# User-Centred Design of Smartphone Augmented Reality in Urban Tourism Context

Zornitza Yovcheva

A thesis submitted in fulfilment of the requirements of Bournemouth University for the degree of Doctor of Philosophy

August 2015

## Copyright Statement

# User-Centred Design of Smartphone Augmented Reality in Urban Tourism Context

**Zornitza Yovcheva**

**Abstract**

Exposure to new and unfamiliar environments is a necessary part of nearly everyone's life. Effective communication of location-based information through various location-based service interfaces (LBSIs) became a key concern for cartographers, geographers, human-computer interaction (HCI) and professional designers alike. Much attention is directed towards Augmented Reality (AR) interfaces. Smartphone AR browsers deliver information about physical objects through spatially registered virtual annotations and can function as an interface to (geo)spatial and attribute data. Such applications have considerable potential for tourism. Recently, the number of studies discussing the optimal placement and layout of AR content increased. Results, however, do not scale well to the domain of urban tourism, because: 1) in any urban destination, many objects can be augmented with information; 2) each object can be a source of a substantial amount of information; 3) the incoming video feed is visually heterogeneous and complex; 4) the target user group is in an unfamiliar environment; 5) tourists have different information needs from urban residents.

Adopting a User-Centred Design (UCD) approach, the main aim of this research project was to make a theoretical contribution to design knowledge relevant to effective support for (geo)spatial knowledge acquisition in unfamiliar urban environments. The research activities were divided in four (iterative) stages: (1) theoretical, (2) requirements analysis, (3) design and (4) evaluation. After critical analysis of existing literature on design of AR, the theoretical stage involved development of a theoretical user-centred design framework, capturing current knowledge in several relevant disciplines. In the second stage, user requirements gathering was carried out through a field quasi experiment where tourists were asked to use AR browsers in an unfamiliar for them environment. Qualitative and quantitative data were used to identify key relationships, extend the user-centred design framework and generate hypotheses about effective and

efficient design. In the third stage, several design alternatives were developed and used to test the hypotheses through a laboratory-based quantitative study with 90 users. The results indicate that information acquisition through AR browsers is more effective and efficient if at least one element within the AR annotation matches the perceived visual characteristics or inferred non-visual attributes of target physical objects.

Finally, in order to ensure that all major constructs and relationships are identified, qualitative evaluation of AR annotations was carried out by HCI and GIS domain-expert users in an unfamiliar urban tourism context. The results show that effective information acquisition in urban tourism context will depend on the visual design and delivered content through AR annotations for both visible and non-visible points of interest. All results were later positioned within existing theory in order to develop a final conceptual user-centred design framework that shifts the perspective towards a more thorough understanding of the overall design space for mobile AR interfaces.

The dissertation has theoretical, methodological and practical implications. The main theoretical contribution of this thesis is to Information Systems Design Theory. The developed framework provides knowledge regarding the design of mobile AR. It can be used for hypotheses generation and further empirical evaluations of AR interfaces that facilitate knowledge acquisition in different types of environments and for different user groups. From a methodological point of view, the described user-based studies showcase how a UCD approach could be applied to design and evaluation of novel smartphone interfaces within the travel and tourism domain. Within industry the proposed framework could be used as a frame of reference by designers and developers who are not familiar with knowledge acquisition in urban environments and/or mobile AR interfaces.

**Keywords**: Augmented Reality browsers; urban environments; mobile Location-Based Services; Tourism; design framework;

## Publications

During the course of this project, a number of publications have been made which are based on the work presented in this thesis. They are listed here for reference.

### Journal Papers

[1] **Yovcheva, Z.**, Buhalis, D., Gatzidis, C. and van Elzakker, C., 2014. Empirical evaluation of smartphone Augmented Reality browsers in urban tourism destination context. *International Journal of Mobile Human Computer Interaction*, 6 (2), 10-31.

[2] **Yovcheva, Z.**, Buhalis, D. and Gatzidis, C., 2011. Overview of smartphone Augmented Reality applications for tourism. *e-Review of Tourism Research (eRTR)*, 10 (2), 63–66.

### Reviewed Conference and Workshop Papers

[3] **Yovcheva, Z.**, Buhalis, D., Gatzidis, C. and van Elzakker, C., 2013. Proposing a design framework for smartphone AR browsers used in unfamiliar urban tourism destinations. *In: Proceedings of the Augmented Reality MobileHCI 2013 workshop*, Munich, 27–30 August, 2013, Germany.

[4] **Yovcheva, Z.**, Buhalis, D. and Gatzidis, C., 2013. Towards meaningful augmentation of the cityscape: new challenges for mobile GeoHCI. *In: Proceedings from the GeoHCI workshop @ CHI2013*, Paris, France, 27–28 April 2013.

[5] **Yovcheva, Z.**, Buhalis, D. and Gatzidis, C., 2013. Engineering augmented tourism experiences. *In:* Cantoni, L. and Zheng, X., eds., *Proceedings of the International Conference on Information and Communication Technologies in Tourism (ENTER2013)*, Innsbruck, Austria, Springer, 24-35.

