do not mention the author; the author or who produced the map is not important'

(Fa, non-geoliterate).

Some respondents were confused about who was the producer of the map mashups. They perceived Google was the producer of the map mashup application because it used Google Map as its base map. This misconception is contradicting the fact that the producer of map mashup may not limited to one source; the sources of map mashup are categorised into two - background and foreground data. Background data is drawn from which base map is used in the application; for example, Google Map, Bing Map; foreground data might be drawn from other sources including the news and local data from the producer who created the map. This misunderstanding has been identified from the excerpts below;

‘On Google Map, they do not mention the author; the author or who produced the map is not important’ (Fa, non-geoliterate).

‘These two maps used similar base maps. There was no specific update on the last updated date. Just 2013. Those maps were produced by Google. I just looked at the ‘powered by’ [Google’s copyright at the bottom of base map]’

(Fik, non-geoliterate)

Therefore, **hypothesis 2 is not supported**.

7.5 Discussion

Results of the survey indicate significant differences between the responses perceiving Mashup A (map labelled with high credibility rating) and Mashup B (map labelled with low credibility rating) as having more believable (credible) information. In a previous experiment, Experiment 3, which provided an experimental context without any credibility label on the map, the ratio of respondents that chose Mashup B and Mashup A was 1:1. Meanwhile, Experiment 4, which provided an experimental context with a credibility rating on the map, demonstrated the odds ratio of respondents that chose Mashup B and Mashup A as 1:3. Statistical testing using the Pearson Chi-Square yielded a significant association between the two contexts – presence of credibility labelling and absence of labelling- with the choice of the ‘high credibility’ map. Hence, **hypothesis 6 is supported.**

Analysis of the total scores that measure the influence of credibility labelling demonstrates a moderate influence on respondents’ judgement. These findings provide useful insights into the positive influence of the credibility ratings element on respondents’ judgement, although the impact falls within the moderate level. This level of impact however was comparable with the influence of visual cues, which in this study comprised ‘colour scheme’, ‘symbol design’ and ‘overall presentation’, in respondents’ judgement.

Analysis on the total scores measuring the influence of visual cues, which consist of 'colour scheme', 'symbol design' and 'overall presentation' elements, indicates that the level of influence falls in moderate category. Comparative analysis in the previous experiment, Experiment 3 demonstrated a significant association between the presence of credibility labelling and whether or not visual cues (colour scheme and symbol design) have an influence on respondents' judgement. The effect based on the odd ratio, the influence of colour scheme, was 8 times higher if the map was not labelled with a 'credibility rating' than when labelled. Likewise, the influence of 'symbol design' was 3.3 times higher if the map was not labelled with a credibility rating than when labelled. Therefore, **Hypothesis 7 is supported where the influence of visual cues were decresed.**

Analysis of the difference in responses between the two groups, which vary in terms of knowledge background- namely geoliterate and non-geoliterate, demonstrate no significant difference in agreement about the influence of visual cues. On average the influence of visual cue elements from geoliterate respondents (Mdn = 61.1) was lower than from those of non-geoliterate .respondents (Mdn = 72.2), U = 521.5, z=-1.5, p>0.12, r = 0.2 (small effect size). The different however not statistically significant; therefore, the **Hypothesis 8 is not supported.**

Analysis of the total scores of response differences between the two groups –geoliterate and non-geoliterate groups- to the influence of a credibility rating on respondents' judgement demonstrates no significant variation. Both groups rated were in slightly agreement (point-5) about the influence of credibility labelling. On average the influence of the credibility labelling from geoliterate respondents (mean = 48.8, SE = 6.6) was lower than from non-geoliterate respondents (mean = 55.67, SE = 4.2). This difference was not significant $t(73) = -0.92$, $p > 0.05$: however, it did represent a small sized effect $r = 0.1$. Therefore, the **Hypothesis 9 is not supported.**

Analysis of the influence of the 'identity of the map producer' indicates the respondents' ratings were, on average, at an undecided/neutral point. Half of the response distribution demonstrated a wide variation, ranging from disagree to strongly agree. The proportion of influence rating on a positive to negative continuum was in a ratio 1:1. This is consistent with the findings from the previous Experiment 3 (see Chart 6.7) and Experiment 2 (Table 5.32), where half of respondents perceived this metadata related to source-authority variable as important or having a positive influence on their judgement, whilst another half perceived it as unimportant. Moreover, comparison analysis between Experiment 3 and Experiment 4 demonstrated no significant association between the presence of credibility labelling and whether or not 'the identity of the map producer' has an influence on respondents' judgement. **There is no significant different between the influence of 'map producer'**

in these two experimental context. The odd of the no influence of the 'identity of the map producer' in respondents' judgement was **1.1 times higher if the map were not labelled with a credibility rating than if labelled.** From Mann-Whitney, the influence of 'the identity of map producer' from geoliterate respondents (Mdn = 16.66; mean = 34.21) did not differ significantly from non-geoliterate respondents (Mdn = 50.0; mean = 40.26), U = 552.0, z=-1.2, p>0.5, r = 0.1 (small effect size). Therefore, the **Hypothesis 2 is not supported**.

It is worthy of note, however that the lack of positive influence of metadata related to source, which in this study was 'the identity of the map producer', may be due to the age of respondents in the sample. On average, the samples in the geoliterate and non-geoliterate groups were aged 21 years. Kruskal-Wallis test was conducted to check this hypothesis. The results (see Table 27 and Table 28 in Appendix C) indicate, however the influence levels of the 'identity of map producer' were not significantly affected by the age of respondents, $H(4) = 3.4$, $p > 0.5$. Analysis of gender demonstrated also no significant difference in the influence of a credibility labelling on respondents' judgement (see Chart 17 in Appendix C). Table 7.19 below presents the results of hypotheses.

Table 7-19 The results of hypotheses in this experiment

	The hypotheses statements in this experiment	Results
Hypothesis 6	Credibility label has a significant influence in respondents credibility assessment	supported
Hypothesis 7	The presence of credibility label has a significant effect to the influence of visual cues in respondents' judgement	supported
Hypothesis 8	There is a significant difference in the influence of visual cue variables between geoliterate and non-geoliterate respondents when making judgement	Not supported
Hypothesis 9	There is a significant difference in the influence of credibility label between geoliterate and non-geoliterate respondents when making judgement	Not supported

This study demonstrates the impact of credibility labelling using colour coded traffic light ratings (CCTL) for online mapping and particularly for map mashup. The probability of respondents making informed judgements by choosing a high credibility map based on this rating label is three times higher than the setting without the label. This finding is in line with the study of Kelly et al. (2009), which demonstrates that the probability of users identifying healthier foods from a traffic light labelling format was five times more likely than from a label using monochrome text information.

The lack of influence of metadata related to sources-authority (i.e. the identitiy of the map producer) in respondents' judgement, whether respondents were from a geoliterate or a non-geoliterate background or from a different gender, emerged as a concern in this study. A few explanations for this finding are;

1) There is a trend to perceive the source or author of a site in a Web environment as of little importance. Warnick (2004) has pointed out changes in credibility assessment that have been applied in a web environment compared to non-digital media. The ubiquitous and lack of format standardisation of the placement of this variable on a website may have led to the low importance of this element when assessing credibility. The extra task needed for checking this variable, as well as the absence of this element in some websites, may gradually influence the low perception of this element in a web environment. This trend may provide an explanation of the perceived lack of importance of the 'identity of the map producer' on map mashup information.

2) This may be due to low motivation on the part of respondents to engage deeply with the experimental task. Analogy of low motivation users have been described by Fogg et al. (2003) as users who browse and surf websites where they occasionally evaluate information critically; but they may be highly motivated when scrutinising information relevant to a specific critical need, such as seeking information to find a cure for cancer. Petty and Cacioppo (1986) identified that users in such a group would rely on peripheral signals (e.g. visual design, aesthetic etc.) when assessing credibility, rather than base their assessments on critical elements (e.g. sources, currency etc.). The proposal of a credibility rating label may reduce the extra checking activity that is inevitably needed when users perform credibility assessment. The strategy that focuses on visual elements to attract users' attention in low motivation groups is well established in marketing and advertising products. A study by Fogg et al. (2003) supports the view that visually related elements are widely used to determine the credibility of online

information. The proposed CCTL ratings label recommends a solution for users who are sometimes low in motivation, enabling them to scrutinise critical elements when judging the credibility of information.

7.6 Conclusion

This experiment provided some useful insights and strengthened the findings generated from previous experiments. It examined the influence of visual cues (i.e. colour scheme, symbol design, and overall presentation), credibility labelling and the identity of the map producer in a context that proposes credibility labelling scheme on map mashup applications. Several useful insights resulted from an experimental comparison between a context that presents credibility labelling and one without labels:

1) The number of responses that chose ‘high credibility map’ (Map A) rather than ‘low credibility map’ (Map B) is three times higher in the context that presents credibility labelling on the map. The context without credibility labelling has demonstrated no significant difference between the responses that chose Map A and Map B, where the ratio was 50:50.
2) There was a decrease in the perceived influence of visual cue variables, which in this study were colour scheme and symbol design, in the responses. Based on the odds ratio, the influence of colour scheme was 8 times higher, and the influence of symbol design was 3 times higher in a context without a credibility label.
3) There was no significant change in the influence of identity of the map producer between the two contexts. The proportion of respondents who rated this as having influence and those who rated no influence was 50:50. This finding strengthens the findings from Experiment 2 and Experiment 3 where half of respondents perceived this metadata related to source element as an important influence on their credibility assessment, whilst the other half perceived this element as unimportant.

Moreover, this study pointed out the limited use of the ‘identity of the map producer’, which is usually recorded as one metadata element, to provide a positive influence on respondents’ judgement when assessing credibility. The trend in the web environment, where the source of information may not be emphasised, may lead to a lessening of importance in this element among web consumers when judging information credibility. This seems implied in the map mashup environment, due to the use of the website as the medium of dissemination.

This study demonstrates the impact of credibility labelling using colour coded traffic light ratings (CCTL) for online mapping and particularly for map mashup. This labelling system could encourage the mashup developers to improve the quality of map contents in order to achieve favourable ratings. However, a standardised design label of CCTL should be applied for all online map mashup products on the Web in order to reduce map users' confusion of label meanings. A public education campaign to accompany the introduction of CCTL labelling (Kelly et al., 2009, p.127) to help users interpret the labels according to the mapping guidelines would be a sensible part of the introduction of such a scheme.

8 DISCUSSIONS

8.1 Introduction

Four experiments have been conducted to achieve the first and second objectives of this research; the first objective which is 'to examine the influence of metadata related map producer and map data supplier on respondents' assessment of credibility was conducted in Experiment 1, 2 and 3. Five hypotheses were tested in these experiments. The second objective of this research which is 'to examine the influence of colour coded traffic light labelling on respondents' assessment of latter experiment. Four hypotheses were tested in this experiment.

The responses in the first experiment (Experiment 1) were collected from open-ended questionnaires. This methodological approach was adopted so that respondents could provide answers in their own words without being directed. In the second experiment (Experiment 2), some changes were made in the approach to collecting information from respondents. This was due to the difficulty of analysing data from open-ended questionnaires, which tend to produce responses that focus on a single item only. Therefore, in Experiment 2 respondents were assumed to have several factors influencing their judgement and, hence, were allowed to rank the influences in order of importance. Spot the differences activity was included in Experiment 2 at the first section of the questionnaires; the purpose was to implicitly suggest respondents to comprehensively analyse and notice the differences between maps before giving answers on the next section. In Experiment 2a, the influence of visual cues on the maps was controlled to implicitly suggest respondents to notice the differences between the foreground data suppliers. In Experiment 3, some changes were made to the tasks given. In this experiment, the tasks were designed to stimulate a sense of deep involvement: a respondent was given a specific role during the experiment and was required to act as in an emergency situation. In Experiment 4, respondents were given a slightly similar task that sharing a few common properties and dataset as in Experiment Three. A major difference between these two experiments was the presence of the visual credibility indicator that was presented as a colour coded traffic light (CCTL) rating label. Table 3.2 presents the experimental settings differences.

As explained in the methodology chapter (Chapter 3), this research applied mixed method research design namely embedded experimental model;

in this framework, think-aloud protocols and observations were conducted to support the results of quantitative data that were collected through online map based questionnaires. Table 3.3 presents the procedure of the think-aloud protocols; Appendix D and E contain the observations and transcripts of the sessions. From the observations, all six respondents seemed carefully read and analyse the maps in order to solve the experimental tasks. They switched between the two maps a few times to make comparison and final decisions. However, two of them did not know the availability of zoom in and out functions and the 'identify' function on the interactive maps. They were too focused at the centre of the map to solve the tasks and insensitive to other parallel elements that could influence credibility of information. In Experiment 3, all six respondents did not look (notice) at the map producer label stamped on the top sidebar of the map. They only noticed that element when the observer highlighted the presence of that element embedded with the maps. Only three of them said that the element might have influenced them if they had noticed it before making decisions. In Experiment 4, only two respondents noticed the presence of and were influenced by the credibility rating label stamped on the top of the map before making decisions. Another two respondents did not notice the presence of the rating label but believed that the element might have influenced them if they had noticed it when making credibility assessment. The other two respondents said it would not have had any influence if they had noticed the label.

Next sections summarise the results from these four experiments according to the main findings.

8.2 The influence of visual cues

From the three experiments that conducted in this research, there was high influence of visual cues, specifically related to colour scheme, symbol design and overall presentation when respondents making judgement related to credibility of map mashup information. In Experiment 1 that applied open-ended question of 'the basis of respondents' judgements in selecting the maps that they will use' indicates the high influence of 'colour scheme' and 'information clarity' emerged in the responses (see Table 4.6). This finding is **not supporting the Hypothesis 1** of this experiment (Experiment 1) that postulates that visual cues have no significant influence in users' credibility assessment. Although the results shows the 'colour scheme' concept was the most dominant keyword found in respondents' answers, the frequency of this concept with the second dominant concept, which is 'information clarity', was statistically insignificant.

The findings from textual analysis on the responses indicate that the ‘colour scheme’ concept that emerged referred to keywords relating to colour selection and colour combination. Meanwhile, the emerged ‘information clarity’ concept referred to keywords that relate to discrimination of the text, features, patterns, colours and the understanding of the meanings of the signs on the maps, such as ‘easy to read’, ‘able to define’ and ‘distinguishable’. As stated by MacEachren (2004, p.213), maps are imbued with meaning by virtue of semiotic relationships. Semiotics is the science of signs, with a sign considered to be a relationship between expression (the sign-vehicles) and its referent (content). Colours, symbols and patterns are the sign-vehicles that represent objects in the real world that are subject to the interpretation of cartographers and map users. The meanings in maps can be interpreted either by reference to a map legend or assumed to be part of the normal readers’ general map schema (e.g. blue is water) (MacEachren, 2004, p.311). Colour schemes used in maps have explicit meanings that represent spatial features. The colour scheme of a map is not like a regular colour used on a textual based medium, but it represents special functions to deliver messages to map readers.

Although ‘colour’ is the dominant keyword found in respondents’ answers, this keyword tends to emerge with other keywords to describe the relation of colour to the clarity of information, combination of schemes used, the influence of overall presentation, information details, the design of the map and individual preferences, as shown in Table 4.7. This indicates that the keyword ‘colour’, found in respondents’ answers, is not a single keyword but has emerged to relate with other concepts. From Table 4.7, ‘clarity’ and ‘combination’ tend to be the dominant keywords used to relate to the colour keyword.

The results shown in Table 4.8 and Table 4.9 also indicate no difference in responses between the group of respondents who have an academic background that is geospatially related (i.e. geoliterate) and the group of respondents drawn from other domains (i.e. non-geoliterate). These surprising results indicate that the majority of geoliterate respondents tend to use the concept of visual cues when making judgements to select the preferred maps in the tasks. This is in line with the non-geoliterate group responses, which show the dominant use of the keyword of ‘colour’ when making judgements in the tasks. Hence, these findings **not supporting the Hypothesis 3** of this experiment that respondents drawn from geoliterate group would be more aware of the critical metadata elements and will use those as the basis of judgements when selecting and rejecting a map.

Further experiments were conducted to confirm these results. In Experiment 2, the ‘colour scheme’ and ‘design look’ which were used to measure the influence of visual cue construct became the dominant element again. The total score of this visual cue construct was at median 75 (see Table

5.15). Wilcoxon test was conducted to compare the score of influence of visual cues construct with data supplier. From this test:

‘the scores of the influence of visual cues was significantly higher than the influence of ‘data supplier’.

Hence, **hypothesis 1 is not supported**. This finding supports the results of previous experiment. Although the influence of visual cues was controlled in Experiment 2a by a similar colour scheme but with a very slight dissimilarity of symbols on both maps, the proportion that perceived the Map A (i.e. the supplied by the Starbucks Coffee) more believable due to the influence of visual cues was 50% of the total respondents. One excerpt from one respondent in open-ended question (Q3) in Experiment 2a that supported this statement is as below;

‘The only downside to Site A was the source of foreground data. The Site B although good lacked the bus sign and also had the coffee house sign board which was less impressive’ (Age 27, non-geoliterate)

In Experiment 3, the experimental context was modified to change the level of tasks. The reviews related to map use levels have been discussed in details in Chapter 2 (see Table 3.2). Liebenberg (1998) has summarised the levels of map use from the studies by Olson (1978), Muehrcke (1979) and Board (1984); level one comprises of map reading tasks such as identification of individual symbols (lines, polygons, points) and the differences of these symbols’ shapes, relative size, level two tasks comprise of recognition of the spatial pattern where at this level, users are still visualise the data in a form of ‘space’. According to Brown and Perry (2001), ‘space’ is a term to describe the object and event that occur in 2D and 3D views but still not relate to the actual event occurs in the real world; ‘place’ is a term to describe the space according to its particular geographic location. Level three comprises of interpretation tasks where map users will use other information and also tend to relate their knowledge and previous experiences to answer the ‘why’ geographic questions at that ‘place’.

In Experiment 1 and 2, the level of tasks was at Level One. The tasks were at the lowest map use level where respondents have to compare the two maps and decide which map they perceived credible to use in their self-campus tour. In Experiment 3, the task was at the high level where respondents were requested to compare the maps and suggest one best route for an ambulance to evacuate trapped victims during earth quake disaster. In this experiment, the emergency context was used to create a sense of critical situation where respondents have to make informed decision before giving their answers. The purpose applying disaster context was not to replicate a situation of a real disaster; it was intended to increase the level of respondent engagement with the task in the experiment. This was also to decrease the gap between the artificiality of the Experiment 1 and 2. A deeper engagement of respondent was required to complete the task in Experiment 3. In Experiment 3 the ‘colour scheme’, ‘clarity of symbol’ and ‘symbol design’ which used to measure the influence of visual cues had a high number of responses (approx. 94%) (see Chart 6.4). The visual cues appeared to be the main factor influencing respondents’ judgement in these three experiments. This finding confirmed the results of Experiment 1 and 2. Therefore, **hypothesis 1 is not supported**.

The results from think-aloud observation confirmed this. The basis of respondents’ judgements on choosing the map was due to its visual cues, specifically the colour scheme and the symbols used. Extraction from the audio transcripts (translated from Malay to English) indicated as below;

‘I chose this map because the presentation was not too crowded and the symbols used were easy to understand’ (Am, geoliterate)

‘I chose this map because the colour [red, black] and presentation looks serious and suit for this critical situation (context)… the another map (Map B) is more suitable for public user’ (Fad, geoliterate)

‘I chose this map because it looks attractive. Another map used black. It blocks the building label’ (Fik, non geoliterate)

‘I chose this map because it used the symbol road block instead of road construction to represent the no access route’ (A, non geoliterate)

In a study by Williams (1967) that examines how a person searches for a visual object, colour became the main basis during a search. In that study, a person had to search for a visual object in a cluttered visual field. It was found that subjects were much better at discriminating and identifying an object in a cluttered visual field using the colour characteristic than the object’s size or shape. This may support the findings in this experiment, where the ‘colour scheme’ concept is dominant in respondents’ answers, whether drawn from the geoliterate or non-geoliterate respondents. ‘Colour scheme’ seems to become the basis of their search for flaws in the comparison maps. The identified flaws then become the basis of their judgement to decide which map is perceived as credible to assist them in their tasks.

David and Jason (2008) argue that judgements based on visual cues would operate within the first few seconds when a respondent makes first time contact with the online medium; in this phase users tend to make judgements based at an intuitive level. Later on, users tend to rely on cognitive judgement when they proceed to scrutinize the contents in depth. At a cognitive level of judgement, users tend to measure more critical elements than visual cues such as the information details, accuracy and the authority of the information. This argument might support the findings of the present study where the factor of ‘usefulness’ tends to emerge as a factor that influences respondents’ judgement. The influence of this factor was as dominant as the influence of ‘colour scheme’ on judgement and indicates that respondents made critical judgements at an intuitive and a cognitive level in the experimental tasks.

8.3 The influence of metadata

As mentioned earlier, the first objective of the four experiments conducted in this research was to examine the influence of metadata related to sources (i.e. map data producer and map data supplier) on respondents’ assessment of the perceived credibility of map mashup information. From the results, metadata elements, which in this study were ‘the identity of map producer’ and ‘map supplier’, do not seem to have much influence on respondents’ judgements. From the comparison analysis between the geoliterate and non-geoliterate group in each experiment, there were no significant differences between the responses to these critical metadata elements.

In Experiment 1, the results show that only an average of 6% respondents use the identity of the map creator/author (map producer) as the main basis for their judgement when selecting maps for the tasks (see Table 4.9). Therefore, **Hypothesis 2 is not supported**. In Experiment 2, the metadata related to sources (i.e. map data supplier) was still not perceived as the important element measured in respondents’ judgement when assessing map information. These metadata were measured by half of the respondents, whilst the other half did not measure these factors in the assessment. From the results in the three experimental conditions, these factors were not ranked as the most important influence (see Table 5.11). There were variations in the order of priority and the differences in response were not significant between the first to fifth orders of importance.

The Table 5.12 presents the results of ‘spot the differences’ activities (in Experiment 2) that conducted before respondents analyse and choose the map that they perceived credible for the given task (Experiment 2). From the results, on average, there was no significant different within the sample that spotted the differences of data supplier parameters with the proportion that measured (ranked) and not measured data supplier. In other words, respondents noticed the ‘data supplier’ of both maps was different; however these did not necessarily influence them to assess credibility of information by using that metadata.

Further online map based questionnaires (i.e. Experiment 2a) was conducted to confirm the findings of previous Experiment 2. In this additional experiment, several changes were made to the experimental design; the intention was to control and minimise all the visual design factors that might affect respondents' judgement in previous Experiment 2, but concentrated on the differences of reputation levels of foreground data suppliers. In this experiment, the data and colour scheme between maps were controlled to be identical. However, the symbols of points of interest (i.e. café and bus stop) were designed to be slightly a little bit different; the intention was to prompt respondents to engage with the maps exercises (i.e. to choose the map that they perceived credible) but with a very minimal influence by the visual cues. The data suppliers of maps were manipulated to be easily distinguishable in terms of their reputations in supplying data related to the campus (i.e. The Starbucks Coffee versus the University of Nottingham).

The results of this experiment (see Table 5.18) demonstrated the number of respondents that chose either Map A and Map were not significant different. Of the sample that chose Map A as more believable (i.e. the foreground map data supplied by the Starbucks Coffee), 70 respondents stated their reasons to choose that map were due to the influence of visual cues (e.g. 'the design of map', 'the readability of labels', 'the clarity of map symbols'). Meanwhile, 68 respondents chose Map B (i.e. the map data supplied by the University of Nottingham). Of this sample, 49 respondents were influenced by 'the foreground data supplier'; whereas other 19 respondents chose this map not due to the data supplier, but had been influenced by the visual cues of the map. There was also no significant different within the geoliterate group that chose either Map A and Map; the results indicated seven geoliterate perceived Map A as more credible and ten geoliterate perceived Map B as more credible. These findings supported previous Experiment 2 results where the proportion that had influenced and not influenced by 'foreground data supplier' was in the ratio of 50:50. Hence, **Hypothesis 2 is not supported**.

In Experiment 3, 51.9% respondents perceived 'the map producer' element to be unimportant in influencing their judgement pertaining to map information credibility. Likewise, a high number of responses rated the importance of 'the map producer' element at the undecided point (9.8%) compared to other manipulated elements. The response differences were statistically significant in the Chi-squared test $\chi 2$ (1, n = 16) = 6.25, $p < 0.05$) (see Chart 6.4). From the think-aloud protocols and observation, two respondents mentioned 'map producer' element as not important in their decision; this element was not considered when they made the judgement to choose the credible map. For example;

'Map producer element has no influence in my judgement. I just focused on the events [task] and on the centred of map [map producer element was displayed at the sidebar, the top corner of map]'

(Am, geoliterate)

‘I did not look at the map producer [which displayed at sidebar, at the top corner of map]. The important is about the clarity of information.... I just looked at the ‘powered by’ label [Google’s copyright at the bottom of base map]’

(Fik, non-geoliterate)

The results between Experiment 3 and 4 were compared to analyse the influence of metadata related to source (i.e. map producer) in two different contexts (with and without credibility label). From the Pearson’s Chi-square test there was no significant association between the presence of a credibility rating and whether or not ‘the identity of the map producer’ influenced respondents’ judgement, $\chi2(1) = 0.021$, $p > 0.8$. From the odds ratio, the odds of the ‘no influence of the identity of the map producer in respondents’ judgement’ was 1.1 times higher if the map were not labelled with a credibility rating than if labelled (see Table 7.19).

This might indicate the ‘map producer’ element is not the dominant element that influences users in credibility assessment. Some users will be influenced by this element; whereas others do not perceive this element as important. With or without stamped credibility rating label on a map mashup did not have significant influence to the perceived important on this element. Excerpt from think-aloud protocol in experiment 4 as below;

‘Who produced the map doesn’t help much. We have to try the map first’ (Fad, geoliterate)

‘I did not look at the map producer. It was not important. What was more important was the clarity and ease of use’ (Fik, non-geoliterate)

‘I did not care about the map author [producer]. On Google Map, they do not mention the author. The author or who produced the map was not important’ (Fa, non-geoliterate)

Therefore, **Hypothesis 2** (i.e. the metadata related to sources have significant influence in respondents’ credibility assessment) **is not supported**.

From the experiments that have been conducted, differences in respondents’ background could not be a reason for the increments and decrements of the level of influence to the tested variables. Difference in backgrounds may not result in the high perceived influence of ‘the identity of the map producer’ in respondents’ credibility assessment. For example in Experiment 1, an average of only 2% geoliterate respondents mentioned they chose the maps due to the influence of ‘map producer’ in their credibility assessment. In Experiment 2 and Experiment 2a, there were no significant differences in the level of priority rated between the two groups. Hence, **the hypothesis 3 is not supported**. The ratio that measured and did not measure within the groups in each condition was 1:1 (see Table 5.13). Although the geoliterate respondents spotted the differences of data suppliers' parameters between both maps (Experiment 2), they were not necessarily influenced by that sources and assessed it in their judgements (see Table 5.12). Moreover reducing visual cues impact in the experimental setting of Experiment 2a did not necessarily able to promote geoliterate respondents to assess metadata related to data supplier in their credibility assessment (see Table 5.12).

In Experiment 3, the number of responses that perceived the element of 'map producer' as important in their judgement however was less than 50% in the two groups (37% and 42%) (see Chart 6.7). Meanwhile, the responses in the two groups that perceived this element as not important in influencing their judgement was comparable, with both groups showing the number of responses to be around 52% (see Chart 6.8). The findings from the think-aloud, confirmed this. Of the two geoliterate respondents, only one of them perceived the important of 'map producer' element. The excerpt from the transcripts as below;

'Ooh, if I had looked at this element [Sarah Smith versus University of Nottingham], it might have influenced me. The map came from [produced by] the University is more credible, more believable' (Fad, geoliterate)

However, another geoliterate respondent responds as below;

'I was not influenced by that element [the producer of map]. I just focused on the events and the symbols used' (Am, geoliterate)

Hypothesis 3 is not supported. Some of geoliterate respondents influenced by the element of 'map producer' and used this element to support their decisions; some of them did not have any influence to this element.

Experiment 4, from Mann-Whitney test, the influence of 'the identity of map producer' from geoliterate respondents (Mdn = 16.66; mean = 34.21) did not differ significantly from non-geoliterate respondents (Mdn = 50.0; mean = 40.26), $U = 552.0$, $z=-1.2$, $p>0.5$, $r = 0.1$ (small effect size). The question is whether the lack of positive influence of metadata related to source-authority, which in this study was the identity of the map producer, may be due to the age of respondents in the sample. On average, the sample in the geoliterate and non-geoliterate groups was aged 21 years. Kruskal-Wallis test was conducted to test this (see Table 36 Appendix C). The results indicate, however the influence levels of the 'identity of map producer' were not significantly affected by the age of respondents, $H(4) = 3.4$, $p > 0.5$. Hence, the influence of tested metadata do not related with gender and age groups of respondents in the sample.

However it is worth to highlight that some respondents might not influence with the ‘map producer’ of the experimental maps due to the reputation of the comparable map sources did not hold a high reputation as the national mapping providers in supplying maps. There was one excerpt from think-aloud protocol supports this as below;

‘No. this element [map producer] will influence me if the producer was the NASA or from mapping department. If it was produced by a university, it is only an academia. Not from an authorised sources’ (Am, geoliterate)

Although only one geoliterate respondent raised this statement, future research should test to further investigate this issue by comparing sources that highly reputable in producing and providing maps and geospatial data such as the national mapping providers. However, results of Experiment 2a demonstrated that the significant different of reputation levels between two comparison maps (which in this case, the Starbucks Coffee versus the University of Nottingham) did not necessarily suggest a dominant influence of ‘foreground data supplier’ in users’ judgements (see Table 5.12).

These findings however were inconsistent with the results found in the non-context, prominence independent setting conducted in the Experiment 3 where respondents made judgement (interpretation) of the influence of element based on statements. Metadata related to source, which in this case were the importance of the ‘website’s affiliation’ (i.e. the affiliation of website that hosted the map) and the ‘foreground and background map data supplier’, were perceived as important in 80% of responses. This is inconsistent with the findings in Experiment 2, 2a and Appendix B (Table 30), where the ratio that perceived the importance of these two metadata elements in their assessment of credibility was about 50:50.

From the non-context, prominence independent setting (Experiment 3) results where interpretation was based on statements, the importance of the ‘currency’ of map data was perceived higher other tested metadata, at about 93% and in line with the responses for the importance of visual cues (92%). This is inconsistent with the finding of the previous experiment where only 5% of responses indicated the element of ‘currency’ in their judgement when selecting or rejecting a map (see Table 7 Appendix A). Nevertheless, the importance of ‘map producer’ was lower than other tested metadata elements, at about 75%. However, this response was still higher than the responses in the context prominence dependent setting, where the importance of the ‘map producer’ element was only rated by 38%. Therefore, the **Hypothesis 4** as following **is not supported**;

‘There is no significant difference between the level of importance of the metadata related to sources between these two contexts - a prominence dependent (i.e. interpretation is based on what looks prominent) and a prominence independent setting (i.e. interpretation is based on a statement)'

One explanation for this inconsistency might be the difference in the ways users interpret the importance of credibility elements. In the context prominence dependent experiment, the applied approach involved judgement about the importance based on prominent elements that a respondent notices. Meanwhile the non-context prominence dependent experiment involved judgements by interpreting statements given in the questionnaire. Credibility assessment via a prominence based approach is supported by the theory of Prominence-Interpretation (Fogg, 2003). This theory posits users' assessment of information credibility occurs in two stages;

1) users would notice element(s) on the map that look prominent;

2) users will make a judgement about information credibility by interpreting the element.

During assessment of online information credibility, users will judge on the basis of what they notice. If they do not notice an element, no judgement will be made. In the context dependent setting, respondents might judge on the basis of what they see and notice when assessing the map. This is in line with the findings found in Fox (2006, p.11) where three-quarters of users (which in this case health seekers) fall into the 'unconcern' category where they check the source and date only 'sometimes, hardly ever, or never'. This finding then generalised into 85 million Americans who gathering health advice online without consistently examining the two key of credibility indicators. To test this, the observer raised a question whether the 'map producer' influence her decision after respondents seem did not notice this element. Excerpts from think-aloud observation supported this statement as below;

'Ooh, if I had looked at this element [Sarah Smith versus University of Nottinham], it might have influenced me. The map came from [produced by] the University is more credible….more believable' (Fad, geoliterate)

'Err…actually I did not look at who produced the map. Yes, Map A is more believable because it produced by the university. Map B is less believable because it produced by nobody, we did not know the background of the author. She might make up the data. So it is less believable. The Map A was produced by an authorised source, we know its [reputation] and we can believe [the source] (Fa, non-geoliterate)

'Ooh, I just realised this element. Yes indeed. Map A was produced by the university. We consider it's from an authorised source compared to Map B that produced by unknown individual' (R, non-geoliterate)

‘I think the map produced by the university is more credible. Map B was produced by private individual isn’t? I think it is better to have a map produced by one organisation compared to an individual’ (A, non-geoliterate)

Nevertheless there were respondents that did not influenced by the ‘map producer’ element although they noticed the element. Excerpt from think-aloud protocol that supported this statement as below;

‘Although if I had noticed, I would not have been influenced. I would look [focus] on the main purpose.’ (Am, geoliterate)

‘I do not care who produced the map. The more important is the map is easy to read and clear. And it is easy to use. It is enough for me to look at who hold the copyright’ (Fik, non-geoliterate)

A study by Elzakker (2004, p.127) that investigated how map users use and read maps by conducting exploratory methods that includes think-aloud protocol observation indicates similar results; there was no significant different between the novice or expert map users in reading maps; they did not put much effort into validating the maps supplied to them; they assumed what was given to them was correct. The participants were also careless about the years (currency) of the maps. Some of them did not notice the mistakes on the map; although they had learnt what is called ‘the grammar of cartographic design’, the participants tended to choose what was the convention rather than what had been documented as a rule of thumb.

In a non-context, prominence independent setting, respondents might judge by interpreting given statements, with their responses likely to be influenced by previous experiences, culture, skills, knowledge (Fogg et al., 2002, p.85) when dealing with a map. Their responses might be based on what they consider to be socially acceptable answers. This issue was raised in a study by Fogg et al. (2003); the findings from that survey, which applied the prominence approach, contradicted the findings from other study (for example see Princeton (2002)) that applied an interpretation based on statements approach in the assessment. In the first study, with regard to the authority related element, 8.8% of respondents measured the identity of site operator element when judging the credibility of online information (Fogg et al, 2002, p.23). Meanwhile, in the latter study, the identity of site operator was rated by 67% of respondents as one of the important factors when choosing websites.

Another study by Parker et al. (2012) that using interpretation based on statements approach via observations and focus groups to examine the criteria users used when choosing location based information during their kayaking trip indicates the perceived importance to the critical elements, including ‘accuracy, clarity, currency, depth and scope and quality of sources’ in their judgement. Whereas in the quantitative study by Parker (2012) confirmed this.

However, these studies were conducted within ‘interpretation based on a statements’ research framework as defined by Fogg et al. (2003) in Prominence-Interpretation (PI). As noted by Fogg et al. (2003) the findings of studies that investigate the influence of credibility elements within ‘interpretation based on statements’ method might differ from studies that conducted within a ‘prominence’ (what looks prominent) based method. The ‘interpretation based on statements’ method may at some extent is biased to what is considered to be socially acceptable answers; it does not necessarily reflect what they actually do. In this framework, users commonly give responses based on statements asked in the questionnaires whereas in the ‘interpretation based on what looks prominent’ method, user decisions are based on what they notice and what looks prominent. This is in line with Morahan-Martin (2004) that argued, there appear to be differences between what people say about how they assess online information and what they observed doing; users are aware the importance of elements of credibility, unfortunately, do not always checking the elements during assessment. Parker (2012, p.265) noted with this issue, hence highlighted the need to further investigates the extent to which respondents fully aware on the presence of the credibility elements that they thought might have influenced their judgements.