### Reviewed Extended Abstracts

[6] **Yovcheva, Z.**, Buhalis, D. and Gatzidis, C., 2013. Usability of geo-referenced information in smartphone context-aware Augmented Reality applications for tourism. *In: 3rd Workshop on Usability of Geographic Information*, Ordinance Survey, Newcastle, UK.

[7] **Yovcheva, Z.**, Buhalis, D., Gatzidis, C. (2012) Overview of smartphone Augmented Reality applications for tourism. *In: Workshop on Innovative approaches to Tourism Marketing and Management research*, Exeter, UK.

**Technical and Industry Reports**

[8] **Yovcheva, Z**. and Buhalis, D., 2013. Augmented Reality in Tourism: 10 unique applications explained. Digital Tourism Think Tank [Online]. Available: http://thinkdigital.travel/wp-content/uploads/2013/04/10-AR-Best-Practices-in-Tourism.pdf

## Acknowledgements

This research would not have been possible without the help of a number of brilliant, creative and thoughtful people. Herewith, I would like to acknowledge those who guided, supported and helped me along the way with bringing this work to a completion.

First and foremost, I would like to extend my gratitude to my first supervisor, Prof. Dimitrios Buhalis. I remain forever thankful for giving me the opportunity to carry out research in the John Kent Institute in Tourism and the expert advice and guidance to realize my ideas. I am sincerely thankful for helping me improve intellectually and professionally during the last four years.

I would also like to thank my second supervisor, Dr. Christos Gatzidis, for his constant support and encouragement. Thank you for your constructive suggestions, recommendations and ideas and methodical assistance. Your guidance was instrumental in helping me to improve my work and rationalize my choices.

I have been extremely fortunate to have the support of Dr. Corne van Elzakker. Thank you for your invaluable feedback, efficient scientific guidance, generous assistance and finding the time to indulge in fruitful and inspiring discussions. I am grateful for the provided technical assistance. Your help with obtaining the mobile field-based equipment was critical for this research. I am also thankful for your help with empirical data collection in Paris, France. Your professional and moral support was essential and contributed substantially towards the realization of this project.

I am very grateful to Dr. Alessandro Inversini who shares my enthusiasm for Human-Computer Interaction research. Thank you for your valuable feedback, ideas and suggestions for data collection and analysis regarding the evaluation of alternative AR annotation designs. I am very grateful for helping me organise and conduct the qualitative expert evaluation in Innsbruck, Austria.

My sincere gratitude also goes to the 112 participants who took out from their personal time and dedicated their effort to help me with my research. Your time and enthusiasm are deeply appreciated.

This research would have not been possible without the generous support of Dr. Ioannis Kent and his passionate dedication to research and development.

I would also like to thank my colleagues and friends at Bournemouth University. I would like to thank Barbara Neuhofer for our fruitful discussions and encouraging plans for future collaborative work. Thanks to Nicolas Gregori, who was a constant inspiration to continue developing my technical skills. A thank you goes to Gregori Kapuscinki, who contributed enormously with ideas and concepts from psychology research, but also kept me smiling during some of the tough times. A special thank you goes to Ivana Rihova who helped me substantially with making my ideas more concrete and providing guidance and help with qualitative research, conceptual framework development and formatting. I would like to express my sincere gratitude to Jessica Weber who continues to share my enthusiasm for user-centred design, user experience and Augmented Reality. Jess, thank you for your support and time! Special thanks to John Fotis for our fruitful discussions that helped me focus my thinking and for making this PhD journey a bit more sunny and happier.

I would also like to thank Bournemouth University staff, especially Karen Ward, Ina Grebliunaite and Sheila Rogers for their administrative support and assistance.

I would also like to thank Mila Luleva who continues to be a constant inspiration in both personal and professional aspects.

I am forever grateful for the constant support of my family, Tzveti, Ivan, Adi, Vasko and its newest member, Angelina. Your unconditional love has made me a better person and given me enormous strength and compassion.

Finally, and above all, I would like to thank Nikolay (Mecho) Dimov for being there for me irrespective of circumstances. Thank you for understanding and knowing me better than anyone else. Echoing the words of Carl Sagan, in the vastness of space and immensity of time, it is my deepest joy to share a planet and an epoch with you.

# Table of Contents

## List of Figures

**List of Tables**

## List of Abbreviations and Acronyms

| | |
|---|---|
| 2D | Two-dimensional |
| 3D | Three-dimensional |
| AR | Augmented Reality |
| ARML | Augmented Reality Markup Language |
| CA | Context-Awareness |
| CoU | Context of Use |
| CTA | Cognitive Task Analysis |
| GPS | Global Positioning System |
| HCI | Human-Computer Interaction |
| HMD | Head-Mounted Display |
| IS | Information System |
| ISMAR | International Symposium for Mixed and Augmented Reality |
| ISO | International Standards Organisation |
| KARML | Keyhole Augmented Reality Markup Language |
| KML | Keyhole Markup Language |
| mLBS | Mobile Location-Based Service |
| mLBSI | Mobile Location-Based Service Interface |
| MobIS | Mobile Information System |
| UCD | User-Centred Design |
| WARM | Winter Augmented Reality Meeting |
| XML | eXtensible Markup Language |

# CHAPTER 1

# INTRODUCTION

**1.1. State-of-the-art Smartphone AR Research**

Summarises briefly the state-of-the-art in current research

**1.2. Problem Definition and Motivation**

The key gaps in literature and drivers that motivated the study

**1.3. Aim and Research Objectives**

Describes the key aim and objectives of the study

**1.4. Research Approach**

Briefly describes the major tenets of User-Centred Design

**1.5. Thesis Contributions**

Positions the major contributions to theory and practice

**1.6. Structure of the Thesis**

Provides brief summaries of all chapters in the thesis

**1.7. Chapter Summary**

Synthesis of the key points in the chapter

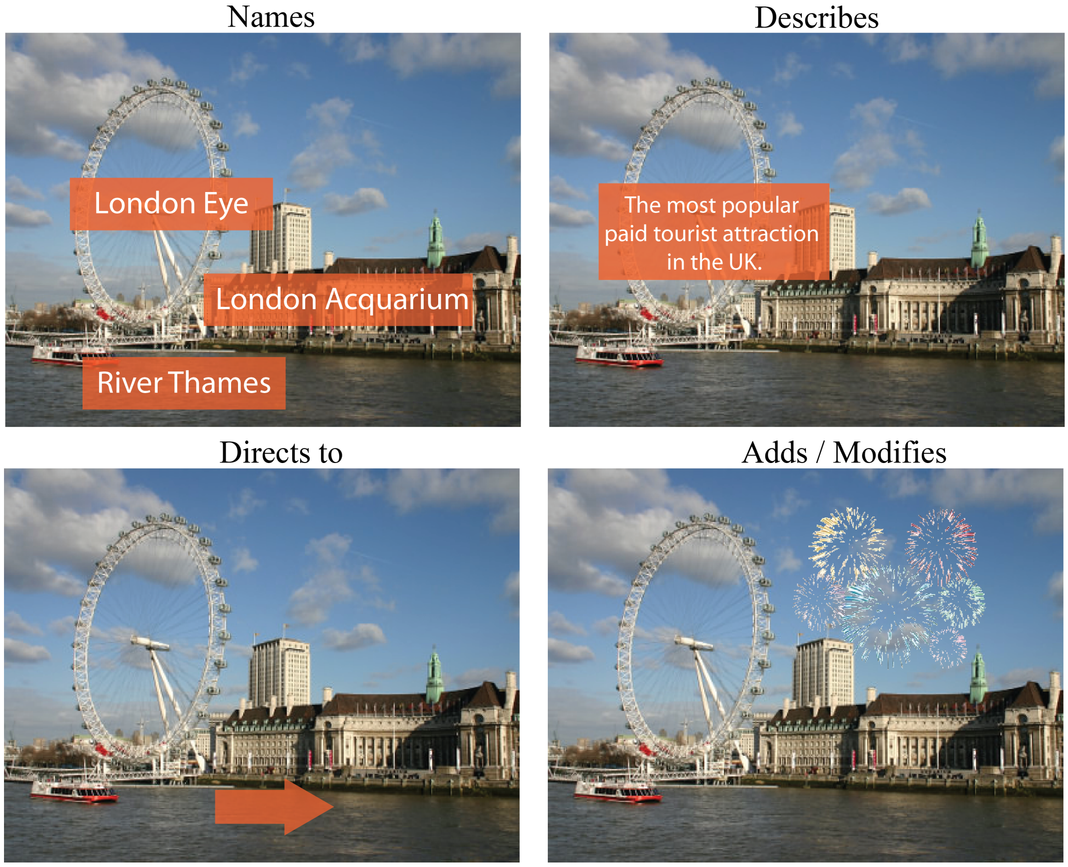
## 1.1. State-of-the Art Smartphone AR Research

Recent advance in mobile hardware and software has led many to believe that eTourism, the field concerned with how technology is used, adopted and applied in Tourism, is on the verge of a paradigm shift. Considering the various new mobile, wearable and ubiquitous computers, substantial research is directed towards smartphone devices. From a business perspective, the "always on, always carried" tenet has opened endless opportunities to reach customers quickly and at any time. From a consumer perspective, the smartphone is the first lightweight portable computer that can provide access to rich hypermedia, anywhere and at any time (Pearce, 2011). This is extremely beneficial for time-pressured tourists who need access to information that can be used "on the go", with minimum physical and mental effort, and minimum influence on their natural activities.

Smartphones are an excellent tool that can externalise and enhance tourists' cognition, helping them deal with the complexity of an unfamiliar urban environment. To this end, context-awareness, or the ability of the smartphone to use on-board physical and virtual sensors in order to adapt to the context where mobile interaction unravels, is considered critical. The development of mobile Location-Based Services (mLBSs) was the first concrete step towards context-awareness, as they "deliver information depending on the location of the device and user" (Raper et al. 2007, p.5). Unsurprisingly, the potential of mLBSs in tourism was quickly harnessed as mobile guides and recommender systems became the largest group of mLBS applications (Raper et al., 2011). Further advance in hardware and software allowed developers to transfer one of the most sophisticated and immersive location-based services to the smartphone: Augmented Reality.