In Experiment 1 and 2, the applied stimuli did not require a deep level of engagement from respondents as they had to choose a map that they perceived credible to assist them in a self-guided campus tour. The minimal responses of the influence of metadata elements in this experiment might result from the low level of engagement of respondents with the experimental tasks. An empirical study by Ferebee (2007) supports this view by demonstrating that, in non-situational involvement (which means a low level of engagement with the context), users tend to notice more of the medium related elements: for example, the design look, design structure, functionality, security and technical capability. Meanwhile, more message related elements, such as information accuracy, usefulness, and clarity, are noticed when there is a high level of involvement.

Therefore, to test this hypothesis, changes were made to the experimental design of Experiment 3. Context was changed to simulate a sense of emergency in order to increase the level of map use tasks from Level Two to Level Three (see Table 3.2 in Chapter 3). In the new experimental design, respondents were required to analyse and suggest the best route from the map they perceived to be more credible. In Experiments 1 and 2, the level of tasks was set at Level One. The tasks were at the lowest map use level where respondents have to compare the two maps and decide which they perceived more credible to use for their self-campus tour. In Experiment 3, the task was at the higher level (i.e. Level Three) where respondents were requested to compare the maps and suggest the best route for an ambulance to evacuate trapped victims during an earth quake disaster. In this experiment, the emergency context was used to create a sense of a critical situation where respondents have to make informed decisions before giving their answers. The purpose of applying a disaster context was not to replicate a situation of a real disaster, but to stimulate respondents to engage with the exercise (experimental task). This was also to decrease the gap between the artificiality of Experiments 1 and 2 and to tackle the issue of respondents possibly becoming indifferent and hence, giving careless answers in the experiment.

Nevertheless, the metadata related to sources (i.e. map producer), tested in Experiment 3, was perceived to be a less important element, with only 38% of respondents perceiving this element as important and 52% of respondents perceiving this element as unimportant to their judgement. The ratio of 50:50 between the proportion that had influenced and not influenced by the metadata related to sources was consistent in the series of experiments in this research. Therefore, **Hypothesis 5** 'there is significant difference between the levels of perceived importance of the metadata variables related to source between these two levels of engagement contexts – low level (Level One) versus high level (Level Three)' **is not supported** (see Table 3.4) . This finding demonstrated that in the high level task (Level Three), the visual cues had more influence than the metadata related to sources (i.e. map producer and map data supplier) in users credibility assessment. This finding was an indicator that some map users were less influenced by the metadata related to sources but more influenced by the visual design and subjective cues on the map mashups. These findings, however were contradictory with the view of Ferebee (2007) who argues that deep engagement with the task appears to be the primary driver for the shift of focus in information processing (i.e. either central or peripheral) as well as the elements being noticed by respondents.

Table 8-1 Experimental setting differences between experiments

Differences between comparison maps (Map A vs. Map B)	**Data**	**Visual cues**	**Sources**	**Added activities**
Experiment 1 (pilot)	identical	different	Within same level of reputation (i.e. map producer)	
Experiment 2	identical	different	Within same level of reputation (i.e. foreground data supplier)	Spot the differences activity

Experiment 2a	identical	Minimal differences	Significant difference of reputation level (i.e. foreground data supplier)	
Experiment 3	Contradictory data	Controlled differences	Significant difference of reputation level (i.e. map producer)	experimental task level increased

Table 8.1 above presents the experimental setting differences of the experiments that demonstrated low influence of metadata related to sources in respondents' credibility assessment. The ratio that had influenced and not influenced by the tested metadata was in 50:50; this ratio also applied within geoliterate group where half of geoliterate respondents perceived the influence of the tested metadata, but another half had no influenced. In Experiment 1 and 2, although limited to surface assessment of static map, the results demonstrated low influence of metadata in conditions of data were identical between both maps, but supplied (or produced) by the sources that hold reputations at the same level; for example Map A was produced by anonymous and Map B produced by Jane Smith; the sources of these two maps might not strong enough to become a basis to select or reject the experimental maps because the suppliers hold a similar level of reputation. 'Spot the differences' activity was conducted in this experiment to implicitly 'suggest' the differences of 'data supplier' between the comparison maps. However, although some of respondents notice the presence of these metadata, but they were not influenced and judged credibility based on it. In Experiment 2a, improvement was made to test this Hypothesis 2 on the simulated maps that supplied by suppliers that hold different levels of reputations. The data between two maps were designed to be identical and very minimal differences of visual cues. The intention was to implicitly highlight the presence of data supplier parameters at the sidebar. The results however demonstrated low influence of data supplier although the impact of visual cues was controlled. In Experiment 3, the data between two maps were designed to contradict each other and increased the level of task to include high level of map use that includes the map analysis and interpretation tasks. However, the results still demonstrated low influence of metadata related to sources.

The lack of influence of metadata related elements, specifically the identity of the 'map producer' and 'map data supplier' in respondents' judgement, whether respondents were from a geoliterate or a non-geoliterate background emerged as a concern in this study. A few explanations for these findings are as below;

i) There is a trend to perceive the source or author of a site in a Web environment as of little importance. Warnick (2004) has pointed out changes in credibility assessment that have been applied in a web environment compared to non-digital media. The ubiquitous lack of format standardisation of the placement of this variable on a website may have led to the low importance of this element when assessing credibility. The extra task needed for checking this variable, as well as the absence of this element in some websites, may gradually influence the low perception of this element in a web environment. The environment of Web 2.0 applications that allow web users to engage with the web contents through third party assessment such as forums, open feedbacks and comments provide rich experiences and platforms for web readers to evaluate and validate products, data and information simply via websites. Due to the lack of uniform layout standard on web based applications, including online maps of the placement of sources related metadata (i.e. map producer) to the users might lead to the less relevance of such elements in users' credibility assessment. The rich of third party assessments available on the World Wide Web has been seen as platform to assist web users to evaluate and validate the online information. Although some of results from think-aloud protocol above support this statement, further research should be conducted to examine up to what extent this hypothesis is true; for example by examining in other contexts of map use. The less relevance of information or metadata related to 'map producer and map data supplier' in experimental maps due to the rich availability of other mechanisms to validate if the data is contradictory has been highlighted by four of the six respondents in the think-aloud protocol as below;

'In the context of navigation, who produce the map/data is not important. What more important is we have to try (test) the application first before we can make any decision…some users might have certain influence of the map producer. But for me, I have to test the application. Just like the TomTom. At first, the brand TomTom was nothing. Then, after try the application, people slowly acknowledge it.….who produced the map is not important because we could search someone that might have reviewed the map. I will rely on the third party reviews in determining the credibility of a map'

(Fad, geoliterate)

'I can validate the data on my own. I can use the map; find my way by trials and errors. Who produced a map is not important. Anyone can produce a map' (Fik, non-geoliterate)

'The more important are the details of information, information clarity and easy to use. If there is inaccurate information on the map, it does not matter, since I will adjust and find my way on my own. Who produced the map is indeed not important' (Fa, non-geoliterate)

‘If I have to face with two maps that displayed contradict information, I will use the satellite navigation device or I will find other map to make comparison. I will use other sources to validate the data’

(Am, geoliterate)

ii) From the results of online questionnaires supported by the think-aloud protocols session discussed above, another reason of the low influence of source related metadata in these experiments were due to the extent of that element being noticed by respondents. Prominence- Interpretation theory describes the process of web users assessing online information relies on the credibility element(s) that noticed by users. If they do not notice the element, there will be no judgement or interpretations based on it. From the think-aloud protocol and observation sessions (Experiment 3), all six respondents did not notice the presence of ‘map producer’ located on the top side bar of the test maps. However, after the observer queried whether that element had any influence on their judgements, only four respondents agreed they might have influenced if they noticed; whereas the other two respondents stated there will be no influence if they noticed that element.

iii) The reason might also due to low motivation on the part of respondents to critically analyse the maps. Analogy of low motivation users have been described by Fogg et. al (2003) as users who browse and surf websites where they occasionally evaluate information critically; but they may be highly motivated when scrutinising information relevant to a specific critical need, such as seeking information to find a cure for cancer. Petty and Cacioppo (1986) identified that users in such a group would rely on peripheral signals (e.g. visual design, aesthetic etc.) when assessing credibility, rather than base their assessments on critical elements (e.g. sources, currency etc.). As argued by Morahan-Martin (2004, p. 502), there are three types of online users – vigilant, concerned and unconcerned; vigilant users are the most methodical in their approach to search online information; concerned users are less diligent compared to vigilant users, but checking the critical elements (source, date) by relying recommendation or trusted results through search engine or seal of approval; unconcerned users are more casual in their approach of seeking information where this group are least likely to have deep engagement with the search topic. The proposal of a credibility rating label may reduce the extra checking activity that is inevitably needed when users perform credibility assessment. The strategy that focuses on visual elements to attract users’ attention in low motivation groups is well established in marketing and advertising products. A study by Fogg et al. (2003) supports the view that visually related elements are widely used to determine the credibility of online information. The proposed CCTL ratings label offers a potential solution for users who are sometimes low in motivation, enabling them to scrutinise critical elements when judging the credibility of information. A study by Elzakker

(2004, p.127) that investigated how map users use and read maps conducted exploratory methods that not focused only on perceptual and cognitive approaches but include think-aloud protocol and observation, indicates similar results; there was no significant difference between the novice or expert map users in reading maps; they did not put much effort into validating the maps supplied to them; they assumed what was given to them was correct. The participants were also careless about the years (currency) of the maps. Some of them did not notice the mistakes on the map; although they had learnt what is called 'the grammar of cartographic design', the participants tended to choose what was the convention rather than what had been documented as a rule of thumb.

It is worth to highlight that this research did not filtered the sample according to users' experience or knowledge on the concept of data sources used on map mashups. From the think-aloud protocol, one respondent was confused about who was the producer of the map mashup. They perceived Google was the producer of the map mashup application because it used Google Map as its base map. Theoretically the data of map mashup application can be classified into two – background and foreground data. The background map commonly drawn through APIs such as Google Map, Bing Map, OpenSpace and OpenStreetMap; the sources of foreground data could be supplied by one or more sources including news and local data; this misconception has been identified from the excerpts below;

'I will look at the map provider. I will just trust the map if I am not familiar with the area. I will look at the date of the last updated. These two maps used similar base maps [by showing the copyright stamped at the bottom of map which is from Google]. There was no specific update on the last updated date. Just in 2013. Those maps produced by Google. I just looked at the 'powered by' label [Google's copyright at the bottom of base map], so the map producer [i.e. Google] had influence my decision' (Fik, non-geoliterate)

When the observer queried whether that respondent familiar with the concept of data sources on map mashup, the excerpt from him as below;

'No. I did not realise about this. What I know the map is produced by Google since there is a Google copyright stamped at the bottom of the map. I did not realise that it was actually from different source'

(Fik, non-geoliterate)

Nevertheless, the experiments in this research had carefully and clearly wording the questions for not to have double meaning related to 'map producer' and 'map data supplier'; which was either referring to supplier of base map or the foreground (top) data layer. For example, in Experiment 2, the question to address the influence of map ‘data supplier’, the statements were used as below;

Q3: What was the basis for your decision in Q2- selecting this set of maps and rejecting the other set of maps? Rank 1 to 5 (i.e. the element below)

The data supplier for the top data layer

Q8: I am familiar with the **data supplier(s) for the top data layer**

Q9: I trust the data supplier(s) that provide the data/information **for the top data layer**

Q10: **The data source(s) of the top layer** for this set of maps is more credible than the other set of maps

In Experiment 3, the statements that used to address the influence of ‘map producer’ and ‘map data supplier’ were used as below;

Q3: On scale of 1-4, indicate how important the following criteria (i.e. element below) is in influencing you to choose the map in Q1 and rejected the other map;

Map producer (map author)

Q5: How important are the following elements in influencing you in assessing a credibility of any online community based map?

The supplier of base (background) map (e.g. Google Map, Yahoo Map, Ordnance Survey OpenSpace map)

The supplier (contributor) of map foreground data (e.g. data supplied by City Council)

The concern issue is that if respondent believed both maps were produced by Google, due to the background maps were supplied by Google Map; the influence of ‘map producer’ in respondents’ judgement to choose either Map A or Map B might not relevant since they perceived the producers of both maps were identical. Although this research had carefully wording the statements to clearly differentiate which layers they referring to, future research should consider on the knowledge of respondents with the concept of the foreground and background data sources used on map mashups. Nevertheless, this might not be the case since:- 1) the methodology of Experiment 2 has included ‘spot the differences’ activity to suggest respondents to critically analyse the comparison maps and then notice the differences of foreground data suppliers, and 2) the experimental design of Experiment 2a, 3 and 4 were designed to compare simulated maps that supplied by different sources that hold different levels of reputations. The findings of these series of experiments confirmed each other on the low influence of metadata related to sources (i.e. foreground data supplier and map producer) in respondents' credibility assessment.

8.4 The influence of credibility labeling (CCTL)

The findings from the Experiment 1, 2 and 3 indicate high influence of visual cues on respondents' judgement of the credibility of map mashup information and low influence of metadata related to source (i.e. map producer and map supplier). The visual cues had more influence than the metadata related to sources (i.e. map producer and map data supplier) in users credibility assessment. This finding indicate that some map users were less influenced by the metadata related to sources but more influenced by the visual design and subjective cues on the map mashups. This finding is important. Online map users may be exposed to misleading, false or inaccurate information and propaganda presented via the map mashup medium. Colour coded traffic light (CCTL) labelling stamped on top of a map mashup offers a possible solution to this problem and so Experiment 4 was conducted to examine the influence of stamped labelling, in the form of a colour coded traffic light scheme, on respondents' judgement. Therefore, the second objective of this research was;

> 'to examine the influence of colour coded traffic light (CCTL) labelling on respondents' assessment of credibility when selecting and rejecting a map mashup'

Comparison analysis between Experiment 3 and Experiment 4 demonstrated a significant association between the presences of credibility labelling and whether or not visual cues, specifically 'colour scheme' and 'symbol design' have an influence on respondents' judgement. Results of the survey indicate significant differences between the responses perceiving Mashup A (map labelled with high credibility rating) and Mashup B (map labelled with low credibility rating) as having more believable (credible) information (see Table 7.9). In Experiment 3, which provided an experimental context without any credibility label on the map, the ratio of respondents that chose Mashup B and Mashup A was 1:1. Meanwhile, Experiment 4, which provided an experimental context with a 'credibility rating' on the map, demonstrated the odds ratio of respondents that chose Mashup B and Mashup A as 1:3. The likelihood of respondents choosing the high credibility map, **if they were given a map with a visual rating indicator (CCTL label), was 3 times higher than when given a map without the visual rating indicator (CCTL label).** Statistical test using the Pearson Chi-Square yielded a significant association between the two contexts – presence of credibility labelling and absence of labelling- with the choice of the 'high credibility' map ($X2$ (1, n = 208) = 12.453, $p < 0.001$). Hence, **hypothesis 6 is supported**.

Analysis of the total scores that measure the **influence of credibility labelling demonstrates a moderate influence** on respondents' judgement (see Table 7.13). These findings provide useful insights into the positive influence of the credibility ratings element on respondents' judgement, although the impact falls within the moderate level. This level of impact however was comparable with the influence of visual cues, which in this study comprised 'colour scheme', 'symbol design' and 'overall presentation', in respondents' judgement. From the think-aloud observation in Experiment 4, of the six respondents that involved four of them chose the Map A (the map that labelled with high credibility label). They had been influenced by the 'credibility rating' label stamped on the Map A. Previous section has highlighted the prominence-interpretation theory that online users only judge credibility based on the element(s) that they notice. The question is whether the respondents will notice the CCTL rating label and make judgement based on the rating? From the think-aloud protocol, two respondents made judgement based on the CCTL rating label stamped on the maps; respondents noticed this element without being highlighted by the observer. The excerpts as below;

'Ok. Since this map has a stamped rating label that indicated 'low rating', so I chose Map A because there is a rating label and indicated as 'high rating' (R, non-geoliterate)

'Because this Map has been reviewed as high credibility rating compared to Map B'

(A, non-geoliterate)

However, another two respondents did not notice the presence of CCTL rating, but agreed that element will influence them if they noticed on the maps. The excerpts as below;

'Ok. If there is a rating, it will influence me. This rating was produced [generated] by whom? If rated by one organisation, then it will influence my decision. So, I chose Map B because it has been rated by one organisation'

(Fad, geoliterate)

'To be honest, before you mentioned about this stamped credibility rating, I did not noticed it at all. But I might have influenced of this element, if I had noticed it at first' (Ami, geoliterate)

Analysis on the total scores measuring the influence of visual cues, which consist of 'colour scheme', 'symbol design' and 'overall presentation' element (variables), indicates that the level of influence falls in moderate category. Comparative analysis in the previous experiment, Experiment 3 demonstrated a significant association between the presence of credibility labelling and whether or not visual cues (colour scheme and symbol design) have an influence on respondents' judgement. The effect based on the odd ratio, **the influence of 'colour scheme', was 8 times higher if the map was not labelled with a credibility rating than when labelled.** Likewise, **the influence of 'symbol design' was 3.3 times higher if the map was not labelled with a credibility rating than when labelled.** Therefore, **Hypothesis 7 is supported** where the influence of visual cues were decreased.

In Experiment 4, from Mann-Whitney test, the influence of visual cue elements from geoliterate respondents (Mdn = 61.1) did not differ significantly from those of non-geoliterate respondents (Mdn = 72.2), $U = 521.5$, $z=-1.5$, $p>0.12$, $r = 0.2$ (small effect size). Excerpts from think-aloud observations (Experiment 4) as below;

'I did not look at the map producer. It was not important. What was more important was the clarity and ease of use' (Fik, non-geoliterate)

'I chose Map B, route 1 because in terms of design, it was more attractive, neat and not too crowded…I did not influenced by the map producer. I just focused on the events and all the symbols' (Am, geoliterate)

The **Hypothesis 8** (i.e. there is significant difference between geoliterate and non-geoliterate respondents in the influence of visual cue variables when making judgement) **is not supported.**

These results might indicate both groups used visual cues (for example the labels, colour scheme, symbol design) to form the basis of their judgement in assessed the credibility of map information. The levels of influence of these elements in their credibility assessment might be different. Some of them might use visual cues as main element to form their judgement, whereas the tested parameters (i.e. credibility rating) as supporting elements. These excerpts also indicate **the moderate influence** of colour coded traffic light (CCTL) rating label in respondents' judgement. Excerpts from think-aloud experiments as below;

'Although if I had noticed, I would not have been influenced. I would look [focus] on the main purpose. This element [rating label] is an additional element to strengthen my decision. I will believe the map more if there is a credibility rating' (Am, geoliterate)

'I had been influenced by the visual attractiveness, symbol design, clarity of symbol, symbol convention……the importance of map producer was just 'important' [if using a scale of important]… the 'very importance' was the symbol…the symbols [number] represent the detail of information'

(Fa, non-geoliterate)

The Mann-Whitney non parametric test was then selected to test the Hypothesis 9 which is ‘there is significant difference between geoliterate and non-geoliterate respondents in the influence of credibility labelling when making judgement’. To test this hypothesis, the mean difference was compared using a parametric test, namely the independent t-test. From this test, on average the influence of the credibility labelling from geoliterate respondents (mean = 48.8, SE = 6.6) was lower than from non-geoliterate respondents (mean = 55.67, SE = 4.2). This difference was not significant $t(73) = -0.92$, $p > 0.05$: it did represent a small sized effect $r = 0.1$ (see Chart 7.5). A few excerpts from four respondents in think-aloud protocols support this statement as below;

‘Ok. If there is a rating, there might be some influenced. This was rated by whom? If it was rated by one organisation, I will be influenced. I chose Map A because there was one organisation which rated the map’ (Fad, geoliterate)

‘My strongest argument is I only influenced by the detail of information. But another reason that might influence my decision is because of the high rating credibility rated on the map’ (Am, geoliterate)

‘Ok. Since this map has a stamped rating label that indicated ‘low rating’, so I chose Map A because there is a rating label and indicated as ‘high rating’ (R, non-geoliterate)

‘Because this Map has been reviewed as high credibility rating compared to Map B’ (A, non-geoliterate)

However, there were two respondents did not had any influence on the presence of the CCTL on map (in Experiment 4). The excerpts as below;

‘Ooh, I do not have any background in mapping. I believe any maps that I found. A rating on map did not help much. I believe any maps on the Internet. I did not look at this rating label. I just focused on the map and the symbol used’

(Fa, non-geoliterate)

‘I think the rating is not too important. I can validate the data on my own. I can use the map; find my way by trials and errors’

(Fik, non-geoliterate)

Therefore, the **Hypothesis 9** (i.e. there is significant difference between geoliterate and non-geoliterate respondents in the influence of credibility labelling when making judgement) **is not supported**.

Further analysis was conducted on the transcripts from think-aloud protocol (in Experiment 4) of the influence of source-metadata (which in this case the ‘map producer’) versus the influence of CCTL rating label on users’ judgement. From the transcripts, two respondents perceived the rating label had more influenced than the ‘map producer’. Excerpts as below;

Observer: Has the element(s) at the side bar influenced your decision?

Respondent: Yes, they had. But the map producer did not help much… because the more important is to test the map first.

Observer: Ok…how about the influence of the credibility rating label on map?

Respondent: Ok. If there is a rating, it will influence me. This rating was produced [generated] by whom?

(Fad, geoliterate)

Observer: How about the producer of the map. Has the element influenced you?

Respondent: No influence. I just influenced by the details of information.

Observer: How about the credibility rating label stamped on the map?

Respondent: Oh, Ok what do you mean by the rating? What is the rating? Ooh do you mean this label [respondent pointed the cursor on the stamped rating label]…. To be honest, before you mentioned about this stamped credibility rating, I did not notice it at all. But I might have influenced of this element, if I had noticed it at first.

(Am, geoliterate)

The implementation of traffic light colour coded rating schemes have been applied to certain food products, electrical appliance energy ratings and in the United Kingdom (UK) car carbon dioxide emissions. This colour coded scheme has been implemented voluntarily in certain supermarkets in the UK (BBC News, 2007) and become a food policy in Western Australian health and school services (Western Australian Department of Health, 2009). A few studies in geospatial domain have proposed visualisation tools to communicate the quality of data to users. For examples, Devillers et al. (2007; 2005) proposes a model and a prototype tool to support a development of visual spatial data quality assessment which can be integrated on a desktop GIS application. Yang and Wang (2004) and Yang (2007) addressed the issues of spatial data quality visualisations in online environments. The authors proposed a more comprehensive framework for visualising spatial data quality using an object-oriented approach specifically for a Web and mobile environment. A study by Mass et al. (2011) proposed the use of a geo-label on scientific maps through GEOSS Geoportal. The label is based on information about data quality which will be extracted from metadata, the data itself and the validation process with in-situ sensors, provenance information, and user-feedback. Yang (2007, p.173) addressed a possibility to provide visual quality information on commercial maps such as Google Map and Microsoft Virtual Earth. This is due to the current trend where web citizens tend to use such maps to browse, locate and query spatial-reference information; this could be implemented by extending its Application Program Interfaces (APIs).

This study demonstrates the moderate impact of credibility labelling using colour coded traffic light ratings (CCTL) for online mapping and particularly for map mashups. The probability of respondents making informed judgements by choosing a high credibility map based on this credibility rating label is three times higher than the setting without the label. From the findings discussed above, the influence of CCTL rating label is higher compared to textual or image based information embedded at the map sidebar to present source-metadata of the maps. This finding is in line with the study of Kelly et al. (2009), which demonstrates that the probability of users identifying healthier foods from a traffic light labelling format was five times more likely than from a label using monochrome text information. As according to Fox (2006, p.12), users tend to pay attention to informative readily available label, such as labelling on food product more than checking the date and source of online information. The diminish diligent to check the sources and date might due to the presence of sources and currency of information on the web pages not clearly presented and not disclosed. This approach may educate and informed the general public on the uncertainty of the data and information they obtained from that medium. The same benefits will be generated if such a visual quality indicator is implemented on top of map mashup applications. Further investigation could be made on how to increase the prominence of the presence of CCTL rating label on the map including the design and the layout as well as on the users' awareness aspects to this new mechanism.

It is worth to highlight that this research did not filtered the samples of respondents according to their familiarity to the environment or area of the map context. As according to Taylor et al. (2008, p.5), navigation within unfamiliar or simple environment often incorporates action based sequences, whereas movement within familiar setting appears to construct a cognitive map; cognitive maps resulted from extended experience of individuals with the environments (or area) where they developed a map like representation in their memory or so called 'maps in the head'. Hanowski et al. (1994) and Kantowitz et al. (1997) indicate when familiarity of one area increased; users are less likely to rely and use information from a system exclusively. In this research, there was a possibility of the factor of respondents' 'familiarity' (i.e. students) to the campus area in their responses influence their judgements; for example, low motivation to engage (involve) with the experimental tasks, since the contexts of tasks was artificial based; the issue of familiarity to the area might had influence them for not completely relied on the map. This is one limitation of this research. Therefore, think-aloud protocol sessions were conducted to examine this issue where the respondents were selected from individuals that resided in different country (i.e. Malaysia) and did not familiar with the University of Nottingham campus. From the analysis, the excerpts of think-aloud protocols (qualitative data) were in line with and supported the results of the online map based questionnaires (quantitative data) that were conducted among respondents that familiar with the University of Nottingham campus; there was no contradictory and inconsistency between the results (statistical based data vs. transcript excerpts) that collected by a sample of respondents that familiar and not familiar with the case study area. For example, in Experiment 3, all six respondents did not look (notice) at the map producer label stamped on the top sidebar of the websites. They only noticed that element when the observer highlighted the presence of that element embedded with the maps. Only three of them said that the element might have influenced them if they had noticed it before making decisions. In Experiment 4, four respondents believed the CCTL rating label might have influenced them when making credibility assessment. The other two respondents said it would not have had any influence if they had noticed the label. The low influence of metadata related to sources and positive influence of CCTL rating label reported within respondents in think-aloud protocols that did not familiar with the campus area was confirmed and supported the findings that demonstrated by respondents that familiar with the campus area. See Appendix D and E for think-aloud protocols complete transcripts.

As have been discussed in section 2.5.2, this research examines the influence of credibility in general (without specific context); the elements identified in general contexts might be useful to understand the common credibility elements and general views of users' perceptions to those elements. However, the findings might not be able to represent specific scenarios or contexts of map uses. As according to Phillips (1984) it is not a fatal flaw for those studies that do not consider the context, including how maps are actually used in real operations; if the results yielded similar findings, a fairly clear picture could be generated by analysing the relative merits of the results in similar studies. This could be seen from the results of the four series of simulated experiments conducted in this research where the findings confirmed each other. The findings demonstrated in this research add significantly to the understanding of individuals' perceptions, largely young adult, on the dominant elements that they judge when assessing the credibility of information on map mashups. Nevertheless, the limitations discussed above should be considered when citing the findings of this research.

8.5 Conclusion

This chapter describes the findings from the four experiments to achieve the Objective 1, which was 'to examine the influence of metadata variables related to map producer and map data supplier', and Objective 2 which was 'to examine the influence of colour coded traffic light labelling on respondents' assessment of credibility. This study has demonstrated the dominant influence of visual cues; including 'colour scheme' and 'symbol design' but the elements tend to relate with the 'clarity of information' and the 'usefulness' of the presented information to respondents when judging the credibility of a map mashup. The findings indicate there is a group of map users that perceived low influence of metadata related to sources that tagged at the side bar of a map mashup application. The final experiment in this study (Experiment 4) has demonstrated the influence of credibility labelling to assist map users making informed judgements about the credibility of a map mashup. This credibility stamped label could assist respondents to make an informed choice and could create awareness among map users concerning the credibility of a map mashup. In the Experiment 4, the use of such a label has been shown to generate a positive moderate influence on respondents when they make judgements. The next chapter of this thesis will therefore discuss the conceptual framework to implement credibility labelling on map mashup applications in a semi-automated or a fully automated manner.

9 A CONCEPTUAL FRAMEWORK OF AUTOMATED CREDIBILITY ASSESSMENT

9.1 Introduction

This section presents an outline model for the automated generation of a credibility rating index. This index would, in turn, be used to provide a ‘colour coded traffic-lights’ (CCTL) credibility label for users of map mash-up applications. The parameters identified for the specification of the model are derived from a review of the literature as well as the findings of this research. The practicalities of calculating the parameters in an automated manner are discussed on the basis of the current state of the art and reviews of expected technology advancement and research directions. An indicative approach to evaluating and assigning weight to the proposed parameters, along with a simple equation to calculate the cumulative score rating, are presented.

9.2 User Credibility Parameters (Elements)

This section discusses the use of parameters (elements) to represent credibility from a general perspective. The parameters are drawn from the literature of various domains. Many parameters and indicators have been discussed and proposed. This section discusses them with a view to later consideration of what will be applicable for an automated credibility index for map mashups.

9.2.1 Users’ Perceived Credibility and Online Static Trust

The studies that discuss elements of trust and trustworthiness in an object (information) are somewhat similar to the focus elements of users’ perceived credibility in information. However, Fogg and Tseng (1999) argue that trust is a subjective and dynamic process that operates between people but not technologies. Several studies have adapted the dynamic of trust between people to the objects that mediate human relationships. For example, see Stewart, 1999 and Golbeck, 2005). The dynamic nature of trust relationships between humans has been developed into the concept of static trust, which is concerned with the ways in which trust is induced from objects.

In the literature, 'trustee attribute' (Skarlatidou et al., 2011) and 'trustworthiness' (Gaudinat et al., 2011) of objects (and information) are the terms used to discuss the conceptualisation of the static nature of trust. The elements that have been discussed correspond to the elements suggested in the research relating to the perceived credibility in a medium, an object and information. Following on from this, the terms 'trust' and 'credibility' have been used interchangeably in the literature. Table 9.1 below presents the studies that have examined similar elements but have discussed them in different terms.

Table 9-1 Comparisons of different studies that examine the elements of users perceived credibility in different terminology

Terminology	Static Trust (Skarlatidou et al., 2011)	Trust (Bishr, 2007)	Trustworthiness Cheskin (1999)
Accuracy	•	x	x
Map Usability	•	x	x
Website usability	•	x	• (navigation)
Map design	•	x	x
Website design	•	x	x
Functionality	•	x	x
Reputation	• (website provider)	• (Individual personality)	• (brand, previous experience)
Logo	•	x	x
Links to online community	•	x	x
Links to privacy policy	•	x	x
Contact details	•	x	x
Seal approval	•	x	•
Customer services	• (chat)	x	• (Help services, order processing indicated)
Testimonials	•	x	x
Feedback mechanism	•	• (Proximity to author's location to the event, outcome of each encounter between actors)	x
Forum	•	x	x
Lineage – how the map was constructed	•	x	x
Aesthetic – background base map	•	x	x
Visual design	•	x	x

(colour scheme)			
Readability	•	x	x
Professionalism	x	x	• Presentation • technology

- = include; x = not include;

9.2.2 Users' Perceived Credibility and Quality

Users' perceived credibility is one of the layers used to determine the quality of an object (or information). Users assess credibility before assessing the quality of an object (or information). The indicators of quality commonly comprise the parameters that determine credibility. Indicators of quality in the literature are generally drawn from research relating to information quality, interaction, service quality and usability. Quality indicators in the Quint2 model (Calero et al., 2005) are based on ISO 9126, which is a standard for evaluating the product quality of software. The Quint2 model extends this standard to include a few elements to adapt to the web product environment, such as availability, degradability, clarity, helpfulness and user friendliness. Another quality model, namely WebQual 4.0 (Barners and Vidgen, 2003), is used to assess the quality of a website environment. This model can be used to evaluate the quality of information-intensive websites and interaction-services websites. The differences between this latter model and the Quint2 model are due to the comprehensiveness with which the life-cycle process of website development and implementation are covered. The dimensions in this model consist of functionality, reliability, usability, efficiency and portability. The quality indicators are more oriented towards the aspects of web development and internal quality. Notwithstanding, WebQual is more specific with regard to the use of a website, where the indicators are obtained from web users and are more general with regard to internal and external quality. Table 9.2 presents a comparison between the elements of quality.

Table 9-2 Comparison between elements of quality models

Quint2 Model (Calero et al., 2005)	**WebQual 4.0 (Barners and Vidgen, 2003)**	**Cappiello et al. (2004, p.6)**
Functionality (F) –	x	x

suitability		
Accuracy (F)	• Accurate information	• accuracy
Interoperability (F)	x	x
Security (F)	x	x
Traceability (F)	x	x
Reliability (R) – maturity	x	x
Fault tolerance (R)	x	x
Recoverability (R)	x	x
Availability (R)	x	• accessibility
Degradability (R)	x	
Usability (U) – understand ability	• Easy to understand	• interpretability
Learnability (U)	• Easy to learn to operate	x
Operability (U)	x	x
Explicitness (U)	• Interaction clear	x
Attractively (U)	x	x
Customisability (U)	x	x
Clarity (U)	• Interaction understandable	x
Helpfulness (U)	x	x
User-friendliness (U)	• Easy to navigate • Easy to use	x
Efficiency (E) – time behaviour	x	x
Resource behaviour (E)	x	x
Portability (P) – adaptability	x	x
Install ability (P)	x	x
Replace ability (P)	x	x
Co-existence (P)	x	x
x	• Believable	x
x	• Timely	• Timeliness
x	• Relevant	x
x	• Information details	x
x	• Reputation	x
x	• Feels safe, secure, confident	x
x	• Sense of personalisation	x
x	• Sense of community	x
x	• Easy to contact	x
x	• Overall view of	x

	Website	
x	x	• Completeness
x	x	• consistency

- = include; x = not include;

Quality indicators for spatial data differ from the indicators of quality on websites. There are several standards that have been produced by organisations, such as the FGDC and the ISO/TC211 committee, as guidelines to measure the quality of spatial data. Oort (2005) summarised the indicators of spatial data quality from several pieces of literature and listed 11 elements of spatial data quality. The latest quality standard is ISO 19157, which is an up-dated combination of ISO 19113, 19114 and 19138 that comprehensively covers the standards relating to spatial data quality and metadata. Data usability is a new element for inclusion in that standard, with the purpose of assisting data consumers to analyse the fitness for purpose of the data in order to define its suitability for new applications (see Danko, 2005). Table 9.3 below presents the indicators of spatial data quality from different sources.

Table 9-3 Indicators of spatial data quality

Indicators	Oort (2005)	ISO 19157 (Danko, 2005)	FGDC (Danko, 2005)
Lineage	•	•	• source information, source time period of content, process date, process contact
Positional accuracy	•	•	•
Attribute accuracy	•	• Thematic accuracy	•
Logical consistency	•	•	•
Completeness	•	•	•
Semantic accuracy	•	•	x
Usage, purpose and constraint	•	• usability	x
Temporal quality	•	•	•
Variation in quality	•	x	x
Meta-quality	•	x	x
resolution	•	x	x

- = include; x = Not include;

9.2.3 Users' Perceived Credibility and Usability

Usability and accessibility are other indicators that have been identified to influence users' perception of credibility of a website. Research related to usability and accessibility (U&A) has been conducted widely in other domains, such as web engineering and human computer interaction (HCI). Several guidelines, including official, unofficial and in-house criteria, have been proposed to assist developers in evaluating the U&A of the software or web applications being developed. Standards guidelines, such as the Web Content Accessibility Guidelines (Rieh and Belkin, 2000), have been released by the World Wide Web (W3C) to suggest general criteria for the design of websites accessible to people with disabilities. In addition, the Americans with Disabilities Act and Rehabilitation Act Amendments, Section 508, have been produced to protect people with disabilities from technology exclusion, which suggests criteria to be considered when designing United States government-affiliated web applications. Unofficial guidelines, such as those by Molich and Nielsen (1990), have identified and suggested several usability criteria for the design of the user interface.