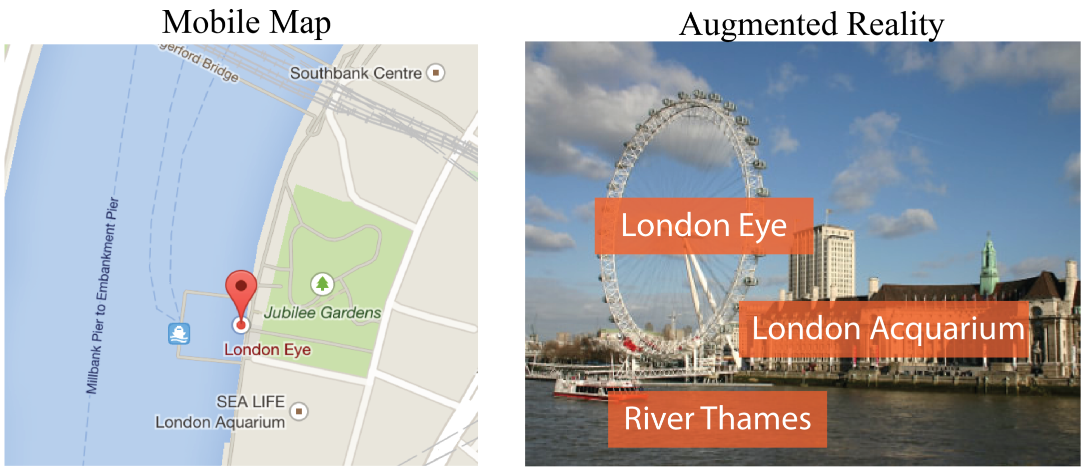
An Augmented Reality (AR) system enhances or augments the (perception of the) surroundings of its user in real-time with virtual (computer generated) information that seems to co-exists with the real world (Azuma et al. 2001). An AR system can perform a number of functions (Figure 1.1) (Wither et al., 2009), including to *name* and *describe* or *direct* the tourist towards a point of interest through virtual arrows. AR can also *modify* how tourists perceive their surroundings through superimposition of special (often 3D or animated) graphics. From the various types of applications, this study is specifically concerned with AR interfaces that deliver virtual content which *names* and *describes* objects or locations of interest. This special type of AR interfaces are called *AR browsers* (Langlotz et al., 2014).

**Figure 1.1. The five different functions of AR annotations.**




Considered "the substitute of the Web browser" (Langlotz et al., 2014, p.155), AR browsers deliver information about locations, objects and points of interest in spatially registered virtual balloons, called *AR annotations*, illustrated in Figure 1.1 (Wither et al. 2009; Madden 2011). In a hypothetical use case scenario, a tourist points a smartphone device towards a building. He is then able to see the name of that building, the year it was built in and, perhaps, explanation about its architectural style. As opposed to spending time to look up this information in a guidebook or on a map, the AR annotation is immediately within the field of view of the user, as it is overlaid on top of (or near to) the physical object. Information delivery in this way would then be associated with much less cognitive and physical effort, as the tourist is not forced to switch back and forth his gaze between information space and the physical world. This becomes evident when we compare on-site information acquisition about physical objects through AR browsers and more traditional map-based interfaces (Figure 1.2).

**Figure 1.2. Difference between map-based and AR representation of information.**




In the last couple of years the popularity of AR browsers grew exponentially. The period between 2009 and 2013 witnessed the creation of more than 700 smartphone AR applications. Substantial resources were dedicated to the design and development of smartphone AR systems within academia and industry. Many special events (e.g. the Winter Augmented Reality Meeting, WARM), symposia (e.g. the International Symposium on Mixed and Augmented Reality, ISMAR) and conferences (e.g. Augmented Reality Summit Conference) were established as platforms for academic and industrial discourse, dissemination and collaboration. The substantial amount of refereed papers in thematic and domain-specific journals, together with the broad range of conference and workshop papers forms convincing proof that the domain is rapidly growing and expanding.

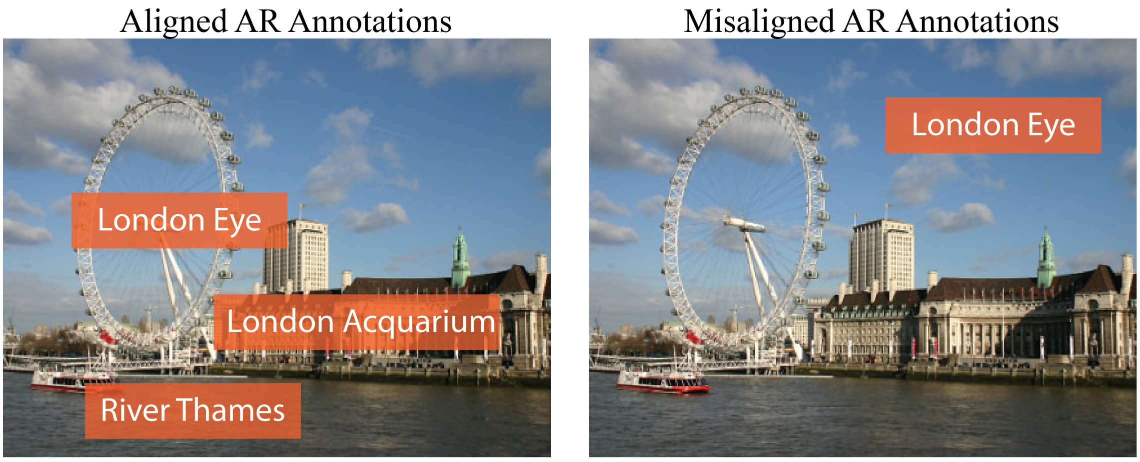
## 1.2. Problem definition and motivation

In spite of its increasing popularity, it seems that AR browsers fail to live up to users' expectations (Grubert et al., 2011; Olsson and Salo 2011; Linaza et al. 2012) as both residents and tourists criticize the usability and utility of such applications. A wide range of studies has consistently documented the technical, content selection and delivery, and design challenges that AR faces.

**Technical Challenges**

The underlying presumption of the majority of research within smartphone AR (e.g. Keil et al., 2011; Grasset et al., 2013; van Krevelen and Poelman, 2010; Geiger et al., 2014; Langlotz et al., 2014) has been that the main challenge preventing wider user adoption has been a technological one, i.e. inaccuracies in registering and tracking the viewpoint of the user. As a result, the system is unable to accurately align physical and virtual worlds (Figure 1.3). Recent studies (e.g. Grasset et al., 2013) have been directed at development of various registration and computer vision techniques to solve such challenges.

**Figure 1.3. Aligned and misaligned representation of information in AR.**




While substantial progress has been made (e.g. Grasset et al., 2013; van Krevelen and Poelman, 2010; Grasset et al., 2013; Langlotz et al., 2013; Langlotz et al., 2014), there is still lack of understanding what are the minimum requirements that a smartphone AR application has to fulfill in order to support its users effectively in different contexts and application domains (Livingston, 2013). For instance, it is logical to assume that more precise tracking and registration is necessary when virtual content augments a heavily built up urban environment. At the same time, while not specifically dedicated to tracking, the results from several studies related to tourism (Turunen et al., 2010; Ganapathy et al., 2011) indicate that users might be able to tolerate inaccurate alignment of physical and virtual worlds. The margin of error that users can tolerate remains unaddressed. More importantly, it is still unclear whether this is the most important or the only requirement that determines effective and efficient use of AR browsers. Considering that both academia (e.g. Luley et al., 2012) and industry (e.g. metaio, 2014) are working towards improving tracking and registration, it is critical that both system and user requirements are further investigated and identified.

**Content Challenges**

Apart from technical tracking and registration, the development of a smartphone AR system is connected with a number of critical decisions related to the type of content it will deliver. Current commercial AR browsers rely on geo-tagged content, such as textual descriptions, pictures or videos (Langlotz et al., 2014). Companies that develop AR browsers typically store such content in proprietary formats in order to protect it.

Developers can also import their own content by geo-tagging it manually, which makes the process very time-consuming and prone to registration errors (Langlotz et al., 2014). In recent years, development has relied on geo-tagged content produced by the general public, also referred to as *user-generated content* (UGC) or *volunteered geographic information* (VGI) (Krumm et al., 2008). The problem with UGC is that it is stored in a format that is not suitable for AR. Developers and designers have to make a number of decisions how to extract, re-format, and more importantly, represent such content through an AR browser interface. It is still unclear, however, what is the content that users need, especially when it comes to urban tourism context. Considering the effort that AR development requires, it is important to elicit user requirements and provide guidelines regarding relevant and useful content delivered through AR.

**Graphical Design and Representation Challenges**

Apart from selecting the right content, it is also important to consider how it is represented to users through graphical AR annotations. AR implemented on smartphone devices is substantially different from traditional graphical user interfaces developed for desktop Information Systems. Delivery of information through smartphone devices is difficult because of various technical (e.g. smaller screen, patchy connectivity, short battery life) and contextual (e.g. dynamic changes in weather and lightning conditions) challenges (Gorlenko and Merrik, 2003; Krogstie et al., 2003). As a result of the dynamic change in context, users can allocate only limited attentional and cognitive resources to the mobile device (Loojie et al., 2007). In addition, in contrast to standard mobile graphical user interfaces, AR combines both physical (real-world) and computer-generated virtual information. This novel user interface metaphor challenges the scope of established human-computer interaction styles (Kjeldskov et al., 2003) and questions the applicability of the scarce range of established mobile design principles. As a result, the design space for AR is very vast and widely unknown (Gabbard and Swan II, 2008). What this means is that there is still little knowledge how content and graphical design decisions impact the effectiveness and efficiency of users and, as a result, design is often sporadic. The main reason for this is that there is still little understanding with respect to the user requirements that have to be fulfilled in order to ensure effective and efficient work with smartphone AR in general, and AR browsers in particular.