9.2.4 Credibility Elements

The previous sections discussed the relation of credibility with other domains that focusing on trust, quality and usability aspects. By considering research interests from these domains, Table 9.4 below presents the general and comprehensive elements related to credibility. The proposed credibility elements of map mashups in this study, therefore considering these elements but specifically selecting the elements that appropriate and practical to be assessed in automated manner. The next section will discuss these elements.

Table 9-4 Overview of credibility related elements from various domains

Element	Descriptions	References
Domain	The address that hosting the website	Fallis and Fricke (2002) *Obtaining url suffix denoting edu. Non-profit or gov. designation* (Wathen and Burkell, 2002)
Currency	How well data are up-to-date	(Fallis and Fricke ,2002) *Timely data* (Longstreet , 2010) *Timeliness, age of data* (Wang and Strong, 1996) (Stark, 2010) *Updated frequently* (Princeton, 2002) *Temporal validity* (ISO19157) (Leibovici et al., 2011)
In-links	a hyperlinks	(Fallis and Fricke ,2002)

	(backlinks) that point towards the websites	*Google search index source* (Google, 2012)
Users collaborations	Users participation with the system (providing comments / reviews)	Gaudinat et al, (2007),
Seal of approval	Endorsement from other organisation (Woodford and Jackson, 2003)	Fallis and Fricke (2002), Cheskin (1999), Fogg et al., (2003) *Display seal of approval from other group* (Princeton, 2002)
advertising	Producing information for promoting sale of product/services (Woodford and Jackson, 2003)	*Distinguish from main content* (Fallis and Fricke, 2002), Fogg et al. 2003)
Privacy policy	A statement that discloses related information of how the party gather, use, manage customer data (Woodford and Jackson, 2003)	Fallis and Fricke (2002)
Sponsorship	A party that fund the activity (Woodford and Jackson, 2003)	*Financial disclosure* (Fallis and Fricke (2002) *What business/organisation support the site* (Princeton, 2002)
Website Usability	Effectiveness of the interaction between human and web applications (Haklay and Tobon, 2003)	(Fallis and Fricke ,2002), Fogg et al., 2003) *Performance* (Fogg et. al 2003) *Information fit to task – functional fit to task* (Longstreet, 2010) *Easy to navigate* (Princeton, 2002) *Response time* (Longstreet 2010) *Accessibility* (Wang and Strong 1996)
Relevancy	The extent of the information meet the current users' tasks and goals (Parker et al., 2012)	*Usefulness of information* (Fogg et.al 2003) *Relevancy* (Wang and Strong 1996)
Referencing	A quotation from or reference from other sources (Woodford and Jackson, 2003)	*Citation based trust* (McGuinness et al., 2006)
Affiliation	connection with a larger organisation (Woodford and	*Site affiliation* (Fogg et al., 2003) *The evidence of connection between one trusted website to one unknown website*

	Jackson, 2003)	*(transference)* (Stewart, 1999) *Highlighting links to other credible site* (Wathen and Burkell, 2002)
Brand	Image of company (Woodford and Jackson, 2003)	*Consistent image – ability to project a company image in other form of media channel* (Longstreet, 2010)
Data sources	Sources of data	*Attribution* (Fallis and Fricke (2002) *Reputation data source and data* (Wang and Strong ,1996) *Able to identify sources* (Princeton, 2002)
Identity of site operator	The background of the creator /producer	*Transparency* (HONcode) http://www.hon.ch/HONcode/Conduct.html *Identification author* (Fallis and Fricke, 2002) *Disclosing site operator, name recognition and reputation* (Fogg et. al 2003) Wang and Strong (1996) *Brand* (Cheskin, 1999) *Knowing who owns the website* (Princeton, 2002) *Able to find important facts about a website* (Princeton, 2002) *Display award and certificates* (Princeton, 2002, Wathen and Burkell, 2002) *Presenting institutional and individual credentials* (Wathen and Burkell, 2002)
Authority	trusted source used in place of a given individual's credibility decisions (Lankes, 2008)	*Authoritative* (Fallis and Fricke (2002),
Accuracy	The closeness to truth or fact (Woodford and Jackson, 2003)	(Fogg et. al (2003)) (Wang and Strong, 1996)
Testimonials	Statements made by other people (Woodford and Jackson, 2003)	*Emotional appeal* (Longstreet 2010) *Ratings* (Bishr and Mantelas, 2008)
Absolute positioning, relative positioning, geometric accuracy	See next field	*Absolute positioning*- The degree to which the digital representation of a real world entity agrees with its true position on the earth's surface *Relative accuracy* – the positional accuracy of a data point in relation to

		other data points *Geometric accuracy* – the trueness of feature data to the shapes and alignment of real world entity they represent (Harding, 2005)(ISO19157)(Leibovici et al., 2011)
Thematic correctness	See next field	*classification correctness, non-quantitative correctness, quantitative attribute accuracy* (ISO19157)(Leibovici et al., 2011) *Spelling error* (Fallis and Fricke, 2002) *Thematic correctness* –the correctness to which attributes in the data record information and classification about the real world entities (Harding, 2005)
Completeness	See next field	*By themes, by neighbourhood, omission, commission* (ISO 19157); (Longstreet 2010) *Breadth, depth, scope contained in the data* (Wang and Strong 1996) *Appropriate amount of data* (Wang and Strong, 1996) *Information focus* (Fogg et al 2003) *Feature completeness, attribute completeness, spatial completeness, temporal completeness, thematic completeness* (Maue and Schade, 2008) *Completeness* – the degree to which data contents corresponds to the real world in accordance with the data coverage (Harding, 2005)
Logical consistency	See next field	*Consistency, conceptual, domain, format, topological* (ISO19157) (Leibovici et al., 2011) *Logical consistency* – the degree to which the data logic complies with the real world features representation (Harding, 2005)
Data usability	The conformance of the dataset for a particular usage within a specific application (Leibovici et al.,	ISO19157 (Leibovici et al., 2011) *Information fit to task – information quality* (Lin, 2010)

	2011)	
Map Usability	Effectiveness of the interaction between human and map applications (Haklay and Tobon, 2003)	*Map usability* (Haklay and Tobon, 2003, Nivala et al., 2008)
Map design	The elements that considered when designing a map	*Symbols, colour scheme, design look* (Skarlatidou et al., 2011)
Website design	The elements that considered when designing a website	(Lin, 2010), *Layout, typography, white spaces, images colour scheme* (Fogg et al., 2003), *Format (representational consistency*) (Wang and Strong, 1996) *Tone of writing* (Fogg et al, 2003)
Design look, visual appeal, aesthetic – overall complexity, layout interface	Overall presentation	*Concise representation* (Wang and Strong, 1996) *Presentation, professionalism quality* (Cheskin, 1999)
Personalised	The ability of the website to be customised to tailor with personal customer information	*Tailored information – personalise, interactivity* (Longstreet 2010) *Innovative – attempting to tailor the info to customer preferences* (Longstreet, 2010) *Empathy* (Lin, 2010)
Underlying motive	The purposes to disseminate information on the website	(Fogg et. al 2003) *Mission, purposes of website* (HONcode) http://www.hon.ch/HONcode/Conduct.html *Objectivity* (Wang and Strong 1996)
Information Clarity	The state of being clear (Woodford and Jackson, 2003)	*Learnability, Understandability* (Lin, 2010) *Ease of understanding* (Wang and Strong, 1996) *Readability of text* (Fogg et.al 2003) *Interpretability* (Wang and Strong, 1996) *Information clarity, text/label clarity* (Fogg et. al 2003)
services	An action to satisfy a need (Woodford and Jackson, 2003)	*Responsiveness* (Lin, 2010) *Enjoyment* (Lin, 2010) *Customer services – how organisation treats customers* (Fogg et. al 2003)
Bias	A preference that influences judgement from being unbalanced	*Information bias* (Fogg et al 2003)

	(Woodford and Jackson, 2003)	

9.3 Automated Credibility Ratings for Map mashup

Previous sections have discussed the comprehensive credibility related elements in various domains. The previous section 2.6 (chapter 2) also has discussed in depth the needs of credibility automated assessment and labelling on online map mashups. The next section will present the proposed parameters that maybe practical to be evaluated in automated manner on map mashup applications.

9.3.1 The practical indicators

The main practical indicators discussed below are elements of credibility suggested for a map mash-up credibility rating. These are indicators that have been drawn from the literature and that may be appropriate to be assessed in an automated manner, possibly in the near to medium future. Recent advances in technologies from various research domains and the possibility that new technologies might emerge in the near future to support the implementation of automated credibility rating of map mash-ups have been taken into account when considering which indicators/elements could be assessed using automated tools. It has been argued that manual assessment of the factors that influence users' perception of credibility could be ineffective due to the high costs and the amount of time involved. The main indicators proposed in the credibility rating model comprise three main components. The main indicators proposed in the credibility rating model are comprised of three main components:

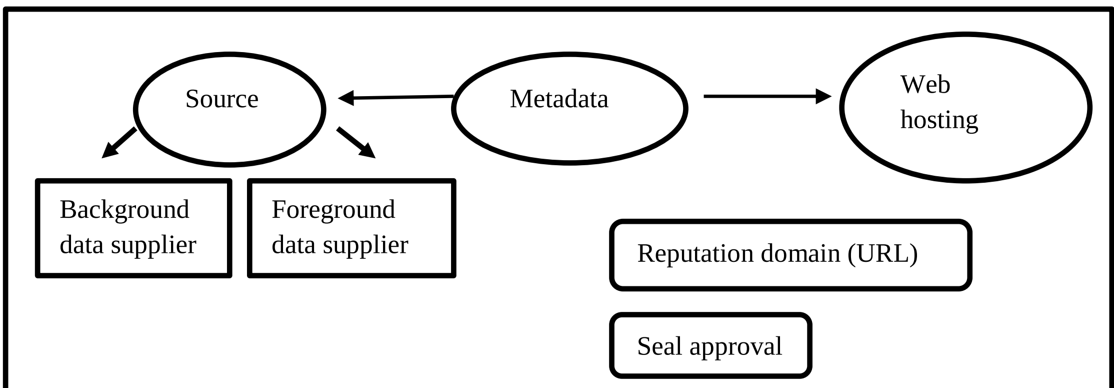
1) Metadata component
2) Data component
3) Usability and Accessibility (U&A) components

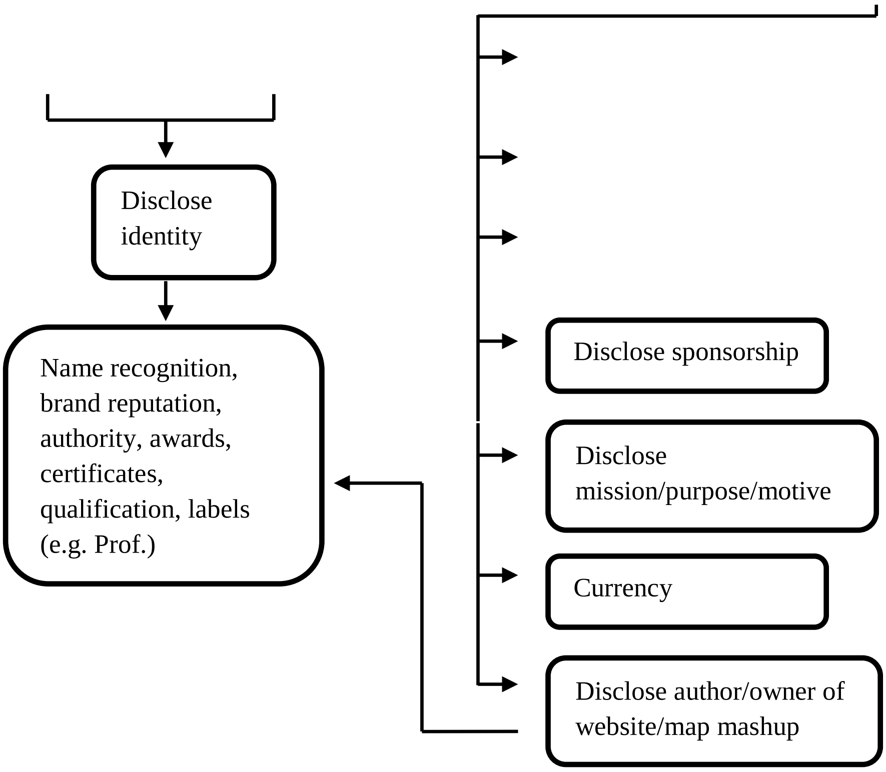
Notwithstanding, it is worth to highlight that usability and accessibility component is not the key in this proposed rating model. The use of this model may exclude this component if the values of the parameters may distort the influence the total credibility ratings. Although design, usability and accessibility are important aspects in users' experience of using the map application, however, these parameters do not positively correlate with the influence of critical metadata parameter. This is due to the high influence of visual cues and low influence of metadata when users assessing the perceived credibility of information, as identified in this research. A usable and accessible website does not imply the correctness of the presented data (information). On the other hand, attention to this component might be needed if the rating assessment is concern with the quality of application.

In the metadata component, there are two sub-components. Figure 9.1 presented the proposed parameters under metadata components. The potential of the parameters to be evaluated in automated ways have been considered before proposing the parameters in the model.

1) Hosting details (i.e. metadata elements of a website that hosted a map mashup)
2) Source details (i.e. metadata elements of data supplier(s) or contributor(s) of information)

Hosting sub-components contains seven parameters including reputation of the domain used to host the mashup's website, seal approval stamped on website, any affiliation or association that hosting the map mashup and the currency of the website and data/information being supplied and updated. Other parameters are on sponsorship, missions or objectives of the maps and websites, and the author(s) or owner(s) of website/map mashup application. In source component, the proposed parameters are the identity of data supplier, including the supplier or contributor of background data (base map) and the foreground data. The values of these parameters might include the name recognition, brand reputation, awards, certificates, professional or academic qualifications or labels e.g. FRICS. Figure 9.1 proposed parameters that could be evaluated from metadata component.





In data component, there are two sub-components. The sub-components are as below:

1) Consistency and correctness component (i.e. the parameters that measure the internal accuracy and correctness of data/information presented on map mashup)
2) Third party reviews component (i.e. the parameters that measure the correctness of data/information on map via reviews from independent and/or dependent sources.

Figure 9.2 presents the proposed indicators or parameters that could be measured under this component. Consistency and correctness sub-components contain five parameters including attribute/thematic correctness, temporal consistency, absolute and relative positioning consistency and logical consistency. Other spatial data quality parameters including completeness and data usability seem less relevant to be measured in automated environment; these parameters require human interpretation to assess its validity.

The next sub-component is third party reviews. Four parameters are suggested under this component including comments/feedbacks from other sources including personal and professional views; the amount of data/information being shared across websites or social media; the number of star ratings and ‘thumb likes’ by viewers that used the data and information or rating agreement on the comments/feedbacks; the number of backlinks, which is the number of other websites that refer or use the data or information.

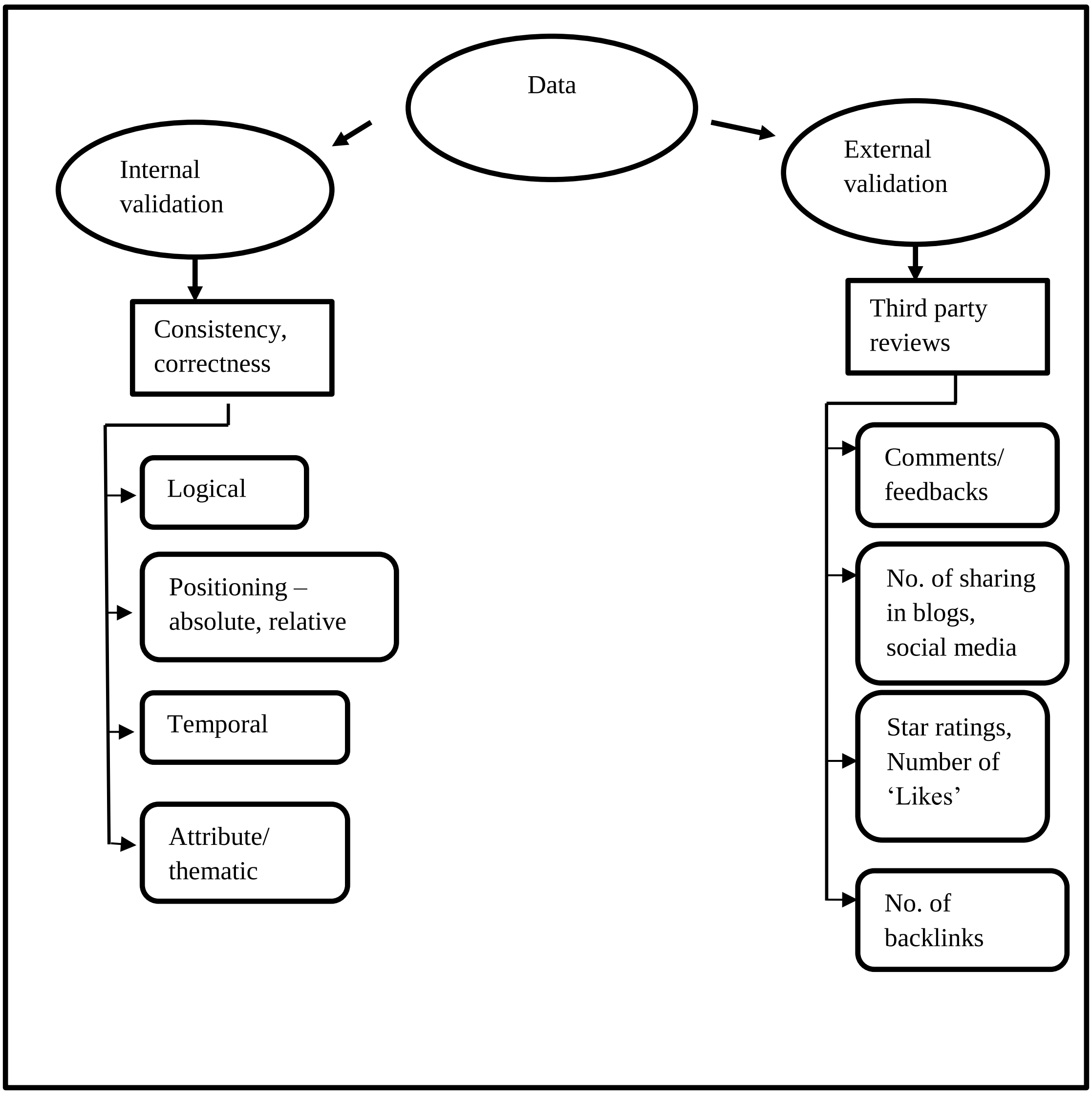


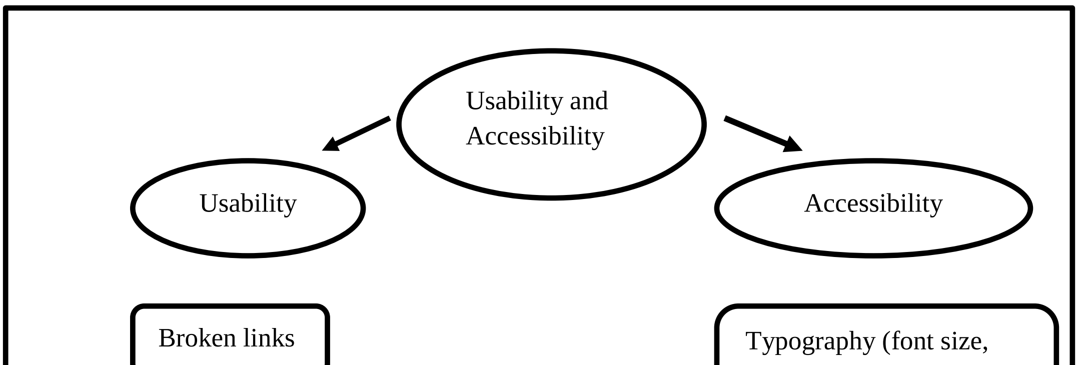
Figure 9-2 The proposed parameters in data component

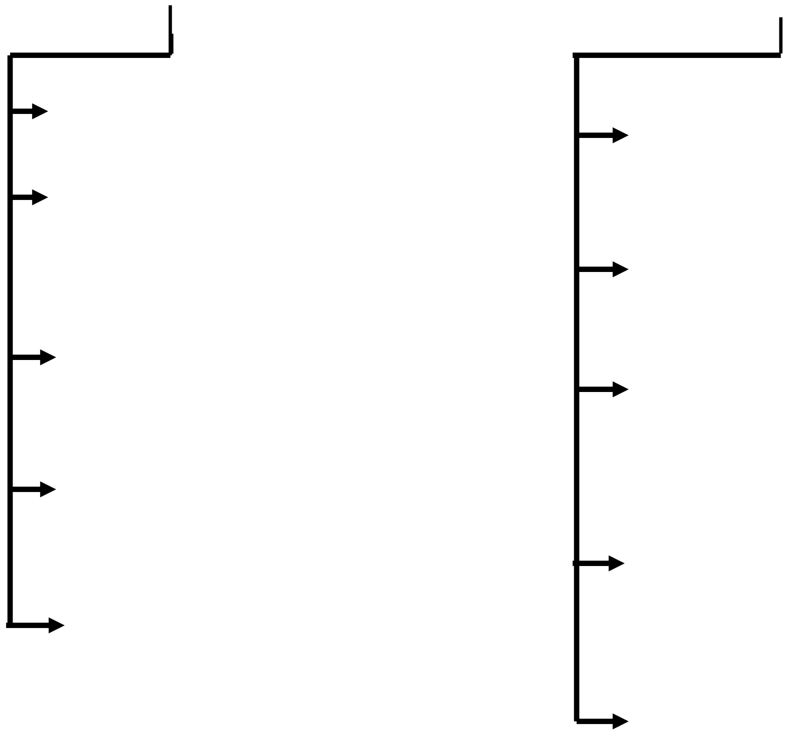
In the usability and accessibility (U&A) component, there are two sub-components. Figure 9.3 presents the proposed indicators or parameters that will be measured under this component.

1) Usability component (i.e. the parameters that measure the usability of map and websites)
2) Accessibility component (i.e. the parameters that measure the accessibility of data/information presented on map mashup)

There are five proposed parameters in usability sub-component. The parameters include;

1) number of broken links and functionality,
2) the structure of tools grouping,
3) the structure of website and map layout positioning.
4) the organisation of data layers in terms of the display setting of on and off layer in relation to zooming level.
5) in accessibility sub-component, the parameters suggested is examining the typography of the text/label in terms of the appropriateness of font size, style and font colours; assess the features size and type, symbol types and sizes, the use of colour scheme on features and the thematic zoning, and the scales of the features versus size of symbol and text.





9.3.2 A Proposed Model for Credibility Rating Index

The section below discusses how to measure the parameters proposed above for a practical credibility rating index. Indicative coded criteria and weighting values are presented as a suggestion for measuring the parameters. The section also proposes indicative formula to produce the cumulative rating.

The implementation issues for managing the automated assessment tool are discussed at the end of this section.

9.3.2.1 Indicative coded criteria and weighting value of the proposed parameters

Table 9.5 presents the indicative criteria and weighted values for parameters in metadata components. Table 9.6 presents the indicative criteria and weighted value for map data/information components and Table 9.7 presents the indicative criteria and scoring value for U&A components. These tables serve as examples of how the proposed parameters might be coded and weighted. The purpose of those tables is to indicate one solution to scoring the parameter values of a rating index. Categorisation of the values, however, should base on the level of risks since each map use context holds different level of credibility assessment; for example a map mashups that will be refer to and look up by users such as news, reports and current events, including crisis map, the value should put more weight on the aspect of ‘expertise’ of the sources; meanwhile for a low risk context such as a map mashup for entertainment or commercial context, the value of the weightage may balance between the ‘trustworthiness’ and ‘expertise’ dimensions.

Empirical research needs to be conducted to determine the criteria/ categories to weight the parameters as well as how to weight the values to score the criteria. The examples of criteria shown below coded values according to the level of reputation and/or the presence of the elements. Further studies to determine the appropriate and feasible criteria with which to measure the parameters are required, particularly when the criteria need to be detected in an automated manner.

Table 9-5 Indicative values for parameters in metadata components

Parameters in Metadata Component	**Possible parameters where the values need to be weighted by future research**

Background data supplier Foreground data supplier Reputation domain (URL)	**Examples parameters:** Authoritative sources (e.g. government mapping provider, governmental data sources) Commercial mapping provider Non-authoritative sources (e.g. news, non-profit organisation) Collaborative crowd source applications (e.g. OpenStreetMap, Wikimapia) High reputation organisation Individual source Low reputation organisation Anonymous source Unidentified source
Disclose identity of map author/creator– name recognition, brand reputation, awards, certificates, qualification (authority), or labels	**Examples parameters:** Disclose identity that indicate good reputation Good reputation Identity not disclosed Not disclosed any identity
Disclose author or owner of website/map mashup application	**Examples parameters:** Disclose author/owner Not disclosed
Seal approval from certified high reputation organisation	**Examples parameters:** Disclose seal approval Not disclosed
Affiliation, association	**Examples parameters:** Disclose affiliation or association with high reputation organisation Not disclose affiliation or association with any organisation Disclose affiliation or association with low reputation organisation
Disclose sponsorship	**Examples parameters:** Non-profit organisation Governmental based sponsorship Commercial sponsor
Disclose mission/purpose/motive	**Examples parameters:** Disclose mission/purpose/motive Not disclose
Currency – date published, last updated	**Examples parameters:** Current updated 6 months last updated 1 year last updated Updated less than 5 years

	Display date published Display date last updated Not updated for more than 5 years No date display

Table 9-6 Indicative weighting values for parameters in map data/information components

Parameters in Map Data/information component	**Possible parameters where the values need to be weighted by future research**
Logical consistency Absolute positioning consistency Relative positioning consistency Attribute/thematic correctness/consistency Completeness Temporal consistency	**Examples parameters:** consistency with other matching dataset
Third party reviews – comments or feedbacks	**Examples parameters:** Positive comments/feedbacks Negative comments/feedbacks
Number of backlinks, number of sharing in blogs or social media	**Examples parameters:** number of sharing, backlinks
Star ratings or number of 'thumb likes'	**Examples parameters:** Number of star ratings, thumb likes

Table 9-7 Indicative weighting values for parameters in usability and accessibility components

Map mashup usability and information clarity	**Possible parameters where the values need to be weighted by future research**
Functionality –website icons, search toolboxes, map functions, broken links	**Examples parameters:** Number of errors, dysfunctions etc.
Tools grouping	**Examples parameters:** grouping order
Data layers organisation – on off layers versus zooming level	**Examples parameters:** Layers organisation
Colour scheme of foreground and background text labels and messages and map features, thematic zoning	**Examples parameters:** Colour conventions Based on cartographic design
Map layout positioning – arrow,	

legend, scale bar, zooming scale	
Scales versus symbol sizes	
Map features size and types Symbol type, size, Labels and messages typography	

9.3.3 The Proposed formula to produce the accumulative ratings

The previous section presents the indicative parameters and their weighting value. The formula to calculate the total rating value is shown below:

Total rating (Equation 9.1)

$= w_1 \sum parameters\ metadata + w_2 \sum parameters\ map\ data +$
$w_3 \sum parameters\ usability$ and accessibility (U&A)

$$= w_1 \sum_{i=1-n_{meta}}^{n_{meta}} meta + w_2 \sum_{i=1-n_{map}}^{n_{map}} data_i + w_3 \sum_{i=1-n_u}^{n_u} usability_i$$

Where

$$w_1 \sum_{i=1-n_{meta}}^{n_{meta}} meta$$

$$= w_1 \sum_{i=1-n_{meta}}^{n_{meta}} (meta_i \times \sum_{j=1}^{k} w_j)$$

w_1 = weighted value of metadata component

w_2 = weighted value of data component

w_3 =weighted value of U&A component

(the weighted value may be determine according to the importance of components within context and over total rating)

Normalisation of the scores might be needed to standardise the total rating scores, especially when the parameters related to usability and accessibility components are included. Further studies have to examine the factors that need to be considered and the contexts in which normalisation of the results is required; for example in terms of the metadata component, there might be a need for normalisation of the total scores when the number of data supplier sources of low reputation is more than that of high reputation sources. The number of data suppliers will influence the accumulative scores; this is to avoid the rating scores from data sources of low reputation to exceed those from data suppliers of high reputation. Another issues that need to be further investigated are how to calculate the total rating and the extent of dynamic presentation of stamped CCTL label due to a variety of metadata from feature to dataset levels; for example how to calculate and stamped rating label if the parameters such as the sources and currency are varied at feature level where each features have different parameters and values; also how if the values of features are varies between zooming scales.

Further investigation is also needed to validate the list of parameters suggested for this model. The parameters proposed are mainly based on the literature. Empirical studies will have to be conducted to evaluate and determine the priority of parameters according to the application context; a map mashup of a disaster rescue operation might have important parameters and requirements that differ to those of a mashup for a consumer based application. Thorough case studies are required to determine the specific and context based parameters.

Future research is required to establish how to categorise the rating scores before label stamping can be produced on a map mashup application. This study has tested and demonstrated the influence of a Colour Coded Traffic Light (CCTL) rating label, whereby scores are coded into three categories – high credibility, intermediate credibility and low credibility. Further empirical research is needed to determine the range of scores among the three categories. An appropriate score measurement to categorise the ratings, using a binary or statistical model (Montero et al., 2004), has to be tested. It is worth to highlight that the proposed CCTL rating label is not intended to become the absolute and rigid approach to assist map mashup users evaluating web mapping applications. The intention is to demonstrate a positive influence in terms of users' perceptions in having a sort of seal approval on web mapping application in order to assist map users making informed judgement of information obtain on the web.

9.4 Implementation Issues of Automated Credibility Rating Index and Labelling

Previous sections have discussed the promising possibility and practicality of proposed parameters that can be assessed in automated form as well as the indicative criteria and calculations needed to produce a credibility rating index. This section discusses the possible management and administration issues that will arise if the automated assessment is to be realised in the near future.

One of the main issues concerns the choice of a suitable organisation that would be able to administer the database index and release the stamp labelling to map mashup applications. The ideal organisations to administer the system and process it at national level would be governmental based organisations, such as a national mapping provider or the Ministry of Information Technology and Communication. At an international level, the United Nations Educational, Scientific and Cultural Organisation (UNESCO), for example, seems to have the potential to become the organisation most suitable to monitor the proposed automated system. Due to their positions of responsibility, these organisations have the potential, to create and promote awareness among web consumers, particularly web based map users, of the risks of large amounts of unfiltered information disseminated on the World Wide Web. Moreover, the release of labelling from such authorised organisations may increase consumers' trust in rating labels stamped on map mashups.

Nevertheless, non-profit organisations are also suited to administer the automated labelling system for map mashup applications. A successful example can be seen on the HONcode stamped labelling system for health information related websites, managed by the Health on the Net Foundation (HON), a non-governmental organisation. The moderators or gatekeeper could be drawn from individuals who actively volunteered in contributing user-generated contents. This strategy has been used in the Wikimapia, Wikipedia and OpenStreetMap where there is a user group namely 'Advanced User', 'sysops' and 'Data Working Group', respectively which have special privilege to deal with the issues related to vandalism, copyright violation, disputes etc. (Goodchild and Li, 2012). In map mashup context, the moderators may draw from the active map mashup developers. For example, in the portal programmableapi.com, (a portal that record the available map mashup applications and application programming interfaces (APIs) that have been developed on the Web, have a record that could identify the active developers.

Driving forces, particularly from governments, are important to promote and support the use of stamped labelling on map mashup applications published or hosted from those countries. Mashup developers need to be nurtured to become responsible for the information being disseminated via their mashup applications, whereas consumers need to be educated about the risks of obtaining unidentified information via map mashup applications. Encouragement from authorised organisations may speed up progress towards the realisation of an automated credibility assessment of map mashup applications. For example, the Digital Earth vision proposed by the United States government in 1998 has produced fruitful progress towards its realisation, in terms of advancements in technology and application developments (Gore, 1998).

Another issue to be considered before the proposed automated system can be implemented concerns data privacy and the permission that has to be granted from mashup developers or owners to allow the automated system to assess their mashup applications. The proposed stamped labelling system differs from the labelling for the HONcode certificate; with the HONcode labelling, the owner of a website will request the HON to apply a certificate to their websites; the release of such a certificate is important for their business purposes. Therefore, in order to get permission from the mashup owner to assess their application, the advantages that they will obtain from the stamped labelling have to be made clear. More visitors to the websites and increased site popularity are two of the possible advantages of using the proposed stamped labelling. In a situation for mashup applications that might be classified as having low or intermediate credibility rating, the proposed automated tool can provide a free assessment for them to improve the design, sources and information being disseminated on their mashup application. Another issue is related to the granting of permission to validate the external data used on the application. In mashup applications, data typically comes from a variety of sources and is embedded on the base map using data APIs; the data might be stored in different servers so permission may be needed to access the origins of the data.

10 CONCLUSION

10.1 Introduction

This research has conducted an empirical study to examine the dominant influence of credibility elements and to support the implementation of automated credibility assessment and labelling on map mashup. There are three objectives of this research which relate to;

Objective 1:

To examine the influence of metadata related elements specifically map producer and map supplier on respondents' assessment of the credibility of map mashup information.

Objective 2:

To examine the influence of colour coded traffic light (CCTL) labelling on respondents' assessment of credibility

Objective 3:

Through literature studies, to propose a conceptual framework to support the implementation of automated credibility assessment and labelling for map mashup applications.

To fulfil these three objectives, the studies in this research have been conducted in three stages using mixed-method approaches, including online map based questionnaires, semi-structured interviews, observations and think-aloud protocols.

STAGE ONE: This stage examined the influence of metadata related to sources on users' judgement when assessing the credibility of map mashup information. The findings from the first stage lead to the study conducted in the second stage.

STAGE TWO: This stage examined the influence of 'seal approval' labelling by the implementation of colour-coded traffic light (CCTL) label on map mashup applications as a means of assisting users to assess the credibility of information. The findings from the second stage then lead to the third stage of this research,

STAGE THREE: This stage proposed a framework to support the implementation of automated credibility labelling on map mashup applications.

Table 10.1 highlights the findings of this research according to the hypotheses that have been tested in the experiments.