**Summary: the challenge to design usable and useful AR browsers**

The issues presented above are all part of the major problem related to how to design an AR browser used in tourism context. Design (both as the properties of a system and the process of creating a product) can relate to different aspects of an information system, such as (1) the logical user interface (e.g. information architecture), (2) the physical user interface (hardware components), or (3) the graphical user interface (e.g. layout, representation of data) (Heo et al., 2009). As the technology matures, the challenge is no longer to only develop algorithms that deliver accurately aligned virtual and physical objects, but provide design guidelines for more useful and usable smartphone AR systems. Coming up with such recommendations is more difficult in urban tourist destination context, as:

1) in any urban destination, many objects can be augmented with information;

2) each object can be a source of a substantial amount of information;

3) the incoming video feed is visually heterogeneous and complex;

4) the target user group is in an unfamiliar environment;

5) tourists have different information needs from urban residents.

In all of these cases, the core principles that drive a design process and the qualities that an IS should possess, are captured in *design knowledge*. Design knowledge is accumulated over time and described in *design theory* (Gregor and Jones, 2007), defined as “a prescriptive theory which integrates normative and descriptive theories into design paths intended to produce more effective information systems” (Walls et al., 2004, p.48). Central to Information Systems Design Theory (ISDT) generation is the concept of identifying (user and system) meta-requirements, or high-level descriptions of what an information system or a particular piece of software should do and look like. Misidentification of user and system requirements is one of the primary causes of customer dissatisfaction and rejection of ISs (Davis, 1982; Jerkins et al., 1984; Avison and Fitzgerald; 2000). In their seminal paper, Nunamaker and Chen (1991) first described the importance of empirical user-based studies, such as observations and experiments, as tools for accurately identifying IS requirements. Smartphone IS designers were also fast to recognize the significance of empirical user research (Fling, 2009). Empirical user-based studies, carried out in actual context of use, are also critical for identifying user requirements, generating design knowledge and developing useful and usable AR interfaces (Gabbard and Swan II, 2008).

Unfortunately, empirical evaluations of smartphone AR browsers with representative users in actual context of use are still very rare, especially when it comes to urban tourism. If designed appropriately, AR browsers can act as extensions of human sensory-motor capabilities and enhance tourists experiences by helping people to learn, think and reason about large-scale environments. Likewise, inappropriate design may cause cognitive overload, difficulties with focusing attention (Price, 2002), confusion and annoyance. The process of knowledge acquisition in large-scale urban environments is complex (Downs and Stea, 1973; Siegel and White, 1975) and many perceptual and cognitive factors need to be considered. The few studies that investigate empirically how AR can enhance on-site experiences in urban environments have focused primarily on navigation (Walther-Franks and Malaka, 2008; Medenica et al., 2011), rather than information acquisition about points of interest through AR browsers. Currently, there is a noticeable lack of discussion what is the role of AR interfaces, and particularly AR browsers, in the overall geospatial knowledge acquisition process in large-scale urban spaces. More importantly, until now, AR research has failed to investigate and discuss the design principles and guidelines that need to be applied with respect to alignment, content, graphical design and representation, in order to make AR browsers truly useful and usable for tourists.

## 1.3. Aim and Research Objectives

User requirements, as well as design principles how to meet those requirements, are considered essential for the success of mobile information systems. User involvement and user-based studies in actual context of use are still rare, especially when it comes to the use of AR browsers in unfamiliar urban tourism destinations. In practice, this leads to lack of wider adoption of AR within the general population and finding meaningful uses for the technology in the field of tourism.

Located within Information Systems Design and Human-Computer Interaction, the main aim of this study is to make a theoretical contribution through generating user-centred design knowledge expressed as the qualities and characteristics that Augmented Reality browsers should possess in order to meet user requirements in urban tourism context.

The following objectives help to further this enquiry:

***Research Objective 1***: Explore the role of AR browsers in supporting (geospatial) information acquisition in large-scale urban tourism destinations.

***Research Objective 2***: Examine the main problems that influence the usability and utility of AR browsers used in urban tourism destinations.

***Research Objective 3***: Investigate how key context of use factors influence the usability and utility of AR browsers.

***Research Objective 4***: Identify the key user requirements that need to be satisfied in order to improve the usability and utility of AR browsers.

***Research Objective 5***: Propose key design parameters that could be used to improve the usability and utility of AR browsers.

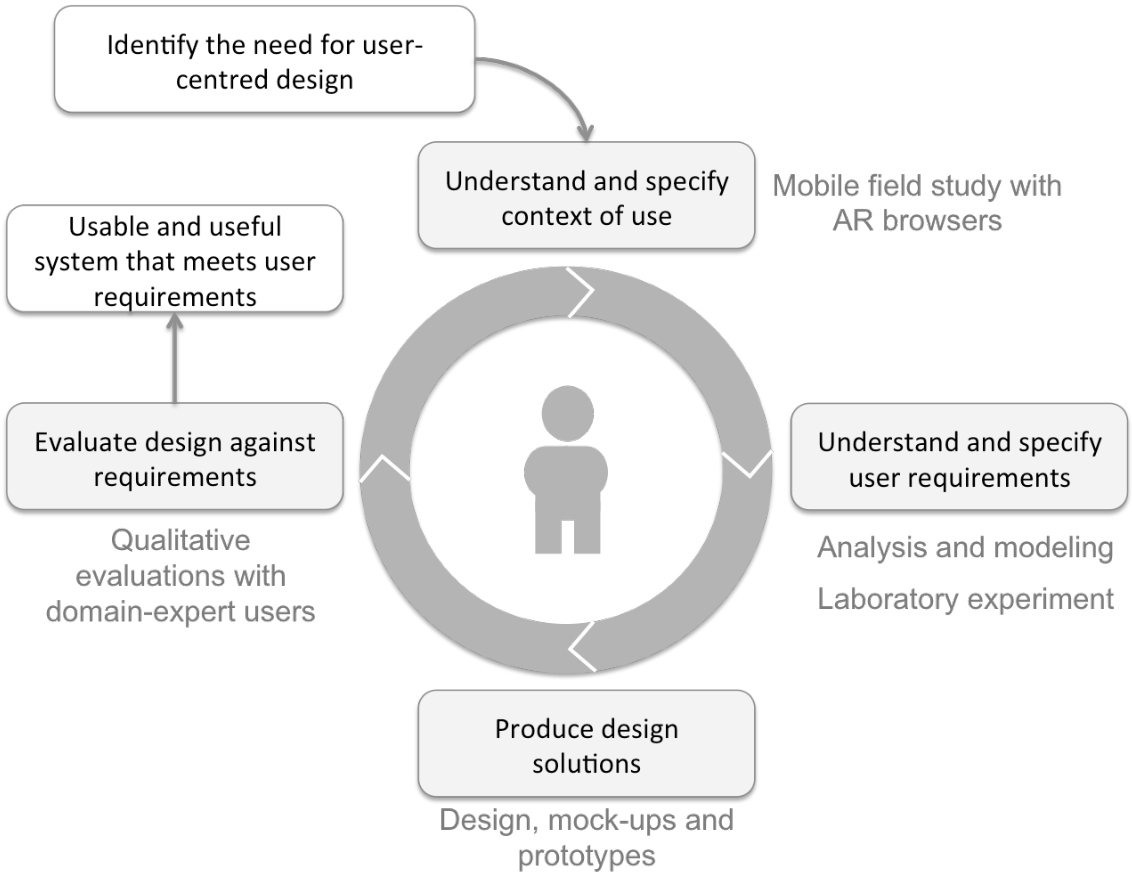
***Research Objective 6***: Capture the key elements and relationships that determine usability and utility of AR browsers in a conceptual user-centred design framework that facilitates the design and evaluation of AR browsers.

***Research Objective 7***: Propose design principles for developing AR browsers used in tourism context.

## 1.4. Research approach

Determining user requirements is a complex task, prone to a number of errors, which have been documented extensively in IS literature (Davis, 1982; Jerkins, 1984; Palmer, 1987; Nielsen, 1993; Hackos and Redish, 1998). A methodology that emphasizes the role of user requirements and their accurate gathering and analysis is ***User-Centred Design*** (UCD). UCD is one of the major concepts that emerged from the early HCI research, describing an approach (and methodology) to design in which the end-users of a product shape out its final outlook (Abras et al. 2004). Since its introduction in the late 1980s, a number of authors have contributed to the initial theory constructs (Nielsen 1993; Mayhew 1999) leading to the recognition that today UCD is "one of the guiding principles for designing usable technologies" (Hacklay & Nivala 2011, p.91). Early user involvement, iterative design and empirical evaluation are the three main principles that underpin the essence of UCD **(**Gould and Lewis, 1985; Maguire, 2001; Abras et al., 2004**)**. Adopting a User-Centred Design (UCD) approach, the research activities in this study were divided in several (iterative) stages (Figure 1.4): (1) theoretical, (2) requirements analysis, (3) consolidation, design and validation and (4) evaluation.

**Figure 1.4. Iterative stages adopted as part of a UCD**




The goals and scope of the current study were considered accordingly, with the decision ultimately made to approach the UCD research methodology from a mixed methods perspective, involving the collection of both qualitative and quantitative data.

## 1.5. Thesis Outcomes

Selecting a User-Centred Design as an overarching methodology offered a particular value for this research, as it provided a structured approach that guided empirical data collection, and thereafter its analysis.

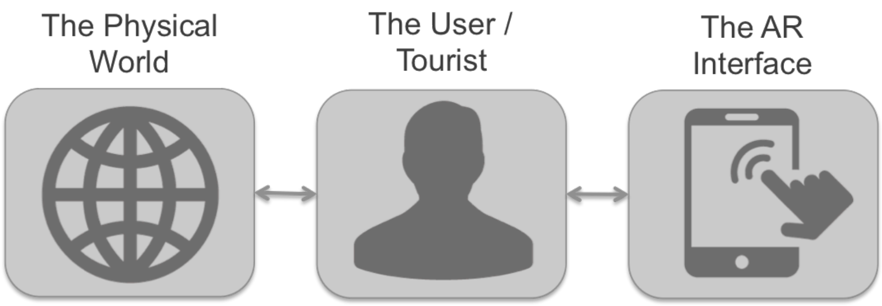
The main goal of generating prescriptive design knowledge is to identify how the proposed new Information System, or a new design, will solve a specific problem. As discussed later in this thesis, understanding the nature of the problem poses a difficulty. The use of Augmented Reality browsers by tourists in unfamiliar environments could be explored and examines from a number of theoretical perspectives not only before empirical data collection, but also once empirical data are gathered and during their analysis.