Table 10-1 Summary of the results of hypotheses revealed in this research

Objective 1	**To examine the influence of metadata related to sources (specifically the identity of map producer and map supplier) on respondents' assessment of the credibility of map mashup information.**		
Hypothesis 1	Visual cues have no significant influence in respondents' credibility assessment on map mashup applications	Not supported	Experiment 1: There were high influence of 'colour scheme' and 'information clarity' emerged in the responses (see Table 4.6). Experiment 2: The scores of the influence of 'visual cues' was significantly higher than the influence of 'data supplier' Experiment 3: The 'colour scheme', 'clarity of symbol' and 'symbol design' which used to measure the influence of visual cues had a high number of responses (approx. 94%) (see Chart 6.4). The visual cues appeared to be the main factor influencing respondents' judgement in these three experiments.
Hypothesis 2	The metadata related to sources (i.e. map	Not supported	Experiment 1: The results show that only

	data supplier, map mashup producer) have significant influence in respondents' credibility assessment		an average of 8% respondents use the identity of the map creator/author (map producer) as the main basis for their judgement when selecting maps for the tasks (see Table 4.9). Experiment 2: The metadata were measured by half of the respondents, whilst the other half did not measure these elements in the assessment. Experiment 2a: Although the respondents spotted (noticed) the metadata, it did not necessarily influence them to assess those elements in their credibility assessments. Although the visual cues were controlled and keep to a minimal influence, the ratio of respondents that had influence and not influence by the metadata was 50:50. Experiment 3: 51.9% respondents perceived the map producer element to be unimportant in influencing their judgement pertaining to map information credibility. Likewise, a high number of responses rated the importance of the map producer element at the undecided point (9.8%) compared to other manipulated elements. The response differences were

			statistically significant in the Chi-squared test X2 (1, n = 16) = 6.25, $p < 0.05$) (see Chart 6.4). Experiment 4: The results between Experiment 3 and 4 were compared to analyse the influence of metadata related to source (i.e. map producer) in two different contexts (with and without credibility label). From the Pearson's Chi-square test there was no significant association between the presence of a credibility rating and whether or not the identity of the map producer influenced respondents' judgement, $\chi2(1) = 0.021$, $p >0.8$. From the odds ratio, the odds of the 'no influence of the identity of the map producer in respondents' judgement' was 1.1 times higher if the map were not labelled with a credibility rating than if labelled (see Table 7.19).
Hypothesis 3	The metadata related to sources have significant influenced within geoliterate respondents	Not supported	Experiment 1: The majority of geoliterate respondents tend to use the concept of visual cues when making judgements to select the preferred maps in the tasks. This is in line with the non-geoliterate group responses, which show the dominant use of the keyword of 'colour' when making judgements in the tasks. Experiment 2: There were no significant differences in the level of

			priority rated between the two groups. The ratio that measured and did not measure within the groups in each condition was 1:1 (see Table 5.11). Experiment 2a: There was no significant different within the geoliterate group that chose either Map A (supplied by Starbucks Coffee) and Map B (supplied by the University of Nottingham); the results indicated seven geoliterate perceived Map A as more credible and ten geoliterate perceived Map B as more credible. (Table 5.18) Experiment 3: In Experiment 3, the number of responses that perceived the element of 'map producer' as important in their judgement however was less than 50% in the two groups (37% and 42%) (see Chart 6.7). Meanwhile, the responses in the two groups that perceived this element as not important in influencing their judgement was comparable, with both groups showing the number of responses to be around 52% (see Chart 6.8). Experiment 4: The influence of the identity of map producer from geoliterate respondents (Mdn = 16.66;

			mean = 34.21) did not differ significantly from non-geoliterate respondents (Mdn = 50.0; mean = 40.26), U = 552.0, z=-1.2, p>0.5, r = 0.1 (small effect size).
Hypothesis 4	There is no significant difference between the level of importance of the metadata related to sources between these two contexts - a prominence dependent (i.e. interpretation is based on what looks prominent) and a prominence independent setting (i.e. interpretation is based on a statement)	Not supported	Experiment 3: The importance of 'map producer' was higher at about 75% in prominence independent context (i.e. interpretation based on statements approach) than the importance of map producer element, which was only rated by 38%, in prominence dependent setting (i.e. interpretation based on what looks prominent).
Hypothesis 5	There is significant difference between the levels of perceived importance of the metadata variables related to source between these two levels of engagement contexts – low level (level one) versus high level (level three)	Not supported	The metadata related to sources (i.e. map producer), tested in the four experiments, was perceived to be less important element. In the low level (Level One) and high level task (Level Three), the visual cues had more influence than the metadata related to sources (i.e. map producer and map data supplier) in users credibility assessment (Table 7.18)
Objective 2	**To examine the influence of colour coded traffic light (CCTL) labelling on respondents' assessment of credibility**		
Hypothesis 6	Credibility label has a significant influence in respondents' credibility assessment	supported	The likelihood of respondents choosing the high credibility map, if they were given a map with a visual rating indicator (CCTL label), was three times higher than when

			given a map without the visual rating indicator (CCTL label). Analysis of the total scores that measure the influence of credibility labelling demonstrates a moderate influence on respondents' judgement (see Table 7.13).
Hypothesis 7	The presence of credibility label has a significant effect to the influence of visual cues in respondents' judgement	supported	The effect based on the odd ratio, the influence of colour scheme, was 8 times higher if the map was not labelled with a credibility rating than when labelled. Likewise, the influence of symbol design was 3.3 times higher if the map was not labelled with a credibility rating than when labelled. (see Table 7.15)
Hypothesis 8	There is significant difference between geoliterate and non-geoliterate respondents in the influence of visual cue variables when making judgement	Not supported	The influence of visual cue elements from geoliterate respondents (Mdn = 61.1) did not differ significantly from those of non-geoliterate respondents (Mdn = 72.2), $U = 521.5$, $z=-1.5$, $p>0.12$, $r = 0.2$ (small effect size) (see Chart 7.4)
Hypothesis 9	There is significant difference between geoliterate and non-geoliterate respondents in the influence of credibility labelling when making judgement	Not supported	In average the influence of the credibility labelling from geoliterate respondents (mean = 48.8, SE = 6.6) was lower than from non-geoliterate respondents (mean = 55.67, SE = 4.2). This difference was not significant $t(73) = -0.92$, $p > 0.05$: it did represent a small sized effect $r = 0.1$ (see Chart 7.5).

In conclusion, the thesis has achieved the objectives and produced several findings, as follows:

1) In the context of objective one, this research has identified a low influence of metadata related sources specifically the identification of the map producer (creator) and the supplier of map data when respondents make informed judgements relating to the credibility of information presented on map mashup applications. Although the respondents noticed the presence of the metadata related to sources, it did not necessarily influence the respondents, including geoliterate respondents, to judge credibility based on that element.

2) For objective two, this research has identified a moderate positive influence of credibility labelling when respondents make informed judgement relating to the credibility of information presented on map mashup applications.

3) For objective three, this research has designed a conceptual framework of practical elements to assess the credibility of a map mashup application in an automated manner; it has proposed an algorithm to produce a credibility rating index and an equation to calculate the accumulative rating value. This research has also identified several possible approaches to the implementation of the automated assessment using state-of-the-art technologies and it highlights several implementation issues that need to be tackled in order to implement such an automated credibility rating on map mashup applications.

10.2 Thesis Contributions

The contributions of this research to the state of the art are therefore as follows:

1) **The low influence of metadata related to sources on users' perceived credibility assessment**

This research has demonstrated the low influence of metadata related to authority elements, specifically map data producer and supplier that influence users when making judgement about the credibility of the information they obtain from map mashup; the research was conducted using a series of simulated experimental maps. The tasks have been designed at two different levels - a low involvement task and a high involvement task. In the low involvement task, the nature of the task was simple: respondents were required to simply choose the best map to guide them in the task. In the high involvement task, the task required a deeper level of engagement where respondents were required to make decisions by playing the given role in the task in a critical situation. This research also identified the differences between what they aware of and what they actually doing; users are aware the importance of elements of credibility, unfortunately, do not always checking the elements during assessment. These findings are important for designers of online maps, particularly map mashups, in considering the elements that users would notice (look prominence) in order to increase the credibility of online maps that they want to publish. However, this research also demonstrated that although the presence of metadata (i.e. foreground data supplier and map producer) noticed by respondents, including geoliterate respondents, it did not necessarily influenced them to judge credibility based on the sources. These findings add significantly to the understanding of individuals' perceptions, largely young adult, on the elements that they judge when assessing the credibility of information on map mashups.

2) The influence of credibility labelling on users' credibility assessment.

This research has identified a positive moderate influence of credibility labelling on users' judgement when assessing map mashup applications. This finding supports the research in the GIS domain that promotes the need for the dissemination of map 'data quality' via visualisation (see Devillers et al., 2007). The use of such a colour coded traffic light label on an online map, particularly on map mashups, is in line with other domains that use colour coded labelling on packaging to help users make an informed judgement about the product. For example, this system is used on food products in supermarkets and is also used for energy labelling of electrical appliances in the European Union countries. This initiative is also in line with the implementation of the 'seal of approval label' on e-commerce and health information websites using the HONcode label and also GEOLabel which is a spatial quality indicator on dataset accessed through GEOSS GeoPortal (Mass et al., 2011).

3) Parameters used in the automated credibility assessment.

This research has proposed the parameters to assess the credibility of a map mashup application. It is suggested that the proposed parameters should be evaluated in an automated manner due to cost considerations. Nevertheless, the proposed parameters could be used in semi-automated and manual assessments as well. The parameters, which can be divided into three major components - metadata, data and usability and accessibility- could be assessed collectively or individually if a scheme prefers to assess credibility in terms of single components.

4) An algorithm to produce a credibility index rating

This research has proposed an algorithm to produce a credibility rating index for map mashup applications; indicative criteria and values, as well as a formula to calculate the rating index, have been proposed. The formula considers the parameters that it is practical to assess in an automated manner.

5) Approaches to implement the credibility assessment in an automated manner

This research has suggested several approaches to the implementation of credibility assessment in automated manner involving current emerging technologies, including web data mining and a linked data infrastructure. Assessment in an automated manner is more efficient in terms of the cost of labour and processing time and the only practical approach for the future. The use of linked data and web data mining technologies makes the assessment by finding appropriate matching sources possible. This approach is more practical and feasible than manual checking and evaluation.

10.3 Limitations

The findings demonstrated from this research should be cited by considering these limitations. Several limitations that have been identified specifically discussed in general and in specific experiments as below;

1) The sample used in this research generally was in the age range of 18 to 35, where the average age was 22 years old. The findings could be generalised to young adult web users and may not be appropriate for other groups of web users. The scope of this study was the collection of responses from one group of web users, campus community members. This group is one of the prospective users of the growing online map mashup applications. The sample may be biased towards the campus community of undergraduate and postgraduate students, research students and research staff due to the distribution of the map based questionnaires within the

university campus; there is also a concern in this research of the familiarity issue of community members that might lead respondents for not being too critical in analysing the experimental maps since they were familiar with the tested environment. Another concern is on the knowledge (or experience) of the respondents with the concepts of map mashups that the background and foreground data could be supplied from various sources into one map.

2) This research examines credibility that bases on the cognitive and perceptual of map use; specific tests were conducted but limited to the respondents' perception and cognitive process within experimental simulated contexts only; this research did not apply more holistic functional map uses such as conducting a test using real context of map uses.

Experiment 1

3) This experiment used an open-ended questionnaire to gather the parameters (elements) that become the basis of map readers' judgements when selecting and rejecting a map. A few drawbacks have been identified when using this type of survey. One drawback was that respondents tend to provide one single response as the basis of their judgement; they might have additional reasons but do not include these in their comments due to restrictions such as time constraints. It is also difficult to verify whether the single element provided is actually the main basis of their judgement or just the element that came to mind when answering the question. Another drawback was the difficulty in supporting the findings statistically; this is due to the textual based analysis that coded the responses according to the concepts that emerged.

Experiment 2

4) This experiment used a simulation map mashup, using a non-interactive static map, and simple experimental tasks to stimulate responses from respondents. This experimental setting might be not sufficient to increase respondents' motivation to give full commitment. Respondents might have made judgements based on a surface assessment without spending enough time to analyse the map information critically. As argued by Petty and Cacioppo (1986), deep engagement with the tasks will promote respondents to make informed judgement by relying on critical elements instead of peripheral elements, such as colour scheme and visual design. The next experiment therefore was designed to create an interactive map setting and to apply a certain extent of cognitive stimuli in the

experimental tasks so that respondents would have to make critical judgement before giving their responses.

5) The main question in the survey might not be sufficient enough to force respondents to analyse the two comparison maps critically in each experiment. The main question, which was '*please choose the map that you will use to assist you in your self-guided campus tour*', might not be sufficient to make respondents step back and evaluate who or what is perceived to be responsible for the information. Some group of respondents might not critically analyse the map and not giving responses based on their perceived credibility to the map. Respondents might simply choose the map based on their preferences and interest, and not based on their perceived credibility to the presented information. In the next experiment (Experiment 3), therefore slight changes will be made to the wording by adding the term 'credibility (believability)' to the question.

6) The experimental variables within each condition might not be sufficient to influence respondents' judgement when assessing the map. The variables might be perceived as comparable with each other. One might not measure the variable because the two manipulated variables were not significantly different in terms of reputation and perceived credibility level. In the next experiment (Experiment 3), therefore, the variables being compared between the two maps are significantly different in terms of the reputation of experimental variable.

Experiments 3 and 4

7) The reliability of the scales used in the questionnaire was poor from the results of Cronbach's alpha, but the results from Intraclass Correlation indicated the correlation of responses from respondents was in large effect. This is because responses measured in this study were based on single item measures, instead of multiple item measures. Although multiple item measures are the dominant approach when measuring a psychological construct, due to the reliability of the scale that can be determined, it should not be considered to be a fatal flaw in one research if single item measures are applied (Wanous et al., 1997, p.247). A few studies have examined the correlations between single item measures and multiple item measures and these indicate good convergent and discrimant validity in single item measures (Hoeppner et al., 2011, p.311; Wanous et al.,1997, p.255)

8) This study tested one design of visual label which was CCTL. A few other designs, such as has been tested by (Kelly et al., 2009) on a food product, can be further tested on online map mashup to examine the influence and suitability of the labelling design.

It is worth to highlight that the findings demonstrated in this research should be cited by considering the limitations stated above.

10.4 Recommendations for future work

From the methodologies, results and analyses presented in this thesis, several opportunities for future work have been identified.

1) Examine the influence of metadata elements using holistic functional approach

   This research has examined the influence on persons making a judgement about the credibility of information from map mashup applications in the aspect of metadata, (namely the supplier of the data and the identity of the map mashup developer/creator). It is recommended for future studies to test and compare data sources (supplier or producer) that clearly distinguish in terms of reputation (e.g. national mapping provider versus anonymous), implies actual map use scenario applications and real map users of map mashup for specific contexts (that not limited to university members); for example actual community that dealing with emergency situation. There are several elements, such as the currency, correctness (accuracy) and completeness of information (data) that could be further evaluated. These elements could be examined at several levels of users' involvement (task) and in several use cases. Experiments could be further conducted using holistic approach such as applying real context of map uses to real population of map uses in order to confirm the influence in a real situation. The further experiment also should concern and filter the respondents according to their familiarity on the context of study including the map viewing environment and their knowledge (or experience) with map mashup application.

2) Development of a credibility index rating

   This research has proposed several parameters that could be assessed in the context of the credibility of a map mashup application. The parameters have been identified through a review of the literature. An algorithm to calculate the cumulative rating value has been proposed. However, several experimental studies have to be conducted, for future work in order:

- To develop a credibility rating index and parameter values for three credibility rating categories – low rating, intermediate rating and high rating. Experimental studies have to be conducted in order to produce benchmark values of the criteria/parameters.
- To evaluate the proposed equation to calculate cumulative rating values and to identify the circumstances in which data normalisation is required.

3) Development of automated credibility assessment

This research has proposed some approaches to assess credibility rating criteria using state-of-the-art technologies, including web data mining and a linked data infrastructure. Although the suggested approaches, particularly data assessment using a linked data infrastructure, are still evolving there is a high possibility that applications based on these technologies will be much more powerful in the next few years. By then, research will have been conducted and the technology will have matured enough to support the implementation of assessment in an automated manner. At present case studies could be conducted to evaluate the practically of evaluating the parameters in the three proposed components -metadata, data and medium- separately, using the suggested approaches. After that, the challenges to implementing the assessment in an automated manner using this technology could be identified in more detail and addressed in future research.

REFERENCES

Abascal, J., Arrue, M., Fajardo, I., Garay, N., and Omás, J. (2004). The use of guidelines to automatically verify Web accessibility. **Universal Access in the Information Society** 3(1): pp.71-79.

Al-Bakri, M. and Fairbairn, D. (2011) User generated content and formal data sources for integrating geospatial data. **25th International Cartographic Conference and the 15th General Assembly of the International Cartographic Association ICC, Palais des Congress, Paris. 3-8 July .**pp.1-8.

Abilock, D. (2008) Applying Information Literacy Skills to Maps. **Knowledge Quest** 36(4): pp.8-12.

Al-Khalifa, H. S. (2011). An Experimental System for Measuring the Credibility of News Content in Twitter. **International Journal of Web Information System** 7(2): pp.130-151.

Albert, W. A., and Van, D. G. T. M. (2011) Color Matters: Color as Trustworthiness Cues in Web Sites. **Technical Communication** 58(2): pp.149-160.

Aphinyanaphongs, Y., and Aliferis, C. (2007) Text Categorization Models for Identifying Unproven Cancer Treatments on the Web. **Studies Health Technology and Informatics** 129(2): pp.968-972.

Au, F. T. W., Baker, S., Warren, I., and Dobbie, G. (2008) Automated Usability Testing Framework. **Proceedings of the 9th Australasian User Interface Conference**. Wollongong Australia: Australian Computer Society, pp.55-64.

Auer, S., Lehmann, J., and Hellmann, S. (2009) LinkedGeoData: Adding a Spatial Dimension to the Web of Data. **The Semantic Web**. 5823: pp. 731-746.

Barnes, S. J., and Vidgen, R. (2001) An Evaluation of Cyber-Bookshops: The WebQual Method. **International Journal of Electronic Commerce** 6(1): pp.11-30.

Barnes, S. J., and Vidgen, R. (2003) Measuring Website Quality Improvement: A Case Study of the Forum on Strategic Management Knowledge Exchange. **Industrial Management and Data System** 103(5): pp. 297-309.

Barnes, S. J., and Vidgen, R. (2004) Interactive E-Government: Evaluating the Website of the UK Inland Revenue. **Journal of Electronic Commerce in Organisations** 2(1): pp.1-22.

Batty, M., Hudson-Smith, A., Milton, R., and Crooks, A. (2010) Map Mashups, Web 2.0 and the GIS revolution. **Annals of GIS** 16(1): pp. 1-13.

BBC News. (2007). **Label wars: GDA vs traffic lights** [online]. Available at http://news.bbc.co.uk/1/hi/uk/6231137.stm[Accessed February 14, 2011].

Berners-Lee, T., Hendler, J., and Lassila, O. (2001). **The Semantic Web. Scientific American: Feature Article** [online]. Available at http://www-sop.inria.fr/acacia/cours/essi2006/Scientific%20American_%20Feature%20Article_%20The%20Semantic%20Web_%20May%202001.pdf [Accessed Jan 2012].

Bishr, M. (2007) Weaving Space into the Web of Trust: An Asymmetric Spatial Trust Model for Social Networks. **Proceedings of the 1st Conference on Social Semantic Web (CSSW)**, Leipzig, Germany, September 26-28.

Bishr, M., and Kuhn, W. (2007) Geospatial Information Bottom-Up: A Matter of Trust and Semantics. In S. I. Fabrikant and M. Wachowicz (Eds.) **The European Information Society Leading the Way with Geo-information**. Springer Berlin Heidelberg, pp. 365-387.

Bishr, M., and Mantelas, L. (2008). A trust and reputation model for filtering and classifying knowledge about urban growth. **GeoJournal** 72(3): pp. 229-237.

Bizer, C., Cyganiak, R., and Gauß, T. (2007a) The RDF Book Mashup: From Web APIs to a Web of Data**. The 3rd Workshop on Scripting for the Semantic Web,** Innsbruck, Austria.

Bizer, C., Cyganiak, R., and Heath, T. (2007b). **How to Publish Linked Data on the Web**. [online] Available at http://wifo5-03.informatik.uni-mannheim.de/bizer/pub/LinkedDataTutorial/ [Accessed Dec 5, 2011].

Black, T. R. (2009) **Doing Quantitative Reseach in the Social Sciences**. London: SAGE.

Board, C. (1978) Map Reading Tasks Appropriate in Experimental Studies in Cartographic Communication. **Cartographica: The International Journal for Geographic Information and Geovisualization** 15(1): pp. 1-12.

Board, C. (1984) Higher Order Map-Using Tasks: Geographical Lessons In Danger Of Being Forgotten. **Cartographica: The International Journal for Geographic Information and Geovisualization** 21(1): pp. 85-97.

Boland, M. R., Rusanov, A., So, Y., Lopez-Jimenez, C., Busacca, L., Steinman, R. C., Bakken, S., Bigger, J. T. and Weng, C. (2013) From expert-derived user needs to user-perceived ease of use and usefulness: A two-phase mixed-methods evaluation framework. **Journal of Biomedical Informatics**. Available at: http://dx.doi.org/10.1016/j.jbi.2013.12.004

Brajnik, G. (2000) Automatic web usability evaluation: what needs to be done?. **The 6th Conference on Human Factors and the Web**, Austin Texas. 19 June. [online]. Available at http://sole.dimi.uniud.it/~giorgio.brajnik/papers/hfweb00.html [Accessed 17 Dec 2011].

Brajnik, G. (2004) Using Automate Tools in Accessibility and Usability Assurance Processes. In Stary, C. and Stephanidis, C (Eds.). **Lecture Notes on Computer Sciences** The 8th ERCIM Workshop on User Interface for All, Vienna,Austria, 3196:pp. 219-234.

Brewer, C. A. (2005) Designing better Maps: **A Guide for GIS Users. California,** USA: ESRI Press.

Brinck, T., and Hofer, E. (2002) Automatically Evaluating the Usability of Websites. **Conference on Human Factors in Computing System**, Minneapolis,USA.

Brown, B., and Perry, M. (2001) Of maps and guidebooks: designing geographical technologies. **ACM SIGGROUP Bulletin**, 22(3): pp. 28-32.

Caquard, S. (2014) Cartography II: Collective cartographies in the social media era. **Progress in Human Geography** 38(1): pp.141-150.

Calero, C., Ruiz, J., and Piattini, M. (2005) Classifying web metrics using the web quality model. **Online Information Review** 29(3): pp.227-248.

Calvo, R. A. (2001) Classifying Financial News with Neural Networks. **The 6th Australasian Document Symposium**, Darlinghurst, Australia.

Cappiello, C., Chiara, F., and Barbara, P. (2004) Data quality assessment from the user's perspective. **Proceedings of the 2004 International Workshop On Information quality in information systems**, Paris, France: ACM New York, pp. 68-73

Carter, J. R. (1988) The Map Viewing Environment: A Significant Factor in Cartographic Design. **The American Cartographer** 15(4):pp.379-385.

Carter, J. R. (2005) The many dimensions of map use. **The Proceedings of the XXII International Cartographic Conference**, 11-16 July. A Coruña. [online] Available at: http://www.cartesia.org/geodoc/icc2005/pdf/oral/TEMA12/Session%203/JAMES%20CARTER.pdf [Accessed Jan 2014].

Carroll, J. J., Bizer, C., Hayes, P., and Stickler, P. (2005) Named Graphs, Provenance and Trust. **The 14th International World Wide Web Conference** Chiba, Japan. ACM [online] Available at.doi: 10.1145/1060745.1060835 [Accessed Feb 2011]

Cartwright, W. (2008) Delivering Geospatial Information with Web 2.0. In M. P. Peterson (Ed.), **International Perspectives on Maps and the Internet** New York: Springer Berlin: pp. 11-28.

Casanova, C. A. (2008) **Discovery and retrieval of geographic data using Google**. MSc. Dissertation. Universidade Nova, Lisboa,Portugal.

Chang, C.-C., and Lin, C.-J. (2011) LIBSVM: A Library for Support Vector Machines. **Journal ACM Transactions on Intelligent Systems and Technology** (TIST), 2 (3): pp.1-27.[online] Available at doi: 10.1145/1961189.1961199

Chang, K.T. (2014) **Introduction to Geographic Information Systems** (Seventh ed.). Singapore: McGraw-Hill.

Charnock, D., Shepperd, S., Needham, G., and Gann, R. (1999) DISCERN: An Instrument for Judging the Quality of Written Consumer Health Information on Treatment Choices. **Journal Epidemiology Community Health**, 53: pp. 105-111.

Cheskin. (1999) **Ecommerce Trust Study. Cheskin Research and Studio Archetype/Sapient Research Project**. [online] Available at http://www.cheskin.com/cms/files/i/articles//17report-eComm%20Trust1999.pdf; [Accessed June 2012]

Christophersen, T., and Kondrat, U. (2010) Reliability, validity and sensitivity of a single-item mesure of online usability. **International Journal of Human Computer Studies**, 69: pp. 269-280. Available at doi: 10.1016/j.ijhcs.2010.10.005

Clarke, D. (2003) Are you functionally map literate? **The Proceedings of the 21st International Cartographic Conference** (ICC), Durban, South Africa.[online] Available at: http://lazarus.elte.hu/cet/publications/088.pdf [Accessed Jan 2014].

Comber, A. J., Fisher, P. F., Harvey, F., Gahegan, M. and Wadsworth, R. (2006) Using Metadata to Link Uncertainty and Data Quality Assessments. In: Riedl, A., Kainz, W. and Elmes, G. (eds.) **Progress in Spatial Data Handling**. Springer Berlin Heidelberg: pp.279-292.

Comber, A. J., Fisher, P. F., and Wadsworth, R. A. (2007) User-Focused Metadata For Spatial Data, Geographical Information And Data Quality

Assessments. **The 10th AGILE International Conference on Geographic Information Science**. Aalborg, Denmark. 8-11 May. pp.1-13. [online] Available at: http://www.agile-online.org/Conference_Paper/CDs/agile_2007/PROC/PDF/71_PDF.pdf [Accessed Dec 2011].

Comber, A. J., Fisher, P. F., and Wards, R. A. (2008) Semantics, Metadata, GI and Users. **Transaction in GIS**, 12(3): pp. 287-291.

Corritore, C. L., Kracher, B., and Wiedenbeck, S. (2003) Online Trust: Concept, Evolving Themes, a Model. **International journal Human Computer Study,** 58: pp.737-758.

Corritore, C. L., Marble, R. P., Wiedenbeck, S., Kracher, B., and Chandran, A. (2005) Measuring Online Trust of Website: Credibility, Perceived Ease of Use, and Risk. **The Americas Conference on Information System (AMCIS) Proceedings**. Paper 370. [online] Available at: http://aisel.aisnet.org/amcis2005/370/ [Accessed March 2010].

Craglia, M., Goodchild, M., Annoni, A., Camara, G., Gould, M., Kuhn, W., et al. (2008) Next-Generation Digital Earth: A position paper from the Vespucci Initiative for the Advancement of Geographic Information Science. **Internationl Journal of Spatial Data Infrastructures Research**, 3: pp. 146-167.

Craglia, M. (2009) From INSPIRE to Digital Earth: Research Challenges for the Next Generation Spatial Data Infrastructure. **The 12th AGILE International Conference on Geographic Information Science**, Hannover German. 2-5 June.[online] Available at: http://ec.europa.eu/dgs/jrc/downloads/jrc_esof_2010_presentation_craglia.pdf [Accessed May 2010].

Creswell, J. W., and Plano Clark, V. L. (2007) **Designing and conducting mixed methods research.** USA: SAGE Publications.

Creswell, J. W. (2009) **Research Design: Qualitative, Quantitative and Mixed Methods Approaches** (Third Edition ed.). United States of America: SAGE Publications.

Danko, D. M. (2005) ISO/TC211: Geographic Information-Metadata ISO 19115. In H. Moellering, H. J. G. L. Aalders and A. Crane (Eds.), **World Spatial Metadata Standards: Scientific and Technical Descriptions, and Full Descriptions with Crosstable** The Netherlands: Elsevier: pp. 535-555.

David, R., and Jason, H. (2008) Aesthetics and credibility in web site design. **Information Processing and Management**, 44(1): pp. 386-399.

DBpedia. (2013) **DBpedia**. [online]. Available at http://dbpedia.org/About [Accessed Feb 3, 2013]

Devillers, R., Bedard, Y., and Jeansoulin, R. (2005) Multidimensional Management of Geospatial Data Quality Information for Its Dynamic Use Within GIS. **American Society for Photogrammetry and Remote Sensing** 71(2): pp. 205-215.

Devillers, R., Bedard, Y., Jeansoulin, R., and Moulin, B. (2007) Towards Spatial Data Quality Information Analysis Tools for Expert Assessing the Fitness for Use of Spatial Data. **International Journal of Geographic Information Science**, 21(3): pp. 261-282.

DOM. (n.d.) **Introduction to DOM**. [online] Available at http://www.w3.org/TR/%20DOM-Level-2-Core/introduction.html [Accessed Dec 5, 2011].

DuVander, A. (2012) **8,000 APIs: Rise of the Enterprise**. [online] Available at: http://blog.programmableweb.com/2012/11/26/8000-apis-rise-of-the-enterprise/ [Accessed June 2013].

Elwood, S., Goodchild, M. F., and Sui, D. Z. (2012) Researching Volunteered Geographic Information: Spatial Data, Geographic Research, and New Social Practices. **Annals of the Association of American Geographers** 102(3): pp.571-590.

Elzakker, C. J. M., and Wealands, K. (2007) Use and Users of Multimedia Cartography. In W. Cartwright, M. Peterson and G. Gartner (Eds.), **Multimedia Cartography** Springer Berlin Heidelberg, pp. 487-504.

Elzakker, C. J. M. (2004) **The use of maps in the exploration of geographic data**. PhD thesis. Utrech University, Netherlandas.

Eysenbach, G., and Kohler, C. (2002) How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and indepth interviews. **British Medical Journal**, 324: pp.573-576.

Eysenbach, G., Kohler, C., Yihune, G., Lampe, K., Cross, P., and Brickley, D. (2001) A Metadata Vocabulary for Self-and Third Party Labelling of Health Website: HIDDEL. **Proceedings AMIA Symposiums**: pp.169:173.

Fallis, D., and Fricke, M. (2002) Indicators of accuracy of consumer health information on the Internet: a study of indicators relating to information for managing fever in children in the home. **Journal of the American Medical Informatics Association: JAMIA**, 9(1): pp.73-79.

Ferebee, S. (2007) An Examination of the Influence of Involvement Level of Web Site Users on the Perceived Credibility of Web Sites. **Persuasive Technology: Lecture Notes in Computer Science** 4744: pp. 176-186.

Fichter, D. (2009) Chapter 1: What is Mashup. In Engard, N. C. (Ed.), **Library Mashups: Exploring New Ways to Deliver Library Data** USA: Information Today, pp. 352.

Field, A. (2009) **Discovering Statistics Using SPSS** (Third ed.). Dubai: SAGE.

Flanagin, A. J., and Metzger, M. J. (2000) Perception of Internet Information Credibility. **Journalism and Mass Communication Quartely,** 77(3):pp. 515-540.

Flanagin, A. J., and Metzger, M. J. (2008) The credibility of Volunteered Geographic Information. **GeoJournal**, 72:pp. 137-148.

Fogg, B. (2002) Chapter 7: Credibility and the World Wide Web**.** In. Fogg, B. (Ed.) **Persuasive Technology: Using Computers to Change What We Think and Do,** USA: Morgan Kaufmann, pp. 312.

Fogg, B. J. (2003) Prominence-interpretation theory: explaining how people assess credibility online. **Proceedings of Computer Human Interaction**, April 5-10, Florida USA: ACM, pp. 722-723.

Fogg, B. J., Cathy, S., David, R. D., Leslie, M., Julianne, S., and Ellen, R. T. (2003) How do users evaluate the credibility of Web sites?: a study with over 2,500 participants. **The Proceedings of the Conference on Designing for User Experiences,** ACM: New York, pp. 1-15.

Fogg, B. J., Marable, L., Stanford, J., and Tauber, E. R. (2002) **How Do People Evaluate a Website's Credibility: Consumer Reports WebWatch** [online] Available at: http://risingline.com/pdf/stanford-web-credibility.pdf [Accessed June 2009].

Fogg, B. J., and Tseng, H. (1999) The elements of computer credibility. **The Proceedings of the SIGCHI conference on Human factors in computing systems**. ACM: New York, pp.80-87.

Fonts Bartolomé, Ó., Huerta Guijarro, J., Díaz Sánchez, L., and Granell Canut, C. (2010) **OpenSearch-geo: the simple standard for geographic web search engines**. MSc. Dissertation. Universitat de Girona.

Fox, S. (2006) **Pew Internet and American Life Project: Online Health Search 2006**. USA. [online]. Available at: http://www.pewinternet.org/2006/10/29/online-health-search-2006/ [Accessed June 2011].

Fu, P., and Sun, J. (2011) **Web GIS: Principles and Applications**. USA: ESRI Press.

Gaudinat, A., Cruchet, S., and Boyer, C. (2011) Enriching the trustworthiness of health-related web pages. **Health Informatics Journal**, 17(2): pp. 116-126.

Gaudinat, A., Grabar, N., and Boyer, C. (2007a) Combination of Heterogeneous Criteria for the Automatic Detection of Ethical Principles on Health Web sites. **The American Medical Informatics Association Symposium Proceedings**. pp. 264-268.

Gaudinat, A., Grabar, N., Boyer, C., Bellazzi, R., Abu-Hanna, A., and Hunter, J. (2007b) Automatic Retrieval of Web Pages with Standards of Ethics and Trustworthiness Within a Medical Portal: What a Page Name Tells Us. **Artificial Intelligence in Medicine**. Springer Berlin Heidelberg. 4594: pp. 185-189.

GeoMAPP. (2009) **FGDC and Dublin Core Metadata Comparison**. [online]. Available at http://www.geomapp.net/docs/MetadataComparison_200903.pdf [Accessed 30 Jan, 2012].

Golbeck, J. A. (2005) **Computing and applying trust in web-based social networks**. PhD thesis, University of Maryland, College Park.

Goodchild, M. F. (2007a) Citizens as Sensors: The World of Volunteered Geography. **GeoJournal**, 69(4): pp.211-221.

Goodchild, M. F. (2007b) Citizens as Voluntary Sensors: Spatial Data Infrastructure in the World of Web 2.0. **International Journal of Spatial Data Infrastructures Research**, 2: pp.24-32.

Goodchild, M. F. (2007c) Beyond Metadata: Towards User Centric Description of Data Quality. **The 5th International Symposium on Spatial Data Quality**, Eenschede Netherlands. [online]. Available at: http://www.geog.ucsb.edu/~good/papers/435.pdf [Accessed April 2010].

Goodchild, M. F. (2008) **Assertation and Authority: The Science of User-Generated Geographic Content**. [online]. Available at www.geog.ucsb.edu/~good/papers/454.pdf [Accessed September 9, 2010].

Goodchild, M. F., and Li, L. (2012) Assuring the quality of volunteered geographic information. **Spatial Statistics**, 1: pp.110-120.

Goodchild, M. F. and Glennon, J. A. (2010) Crowdsourcing geographic information for disaster response: a research frontier. **International Journal of Digital Earth**, 3: pp. 231-241.

Goodwin, J., Hart, G., and Dolbear, C. (2008) Geographical Linked Data: The Administrative Geography of Great Britain on the Semantic Web. **Transaction in GIS**, 12: pp.19-30.

Google. (2012) **Link Schemes**. [online]. Available at: http://support.google.com/webmasters/bin/answer.py?hl=enandanswer=66356 [Accessed June, 2012].

Gore, A. (1998) **The Digital Earth: Understanding our planet in the 21st Century**. Speech for California Science Center. [online] Available at: http://portal.opengeospatial.org/files/index.php?artifact_id=6210andversion=1andformat=htm [Accessed Jan 2009].

Griffiths, K. M., Tang, T. T., Hawking, D., and Christensen, H. (2005) Automated Assessment of the Quality of Depression Websites. **Journal of Medical Internet Research** 7(5): pp.e59.

Grossner, K. E. (2006) Is Google Earth, Digital Earth? - Defining a Vision. **The UCGIS Conference**, Vancouver, WA.[online]. Available at: http://kgeographer.com/assets/Grossner_Clarke_IsGoogleEarthDigitalEarth.pdf [Accessed Feb 2009].

Hahmann, S., and Burghardt, D. (2010) Connecting LinkedGeoData and Geonames in the Spatial Semantic Web. **The 6th International GIScience Conference**. [online] Available at: http://www.giscience2010.org/pdfs/paper_165.pdf [Accessed June 2012].

Haklay, M. (2010) How Good is OpenStreetMap Information? A Comparative Study of OpenStreetMap and Ordinance Survet datasets for London and the rest of England. **Environment and Planning B: Planning and Design**, 37(4): pp. 683-703.

Haklay, M., Singleton, A., and Parker, C. (2008) Web Mapping 2.0: The Neogeography of the GeoWeb. **Geography Compass**, 2(6): pp. 2011–2039.

Haklay, M., and Tobon, C. (2003) Usability Evaluation and PPGIS: Towards a user-centered design approach. **International Journal of Geographic Information Science,** 17(6): pp. 577-592.

Haklay, M., and Weber, P. (2008) OpenStreetMap: User-Generated Street Maps. **Pervasive Computing, IEEE**, 7(4): pp.12-18.

Hanowski, R. J., Kantowitz, S. C. and Kantowitz, B. H. (1994) Driver Acceptance of Unreliable Route Guidance Information. **Proceedings of the Human Factors and Ergonomics Society Annual Meeting** 38**:** pp.1062-1066.

Harding, J. (2005) Chapter 8: Vector Data Quality : A Data Provider's Perspective. In R. Devillers and R. Jeansoulin (Eds.) **Fundamentals of Spatial Data Quality**. Great Britain: ISTE, pp. 141-159.

Hedden, H. (2010) **User Interfaces with taxonomies**. [online]. Available at http://www.hedden-information.com/User%20Interfaces%20with%20Taxonomies.pdf [Accessed Jan 5, 2012].

Herbig, P., and Milewicz, J. (1995) The Relationship of Reputation and Credibility to Brand Success. **Journal of Consumer Marketing.** 12(4): pp. 5-10.

Hill, L. (Ed.). (2000) **Core elements of digital gazetteers: placenames, categories, and footprints**. [online]. Available at http://www.alexandria.ucsb.edu/~lhill/paper_drafts/ECDL2000_paperdraft7.pdf [Accessed Jan 2011].

Hilligoss, B., and Rieh, S. Y. (2007) Developing a Unifying Framework of Credibility Assessment: Construct, Heuristics and Interaction in Context. **Information Processing and Management**, 44: pp. 1467-1484.

Hoeppner, B. B., Kelly, J. F., Urbanoski, K. A. and Slaymaker, V. (2011) Comparative utility of a single-item versus multiple-item measure of self-efficacy in predicting relapse among young adults. **Journal of Substance Abuse Treatment**, 41: pp.305-312.

Hong, T. (2006) The Influence of Structural and Message Features on Web Site Credibility. **Journal of The American Society for Information Science and Technology**, 57(1): pp.114-127.

Hovland, C. I., and Weiss, W. (1951) The Influence of Source Credibility on Communication Effectiveness. **Public Opinion Quarterly**, 15(4):pp. 635-650.

Hsu, C.-W., Chang, C.-C., and Lin, C.-J. (2000). **A Practical Guide to Support Vector Classification**. [online] Available at: http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf [Accessed 10 July 2011]

Hurst, P., and Clough, P. (2013) Will we be lost without paper maps in the digital age? **Journal of Information Science**, 39(1):pp. 48-60. Available at doi: 10.1177/0165551512470043.

Iding, M. K., Crosby, M. E., Auernheimer, B., and Klemm, E. B. (2009) Website Credibility: Why Do People Believe What They Believe. **Instructional Science**, 37: pp.43-63.

Jaaskelainen, R. (2010) Think-aloud protocol. In Gambier, Y. and Doorslaer, L. V. (Eds.) **Handbook of Translation Studies**. John Benjamins Publishing. 2: pp. 371-373.

Jackson, M., Schell, D., and Taylor, F. (2009) NSDI- Coordinating Framework or Battleground for the Management of Geospatial Data. **GIS Professional Magazine**, June 2009: 20-22.

Janowicz, K., Wilkes, M., and Lutz, M. (2008) Similarity-Based Information Retrieval and Its Role within Spatial Data Infrastructures. **Geographic Information Science: Lecture Notes in Computer Science.** 5266: pp.151-167.

Juffinger, A., Granitzer, M., and Lex, E. (2009) Blog Credibility Ranking by Exploiting Verified Content. **The Proceedings of the 3rd workshop on Information Credibility on the Web**, Madrid, Spain.

Kantowitz, B. H., Hanowski, R. J. and Kantowitz, S. C (1997) Driver Acceptance of Unreliable Traffic Information in Familiar and Unfamiliar Settings. Human Factors: **The Journal of the Human Factors and Ergonomics Society**. 39: pp. 164-176.

Kelly, B., Hughes, C., Chapman, K., Louie, J. C.-Y., Dixon, H., Crawford, J., et al. (2009) Consumer testing of the acceptability and effectiveness of front-of-pack food labelling systems for the Australian grocery market. **Health Promotion International**, 24(2): pp.120-129.

Kim, M. and Bednarz, R. (2013) Development of critical spatial thinking through GIS learning. **Journal of Geography in Higher Education**, 37: 350-366.

Kim, H.-L., Scerri, S., Passant, A., Breslin, J. G., and Kim, H.-G. (2011) Integrating Tagging into the Web of Data: Overview and Combination of Existing Tag Ontologies. **Journal of Internet Technology**, 12(4): pp. 561-572.

Kimberling, A. J., Buckley, A. R., Muehrcke, P. C., and Muehrcke, J. O. (2012). **Map use: reading, analysis, interpretation** (7th ed.). Redlancs, California: ESRI Press Academic.

Knight, K. (2010). **Report: Consumers react to seals of approval**. BizRepor: Ecommerce, [online]. Available at: http://www.bizreport.com/2010/08/report-consumers-react-to-seals-of-approval.html[Accessed August 23].

Kobilarov, G., Scott, T., Raimond, Y., Oliver, S., Sizemore, C., Smethurst, M., et al. (2009) Media Meets Semantic Web .How the BBC Uses DBpedia and

Linked Data to Make Connections. **The Semantic Web: Research and Applications Lecture Notes in Computer Science.** 5554: pp.723-737.

Kohring, M., and Matthes, J. (2007). Trust in news media: Development and validation of a multidimensional scale. **Communication Research**, 34(2): pp. 231-252.

Lankes, R. D. (2008) Credibility on the Internet: Shifting from Authority to Reliability. **Journal of Documentation**, 64(5): pp.667-686.

Lassila, O. (1997) **Introduction to RDF Metadata**. [online] Available at http://www.w3.org/TR/NOTE-rdf-simple-intro[Accessed July 4, 2012]

Lautenschutz, A.-K. (2012). Map Readers' Assessment of Path Elements and Context to Identify Movement Behaviour in Visualisations. **The Cartographic Journal: Cognitive, Behaviour, Representation Special Issue**, 49(4): pp.337-349.

Leibovici, D. G., Pourabdollah, A., and Jackson, M. (2011) Which Spatial Data Quality can be Meta-Propagated. **The 7th International Symposium on Spatial Data Quality**, Coimbra Portugal. INESC Coimbra, pp.71-76.

Lewis, D. D. (1998) Naive (Bayes) at Forty: The Independence Assumption in Information Retrieval. **Proceedings of the 10th European Conference on Machine Learning.** Springer-Verlag, pp. 4-15.

Li, S., and Gong, J. (2008) Mashup A New Way of Providing Web Mapping /GIS Services. **The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences**. XXXVI (part 4): pp. 639-648.

Liebenberg, E. C. (1998) Teaching Map Use in a Multicultural Environment. **South African Geographic Journal**, 80(2): pp. 111-117.

Lin, H. F. (2010) An application of fuzzy AHP for evaluating course website quality. **Computers and Education**, 54: pp. 877-888.

Linden, M., and Sheehy, N. (2004) Comparison of Verbal Questionnaire and Map in Eliciting Environmental Perceptions. **Environment and Behaviour**, 36(1): pp. 32-40.

Litwin, M. S. (1995) **How to Measure Survey Reliability and Validity**. USA: SAGE Publications.

Liu, S. B., and Palen, L. (2010) The New Cartographers: Crisis Map Mashups and the Emergence of Neographic Practice. **Cartograpy and Geographic Information Science**, 37(1): pp. 66-90.

Longstreet, P. (2010) Evaluating Website Quality: Applying Cue Utilization Theory to WebQual. **Proceedings of the 43rd Hawaii International**

Conference on System Sciences, 5-8 Jan. Hawaii. IEEE Computer Society, pp.1-7.

Lopez-Pellicer, F., Silva, M., Chaves, M., Javier Zarazaga-Soria, F., and Muro-Medrano, P. (2010) Geo Linked Data Database and Expert Systems Applications, **Lecture Notes in Computer Science**. Springer Berlin / Heidelberg, 6261: pp. 495-502)

Luyt, B., and Tan, D. (2010) Improving Wikipedia's credibility: References and citations in a sample of history articles. **Journal of the American Society for Information Science and Technology,** 61(4): pp.715–722.

MacEachren, A. M. (2004) **How Maps Works: Representation, Visualisation, and Design.** New York: The Guilford Press.

Mass, J., Serral, I., and Pons, X. (2011) GeoVIQUA: a FP7 Scientific Project to Promote Spatial Data Quality Usability: Metadata, Search and Visualisation. **The 7th International Symposium on Spatial Data Quality**. Coimbra Portugal: INESC Coimbra. pp. 71-76.

Maue, P., and Schade, S. (2008) Quality of Geographic Information Patchwork. **The 11th International Conference on Geographic Information Sciences**, University of Girona Spain, pp. 1-8.

Mayer, M. A., Karkaletsis, V., Stamatakis, K., Leis, A., Villarroel, D., Thomeczek, C., et al. (2006). MedIEQ - Quality Labelling of Medical Web Content Using Multilingual Information Extraction. **Studies in Health Technology and Informatics.** 121: pp. 183-190.

McGuinness, D. L., Zeng, H., Silva, P. P. d., Ding, L., Narayanan, D., and Bhaowal, M. (2006). Investigations into Trust for Collaborative Information Repositories: A Wikipedia Case Study. **The Proceeding of the WWW'06 Workshop on Models of Trust for the Web** (MTW'06), Edinburgh, Scotland, pp. 22-30.

McKnight, D. H., and Kacmar, C. J. (2007) Factors and Effects of Information Credibility. **The Proceedings Of The Ninth International Conference On Electronic Commerce**, pp. 423-432.

Metzger, M. J. (2007) Making sense of credibility on the Web: Models for evaluating online information and recommendations for future research. **Journal of American Society of Information, Science and Technology**. 58(13): pp.2078-2091.

Metzger, M. J., and Flanagin, A. J. (2011) Using Web 2.0 Technologies to Enhance Evidence-Based Medical Information. **Journal of Health Communication** 16, 45-58.

Metzger, M. J., Flanagin, A. J., and Medders, R. B. (2010) Social and Heuristics Approaches to Credibility Evaluation Online. **Journal of Communication** 60: pp. 413-439.

Molich, R., and Nielsen, J. (1990) Improving a human-computer dialogue. **Magazine Communications of the ACM** 33: pp.339-348.

Montero, F., Gonzalez, P., Lozano, M., and Vanderdonckt, J. (2004) Quality Models for Automated Evaluation of Web Sites Usability and Accessibility. **The Proceedings of the International Conference on Web Engineering**, ICWE'2004. Munich, pp.28-30.

Monmonier, M. S. (1991) **How to Lie with Maps** (First ed.). US: University of Chicago Press.

Morahan-Martin, J. M. (2004) How Internet Users Find, Evaluate, and Use Online Health Information: ACross-Cultural Review. **Cyberpsychology and Behaviour** 7(5): 497-510.

Muehlenhaus, I. (2012) If Looks Could Kill: The Impact of Different Rhetorical Styles on Persuasive Geocommunication. **The Cartographic Journal: Cognitive, Behaviour, Representation Special Issue** 49(4): pp. 361-375.

Mummidi, L., and Krumm, J. (2008) Discovering points of interest from users' map annotations. **GeoJournal** 72(3): pp.215-227.

Nagura, R., Seki, Y., Kando, N., and Aono, M. (2006) A Method of Rating the Credibility of News Documents on the Web. **The Proceedings of the 29th annual international ACM (SIGIR) conference on Research and Development in Information Retrieval**. Washington USA:ACM, pp. 683-684.

Nakamura, S., Shimizu, M., and Tanaka, K. (2008) Can Social Annotation Support Users in Evaluating the Trustworthiness of Video Clips? **The Proceedings of the 2nd ACM workshop on Information credibility on the Web**, California USA: ACM, pp. 59-62.

National Research Council (2006) **Learning to think spatially: GIS as a support system in the K-12 curriculum**. Washington DC: The National Academies Press. [online]. Available at http://esrik-12gis.emich.edu/k12/PDF/Learning%20to%20Think%20Spatially.pdf [Accessed Jan 2014].

Nivala, A.-M., Brewster, S., and Sarjakoski, L. T. (2008) Usability evaluation of Web Mapping Sites. **The Cartographic Journal: Use and Users Special Issue** 45(2): pp.129-138.

Nottingham. (2012) The University of Nottingham Official Website: **Student Population 2011/2012**. [online]. Available at: http://www.nottingham.ac.uk/about/facts/studentpopulation20112012.aspx[Accessed March 2013].

Nuth, K., Mitchell, N., and Schaab, G. (2007) Judging and Visualising the Quality of Spatio-Temporal Data on the Kakamega in West Kenya. **The International Symposium of Spatial Data Quality.** Enschede, The Netherlands, pp. 1-8.

Nyerges, T. L. (1991) Analytical Map Use. **Cartography and Geographic Information Systems**, 18(1): pp.11-22.

Ochoa, X., and Duval, E. (2008) Quantitative Analysis of User-Generated Content on the Web. **The Proceedings of the Web Science Workshop**, Beijing, China, April 22, pp. 30-36.

O'Donovan, J., Kang, B., Meyer, G., Hollerer, T., and Adali, S. (2012) Credibility in Context: An Analysis of Feature Distributions in Twitter. **The ASE/IEEE International Conference on Social Computing andInternational Conference on Privacy, Security, Risk, Trust.** [online] Available at: http://www.wigis.net/documents/papers/Credibility%20in%20Context%20An%20Analysis%20of%20Feature%20Distributions%20in%20Twitter.pdf [Accessed June 2012].

Ogrinz, M. (2009). **Mashup Patterns: Design and Examples for the Modern Enterprise**. USA: Addison-Wesley Professional.

Oort, P. V.(2005) **Spatial data quality: from description to application**. Delft, The Netherlands: NGC .

Ordnance Survey. (2013a) **Code-Point Open**. [online] Available at: http://www.ordnancesurvey.co.uk/oswebsite/products/code-point-open/[Accessed Feb 3, 2013].

Ordnance Survey. (2013b) **1:50000 Scale Gazetteer**. [online] Available at: http://www.ordnancesurvey.co.uk/oswebsite/products/50k-gazetteer/index.html [Accessed Feb 3, 2013].

Ordnance Survey. (2013c) **The linked data Web**. [online] Available at: http://www.ordnancesurvey.co.uk/oswebsite/education-and-research/research/linked-data.html [Accessed Feb 3, 2013].

Orphanet. (2013) **Orphanet: The Portal for Rare Disease and Orphan Drugs**. [online]. Available at http://www.orpha.net/consor/cgi-bin/index.php [Accessed March 4, 2013].

Papadias, D., Kalnis, P., Zhang, J., and Tao, Y. (2001) Efficient OLAP Operations in Spatial Data Warehouses. **Advances in Spatial and Temporal Databases: Lecture Notes in Computer Science**, 2121: pp. 443-459.

Park, H. W., Barnett, G. A., and Nam, I.-Y. (2002) Hyperlink–Affiliation Network Structure of Top Web Sites: Examining Affiliates with Hyperlink in Korea. **Journal of The American Society for Information Science and Technology**, 53(7): pp. 592-601.

Parker, C. J. (2012) **A Human Factors Perspective on Volunteered Geographic Information**. PhD. Thesis, Loughborough University.

Parker, C. J., May, A., and Mitchell, V. (2012) Understanding Design with VGI using an Information Relevance Framework. **Transaction in GIS**, 16(4): pp.545-560.

Parker, C. J., May, A., and Mitchell, V. (2013) The role of VGI and PGI in supporting outdoor activities. **Applied Ergonomics**, 44(6): pp.886-894. Available at doi: http://dx.doi.org/10.1016/j.apergo.2012.04.013

Perkins, C. (2008) Cultures of Map Use. **The Cartographic Journal**, 45(2): pp.150-158.

Peng, Z. R., and Tsou, M. H. (2003) **Internet GIS: Distributed Geographic Information Services for the Internet and the Wireless Network**. US: Wiley and Sons.

Petty, R. E., and Cacioppo, J. T. (1986) The Elaboration Likelihood Model of Persuasion. **Advances in Experimental Social Psychology**, 19: pp.179-190.

Phillips, R. J. (1984) Experimental Method in Cartographic Communication: Research on Relief Maps. **Cartographica: The International Journal for Geographic Information and Geovisualization**, 21(1):pp.120-128.

Poore, B., and Wolf, E. (2010) The Metadata Crisis: Can geographic information be made more usable?. **The Second Workshop of Usability of Geographic Information**. University College London, pp.1-8.

Popping, R. (2000). **Computer-assisted Text Analysis**. London: SAGE.

Price, S. L., and Hersh, W. R. (1999) Filtering Web Pages for Quality Indicators: An Empirical Approach to Finding High Quality Consumer Health Information on the World Wide Web. **The American Medical Informatics Association Symposium**. PubMed, pp. 911-915.

Princeton (2002) **A matter of trust: what users want from websites**.[online]. Available at: http://www.consumerwebwatch.org/dynamic/web-credibility-reports-a-matter-of-trust-abstract.cfm [Accessed November, 2011].

Provisu (2013) **Trustworthy Internet Documents on Vision Disorders, Eye Diseases and the prevention of blindness**. [Online]. Available at: http://www.provisu.ch/PROVISU/index_en.html [Accessed Jun 1, 2013].

Provost, M., Koompalum, D., Dong, D., and Martin, B. C. (2006) The initial development of the WebMEdQual scale: Domain assessment of the construct of quality of health website. **International Journal of Medical Informatics**, 75: pp.42-57.

Quackwatch. (2010). **A Special Message for Cancer Patients Seeking Alternative Treatment**.[online]. Available at: http://www.quackwatch.org/00AboutQuackwatch/altseek.html[Accessed Dec 6, 2011]

Rieh, S. Y. (2002) Judgement of Information Quality and Cognitive Authority in the Web. **Journal of The American Society for Information Science and Technology**, 53(2): pp.145-161.

Rieh, S. Y., and Belkin , N. J. (1998) Understanding Judgement of Information Quality and Cognitive Authority in the WWW. **Proceeding of the 61st Annual Meeting of the American Society for Information Science** Medford, NJ: Information Today 35: pp. 279-289.

Rieh, S. Y., and Belkin, N. J. (2000) Interaction on the Web: Scholars' Judgement of Information Quality and Cognitive Authority. **The Proceeding of the 63rd Annual Meeting of the Annual Meeting of the American Society for Information Science** (ASIS), Chicago.pp. 25-38.

Rieh, S. Y., and Danielson, D. R. (2007) Credibility: A multidisciplinary framework. **The Annual Review of Information Science and Technology**.Medford, NJ: Information Today, pp.307-364.

Rieh, S. Y., and Hilligoss, B. (2008) College Students' Credibility Judgments in the Information-Seeking Process: Digital Media, Youth, and Credibility. in Metzger, M. J., Flanagin, A.J., John, D. and MacArthur, C. **Foundation Series on Digital Media and Learning**. Cambridge, MA: The MIT Press,pp. 49–72. doi: 10.1162/dmal.9780262562324.049

Roberts, C. W. (2000). A Conceptual Framework for Quantitative Text Analysis. **Quality and Quantity**, 34(3): pp.259-274.

Roche, S., Propeck-Zimmermann, E., and Mericskay, B. (2013) GeoWeb and Crisis Management: Issues and Perspectives of volunteered geographic information. **GeoJournal**, 78: pp.21-40.

Rossukorn, S. (2011). **The influence of communication modality and shared visual information on collaboration in virtual teams**. MSc Dissertation. University of Nottingham, UK.

Scholz-Crane, A. (1998) Evaluating the Future: A Preliminary Study of the Process of How Undergraduate Students Evaluate Web Sources. **Journal Reference Services Review**, 26(3/4): pp.53-60.

Schutzberg, A., and Francica, J. (2005) **Geospatial Technology Offers Katrina Response Much, Delivers Some. Editorial**. [online]. Available at: http://www.directionsmag.com/editorials.php?article_id=1947[Accessed April 2012].

Schweiger, W. (2000) Media Credibility -experience or Image? A Survey on the Credibility of the World Wide Web in Germany in comparison to other Media. **European Journal of Communication**, 15(1): pp.37-59.

Selnes, F. (1996) Antecedents and consequences of trust and satisfaction in buyer-seller relationships. **European Journal of Marketing**, 32(3/4): pp. 305-322.

Shek, S. P. W., Sia, C.-L., and Lim, K. H. (2003) A Preliminary Assessment of Different Trust Formation Models: The Effect of Third Party Endorsement on Online Shopping. **The Proceedings of the 36th International Conference on System Sciences**, Hawaii. 6-9 Jan. Available at doi: 10.1109/HICSS.2003.1174443.

Sheridan, J., and Tennison, J.(2010) **Linking UK Government Data**: ACM Press.

Sidda, N. K. (2009) **A Framework for the Management of Spatial Data Quality Information.** MSc Dissertation. ITC Enschede, The Netherlands.

Skarlatidou, A. and Haklay, M. (2006) Public Web Mapping: Preliminary Usability Evaluation. **In Proceedings GIS Research UK**, April 5-7 Nottingham, UK, pp. 1-5.

Skarlatidou, A., Haklay, M., and Cheng, T. (2011) Trustee in Web GIS: the role of the trustee attributes in the design of trustworthy Web GIS applications. **International Journal of Geographic Information System**, pp.1-18.

Skarlatidou, A., Haklay, M., Cheng, T., and Francis, N. (2010a) Trust in Web GIS: A Preliminary Investigation of the Environment Agency's WIYBY Website with non-expert users. In Haklay, M., Morley, J. and Rahemtulla, H (eds). **Proceedings of the GIS Research UK 18th Annual Conference GISRUK 2010**, University College London (UCL).

Skarlatidou, A., Haklay, M., and Cheng, T. (2010b) Preliminiary Investigation of Web GIS Trust: The Example of the 'WIYBY' Website The **International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences**, 38(Part 11): pp.411-416.

Slocum, T. A., Blok, C., Jiang, B., Koussoulakou, A., Montello, D. R., Fuhrmann, S., and Hedley, N. R. (2001) Cognitive and Usability Issues in Geovisualisation. **Cartography and Geographic Information Systems**, 28(1): pp.61-75.

Spinsanti, L. and Ostermann, F. (2013) Automated Geographic Context Analysis for Volunteered Information. **Applied Geography**, 43: pp.36-44.

Stadler, C., Lehmann, J., Höffner, K., and Auer, S. (2012) LinkedGeoData: A Core for a Web of Spatial Open Data. **Semantic Web**, 3(4): pp. 333-354.

Stanford, J., Tauber, E. R., Fogg, B. J., and Marable, L. (2002) **Experts vs. Online Consumers: A Comparative Credibility Study of Health and Finance Web Sites. A Consumer Reports WebWatch research report**, pp.1-59. [online] Available at: http://www.consumerwebwatch.org/dynamic/web-credibility-reports-experts-vs-online-abstract.cfm [Accessed Jan 2009].

Stark, H.J. (2010) Quality assessment of VGI based on Open Web Map Services and ISO/TC 211 19100-family standards. **Free and Open Source Software for Geomatics Conference** FOSS4G 2010, Barcelona, pp: 101-110. [online] Available at: http://gispoint.de/fileadmin/user_upload/paper_gis_open/537509015.pdf [Accessed Jan 2011].

Stewart, K. J. (1999) Transference as a means of building trust in World Wide Web sites. **The Proceedings of the 20th International Conference on Information Systems**. North Carolina, US, pp. 459-464.

Strickland, J. (nd) **How Google Docs Works**. [online]. Available at: http://computer.howstuffworks.com/internet/basics/google-docs3.htm [Accessed Nov, 2012].

Suchan, T. A., and Brewer, C. A. (2000) Qualitative methods for research on mapmaking and map use. **The Professional Geographer**, 52: pp.145-154.

Sudman, S. (1976) **Applied Sampling**. United Kingdom: Academic Press.

Tiits, K. (2003) **Usability of Geographic Information System in Internet: A case study of Journey Planners**. MSc Dissertation. Tartu University.

Timpf, S., Volta, G. S., Pollock, D. W., and Egenhofer, M. J. (1992) A Conceptual Model of Wayfinding Using Multiple Levels of Abstraction Vol.

63. A. Frank, I. Campari and U. Formentini (Eds.), **Lecture Notes in Computer Scienc**e, pp 348-367.

Turner, A. J. (2006) **Introduction to Neoegeography**. O'Reilly Shortcut, pp.1-54. [online] Available at: http://pcmlp.socleg.ox.ac.uk/sites/pcmlp.socleg.ox.ac.uk/files/Introduction_to_Neogeography.pdf [Accessed Jan 2011].

Ulicny, B., and Baclawski, K. (2007) New Metrics for Newsblog Credibility. **The International Conference on Weblogs and Social Media**. Colarado, USA, pp.1-3.

Vaibhav. (2012) **Landscape of Digital Maps: Facts, Usage Trends and Users Statistics.**[online]. Available at: http://onvab.com/blog/digital-maps-facts-usage-trends-users-statistics/ [Accessed 23 January, 2013].

Vaus, D. D. (2008) **Analyzing Social Science Data**: SAGE

Wang, Y., and Richard, R. (2007) Rule-based Automatic Criteria Detection for Assessing Quality of Online Health Information. **The International Conference Addressing Information Technology and Communication in Health** (ITCH). Victoria, Canada, pp. 123-145.

Wang, Y. D., and Emurian, H. H. (2005) An Overview of Online Trust: Concepts, Elements, and Implications. **Computers in Human Behaviour**, 21: pp.105-125.

Wang, R. Y. and Strong, D. M . (1996) Beyond accuracy: what data quality means to data consumers. **Journal of Management and Information System**, 12: pp. 5-33.

Wanous, J. P., Reichers, A. E., and Hudy, M. J. (1997) Overall Job Satisfaction: How Good Are A Single-Item Measures?. **Journal of Applied Psychology,** 82(2): pp.247-252.

Warnick, B. (2004) Online Ethos: Source Credibility in an "Authorless" Environment. **American Behavioral Scientist**, 48(22): pp.256-265.

Wathen, C. N., and Burkell, J. (2002) Believe It or Not: Factors Influencing Credibility on the Web. **Journal Of The American Society For Information Science And Technology**, 53(2): pp.134–144.

Weibel, S. L. (2005) The Dublin Core Metadata Element Set In H. Moellering, H. J. G. L. Aalders and A. Crane (Eds.), **World Spatial Metadata Standards: Scientific and Technical Descriptions and Full Descriptions with Crosstable**. The Netherlands: Elsevier. pp. 493-513.

Weisberg, H. F., and Bowen, B. D. (1977) **An Introduction to Survey Research and Data Analysis**. San Francisco, United States of America: W.H. Freeman and Company.

Weisberg, H. F., Krosnick, J. A., and Bowen, B. D. (1996) **An Introduction to Survey Research, Polling, and Data Analysis** (Third Edition ed.). United States of America: SAGE publications, Inc.

Western Australian Department of Health. (2009) **Healthy Options WA: Food and Nutrition Policy for WA Health Services and Facilities**.[online]. Available at http://www.health.wa.gov.au/circularsnew/attachments/390.pdf [Accessed February 27, 2011].

Williams, K., and Calvo, R. A. (2002) A Framework for text categorisation. **The Proceedings of the 7th Australasian Document Computing Symposium**. Sydney, Australia, pp.1-7.

Williams, L. G. (1967) The Effects of Target Specification on Objects Fixated during Visual Search. **Acta Psychologica**, 27: pp. 355-360.

Wilson, M. W., and Graham, M. (2013) Neogeography and volunteered geographic information: a conversation with Michael Goodchild and Andrew Turner. **Environment and Planning A**, 45(1): pp.10-18.

Woodford, K., and Jackson, G. (2003) **Cambridge Advanced Learner's Dictionary.** United Kingdon: Cambridge University Press.

WRAPIN. (2012). **Worldwide Online Reliable Advice to Patients and Individual.** [online] Available at: http://www.wrapin.org/[Accessed May 4, 2012].

Wu, Y., and Winter, S. (2009) Inferring relevant gazetteer instances to a placename**. The 10th International Conference on GeoComputation**, University of New South Wales Sydney Australia, pp.1-7.

Yang, T. (2007) **Visualisation of Spatial Data Quality for Distributed GIS**. PhD thesis. University of New South Wales, Australia.

Yang, T., and Wang, J. (2004) Visualisation of Spatial Data Quality for Internet and Mobile GIS Applications. **Journal of Spatial Science**, 49(1): pp.97-107.

Yang, Y., and Liu, X. (1999) A re-examination of text categorisation methods. **The Proceedings Of The 22nd Annual International Conference On Research And Development In Information Retrieval**: ACM, pp. 42-49.

APPENDIX A

<u>Descriptions of Map Style used in Experiment 1</u>

Table 1 Descriptions of Map Style used in Experimental Task 1 (T1)

Variable 2	Currency	
Condition	1 (see Figure 4.2)	
	Map A	Map B
Currency parameter	No update information available	1999
Webpage background	Black	Grey
Building colour	Black grey	Black grey
Park/garden colour	Dark green	Light green
Land colour	Light yellow	Light grey
Road colour	Grey	Grey
Text attribute colour	White and black	White and black

Table 2 Descriptions of Map Style used in Experimental Task 2 (T2)

Variable 1	Identity of Mashup producer	
Condition	2 (see Figure 4.3)	
	Map A	Map B
Author/creator's identity	Undergraduate student	Member of outside community
Webpage background	Blue	Brown
Building colour	Red	Blue
Park/garden colour	Light green	Yellow, green
Land colour	Light yellow	White
Road colour	Red	Black
Text attribute colour	Black	White, black

Table 3 Descriptions of Map Style used in Experimental Task 3 (T3)

Variable 2	Currency	
Condition	3 (see Figure 4.4)	
	Map A	Map B
Currency parameter	12 January 2010	3 January 2009
Webpage background	Black	Yellow
Building colour	Light maroon	Light maroon
Park/garden colour	Light green	Light green
Land colour	Light brown	Light brown
Road colour	Red and black	Red and white
Text attribute colour	Black	Black

Table 4 Descriptions of Map Style used in Experimental Task 4 (T4)

Variable 2	Currency	
Condition	4 (see 4.5)	
	Map A	Map B
Currency parameter	25 June 2005	30 April 2007
Webpage background	Blue	Light purple
Building colour	Brown	Brown
Park/garden colour	Light green	Light green
Land colour	Light yellow	Purple
Road colour	White	Grey and white (secondary road)
Text attribute colour	Black	Black

Table 5 Descriptions of Map Style used in Experimental Task 5 (T5)

Variable 1	Identity of Mashup producer	
Condition	1 (see Figure 4.6)	
	Map A	Map B
Author/creator's identity	University researcher	Mapping Agency
Webpage background	Grey	Black
Building colour	Dark blue	Orange
Park/garden colour	Green	Bright green
Land colour	Light yellow	Light yellow
Road colour	Yellow, black line, black dotted line	Yellow, black line, black dotted line
Text attribute colour	Black	Black

Table 6 Descriptions of Map Style used in Experimental Task 6 (T6)

Variable 1	Identity of Mashup producer	
Condition	3 (see Figure 4.7)	
	Map A	Map B
Author/creator's identity	No information provided	Administrator, Level 2 NVQ
Webpage background	Light grey	Black
Building colour	Black	Dark purple
Park/garden colour	Yellow, green	Light green
Land colour	Light yellow	Light yellow
Road colour	Grey	Orange
Text attribute colour	White	Black

Additional Results: Currency

Table 7 Frequencies analysis of the map that respondents chose or rejected in each experimental task

	Most voted	%	Count	Least voted	%	Count	Chi-square test
Task 1	B	70.37	57	A	29.63	24	X^2 (1, n = 81) = 13.4, p <0.001)
Task 3	A	64.79	46	B	35.21	25	X^2 (1, n = 71) = 6.21, p <0.05)
Task 4	A	55.07	38	B	44.93	31	X^2 (1, n = 69) = 0.71, p <0.40)

Table 8 The occurrence of concepts in the responses to 'what was the basis of you choosing the map and rejecting the other map?'

Rank	**Concepts**	**T1**	**T3**	**T4**	**average**
1	Colour scheme	44	16	30	30
2	Information Clarity	38	31	21	30
3	Overall (Design look)	22	14	9	15
5	Information Details	21	11	4	12
6	Map design	23	5	4	11
8	currency	3	6	7	5

Table 9 A matrix of association between the concepts of "colour scheme" with other concepts

	Clarity	Design look	Combination (scheme)	Details	Individual preferences	Map design
T1 Colour	25	11	17	10	7	11
T3 Colour	5	5	6	5	0	1
T4 Colour	8	5	10	1	7	2

Table 10 Analysis of academic background of those who indicated concept of colour in the responses

	Geography, GIS, Land Survey, Cartography, Remote Sensing	Engineering	Sciences	Social Sciences, Law Education	Medical health sciences	others	Not respond
T1	11	11	6	4	1	5	5
T3	1	5	2	1	1	3	3
T4	10	5	5	4	2	3	1

Table 11 Analysis of academic background of those who indicated the concept of metadata in the responses

	Geography, GIS, Land Survey, Cartography, Remote Sensing	Engineering	Sciences	Social Sciences, Law Education	Medical health sciences	others	Not respond
T1	1	1	1	0	0	0	0
T3	2	1	3	0	0	0	0
T4	2	1	4	0	0	0	0

The number of geoliterate respondents that use these critical metadata elements (i.e. identity of author or currency) was very low. Hence, **Hypothesis 3 is not supported.** Example responses from the questionnaires are given below;

'Slightly out of date map compared to the other one' (F, geoliterate)

'...and the Maps B have been updated in 1999 which is no update information available in maps A' (F, geoliterate)

The results also have shown that only 3% to 10% of the total respondents use the currency of maps as the main basis for their judgement when selecting maps for the tasks. The respondents' percentage was higher (i.e. 10%) in the task that provides a comparison by stating that the map data was last updated in 2005 versus map data last updated in 2007 (experiment label: T4). Of this percentage, five respondents chose the map that stated 'the map last updated in 2007'; the reason for choosing this map was that it was more up to date than the second map. Two respondents chose the map that had been last updated in 2005; the reasons mentioned concerned the positive influence of the map appearance on their judgement, as below:

'Though the date is older than set B map, the style and map appearance looks more recent. (F, 25-30, geoliterate)'

‘The cartography is clearer although updated in 2005’ (F, 25-30, geoliterate)

In Task 3 (experiment label: T3), 8.5% of respondents used map currency as the basis on which to make their judgement when selecting or rejecting the map. This group of respondents chose the map updated in 2010 as their preferred map, rather than a map that had been updated in 2009. The most updated map was chosen in this task. Meanwhile, in Task 1 (experiment label: T1), 3.7% of respondents used map currency as their basis on which to select the map. In this task two respondents chose the map that provided the date it was last updated (i.e. 1999) while one (1) respondent chose the map that not did provide update information; the latter respondent chose that map because the year 1999 seems too outdated compared to a map that did not provide any update information. The map that does not provide update information has a possibility of being more current than the other map

Additional Results: The level of perceived credibility of the map

Descriptive analysis was conducted on the credibility related question of ‘how much you perceived the selected map is credible’ and ‘how much you perceived the rejected map is credible’. 5- point Likert scale was used to measure these questions. Table 12 presents the analysis of the proportions that chose the map that yields a high frequency of votes.

Table 12 Descriptive analysis on the level of users’ perceived credibility

		Mean	Standard Deviation (SD)	Standard Error (SE)	Skewness
Task 1 (n=81) No update vs. 1999	Most voted =B	3.86	0.082	+-0.082	-0.737
	Least voted = A	2.72	0.877	+-0.097	0.309
Task 2 (n=75) Student vs. outside community	Most voted =B	3.89	0.798	+-0.092	-0.787
	Least voted = A	2.72	0.924	+-0.107	0.065
Task 3 (n=71) 2010 vs. 2009	Most voted =A	4.13	0.675	+-0.080	-0.443
	Least voted = B	2.96	0.917	+-0.109	-0.258
Task 4	Most	3.87	0.803	+-0.097	-0.811

(n=69) 2005 vs. 2007	voted =A				
	Least voted = B	2.80	0.948	+-0.114	0.316
Task 5 (n=68) University researcher vs. mapping agency	Most voted =B	4.06	0.808	+-0.098	-1.157
	Least voted = A	2.99	0.872	+-0.106	0.168
Task 6 (n=67) No information vs. administrator	Most voted =B	3.91	0.848	+-0.104	-1.362
	Least voted = A	2.79	0.930	+-0.114	0.202

From the Table 12, only two tasks the levels of perceived credibility between two comparison maps were at the 4-point rating scale (high credibility), which were in Task 3 (T3) and Task 5 (T5). In the other tasks, the levels of perceived credibility to the selected maps were below 4-point rating scale and seem dominant at the 3-point rating scale (neutral). Meanwhile the levels of perceived credibility to the rejected map, generally the mean values were almost 3.0, which indicate their perceived credibility level were neutral.

Analysis on the perceived credibility level of the map they chose or rejected, as in Table I above demonstrates respondents perceived the rejected map as having lower credibility than the selected map. For example in Task 4, the perceived credibility levels among the group that rejected Map B were dominant at the neutral level (55%) (see Figure 1). The perceived credibility levels among the group that rejected Map A were dominant at neutral (35%) and low credibility (48%) levels (see Figure 2). This may indicate that, for certain tasks, although rejections were made, the respondents were mostly undecided about the level of perceived credibility of the rejected maps.

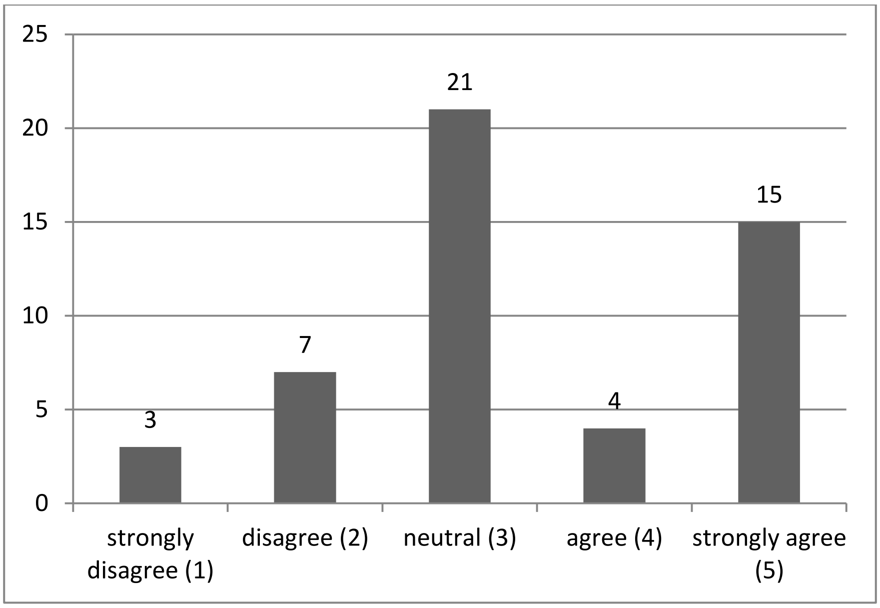


Figure 1 The perceived credibility levels among the group that rejected Map B in Task 4 was dominant at point-3

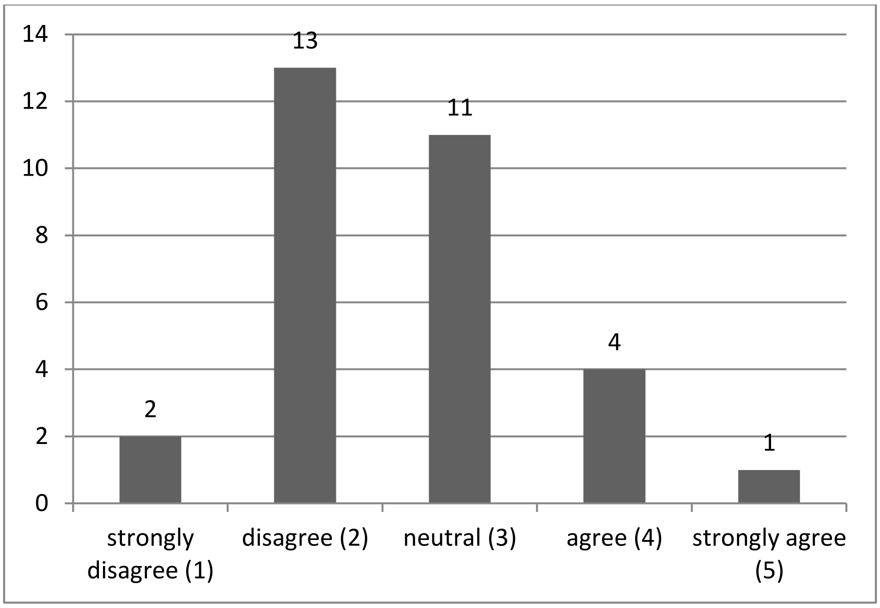


Figure 2 The perceived credibility levels among the group that rejected Map A in Task 4 were dominant at point-3 and point-2

The results in Table 14 also show that only 5.3% to 13% of the total respondents use the identity of the map creator/author (map producer) as the main basis for their judgement when selecting maps for the tasks. The response percentage was higher (13%) in the task that discloses the identity of the map creator as 'an administrator of a company and has National Vocational Qualification' versus a map without the map creator's background (experiment label: T6). This may indicate the strong influence of a map that disclose the author/creator's identity on respondents' perceptions. Nevertheless, this percentage is not significant compared to the percentages from other tasks (i.e. Task 2 (5.3%) and Task 5 (8.8%)).

In Task 2 (experiment label: T2), the respondents who chose the metadata element as the main basis of their judgement seem divided on the selection of maps: 2.7 % of them perceived the map created by the local university student to be more credible than the map created by the outside community; meanwhile 2.7% of respondents perceived the reverse. In Task 5 (experiment label: T5), 8.8% of respondents perceived the map created by a local person in that location as more credible than the map created by someone who works for the mapping providers. This seems to indicate the positive influence of local persons in communicating a message that relates to their local area via the map medium on the perceived credibility of map readers.

Additional Results: Analysis related to gender

Table 13Analysis of gender of those who indicated concept of colour in the responses

Task	Female	Male
T1	22	21
T2	26	23
T3	6	10
T4	14	16
T5	17	12
T6	19	10

Table 14 Analysis of gender that indicated the concept of metadata in the responses

	Female	Male	Total	Total (%)
T1 (currency) No update vs. 1999	2	1	3	3.7
T3 (currency) 2010 vs. 2009	3	3	6	8.5
T4 (currency) 2005 vs. 2007	4	3	7	10.1
T2 (identity of author) Student vs. outside community	2	2	4	5.3
T5 (identity of author) University researcher vs. mapping agency	3	3	6	8.8
T6 (identity of author) No information	4	5	9	13.4

vs. administrator				
Mean	3.00	2.83	5.20	

APPENDIX B

Descriptions of Map Style used in Experiment 2

Table 15 Descriptions of Map Style used in Condition 1

Variable 1	**Supplier of foreground data**	
Condition	1 (see Table 5-1)	
	Map A	Map B
Supplier	Nottingham city council	Student Union
Webpage background	white	Light blue
Building colour	Light grey	Light yellow
Park/garden colour	Light brown	Light brown
Land colour	Light brown	Light brown
Road colour	white	white
Text colour	black	Black

Table 16 Descriptions of Map Style used in Condition 2

Variable 1	Supplier of foreground data	
Condition	2 (see Table 5-1)	
	Map A	Map B
Supplier	BBC	Mix Bistro Cafe
Webpage background	Light green	Light beige
Building colour	yellow	green
Park/garden colour	Light brown	Light brown
Land colour	Light brown	Light brown
Road colour	white	white

Table 17 Descriptions of Map Style used in Condition 3

Variable 1	**Supplier of foreground data**	
Condition	3 (see Table 5-1)	
	Map A	Map B
Supplier	Anonymous	Jane Smith
Webpage background	Black	Yellow
Building colour	Brown	Red
Park/garden colour	Light brown	Light brown
Land colour	Light brown	Light brown
Road colour	white	white
Text colour	black	Black

Table 18 Descriptions of Map Style used in Condition 4

Variable 2	**Website's affiliation**	
Condition	4 (see Table 5-1)	
	Map A	Map B
Affiliated with	Ordnance Survey	University of Nottingham
Webpage background	Bl ack	White
Building colour	Golden	Golden
Park/garden colour	Light green	Bright green
Land colour	Light brown	Light brown
Road colour	white	white
Text attribute colour	black	Black

Table 19 Descriptions of Map Style used in Condition 5

Variable 2	**Website's affiliation**	
Condition	5 (see Table 5-1)	
	Map A	Map B
Affiliated with	Starbuck Cafe	Google
Webpage background	White	Grey
Building colour	Light grey	Brown
Park/garden colour	Green	Light green
Land colour	Light brown	Light brown
Road colour	white	white
Text colour	black	black

Table 20 Descriptions of Map Style used in Condition 6

Variable 2	**Website's affiliation**	
Condition	6 (see Table 5-1)	
	Map A	Map B
Affiliated with	CRAC (unknown company)	No affiliation provided
Webpage background	Light blue	black
Building colour	Green blue	Green blue
Park/garden colour	brown	brown
Land colour	beige	beige
Road colour	white	white
Text colour	Black	Black

Additional Results: Analysis of the perceived credibility to the 'data supplier'

This study measured the perceived credibility to the tested - variables of data supplier in each of the experimental conditions using two items (Q9a and Q10a). Table 21 presents the descriptive analysis of respondents' level of agreement to the statement of '*I perceived the information supplied by the data supplier(s) of the top data layer is credible*' (item Q9a). Chart 1 presents the results in visual form.

1) A comparison between the variables in Condition 1 (displays in blue line) indicates significant differences on the variable '*City Council*' than '*Student Union'* on point 4 (agree) and point 5 (strongly agree).

2) In Condition 2 (displays in green line), major differences between the variable of '*BBC*' and '*Mix Bistro Café*' were indicated on point 3 (neutral) and point 5 (strongly agree). Within this condition, respondents perceived credibility in the variable '*BBC*' more than the '*Mix Bistro Café*' variable.

3) From the chart, the responses between variable '*anonymous*' and '*Jane Smith*' in Condition 3 (displays in yellow line) shows significant difference, particularly on point 3 (neutral) and point 4 (agree) where a majority of responses tended to settle at these two agreement levels. Respondents tend to perceived the credibility of variable of '*anonymous*' more than '*Jane Smith'* in Condition 3.

Table 21 Responses of '*I perceived the information supplied by the data supplier(s) of the top data layer is credible*' (item Q9a)

		1 strongly disagree	**2**	**3**	**4**	**5 strongly agree**	**mean**	**SD**
Condition 1	**City Council**	3.1% (2)	4.7% (3)	26.6% (17)	42.2% (27)	23.4% (15)	3.78 n=64	0.967
	Student Union	2.1% (1)	12.8% (6)	29.8% (14)	36.2% (17)	19.1% (9)	3.57 n=47	1.016
Condition 2	**BBC**	5.3% (4)	6.7% 5)	36.0% (27)	25.3% (19)	26.7% (20)	3.61 n=75	1.114
	Mix Bistro Cafe	0.0% (0)	9.4% (3)	25.0% (8)	53.1% (17)	12.5% (4)	3.69 n=32	0.821
Condition 3	**Anonymous**	4.3% (4)	9.7% (9)	41.9% (39)	37.6% (35)	6.5% (6)	3.32 n=93	0.899
	Jane Smith	4.2% (2)	16.7% (8)	35.4% (17)	37.5% (18)	6.2% (3)	3.25 n=48	0.957

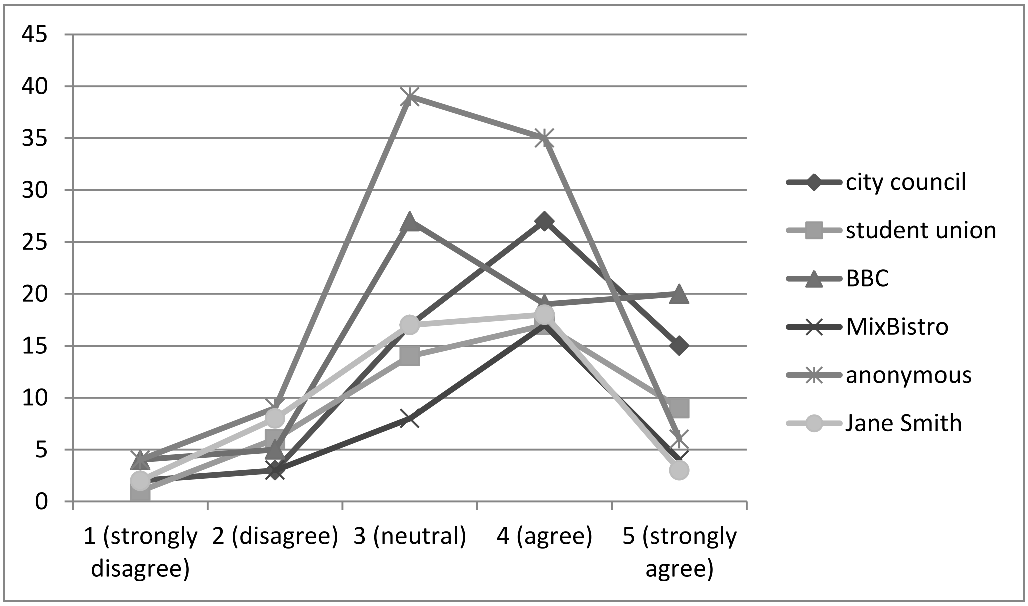


Chart 1The level of agreement with the statement of ' I perceived the information supplied by the data supplier of the top data layer is credible'

Table 22 presents the distribution of responses to the statement 'the data source of the top layer looks more credible than the other set of maps' (Q10a). In this table, respondents perceived sources from '*City Council*' as more credible than sources from '*student union*'; sources from *'BBC*' were perceived as more credible than '*Mix Bistro Café*' and sources from *'Anonymous*' were perceived as more credible than data from '*Jane Smith'*.

Generally, the differences between the two comparable variables in each condition were consistent, and indicate significant differences at point 3 (neutral) and point 4 (agree), particularly at Condition 2 and Condition 3. The majority of responses tended to settle at point 3 (neutral) and point 4 (agree). Chart 2 summarises the results in graphical form.

Table 22 Responses of '*The data source of the top data layer looks more credible than the other set of maps*' item Q10a

		1 strongly disagree	2	3	4	5 strongly agree	mean	SD
Condition 1	City Council	7.8% (5)	9.4% (6)	37.5% (24)	31.2% (20)	14.1% (9)	3.34 n=64	1.087
	Student Union	6.4% (3)	14.9% (7)	40.4% (19)	29.8% (14)	8.5% (4)	3.19 n=47	1.014
Condition 2	BBC	4.0% (3)	6.7% (5)	37.3% (28)	30.7% (23)	21.3% (16)	3.59 n=75	1.028
	Mix Bistro Café	9.4% (3)	18.8% (6)	31.2% (10)	25.0% (8)	15.6% (5)	3.19 n=32	1.203
Condition 3	Anonymous	6.5% (6)	18.3% (17)	40.9% (38)	31.2% (29)	3.2% (3)	3.06 n=93	0.942
	Jane Smith	8.3% (4)	10.4% (5)	43.8% (21)	35.4% (17)	2.1% (1)	3.12 n=48	0.937

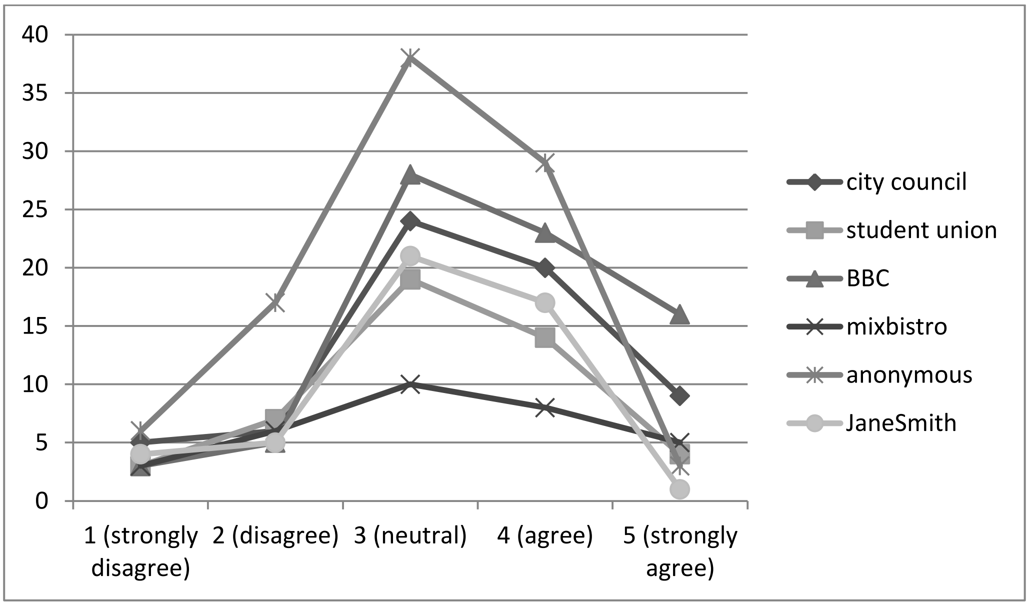


Chart 2 The level of agreement with the statement, 'The data source looks more credible than the other set of maps' item Q10a

The correlation coefficient was used to measure the correlation between responses in the two items that measure the same construct (i.e. perceived credibility to the data supplier) in each experimental condition. A Kendall's tau correlation test was selected because it is appropriate for a dataset with a large number of tied ranks. This test was used to measure the correlation between two variables without controlling the effect of one or more variables; bivariate correlation was therefore used. Table 23 indicates the correlation coefficient between the two items (Q9a and Q10a) that measure the perceived credibility of data supplier variables.

Table 23 Correlation Coefficient between respondents perceived credibility to the data supplier

Condition	sample	Correlation
Condition 1	Select City Council	0.289**
	n= 64	
	Select Student Union	0.442***
	n=47	
Condition 2	BBC	0.638***
	n= 75	
	Mix Bistro	0.277^{ns}
	n=32	
Condition 3	Anonymous	0.384***
	n=93	

	Jane Smith	0.471***
	n=48	

Ns= not significant (p>.05), *p<.05, **p<.01, ***p <.001

From the Table 23, the correlation of the two items (Q9a and Q10a) that measure respondents perceived credibility to the data supplier was significantly related in each experimental condition at $p < 0.05$ and $p < 0.001$, respectively; except under Condition 2 at variable *Mix Bistro Café*, $\tau = 0.277$, $p > 0.068$, where the correlation coefficient was not significantly different.

Table 24 presents the responses of '*How much you perceived the selected map as credible*' and '*How much you perceived the rejected map as credible*?*'* These responses were selected from the sample that chose the '*data supplier'* variable as a basis for their decision in selecting and rejecting the map. The table presents the results according to which map respondents chose in the experimental tasks. From the results;

1) The level of perceived credibility to the map selected in each condition generally settled at point 4 (neutral), point 5 and point 6 (slightly credible to credible).

2) In each condition, around 46% to 71% of responses had levels of perceived credibility at point 5 and point 6 (slightly credible to credible) towards the map.

3) There were 25% to 35% of respondents perceived the map they selected as very high credible (point 7), particularly the maps that manipulated the variable of '*City Council'*, '*Student Union*' and the '*BBC*'.

4) The sample that chose the map that evaluate the variable of '*Jane Smith*' in Condition 3 was not analysed, due to the number of sample of below five 5 respondents.

Table 24 The responses of ' *How much you perceived the selected map as credible*?' and ' *How much you perceived the rejected map as credible*?' in the sample that chose the factor of data supplier

	1 very low credible	**2 low**	**3 slightly**	**4 neutral**	**5 Slightly**	**6 high credible**	**7 very high credible**	**mean**	**SD**
					Condition 1				
Select Map A (City Council)	0.0% (0)	0.0% (0)	0.0% (0)	8.8% (3)	29.4% (10)	35.29% (12)	26.5% (9)	4.18 n=34	0.576
reject Map B (Student Union)	0.0% (0)	8.8% (3)	5.9% (2)	35.3% (12)	17.6% (6)	20.6% (7)	11.8% (4)	3.47 n=34	0.896
Select Map B (Student Union)	0.0% (0)	3.8% (1)	3.8% (1)	11.5% (3)	23.1% (6)	23.1% (6)	34.6% (9)	4.08 n=26	0.891
Reject Map A (City Council)	3.8% (1)	3.8% (1)	7.7% (2)	42.3% (11)	11.5% (3)	19.2% (5)	11.5% (3)	3.35 n=26	0.977
					Condition 2				
Select Map A (BBC)	0.0% (0)	0.0% (0)	0.0% (0)	13.5% (5)	24.3% (9)	27.0% (10)	35.1% (13)	4.22 n=37	0.672
Reject Map B (Mix Bistro)	0.0% (0)	13.5% (5)	13.5% (5)	37.8% (14)	10.8% (4)	10.8% (4)	13.5% (5)	3.22 n=37	1.004
Select Map B	0.0% (0)	0.0% (0)	0.0% (0)	14.3% (2)	35.7% (5)	35.7% (5)	14.3% (2)	4.00	0.555

(Mix Bistro)								n=14	
reject Map A (BBC)	7.1% (1)	7.1% (1)	0.0% (0)	42.9% (6)	21.4% (3)	21.4% (3)	0.0% (0)	3.21 n=14	0.893
Condition 3									
Select Map A (Anonymous)	0.0% (0)	2.2% (1)	0.0% (0)	23.9% (11)	26.1% (12)	28.3% (13)	19.6% (9)	3.91 n=46	0.725
Reject Map B (Jane Smith)	2.2% (1)	13.0% (6)	10.9% (5)	60.9% (28)	4.3% (2)	4.3% (2)	4.3% (2)	2.89 n=46	0.767
Select Map B (Jane Smith)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	33.3% (1)	33.3% (1)	33.3% (1)	4.33 n=3	0.577
Reject Map A (Anonymous)	33.3% (1)	0.0% (0)	0.0% (0)	33.3% (1)	0.0% (0)	33.3% (1)	0.0% (0)	2.67 n=3	1.528

Table 25 summarises the responses in Table 24 by focusing on the positive credibility responses, which at point 5, point 6, point 7 and neutral as well as the differences between the two comparable variables. The findings from the table as follows;

1) Generally, about 74% to 90% of respondents had a degree of perceived credibility to map they selected.

2) However, the numbers of respondents with a degree of perceived credibility in the rejected map were within 35% to 50%, in Conditions 1 and 2 respectively but not Condition 3.

3) The number of neutral responses was higher when respondents were asked to rate their level of perceived credibility in the map they had rejected rather than the map they had selected, with differences of about 24% to 31% among respondents.

4) In Condition 3, the numbers of respondents perceived the credibility of the rejected map was only 13% whereas 61% of responses settled at the undecided point. This might indicate that rejecting a map in the experimental condition did not mean that they did not perceived the map as having a credibility, but the number of respondents that perceived the rejected maps have credibility decreased by 40% to 60 % in each respective condition.

Table 25 Summary of the responses to the statement of *How much you perceived the selected map as credible* ' and *How much you perceived the rejected map as credible*?' at point 5, point 6, point 7 and neutral (point 4)

Perceived credibility level	**Combination of Point 5 (slightly credible), Point 6 and Point 7 (very high credible)**	**Differences**	**neutral**	**Differences**
Condition 1				
Select Map A (City Council)	91.2%	41.2%	8.8%	26.5%
reject Map B (Student Union)	50%		35.3%	
Select Map B (Student Union)	80.8%	38.5%	11.5%	30.8%
Reject Map	42.3%		42.3%	

A (City Council)				
		Condition 2		
Select Map A (BBC)	86.5%	51.4%	13.5%	24.3%
Reject Map B (Mix Bistro)	35.1%		37.8%	
Select Map B (mix bistro)	85.7%	42.8%	14.3%	28.6%
reject Map A (BBC)	42.9%		42.9%	
		Condition 3		
Select Map A (Anon)	73.9%	60.9%	23.9%	37%
Reject Map B (Jane Smith)	13.0%		60.9%	

<u>Additional Results: The perceived credibility to the map they selected and rejected</u>

Of the sample that chose the factor of *affiliation* in Q2, Table 26 presents their responses to '*How much you perceived the selected map as credible*?*' and 'How much you perceived the rejected map as credible*?' From the table;

1) On average respondents perceived credibility levels tended to settle at point 5 (slightly credible) and point 6 (credible) when they rated their level of perceived credibility to the map they had selected.

2) In Condition 4, however, a high number of credibility responses rated (47%) at point 7 (very high credibility) more than point 6 and point 5 (35%) in the sample that selected Map B, which was the map affiliated with the variable '*The University of Nottingham'*.

3) In Condition 5, only six respondents selected the map that affiliated with the variable of '*Starbucks*' whilst 51 respondents selected the map that affiliated with the variable of '*Google*'. In this condition, respondents perceived levels of credibility tended to settle at point 4 to point 7 (82.4%).

Table 26 Distribution of the responses of ' *How much you perceived the selected map as credible* ' and ' *How much you perceived the rejected map as credible?*' in the sample that chose the factor of affiliation

Sample	**1 very low credible**	**2 low credible**	**3 slightly low credible**	**4 neutral**	**5 slightly credible**	**6 credible**	**7 very high credible**	**mean**	**SD**
				Condition 4					
Select Map A (OS)	0.0% (0)	0.0% (0)	0.0% (0)	7.7% (2)	34.6% (9)	34.6% (9)	23.1% (6)	4.15 n=26	0.54
reject Map B UoN	0.0% (0)	3.8% (1)	3.8% (1)	46.2% (12)	15.4% (4)	23.1% (6)	7.7% (2)	3.46 n=26	0.76
Select Map B UoN	0.0% (0)	0.0% (0)	0.0% (0)	18.8% (6)	15.6% (5)	18.8% (6)	46.9% (15)	4.28 n=32	0.77
Reject Map A (OS)	0.0% (0)	6.3% (2)	9.4% (3)	43.8% (14)	15.6% (5)	15.6% (5)	9.4% (3)	3.34 n=32	0.87
				Condition 5					
Select Map A (Starbucks)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	33.3% (2)	33.3% (2)	33.3% (2)	4.33 n=6	0.52
Reject Map B (google)	0.0% (0)	0.0% (0)	0.0% (0)	66.7% (4)	16.7% (1)	16.7% (1)	0.0% (0)	3.33 n=6	0.52
Select Map B (Google)	0.0% (0)	0.0% (0)	0.0% (0)	17.6% (9)	23.5% (12)	21.6% (11)	37.3% (19)	4.20 n=51	0.72
reject Map A (Starbucks)	5.9% (3)	11.8% (6)	11.8% (6)	51.0% (26)	2.0% (1)	5.9% (3)	11.8% (6)	2.96 n=51	1.02
				Condition 6					
Select Map A	0.0% (0)	0.0% (0)	4.0% (1)	44.0% (11)	16.0% (4)	24.0% (6)	12.0% (3)	3.60 n=25	0.76

(CRAC)									
Reject Map B (No affiliation)	4.0% (1)	24.0% (6)	24.0% (6)	48.0% (12)	4.0% (1)	8.0% (2)	12.0% (3)	3.04 n=25	1.02
Select Map B (no affiliation)	0.0% (0)	5.0% (1)	0.0% (0)	15.0% (3)	25.0% (5)	30.0% (6)	25.0% (5)	4.00 n=20	0.8
Reject Map A (CRAC)	0.0% (0)	15.0% (3)	15.0% (3)	35.0% (7)	10.0% (2)	15.0% (3)	10.0% (2)	3.15 n=20	0.99

Table 27 presents the summary of the responses to the statement '*How much you perceived the rejected map as credible*?' in Table 28. From the table;

1) In the responses of perceived credibility to the rejected map, respondents tend to rate their perceived credibility level at the undecided point; the responses of the sample that rejected the map that affiliated with the '*Google*' variable are not included due to the sample size of 6 respondents only.
2) In Conditions 5 and 6, 35% to 50% of responses were rated at the undecided point where ¼ of responses were between 'slightly credible to high credible' and another ¼ were between 'slightly low credible to low credible'. However Condition 1 indicates a different pattern; the proportions that indicated as 'undecided' and had a degree of perceived credibility in the rejected map were not much different.

Table 27 Summary of the responses to: *'How much will you perceive the rejected map as credible*?'

	Low credible (%)	**Neutral (%)**	**Credible (%)**
	Condition 4		
Reject Map B (UoN)	7.7	46.2	46.2
Reject Map A (OS)	15.6	43.8	40.6
	Condition 5		
Reject Map B (Google)	0.0	66.7	33.3
reject Map A (Starbucks)	29.4	51.0	19.6
	Condition 6		
Reject Map B (No affiliation)	28.0	48.0	24.0
Reject Map A (CRAC)	30.0	35.0	35.0

Table 28 The responses of ' *How much you perceived the selected map as credible?*' and ' *How much you perceived the rejected map as credible?*' in the sample that chose the factor of data supplier

	1 very Low credible	2 low	3 slightly	4 neutral	5 Slightly	6 high	7 very high credible	mean	SD
Condition 1									
Select Map A (City Council)	0.0% (0)	0.0% (0)	0.0% (0)	8.8% (3)	29.4% (10)	35.29% (12)	26.5% (9)	4.18 n=34	0.576
reject Map B (Student Union)	0.0% (0)	8.8% (3)	5.9% (2)	35.3% (12)	17.6% (6)	20.6% (7)	11.8% (4)	3.47 n=34	0.896
Select Map B (Student Union)	0.0% (0)	3.8% (1)	3.8% (1)	11.5% (3)	23.1% (6)	23.1% (6)	34.6% (9)	4.08 n=26	0.891
Reject Map A (City Council)	3.8% (1)	3.8% (1)	7.7% (2)	42.3% (11)	11.5% (3)	19.2% (5)	11.5% (3)	3.35 n=26	0.977
Condition 2									
Select Map A (BBC)	0.0% (0)	0.0% (0)	0.0% (0)	13.5% (5)	24.3% (9)	27.0% (10)	35.1% (13)	4.22 n=37	0.672
Reject Map B (Mix Bistro)	0.0% (0)	13.5% (5)	13.5% (5)	37.8% (14)	10.8% (4)	10.8% (4)	13.5% (5)	3.22 n=37	1.004
Select Map B (Mix Bistro)	0.0% (0)	0.0% (0)	0.0% (0)	14.3% (2)	35.7% (5)	35.7% (5)	14.3% (2)	4.00 n=14	0.555
reject Map A (BBC)	7.1% (1)	7.1% (1)	0.0% (0)	42.9% (6)	21.4% (3)	21.4% (3)	0.0% (0)	3.21 n=14	0.893

Condition 3									
Select Map A (Anonymous)	0.0% (0)	2.2% (1)	0.0% (0)	23.9% (11)	26.1% (12)	28.3% (13)	19.6% (9)	3.91 n=46	0.725
Reject Map B (Jane Smith)	2.2% (1)	13.0% (6)	10.9% (5)	60.9% (28)	4.3% (2)	4.3% (2)	4.3% (2)	2.89 n=46	0.767
Select Map B (Jane Smith)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	33.3% (1)	33.3% (1)	33.3% (1)	4.33 n=3	0.577
Reject Map A (Anonymous)	33.3% (1)	0.0% (0)	0.0% (0)	33.3% (1)	0.0% (0)	33.3% (1)	0.0% (0)	2.67 n=3	1.528

Table 29 summarises the responses in Table 28 by focusing on the positive credibility responses, which at point 5, point 6, point 7 and neutral as well as the differences between the two comparable variables. The findings from the table as follows;

1) Generally, about 74% to 90% of respondents had a degree of perceived credibility for the map they selected.

2) However, the numbers of respondents with a degree of perceived credibility in the rejected map were within 35% to 50%, in Conditions 1 and 2 respectively but not Condition 3.

3) The number of neutral responses was higher when respondents were asked to rate their level of perceived credibility in the map they had rejected rather than the map they had selected, with differences of about 24% to 31% among respondents.

4) In Condition 3, the numbers of respondents perceived the credibility of the rejected map was only 13% whereas 61% of responses settled at the undecided point. This might indicate that rejecting a map in the experimental condition did not mean that they did not perceived the map as having a credibility, but the number of respondents that perceived the rejected maps have credibility decreased by 40% to 60 % in each respective condition.

Table 29 Summary of the responses to the statement of How much you perceived the selected map as credible ' and How much you perceived the rejected map as credible?' at point 5, point 6, point 7 and neutral (point 4)

Perceived credibility level	**Combination of Point 5 (slightly credible), Point 6 and Point 7 (very high credible)**	**Differences**	**neutral**	**Differences**
		Condition 1		
Select Map A (City Council)	91.2%	41.2%	8.8%	26.5%
reject Map B (Student Union)	50%		35.3%	
Select Map B (Student Union)	80.8%	38.5%	11.5%	30.8%
Reject Map	42.3%		42.3%	

A (City Council)				
Condition 2				
Select Map A (BBC)	86.5%	51.4%	13.5%	24.3%
Reject Map B (Mix Bistro)	35.1%		37.8%	
Select Map B (mix bistro)	85.7%	42.8%	14.3%	28.6%
reject Map A (BBC)	42.9%		42.9%	
Condition 3				
Select Map A (Anonymous	73.9%	60.9%	23.9%	37%
Reject Map B (Jane Smith)	13.0%		60.9%	

Summaries of the results of the respondents' perceived credibility presented above are as follow:

1) strong perceived credibility in the data supplied by the '*City Council*' rather than data supplied by '*Student Union'*;
2) strongly perceived credibility in the data supplied by the '*BBC*' rather than the data supplied by '*Mix Bistro Café'*;
3) moderate credibility in the data supplied by '*Anonymous*' rather than in data supplied by '*Jane Smith'*.

General analysis of the respondents' perceived credibility in the experimental company affiliated with the map yielded strong perceived credibility to the affiliated company brand '*Google*' than '*Starbucks*' and a comparable strong perceived credibility to the affiliated company brand between '*Ordnance Survey*' and the '*University of Nottingham*'. The response differences between the map that affiliated with '*Ordnance Survey*' and the '*University of Nottingham*' showed no significant difference. In other words, the perceived credibility was comparable between these two variables. But the perceived credibility level to the '*University of Nottingham*' indicated a high number of responses at scale 7 (very high credibility) (46.9%) compared to the perceived credibility level in the '*Ordnance Survey'* (23.1%)

Meanwhile, there was a moderate perceived credibility level in the affiliated company brand '*CRAC*' where the differences in the response to the map that affiliated with the '*CRAC*' and the map that manipulated 'no affiliation' showed no significant difference. This might indicate that a map affiliated a company with a low reputation (e.g. CRAC) and affiliated with no company might produce comparable results in terms of respondents' judgement.

Analysis of the perceived credibility level in the selected map and a rejected map indicates respondents perceived level to the rejected map; the results showed that there was a certain degree of perceived credibility in the rejected map, which might demonstrate that rejecting the map did not mean lack of perceived credibility in the map; a group of respondents settled on the undecided point when giving their responses about their perceived level of credibility to the rejected map.

Results: The proportion that measured 'affiliation' variable

Hypothesis 3 is:

The metadata related to sources have significant influenced within geoliterate respondents

Frequency analysis was conducted on the variable of 'web producer's affiliation' on responses to 'what was the basis of your decision in selecting this set of maps and rejecting the other set of maps'; please choose up to five reasons from the list by ranking the reasons according to your order of priority'. Table 30 shows the results in rank order. From the results;

1) there was quite a high proportion of respondents (23.1%) in Condition 5, compared to the other two conditions, which measured the 'website affiliation' factor; In Condition 5, the mashup web pages manipulated the variable of '*Google*' and '*Starbucks Café*' as their affiliate members. In this condition, 47.2% of respondents did not measure this factor in their assessment.

2) A low number of respondents (8.3%) measured the 'affiliation' factor in Condition 6, where the variables of '*CRAC*' and '*no affiliation*' were manipulated. 56.7% of respondents did not measure this factor in their assessment under this condition.

Table 30 The number of respondents (as a percentage) that measured the 'affiliation element according to the ranking order

Ranked factor	**Condition 4 (%)**	**Condition 5 (%)**	**Condition 6 (%)**
Variables	Ordnance Survey vs. University of Nottingham	Google vs. Starbuck cafe	CRAC (unknown company) vs. no affiliation
First	18.2	23.1	14.4
Second	10.7	9.3	12.5

Third	11.6	10.2	11.5
Fourth	5.8	7.4	1.9
Fifth	1.7	2.8	2.9
Not measured	52.1 (63) n=121	47.2 (51) n=108	56.7 (59) n=94
Respondents measured	47.9 (58)	52.8 (57)	43.3 (35)
Average ratio	1:1		

Overall analyses as following;

1) the highest proportion of respondents who measured the 'affiliation' factor in this assessment were in Condition 1 (54.1%) with the number of respondents that did not measure the factor at 45.9%. In Condition 1, the manipulated 'affiliation' factor was the variable of '*City Council*' versus '*Student Union*' as the data supplier.

2) the lowest proportion of respondents that measured the 'affiliation' factor was in Condition 3 (35.46%) with the number of respondents that did not measure the factor at 64.54%. In Condition 3, the -tested metadata was the variable of data supplier '*Anonymous*' versus '*Jane Smith*'.

3) Overall, the proportion of respondents who measured the 'affiliation' factor and did not measure the factor was probably 50:50 (1 to 1). About 50% to 60 % of respondents did not measure the critical factor, and the other half measured this factor in the experimental tasks.

4) On average, the ratio that assessed and not assessed metadata related to source (i.e. affiliation) was 1:1.

Table 31 The proportion that had spotted the differences of 'affiliation of map producer' between maps

Conditions	Spotted differences of affiliation parameters	Measured (ranked) affiliation	Not measured affiliation
Condition 4 (T2)	Yes	28	17
Condition 5 (T4)	Yes	29	24
Condition 6 (T6)	Yes	25	27
Average		27.3 ~ 27	22.7 ~ 23

Specific analysis was conducted to the sample that spotted the differences of parameters of map producer 'affiliation' in the spot the differences section. From Table 31, there were no significant differences between the proportion that assessed and not assessed affiliation in their judgements.

Table 32 below presents the results of the Mann Whitney-U test that compares the differences in scores between the geoliterate and non geoliterate groups in Conditions 4 to 6 to the sample that measured the factor of ‘affiliation’ in their assessment. From the test, there were no significant differences in the level of priority rated between the two groups. Hence, the **hypothesis 3 is not supported**. The ratio that measured and did not measure within the groups in each condition was 1:1.

Charts 3 to 5 present the score distributions in Conditions 4 to 6, respectively. The charts demonstrate the distribution of scores that spread from position 0 to position 5. In Conditions 4 and 5, a high number of responses within the non-geoliterate group rated the factor at the first priority level (point 1=highest rank) rather than other ranks; this number is significantly different, particularly in Condition 5 between scores at point 1 and point 2 at X^2 (1, n = 23) = 11.5, $p < 0.05$). In the geoliterate group, there were no significant differences in the number of responses that rated the priority level from point 1 to point 5.

From the Table 32, analysis between groups (geoliterate vs. non-geoliterate) demonstrated no significant different between the results among the sample that noticed (spotted) the metadata related to sources (i.e. affiliation of map producer). Analysis within group also showed no significant different between the proportion that assessed and not assessed the affiliation parameters among geoliterate group as well as non-geoliterate group that spotted the metadata.

Table 32 Results comparison on the differences of ranked scores between groups based on Mann Whitney U test in Conditions 4 to 6

	Geoliterate	Non-geoliterate		
Condition 4	Mean rank = 57.61 Mdn = 0 , n = 52	Mean rank = 59.23 Mdn = 1, n = 64	U = 1.61, z = -0.279, p = 0.781, r = 0.03	Not significant at p > 0.05, r= small effect
Measured the factor	24	33		
not measured	28	31		
Spotted ‘affiliation’				
Measured the factor	11	17		
not measured	8	9		
Condition 5	Mean rank =	Mean rank =	U = 1.139,	Not

	55.51 Mdn = 1, n = 43	49.48 Mdn = 1, n =60	z = -1.071, p = 0.284, r = 0.11	significant at p > 0.05, r= small effect
Measured the factor	24	32		
not measured	19	28		
Spotted ‘affiliation’				
Measured the factor	12	17		
not measured	10	12		
Condition 6	Mean rank = 50.83 Mdn = 0, n = 41	Mean rank = 49.41 Mdn = 0, n = 58	U = 1.155, z = -0.268, p = 0.789, r = 0.03	Not significant at p > 0.05, r= small effect
Measured the factor	23	25		
Factor not measured the factor	18	33		
Spotted ‘affiliation’				
Measured the factor	12	13		
not measured	10	15		

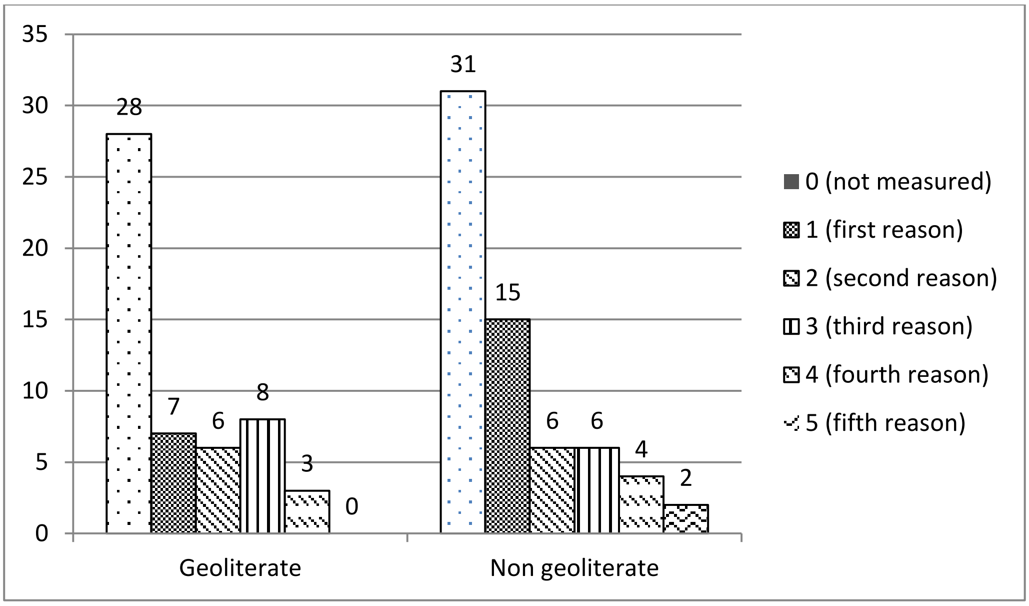


Chart 3 Score distribution between the geoliterate and non-geoliterate groups in experimental Condition 4 (Ordnance Survey vs. University of Nottingham)

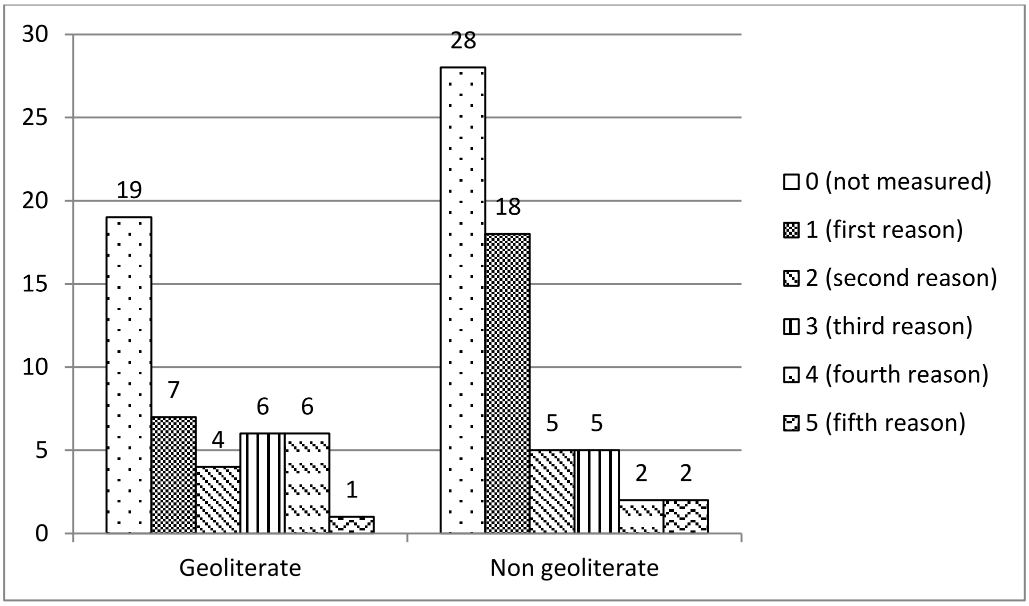


Chart 4 Score distribution between the geoliterate and non-geoliterate groups in experimental Condition 5 (Google vs. Starbuck café)

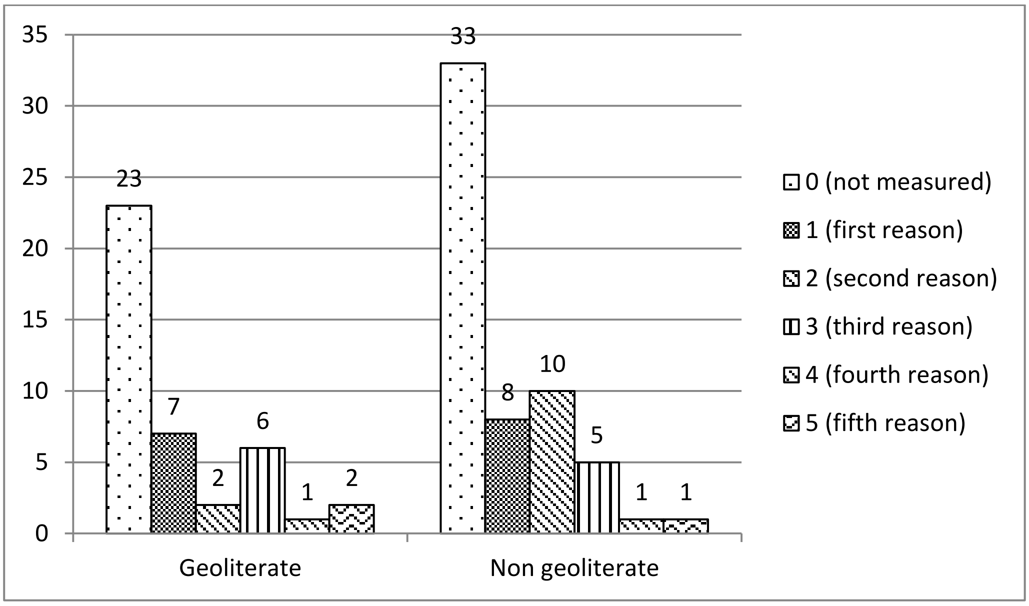


Chart 5 Score distribution between the geoliterate and non-geoliterate groups in experimental Condition 6 (CRAC (unknown company) vs. no affiliation)

APPENDIX C

Results of the extent of respondents' perceived credibility in the map they had chosen in the main question

The first section of the questionnaire was designed to examine respondents' perceived credibility in the map they had chosen in the main question. Descriptive analyses were conducted on the five individual items. Table 33 presents the individual analysis of the mean and standard deviation of the items. The conducted descriptive analysis indicated the mean of respondents' agreement with the statements for each item were on average at point 5, which indicated slight agreement. In other words, respondents seemed the map they had chosen had intermediate level of perceived credibility.

Table 33 The Mean and Standard Deviation of the multi-item measure

	Total responses (n=75)		Of responses that chose Map A (n= 55)		Of responses that chose Map B (n=20)	
	Mean	SD.	Mean	SD.	Mean	SD.
Q2(a) On a scale of 1-7 indicate how much you perceived the map mashup you chose in AQ as believable	5.29	0.85	5.3	0.90	5.25	0.72
Q2(b) On a scale of 1-7 indicate how much you perceived the map mashup you chose in AQ as trustworthy	5.11	1.01	5.18	1.00	4.90	1.02
Q2(c) On a scale of 1-7 indicate how much you perceived the map mashup you chose in AQ as credible	5.17	1.13	5.31	1.07	4.80	1.24
Q2(d) On a scale of 1-7 indicate the competency level of the source(s) of information on the map you chose in Q1	4.92	1.148	5.00	1.20	4.70	0.98

Q2(e) On a scale of 1-7 indicate the expertise label of the source(s) of information on the map you chose in Q1	4.77	1.23	4.93	1.32	4.35	0.81
Total scores	25.26		25.72		24	

The multi-item measure comprised of facet believability, trustworthy, credibility, source's competency and expertise to measure respondents' perceived credibility of the map they had chosen in the main question. The scores from each respondent for all of these five items were accumulated in a new variable, namely 'respondents' perceived credibility'. The total scores were then divided into three equal-sized groups and categorised as 'low to high perceived credibility'. Rating at point 4 (undecided point) in each item was recoded into value 0; this adjustment may affect those who rated point 4.

For example, if a respondent rated point 4 in four of the items, their total scores may fall within the lower category. Therefore, the percentage of scores in the low category may not only indicate the responses of low perceived credibility, but may include those who rated undecided/neutral responses. Table 34 presents the equal-sized classifications of scores according to the respondents' perceived credibility of the map they had chosen in the main question. The mean of the total scores was 22.5, standard deviation was 7.14; on average, the level of respondents' perceived credibility of the map they had chosen fell in the category of intermediate level.

Table 34 Classifications of scores according to low, intermediate and high perceived credibility

Total scores	Frequency (n=75)	Percentage	Categories
6-19	25	33.3	Low
20-27	27	35.9	Intermediate
28-35	23	30.7	High

Results of the influence of the credibility labelling on their judgement of credibility

Of the responses favouring the ‘high credibility map’ (Map A), respondents’ agreement with the influence of these elements on the map was at point 5 (median value) (slightly agree) (as in Chart 6 and 8). Chart 7 and 9 present the distribution of the frequency of these responses at each scale point. These charts demonstrate the influence of the credibility rating on respondents’ judgement. The responses favouring the ‘low credibility map’ (Map B), however, yielded respondents’ agreement in median value at point 2 (disagree) (as in Chart 6) and point 3 (as in Chart 8). A negative influence was indicated on the group that chose the ‘low credibility’ map by the response variation that settled at a negative continuum. These findings indicate a positive influence on the credibility rating indicator in the group that chose the ‘high credibility’ map and a negative impact in the group that chose the ‘low credibility’ map.

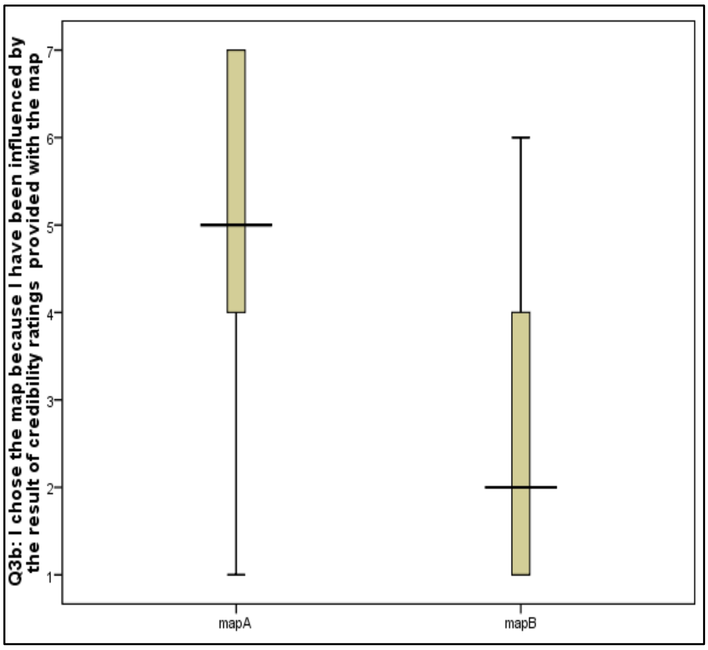


Chart 6 Distribution of the responses: 'I chose the map because I have been influenced by the label of credibility ratings on the map (Q3d)

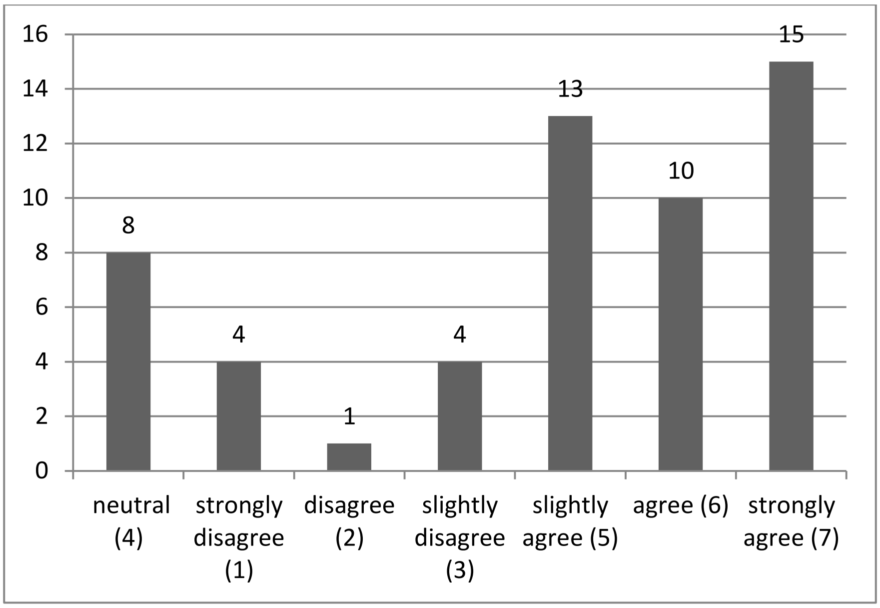


Chart 7 Distribution of responses frequency (Q3d) at each scale point

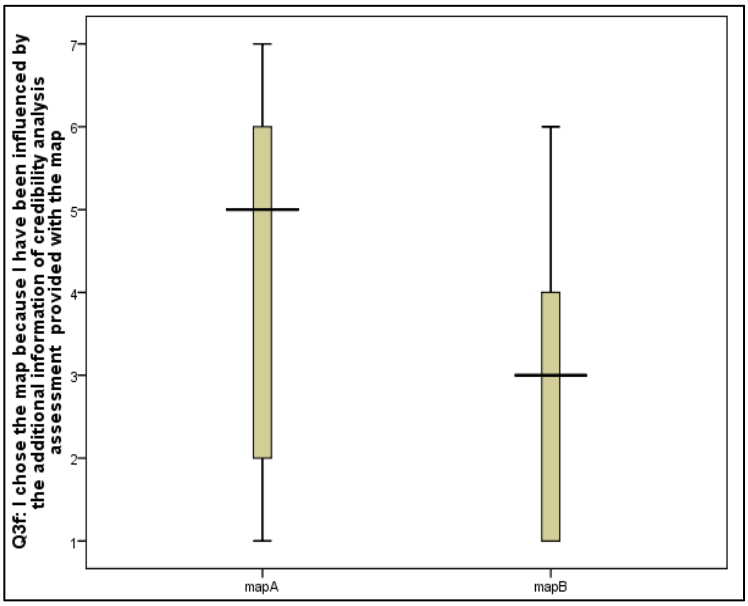


Chart 8 Distribution of the responses 'I chose the map because I have been influenced by the additional information of credibility rating assessment provided with the map'(Q3e)

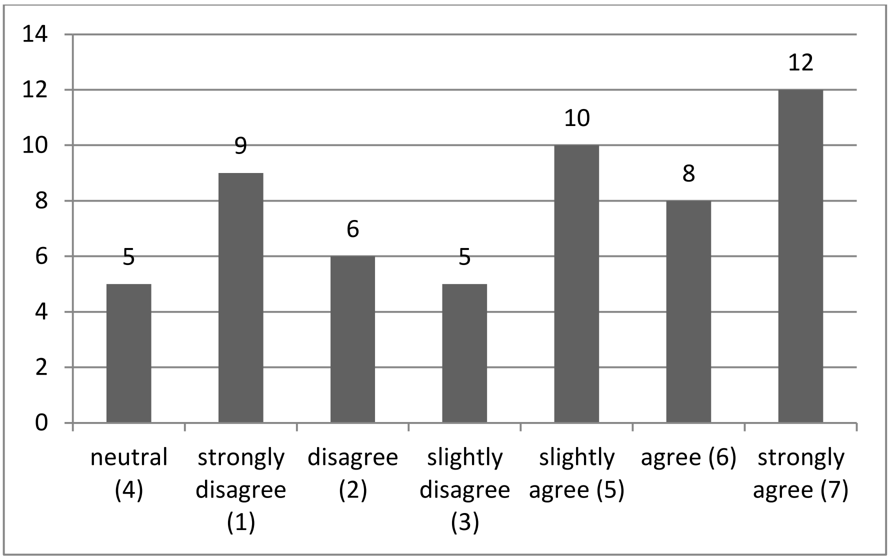


Chart 9 Distribution of responses frequency (*Q3e: I chose the map because I have been influenced by the label (result) of credibility ratings provided with the map*) at each scale point

<u>Additional Results: Responses distribution between groups</u>

The figure below presents the median responses for each element, according to geoliterate and non-geoliterate groups. The pattern of responses seem indicate no significant difference between the two groups; on average both groups rated the 'symbol design' and 'overall presentation' at point 5 (slightly agree) with the influence of these elements, according to their judgement. The influence of the 'colour scheme' element within the geoliterate group, however, was rated at point 4 (undecided) with half proportion variations ranging from point 5 (slightly agree) to point-2 (disagree) (See Chart 10). In contrast, the influence of this element within the non-geoliterate group tended to settle at point 5 (slightly agree) with less variation ranging from point 5 (slight agree) to point 3 (slightly disagree). See Charts 11 and 12.

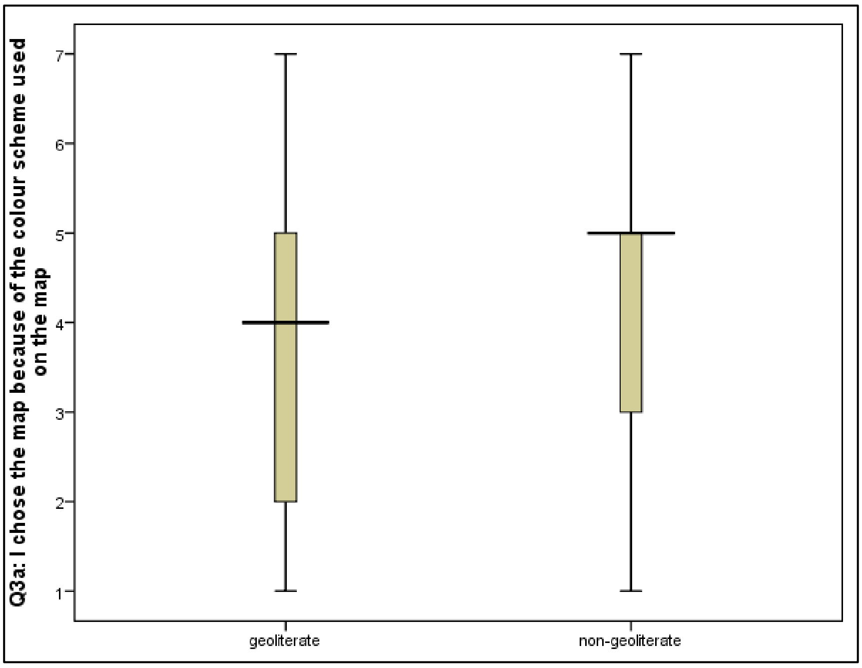


Chart 10 Response distribution of the influence of 'colour scheme' element between groups

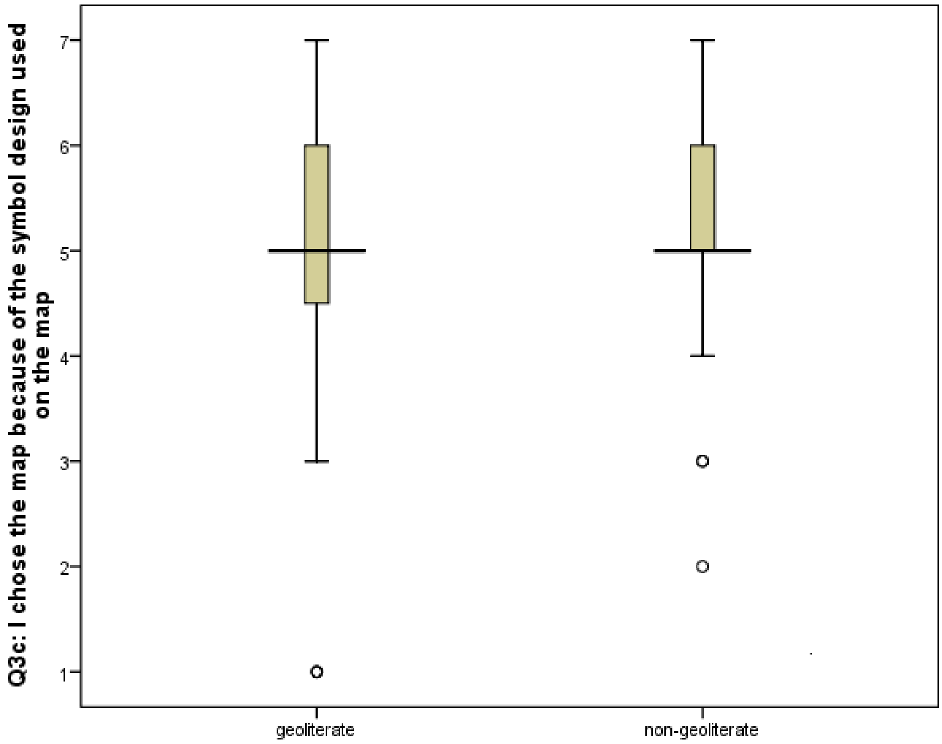


Chart 11 Response distribution of the influence of 'symbol design' element between groups

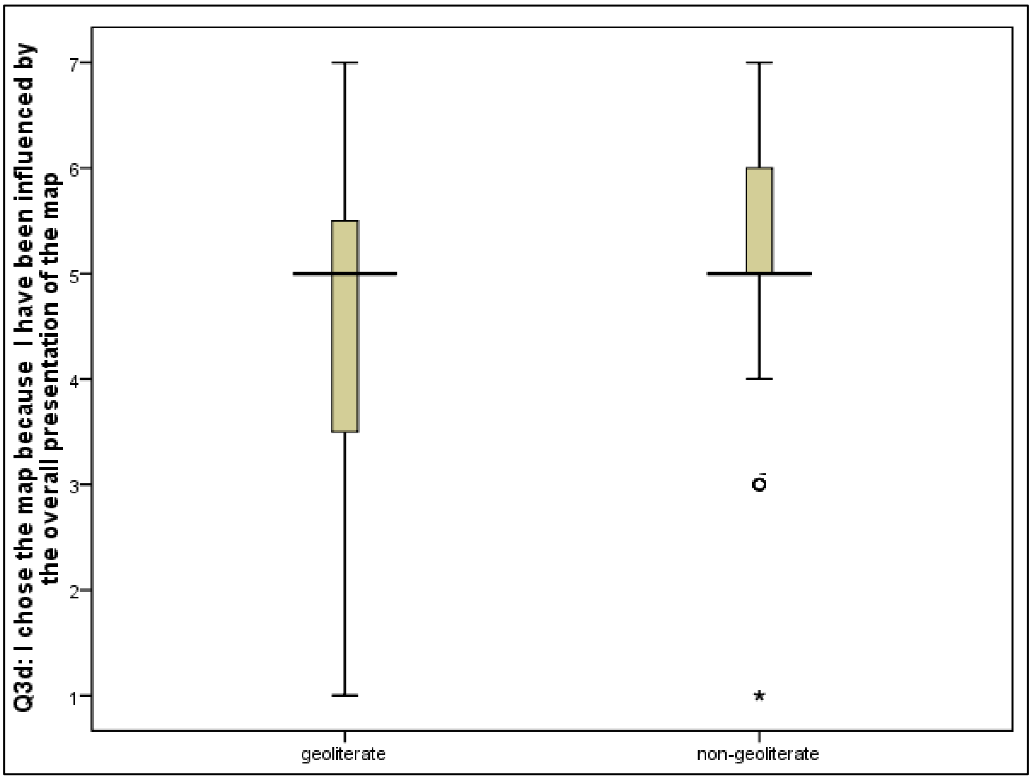


Chart 12 Response distribution of the influence of 'overall presentation' element between groups

The response distribution between groups seems to indicate that agreement with the influence of symbol design in respondents' judgement was at a positive level. Charts 11 and 12 indicate small variations in both groups, with half proportions within the positive spectrum. Response distribution between groups in Chart 11 tends to indicate small variations within the half proportion, particularly in the non-geoliterate group. These findings may indicate the tendency for response homogeneity and agreement at a positive level in the responses from the non-geoliterate group to the influence of symbol design and overall presentation elements in respondents' judgement.

Additional Results: Responses distribution of the influence of map producer

Descriptive analysis of the influence of the identity of the map mashup producer (author) was conducted. This was to examine the influence of single-item metadata in respondents' assessment when judging map information credibility. Of the responses that chose 'high credibility map', descriptive analysis of the results yielded mean responses M=3.91, standard deviation (SD) = 2.44, and median =4 (undecided). Of the responses that chose 'low credibility map', descriptive analysis of the results yielded mean responses M=2.30, standard deviation (SD) = 1.87, and median =1.5 (undecided). Chart 13 presents the response distribution of this item according to the group that either chose ‘high credibility map’ (Map A) or ‘low credibility map’ (Map B). Chart 14 presents the frequency of responses at each scale point.

The high variation from a positive to negative continuum of responses that chose ‘high credibility’ map (Map A) seems to indicate lack of agreement homogeneity concerning the influence of this metadata element in respondents’ credibility assessment. The responses that chose Map B demonstrate a low variation at a negative scale continuum; this may indicate the lesser importance of this element in influencing respondents’ judgement to choose Map B.

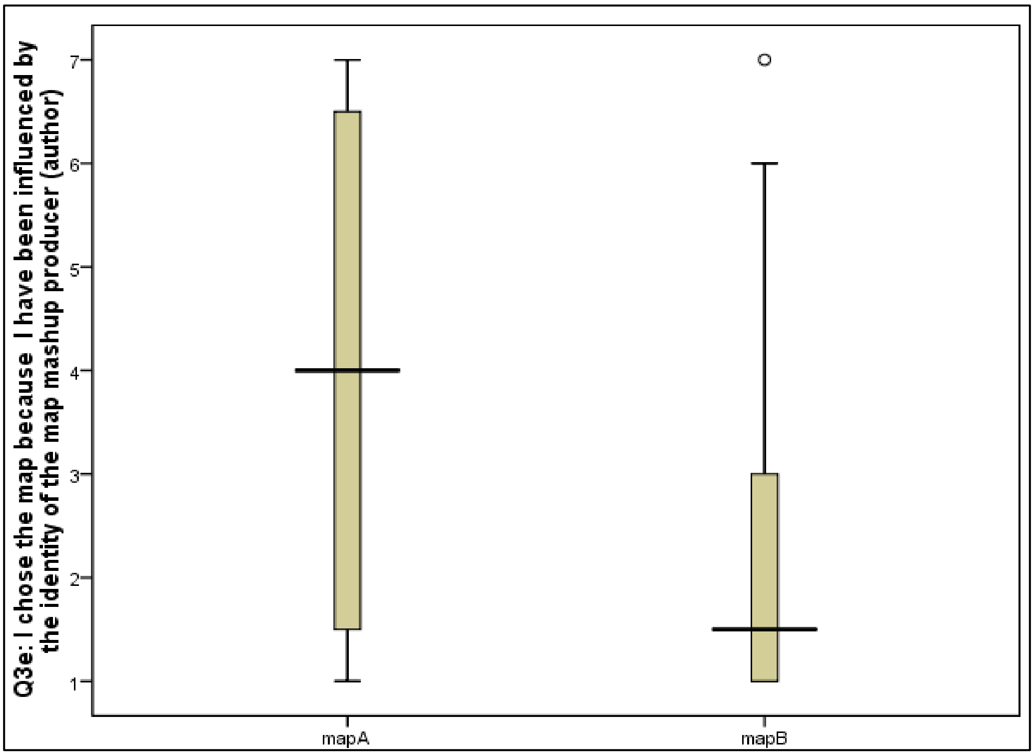


Chart 13 Distribution of the responses: 'I chose the map because I have been influenced by the identity of the map mashup producer (author)’ (Q3f)

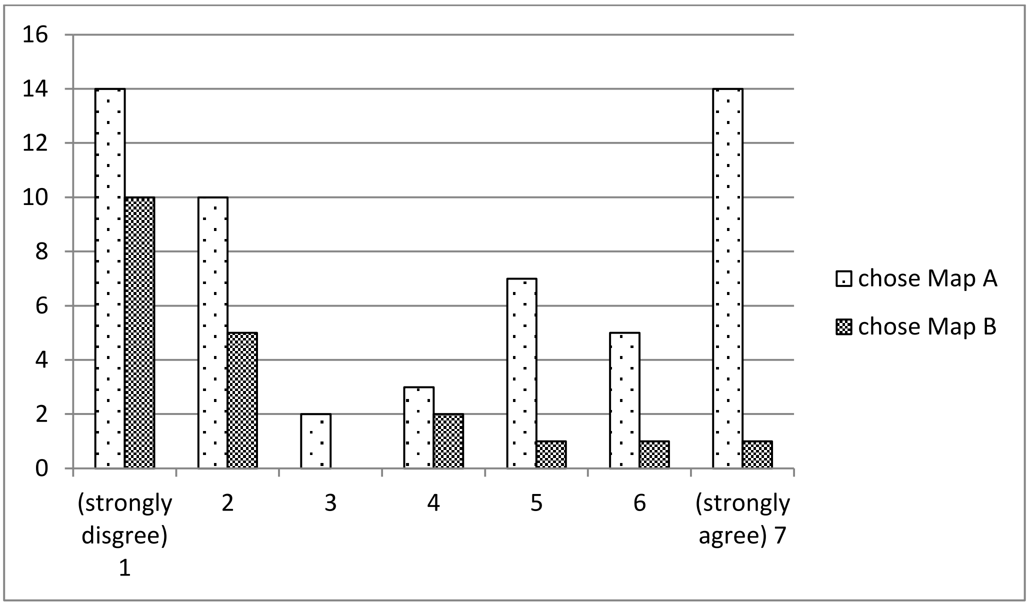


Chart 14 Distribution of responses frequency (Q3f: the influence of the identity of map producer) at each scale point

Additional Results: Responses distribution of the influence of map producer and credibility ratings between groups

Chart 15 and Chart 16 present the response variations of the influence of map producer and credibility rating label, respectively. The distribution of scores of the two groups indicated considerable variations from low to high influence.

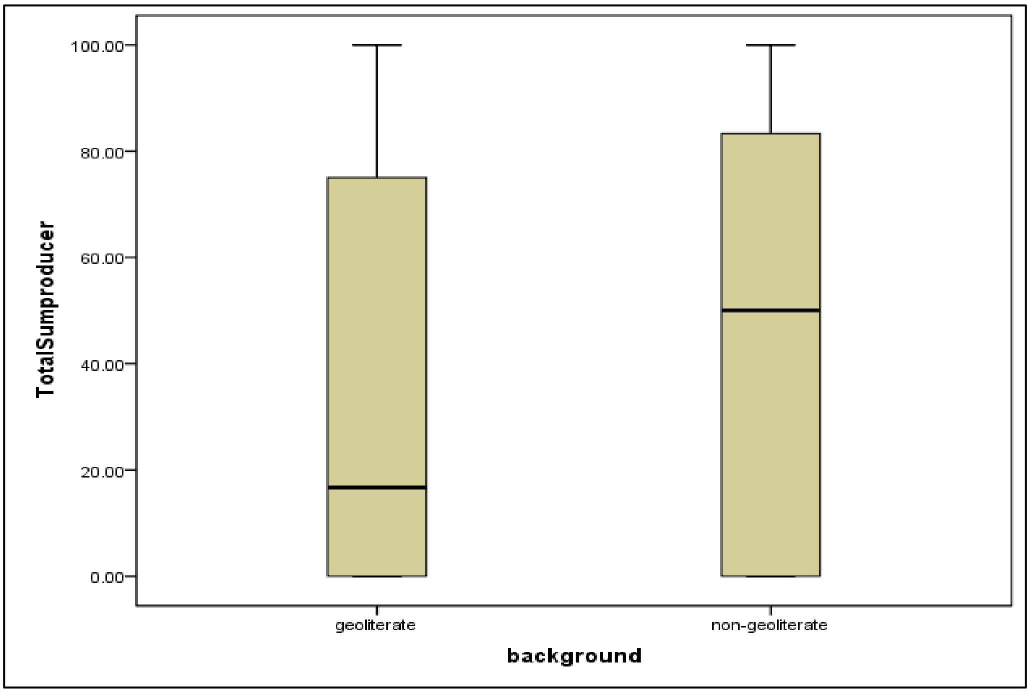


Chart 15 Response distribution of the influence of map producer between groups

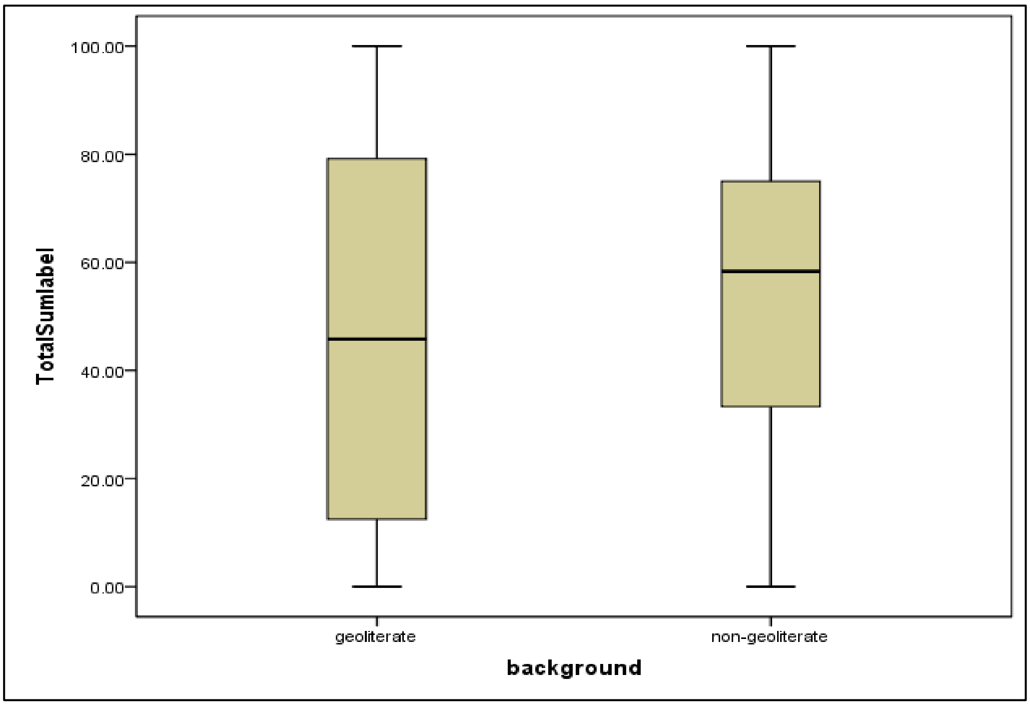


Chart 16 Response distribution of the influence of credibility labelling between groups

Additional Results: The influence of metadata between age groups

Exploratory analysis was conducted to examine the distribution of the data within age groups. Shapiro-Wilk test indicates the data for each age group were significantly different from normal distribution, except for the age group between 31 to 35; hence, the assumption of homogeneity of variance has been violated for the other four age groups. Table 35 presents these results.

Table 35 Results from Shapiro-Wilk test

Groups	Shapiro-Wilk (D)	Significant (p)	Deviation from normal
19 and below (n = 22)	0.81	p < 0.001	significant
20 to 21	0.87	p < 0.003	significant
22 to 24	0.76	p < 0.001	Significant
25 to 30	0.75	p < 0.001	Significant
31 to 35	0.33	P > 0.5	Not significant

As such, these data are generally not normally distributed. To test whether the influence of metadata in this research was influenced by the age of respondents, Kruskal-Wallis was used. This test was conducted to check whether the age significantly affect the influence level. The results indicate that the influence levels of the 'identity of map producer' were not significantly affected by the age of respondents, $H(4) = 3.4$, $p > 0.5$. Hence, the results of this research did not influenced by the age of respondents. Table 36 presents the means between age groups.

Table 36 Results of mean between groups

Age Group	
Geoliterate	
19 and below (n = 7)	Mean = 47.61
20 to 21(n = 11)	Mean = 24.24
22 to 24 (n =9)	Mean = 27.78
25 to 30 (n = 1)	Mean = 26.33
31 to 35 (n = 0)	-
Non-geoliterate	
19 and below (n = 15)	Mean = 38.88
20 to 21 (n = 15)	Mean = 47.77
22 to 24 (n = 11)	Mean = 46.96
25 to 30 (n = 2)	Mean = 66.67
31 to 35 (n = 4)	Mean = 61.11

Additional Results: Analysis between gender

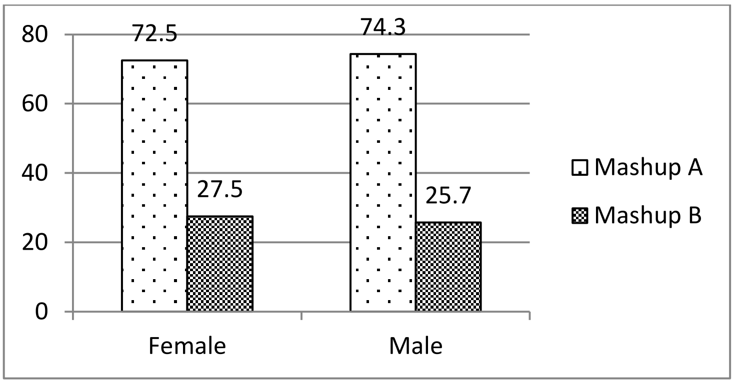


Chart 17 Results comparison of the responses that chose Map A and Map between genders

From the chart above, both respondents from the female group (72%) and the male group (74%) perceived Mashup A, the map labelled 'high credibility map' as having more credibility than the Mashup B, the map labelled 'low credibility map'.A low number of responses, 28% and 26% respectively of the geoliterate and non-geoliterate groups, indicated perceived credibility in Mashup B.

The total scores of the two items measuring the influence of credibility labelling were then analysed according to the gender of respondents. Exploratory analysis was conducted on the sampling distribution of the total scores between the two groups to check normality. From the Shapiro-Wilk significance test, the data from the female and male groups appeared to be non-normal, $D(40) = 0.95$, $p < 0.05$ and $D(35) = 0.90$, $p < 0.05$, respectively. The significant value of less than 0.05 indicates deviation from normality and resulted in rejection of the assumption to apply the independent t-test.

The Mann-Whitney non parametric test was then selected to test the hypothesis. From Mann-Whitney, the influence of the credibility labelling from female respondents (Mdn = 62.5) did not differ significantly from male respondents (Mdn = 50.0), $U = 584.5$, $z=-0.81$, $p>0.5$, $r = 0.1$ (small effect size). From Mann-Whitney, the influence of the 'map producer' from female respondents (Mdn = 41.66) did not differ significantly from male respondents (Mdn = 16.67), $U = 552.0$, $z=-1.2$, $p>0.5$, $r = 0.1$ (small effect size).

Analysis of gender demonstrated no significant difference in the influence of a credibility labelling on respondents' judgement. From Mann-Whitney, the influence of the credibility labelling from female respondents (Mdn = 62.5) did not differ significantly from male respondents (Mdn = 50.0), $U = 584.5$, $z=-0.81$, $p>0.5$, $r = 0.1$ (small effect size). On the other hand, from Mann-Whitney, the influence of the 'map producer' from female respondents

(Mdn = 41.66) did not differ significantly from male respondents (Mdn = 16.67), $U = 552.0$, $z=-1.2$, $p>0.5$, $r = 0.1$ (small effect size). From the median, it demonstrates that the female group perceived these two elements had higher influence compared to male group, although these were not statistically different.

APPENDIX D

Experiment 3: Transcript of the Think-aloud protocols

Respondent id:	Fad (not a real name)
Background:	Geoliterate
Age:	25
Gender:	Male
Decision	Map A
Before	Before the session started, respondent was requested to sign in the consent form after brief explanation were given particularly the aims of the experimental tests, who will benefit and what will become the research materials. Explanation of the method 'think-aloud protocols' that will be used throughout the session was also given. Respondent was requested to verbalise what they thought and why they chose [the answers] during the session
Respondent	Respondent read the experimental task and the main question loudly.
Observer	Requesting respondent to understand the main question, Q1 before proceed with the task
Respondent	Respondent asked where was the trapped victims?
Observer	Requesting respondent to look back at the descriptions of the experimental task. The trapped victims were at the red-coloured building. Stressing respondent to determine the safest route for ambulance to rescue
Respondent	Respondent switched between the two maps pages. 'I chose Map A…. because the colour looks more serious. The Map A is not suitable for navigation, journey….[the map seem] not for safety purposes. We can know it from the colour, and road block symbols.
Observer	Was the elements on the side bar influence your decision?
Respondent	mmmm……[thinking], which side bar? Oh [when he noticed the element at the side bar]….
Observer	Was it influence your decision?
Respondent	No. it did not help much.
Observer	How you want to determine the Map A [map he chose] as more credible than Map B [map he rejected]
Respondent	Because of the colours, the way of presentation was more likely for [something] serious application
Observer	The influenced of symbol design?
Respondent	Not really ok…but it presented more neatly.
Observer	The influenced of map producer?
Respondent	mmmm.. [thinking] map producer?
Observer	At the top sidebar
Respondent	Ok…but it did not influence me
Observer	Either Sarah Smith or the University of Nottingham

	[parameters of map producer], which one is more credible?
Respondent	Oh if had looked at this element…it might have influenced me… the University is more credible….more believable
Observer	This question is out of context: [the intended purpose was to avoid respondent felt offence of their previously respond] In a real world, how you would decide this map can be trusted, more believable than another map?
Respondent	First, I have to try the application. For example, we have to use the Garmin, TomTom applications…from the test, and then we can decide which one can be trusted.
Observer	Was who produced a map will influence your decision?
Respondent	In the context of navigation, who produce the map/data is not important. What more important is we have to try (test) the application first before we can make any decision.
Observer	So, was the map producer influence your decision in the experimental task?
Respondent	Some users might have certain influence of the map producer. But for me, I have to test the application. Just like the TomTom. At first, the brand TomTom was nothing. Then, after try the application, people slowly acknowledge it. The name of producer is not important
Observer	How if the map producer is the University of Nottingham
Respondent	My perception is surely the map is good
Observer	Had you used Google Map before?
Respondent	Yes
Observer	Have you used Map mashup; for example a map that used Google Map as a background, but foreground drawn from other sources
Respondent	No
Observer	Respondent's cooperation in participating the survey was duly appreciated

Respodent id:	Fa (not a real name)
Background:	Geoliterate
Age:	25
Gender:	Female
Decision	Map A (the map producer was the University of Nottingham)
Observer	Requested respondent to understand the main question.
Respondent	Respondent read the experimental task and the main question loudly. Respondent asked the meaning of the term 'panning' in the experimental task descriptions

	Respondent play around with the map – panning, identify point features, click the legend, zooming the map
Observer	[comment] – respondent was too focused to determine the best route for the ambulance by looking at the options of routes on the map. - respondent focused on the central of the map layout. She still did not notice the metadata [the map producer]
Respondent	'I chose Map A. It looks more believable…. It provides more information [compared to other map] and more detail. This feature is important in suggesting the best way for the ambulance.
Observer	How about the 'design look' of the map. Was it important?
Respondent	Oh, I just looked at the detail of the information. Colour scheme did not influence me. Symbol, colour, size, and design did not influence my decisions because both maps used symbols that easy to understand. The colour coding did not influence me too. Although Map A is black and white colour, whereas Map B is quite colourful, but these did not influence my decisions. Information at the side bar [the legend] was very helpful in guiding me to read the symbols of landslides.
Observer	Was there any elements influence your decision?
Respondent	The numbers [points] of landslide events between the two maps also influence my decision.
Observer	[comment] -respondent still did not mention any statement related to the producer of the map. How about the influence of zoom in and out, scale bar functions…did these elements helping you in making decision? [the intended purpose to ask this question was to let respondent focused at the peripheral map [not at the central map, since the map producer was stamped at the peripheral side of map]
Respondent	Yes, these functions did help me.
Observer	[comments] -respondent still did not notice the map producer at the side bar. -She only used the clicking functions at the legend that linking the legend with its associate features. -She did not even used the mouse cursor to click the map producer element that stamped on top of the legend. Have you used any Web GIS application before?

Respondent	No
Observer	Were these elements – visual attractiveness, symbol design, clarity of symbols, symbol convention, colour coding, colour convention had any influence in your decision? [respondent was given a list of questions (a similar question in the questionnaire i.e. Question 3) where she have to rate in 1 to 5 of the level of influence of these elements in her decision]
Respondent	Yes. The visual attractiveness, symbol design, clarity of symbols, symbol convention had influenced me, but not influence by the colour coding and the colour convention. The map producer also was important and had influenced me.
Observer	Respondent was requested to conclude again the elements on map that influence her decision
Respondent	Symbol design because we can understand it easily
Observer	Was the map producer important and did it influence your decision?
Respondent	Err…actually I did not look at who produced the map
Observer	[comment] [respondent then quickly search where was the map producer element on the map]
Respondent	Yes, Map A is more believable because it produced by the university. Map B is less believable because it produced by nobody, we did not know the background of the author. She might make up the data. So it is less believable. The Map A was produced by an authorised source, we know its [reputation] and we can believe [the source] However, the important of map producer [if based on the rating scale in the question] was just at the level of 'important'. The 'very important' element that influence me was the symbols because its present the detail of the information
Observer	Respondent's cooperation in participating the survey was duly appreciated

Respodent id:	R (not real name)
Background:	Non-Geoliterate
Age:	21
Gender:	Female
Decision	Map A
Before	Before the session started, respondent was requested to sign in the consent form after brief explanation were given particularly the aims of the experimental tests, whom will benefit from this and what are the output of the research materials. Explanation of the method 'think-aloud protocols' that will be used throughout the session

	was also given. Respondent was requested to verbalise what they thought and why they chose [the answers] during the session.
Respondent	Respondent read the experimental task and the main question loudly.
Observer	Requesting the respondent to understand the main question, Q1 before proceed the task
Respondent	Respondent asked what were the meaning of each symbol
Observer	Instructed respondent to use the legend the sidebar
Respondent	Respondent switched between the two maps pages. ‘I chose Map B…. because the points of roadblocks. In the Map A the symbols used did not really help’.
Observer	Was there any reason instead of symbol design?
Respondent	mmmm……[thinking] ‘the black colour used on the building features block the texts’.
Observer	Was there any influence of the elements at the sidebar?
Respondent	‘Ooh…I just realised the data between this two maps are different. Is it this represent similar or different event of landslide?
Observer	‘The maps were used different sources of data’.
Respondent	‘But, which map presents more landslide points?’ ‘Ok, I chose Map A because the more important one is the information that we can read through the symbols used on the map’.
Observer	How is it actually we can determine the correctness of information. How we can tell the data in either Map A or Map B is not misleading?’
Respondent	‘mmmm..I do not know! because for me there are not so much different’
Observer	Was there any influence of map producer?
Respondent	‘Ooh, I just realised this element. Yes indeed. Map A was produced by the university. We consider it’s from an authorised source compared to Map B that produced by unknown individual’.
Observer	Ok, could you conclude your final decision?
Respondent	I chose the map, because it was more detail and the map was produced by an authorised source. I assumed the map has been verified by so many levels. If the map produced by an individual, we could not confirm the source of data; she/he might make up the data.
Observer	Respondent’s cooperation in participating the survey was duly appreciated

Respodent id:	Am (not real name)
Background:	Geoliterate
Age:	28

Gender:	male
Decision	Map B
Before	Before the session started, respondent was requested to sign in the consent form after brief explanation were given particularly the aims of the experimental tests, who will benefit and what will become the research materials. Explanation of the method 'think-aloud protocols' that will be used throughout the session was also given. Respondent was requested to verbalise what they thought and why they chose [the answers] during the session
Respondent	Respondent read the experimental task and the main question loudly.
Observer	[observation] – Respondent played around with the map. Click the features on map. Click the building. Zoom in and out, panning the map. Respondent switched between the two maps pages. Clicked the points (roadblocks, landslides) -respondent seems too focus on the centre of the map. Not even clicked the items at the sidebar.
Respondent	'I chose Map B'
Observer	Was there any reason of this decision?
Respondent	'First, the information on this map is not too crowded. Second, the symbols used on the map are easy to understand. There are some points in Map A not appeared on Map B But the appearance on Map B is clearer. The design of symbols is clear and easy to understand the meanings of each symbol'
Observer	'If you look at the map legend, at the sidebar, was there any element influence your decision?
Respondent	'Oooh, if I had looked at the legend, then I can understand the meaning of Point 1, Point 2 …they represented each roadblock. 'So I have to look at Map B because I had chosen this map at the first place'.
Observer	Not necessarily.
Respondent	'I suggest Map B, route 1 for the ambulance. The design on Map B is neater, looks more presentable, less crowded'. 'If Map A, the design is not too good. The colours used were not attractive'. 'Both map used symbols that not easy to understand'. 'Are these two map displayed similar data?'
Observer	'No. The maps displayed different data. The points of landslides, the locations of roadblocks are from different sources'
Respondent	'mmmm..At first, I thought the maps supplied by same data sources. But actually, they were different'. 'But I still chose Map B'

Observer	‘So actually in a real world situation, how you would actually evaluate the correctness of information on map. How you want to determine that this map can be trusted?’
Respondent	‘Ok. My first impression was these two maps are identical. The different was at the number of landslides. So, I think both of these map are believable. But for the aspect of easy to use, I chose Map B because it is not too crowded and the symbol designs more neat’. ‘In a real world, if I have to face with two maps that displayed contradict information, I will use the satellite navigation device or I will find other map to make comparison. I will use other sources to validate the data’.
Observer	So was the map producer influence in your decision during the experimental task?
Respondent	What do you mean by map producer?
Observer	Map producer is like this [showing the stamped map producer at the sidebar], the author of the map.
Respondent	Not influence! Because I just focused on the event and the symbols used. Since I perceived the two maps were identical, with not so many significant differences, so I assumed the producer of the both two maps are same person [sources] too.
Observer	Respondent’s cooperation in participating the survey was duly appreciated

Respodent id:	Fik (not real name)
Background:	Non-geoliterate
Age:	25
Gender:	male
Decision	Map B
Before	Before the session started, respondent was requested to sign in the consent form after brief explanation were given particularly the aims of the experimental tests, who will benefit and what will become the research materials. Explanation of the method ‘think-aloud protocols’ that will be used throughout the session was also given. Respondent was requested to verbalise what they thought and why they chose [the answers] during the session
Respondent	Respondent read the experimental task and the main question loudly.
Observer	[observation] – Respondent played around with the map. Click the features on map. Zoom in and out, panning the map. Respondent switched between the two maps pages.
Respondent	‘These two maps are similar. So I have to choose the best map. The map that fit for the ambulance officer to suggest the safest route? So where the entrance of the ambulance?

Observer	I pointed out at the entrance gate near the lake side.
Respondent	So what are these symbols mean? Is it roadblock indicates the road is closed?
Observer	[observation] respondent panned the map
Respondent	‘So where the entrance of the ambulance?’
Observer	The ambulance might enter the campus from this gate [entrance near the lakeside] and use the main road coloured in yellow. So you have to suggest which route for the ambulance to access.
Respondent	‘I chose Map A. Since I am not familiar with the campus. I have not been to Nottingham yet.
Observer	[Comment] respondent too focus on the line, points, polygon features in order to suggest the best route for the ambulance. Did not care of the source(s) of map.
Respondent	‘mmmm..The number of landslide points differs. But they look no different. I did not see any different between these two maps. So I chose Map B’ because it more attractive. ‘Map A used black colours. Map B is more transparent. The text labels are clearer. And it seems easy to use. So it will easy to be used by the ambulance officer, since it can be read easily and clearly.
Observer	‘How about the design look of these two maps in terms of either it look professional or amateur design?
Respondent	They are not too professional or too amateurish. There are more symbols of landslide on Map A, so it is look more details. But Map B is not too details. But I think, the more important is on the clarity of information. Because user want to use the map to navigate right? So I still stick and chose Map B due to the clarity of information and the colour applied on map.
Observer	So how you want to determine whether this map is correct or incorrect? Did you influence by the producer of the map?
Respondent	‘I will look at the map provider. I will just trust the map if I am not familiar with the area. I will look at the date of the last updated. These two maps used similar base maps. [by showing the copyright stamped at the bottom of map which is from Google]. There was no specific update on the last updated date. Just in 2013. Those maps produced by Google. I just looked at the ‘powered by’ label [Google’s copyright at the bottom of base map], so the map producer had influence my decision’
Observer	Have you used map mashup before?

Respondent	‘Yes, I used to use mashup to find a restaurant’
Observer	Are you familiar with the concept of map mashup where the base map commonly supplied by commercial map provider such as Google Map, Bing Map but actually the foreground data are supplied by other sources?
Respondent	‘No. I did not realise about this. What I know the map is produced by Google since there is a Google copyright stamped at the bottom of the map. I did not realise that it was actually from different sources.
Observer	[comment] there is a misunderstanding on the concept of who produced the data, particularly on map mashup environment where the sources might draw from a few sources. The respondent still perceived the map is produced by Google, because it used the base map from Google, and not realised that the foreground data on map mashups are actually could be supplied by different sources.
Observer	So other than Google Map, who do you think was the producer of this Map B? And did it influence you?
Respondent	‘I do not care who produced the map. The more important is the map is easy to read and clear. And it is easy to use. It is enough for me to look at who hold the copyright.
Observer	You are given an option to re-think your decision.
Respondent	I will stick to choose Map B.
Observer	Respondent’s cooperation in participating the survey was duly appreciated

Respodent id:	A (not real name)
Background:	Non-geoliterate
Age:	33
Gender:	Female
Decision	Map A
Before	Before the session started, respondent was requested to sign in the consent form after brief explanation were given particularly the aims of the experimental tests, who will benefit and what will become the research materials. Explanation of the method ‘think-aloud protocols’ that will be used throughout the session was also given. Respondent was requested to verbalise what they thought and why they chose [the answers] during the session
Respondent	Respondent read the experimental task and the main question loudly.
Observer	[observation] – Respondent played around with the map. Click the features on map. Zoom in and out, panning the map. Respondent switched between the two maps pages. [observation] respondent too focus on the centre of the map; to find the best route.

Respondent	‘Map A looks fit for this purpose because the route is more accessible.
Observer	Any reasons to choose this Map?
Respondent	These symbols – roadblock and construction- mean we can use the road or no access at all? Since the symbols on route 1 and 2 clearly shown the routes are accessible, due to the locations of these symbol and the design, so I chose Map A
Observer	How about colour coding used on the map?
Respondent	No influence.
Observer	How about the symbol design?
Respondent	Yes, it influences me.
Observer	How about the appearance. The design whether looks professional or amateurish?
Respondent	‘No influence’.
Observer	‘How about elements at the sidebar?
Respondent	Not so much influence me. Since the function is just for user to click, and it will link to which points features on the map.
Observer	How do you want to evaluate either Map A or Map B is incorrect?
Respondent	‘Oh, which one is more believable? Which one we can trust? mmm… how is it actually to evaluate the credibility of a map? Because this map locates the road with a symbol of road construction image. And another map locates the symbol using roadblocks image which means no access at all.
Observer	So how you want to evaluate this map correct or misleading?
Respondent	‘mmm…since it was from Google. So I trust the map because it was created by Google.
Observer	Map A was produced by the university and Map B was produced by Sarah Smith. Did these elements of map producer at the top side bar, influence your decision? [comment] respondent did not notice the element of map producer at the top at the sidebar.
Respondent	‘I think the map produced by the university is more credible. Map B was produced by private individual isn’t? I think it is better to have a map produced by one organisation compared to an individual.
Observer	So did it influence you?
Respondent	Yes, I chose Map A because the data are more detailed, the symbol used, and because of the producer of the map. In Map B, they did not mention who Sarah Smith was?
Observer	Respondent’s cooperation in participating the survey was duly appreciated

APPENDIX E

<u>Experiment 4: Transcript of the Think-aloud protocols</u>

Respondent id:	Fad (not a real name)
Background:	Geoliterate
Age:	25
Gender:	Male
Decision	Map A
Observer	Requested respondent to understand the main question.
Respondent	Respondent read the experimental task and the main question loudly.
Observer	Highlighted respondent that he has to determine the shortest route from purple coloured to red coloured building
Respondent	'I chose Map B'
Observer	Any reason? You can use the zoom in, zoom out functions
Respondent	How to zooming?
Observer	Click at the '+' function
Respondent	Respondent pan, zooming in and out, then click [identify] the features on map to find the shortest route 'Map B is more fit for navigation purposes'
Observer	Was it due to ease of use?
Respondent	'It's more creative…easy to use… Map A looks more professional. Map B more fit for public use'
Observer	How to know which map is more believable…because the data in these two maps are contradict between each other… the producers were different. What element(s) influenced you?
Respondent	Due to the presentation
Observer	Were the elements at the side bar influence you?
Respondent	Yes, they were. But the map producer did not help much… because the more important is to test the map first.
Observer	Ok…how about the influence of the credibility rating label on map?
Respondent	Ok. If there is a rating, it will influence me. This rating was produced [generated] by whom?
Observer	If we look at the CCTL label carefully, there was one organisation rated the maps
Respondent	If rated by one organisation, then it will influence my decision. So, I chose Map B because it has been rated by one organisation

Respondent	Who produced the map is not important because we could search someone that might have reviewed the map. I will rely on the third party reviews in determining the credibility of a map
Observer	Respondent's cooperation in participating the survey was duly appreciated

Respodent id:	Fa(not a real name)
Background:	Non-Geoliterate
Age:	25
Gender:	Female
Decision	Map B (Sarah Smith)
Observer	Requested respondent to understand the main question.
Respondent	Respondent read the experimental task and the main question loudly. She used the zoom in function. She panned the map to the red coloured building. She browse the Map A 'the buildings on the map were black coloured, hence the texts on map. It is difficult to read the label'. She browses the Map B. 'In Map B, the text labels are easy to read. We can know which way to go by read the labels on features. This Map B is more has more details information. It has more points of roadblocks and landslides compared to Map A' 'The used colours on the map indeed influenced my decision because the buildings presentation was clearer. Because if we want to ride a bike, we will use the landmark on our route to find the way. If we know the name of building from the label on a map, we can use it to find the route and know which junction we are' 'The legend on map is also helping me to understand the meaning of symbols on the map'
Observer	Was the credibility rating on the map influence your decision?
Respondent	'ooh, I do not have any background in mapping. I believe any maps that I found. A rating on map did not help much. I believe any maps on the Internet' 'I did not look at this rating label. I just focused on the map and the symbol used'.
Observer	Were you influence by the symbol design?
Respondent	Both maps used the symbols that quite similar and easy to understand. Hence, it did not influenced my decision
Observer	Were you influence by the overall presentation?
Respondent	'Yes, the colours and symbols used help me in making decision'
Observer	About the identity of map producer, was this element

	influence you?
Respondent	'No. It did not influence me. Just like the credibility rating label, both of them did not have any influence.
Observer	So, can you conclude which map you perceived more believable for this experimental task?
Respondent	Map B, because it more relevant and easy to read the label
Observer	This is out of experiment context, in a real world, how you are going to use the online maps. Will you read a map using a laptop or you will print out the map?
Respondent	'It depends on a situation. If there is a WIFI, I will access Google Map via my smart phone. If there is no connection, I will print out the map to use during the journey. 'I do not care about who is the author/producer of a map. That element is not important. I just focus on the map and just use the one that more accurate'. 'I used to use Google Map to find one place. Unfortunately, the map of that place was not updated. When I encountered with this problem, I tried to adjust the situation. I navigated the area independently not relied on the map, but by using the local landmarks and sign post. I will find the location on my own' 'I did not care about the map author [producer]. On Google Map, they do not mention the author. The author or who produced the map was not important' 'And now I prefer to use satellite navigation device to navigate since it has sound and instruct us where to go, next and next'. 'I will only use Google Map if it is only the option that I have'. 'The more important are the details of information, information clarity and easy to use' 'If there is inaccurate information on the map, it does not matter, since I will adjust and find my way on my own' 'who produced the map is indeed not important'.
Observer	Respondent's cooperation in participating the survey was duly appreciated

Respodent id:	R (not a real name)
Background:	Geoliterate
Age:	25
Gender:	Female
Decision	Map A
Observer	Requested respondent to understand the main question.
Respondent	Respondent read the experimental task and the main

	question loudly.
Observer	[Observation] respondent clicked the building to identify the features. 'you can use the zoom in and out to look for more details features'
Respondent	'Ooh, Map A is less detail. Both maps presented slight different data.
Observer	Which map you will choose for the experimental task?
Respondent	'ooh, is it we have to you use and follow the proper bike route; we cannot cross the hills and the park to arrive the red coloured building?
Observer	Yes, you have to follow the mapped route only.
Respondent	'The Map A is not so detail'. 'So how to ride a bike. Since there are so many roadblocks because this map shown so many roadblock points. Whereas this [another map] did not shown that point.
Observer	[observation] respondent switching between the two maps, zooming in and out of the map
Respondent	'Ok. Since this map has a stamped rating label that indicated 'low rating', so I chose Map A because there is a rating label and indicated as 'high rating'.
Observer	How to know which map is more believable…because the data in these two maps are contradict between each other… the producers were different. What element(s) influenced you?
Respondent	Due to the presentation
Observer	Were the elements at the side bar influence you?
Respondent	Yes, they were. But the map producer did not help much… because the more important is to test the map first.
Observer	Ok…how about the influenced by the credibility rating label on map?
Respondent	Ok. If there is a rating, it will influence me. This rating was produced [generated] by whom?
Observer	If we look at the CCTL label carefully, there was one organisation rated the maps
Respondent	If it was rated by one organisation, I will be influenced.
Respondent	So, I chose Map B because it has been rated by one organisation. Who produced the map is not important because we might find someone reviews the map. I will rely on the other sources in determining the credibility of a map
Observer	[comment] – Respondent finally noticed the credibility rating label on her own. Respondent's cooperation in participating the survey was duly appreciated

Respondent id:	Am (not real name)
Background:	Geoliterate
Age:	28
Gender:	male
Decision	Map A
Observer	Requested respondent to understand the main question.
Respondent	Respondent read the experimental task and the main question loudly.
Observer	[Observation] respondent panned and clicked the building to identify the features. Respondent switched between the two maps.
Respondent	'So I have to choose the main road that accessible for a bike, right? Is this small lane is for pedestrian? So there are so many constructions. There are roadblock everywhere, so we cannot use the route?
Observer	Yes, you have to follow the routes that had been mapped only. [observation] respondent did not care on the rating label on top of the map.
Respondent	'I chose Map A because during the landslides, we can use the routes at the back of the library to go to the red coloured building, although the distances that have to take are quite long. But it is better since there are no roadblocks and landslides occurred there.
Observer	[observation] respondent perceived the Map A was more credible and not due to the sources (i.e. the university) but due to the data presented.
Respondent	'If Map B, it is going to be more difficult since it shown so many roadblocks and landslides.
Observer	Were the elements at the sidebar influence your decision?
Respondent	No, I just influence by the extent of information detail on the map.
Observer	How about the producer of the map. Was the element influence you?
Respondent	No influence. I just influence by the details of information.
Observer	How about the credibility rating label stamped on the map?
Respondent	Oh, Ok what do you mean by the rating? What is the rating? Ooh do you mean this label [respondent pointed the cursor on the stamped rating label].
Observer	[comment] – at first, respondent did not noticed the stamped credibility rating label.
Respondent	'Ok, I did not influenced by this rating label. I just focused on the map to go from point A to point B.
Observer	If there is a credibility rating label on map, would you rely on it?

Respondent	To be honest, before you mentioned about this stamped credibility rating, I did not noticed it at all. But I might have influenced of this element, if I had noticed it at first. ‘My strongest argument is I only influence by the detail of information. But another reason that might has influence my decision is because of the high rating credibility rated on the map’. ‘At first, I did not notice. But if I had noticed, it still would not influence me. I will look at the purpose of map use. And this rating label could be additional element to support my decision’ ‘I will believe the map more, if there is a credibility rating stamped on the map’.
Observer	How about the map producer, was this element influence you?
Respondent	‘No. this element will influence me if the producer was the NASA or from mapping department. If it was produced by a university. It is only an academia. Not from an authorised sources.
Observer	Respondent’s cooperation in participating the survey was duly appreciated

Respodent id:	Fik (not real name)
Background:	Non-geoliterate
Age:	25
Gender:	male
Decision	Map B
Observer	Requested respondent to understand the main question.
Respondent	Respondent read the experimental task and the main question loudly.
Observer	[Observation] respondent panned and zoom in and out the maps Respondent switched views between the two maps.
Respondent	‘The Map A has a stamped high credibility rating. But in terms of the colour used, I did not know the name of the buildings. Map B is clearer. But back into the question is to suggest the safest route from purple coloured to red coloured buildings. So I chose Map A because there is one organisation that reviewed this map and rated this map as high credible compared to the Map B.
Observer	In terms of the data presented on map, which map has more information?
Respondent	‘Map B has more information. Map A rated as high credible but it has less data compared to Map B.
Observer	[observation] Respondent looks confuse when making

	decision which map to choose.
Respondent	‘In terms of the information, Map B has more details and clear. In Map A the map used black coloured, some of building names are not really clearly displayed’ ‘I chose Map B’.
Observer	Although, the credibility rating is low? [comment] although there were more data in Map B, how to make sure that the data is correct, and not misleading.
Respondent	Yes, I think the rating is not too important. I can validate the data on my own. I can use the map; find my way by trials and errors. Who produced a map is not important. Anyone can produce a map.
Observer	Respondent’s cooperation in participating the survey was duly appreciated

Respodent id:	A (not real name)
Background:	Non-geoliterate
Age:	33
Gender:	Female
Decision	Map A
Observer	Requested respondent to understand the main question.
Respondent	Respondent read the experimental task and the main question loudly.
Observer	[Observation] respondent panned and clicks the features on map. Respondent switched views between the two maps.
Respondent	‘Where is the purple coloured building’?
Observer	You can use the zoom in and out function to view the map.
Respondent	‘The colour on Map A is black and make the text label not easy to read’ ‘Ok. I chose Map A’
Observer	Any reason?
Respondent	Because this Map has been reviewed as high credibility rating compared to Map B.
Observer	Was there any influence of the elements at the sidebar?
Respondent	Not really. Not influenced by the symbols design. Colour used in Map A is not attractive. So colours did not influence my decision. But since the producer of this map is the University of Nottingham, it supports my decision to choose Map A.
Observer	Respondent’s cooperation in participating the survey was duly appreciated