**Table 1.1. The table describes how empirical and theoretical work informed the final user-centred design framework.**

| Stage | The Physical World | The User / Tourist | The AR Interface |
|---|---|---|---|
| **Stage A - Theoretical Framework Development** | ***Literature***- Siegel and White (1975) – Types of spatial knowledge tourists needs in large environments<br><br>***Literature*** - Landmarks as anchoring points for obtaining spatial knowledge (Hart and Moore, 1973) | ***Literature*** - The formation of information needs as consciously identified cognitive gaps (Dervin, 1984).<br><br>***Assumption 1*** - Acquired knowledge about landmarks will influence the way tourists interact with an AR browser. | ***Literature*** - Different interfaces can communicate different types of knowledge (Oulasvirta et al., 2009).<br><br>***Literature*** - Spatial interfaces are composed of two layers – base and attribute layer (Longley et al., 2011). |
| **Stage B – Mobile Field-Based Evaluation of AR Browsers** | ***Empirical Data*** – Acquired knowledge influences information needs (validates Assumption 1).<br><br>***Empirical Data*** – Success with AR browsers is influenced by the type of physical object tourists are trying to find information about.<br><br>***Literature*** – The above observation is due to the influence of visual salience and legibility (Craik and Appleyard, 1980; Nasar et al., 2005).<br><br>***Empirical Data*** – Visually salient features trigger interaction with the AR browser.<br><br>***Literature*** – Visual salience is a unique property dependent on the feature, the surrounding environment and the user (Caduff and Timpf, 2008).<br><br>***Assumption 3*** – The type of physical environment will directly influence the information needs of tourists and interaction with the AR browser. | | ***Empirical Data*** – Current AR browsers do not deliver the information that tourists need.<br><br>***Empirical Data*** – Current AR browsers are difficult to use, as tourists cannot associate virtual and physical worlds.<br><br>***Assumption 4*** – Effectiveness and efficiency of using AR browsers depends on whether AR annotations match directly or indirectly the physical world. |
| **Stage C – Laboratory Evaluation of AR Browsers** | ***Empirical Data*** - Confirmed Assumption 4. Users take less time and make fewer errors when the design of AR annotations matches directly the characteristics of the physical object for both precise and imprecise placement. | | |
| **Stage D – User Feedback and Qualitative Evaluation of AR** | ***Empirical Data*** – Confirmed Assumption 3. Information needs formation is hindered in visually non-salient (uniform, residential) environments.<br>***Empirical Data*** – Tourists want to acquire information about both visible and non-visible physical objects. | | |
| **Stage E – Conceptual User Centred Design Framework** | ***Consolidation of theory and empirical observations in order to:***<br>***1) explain the main interactions among the three components and elicit user requirements***<br>***2) prescribe design of AR browsers to satisfy user requirements*** | | |

There are a number of factors that could influence the use of AR browsers by tourists. Therefore, it was important to identify the main components of interaction and the relationships among them at least on a very high level. As Chapter 4 describes, there are three main interaction components: 1) the physical world, 2) the tourist and 3) the AR interface which mediates the relationship between the tourist and the physical world (Figure 1.5).

**Figure 1.5. The three main framework components**




**Stage A**: In line with the iterative process of Information Systems Design theory generation (Walls et al., 1992; Hevner et al., 2004; Arazy et al., 2010), the first step was the identification of kernel theories, or the most relevant empirical research concepts, frameworks and models that could help in understanding the design space for AR browsers and the nature of different contexts of use.

**Stage B, C and D**: Empirical research was then carried out where quantitative and qualitative data were collected in order to gauge insights on the relationships among the three primary components and their interaction.

Table 1.1 shows in a concise format the evolution of the design framework throughout the various stages. It also maps how theory and empirical data came together in order to inform the outcomes of the thesis.

## 1.6. Thesis contributions

This study aims to contribute to the wide field of Information Systems Design by investigating AR browsers as (visual) tools that can enhance and support (geo)spatial information acquisition in large-scale environments. In line with the general process of Information Systems Design theory generation (Nunamaker and Chen, 1991; Walls et al., 1992; Markus et al., 2002; Gregor, 2009), the main theoretical contribution of the study is the development of a new user-centred design framework, which places tourists at the centre of design. As such, the primary contribution of this thesis is theoretical in nature. As Wobbrock (2012) clarifies, a theoretical contribution can consist of concepts, models, principles, frameworks or a variation of those. Conceptual models and frameworks can be qualitative and quantitative in nature. While quantitative data were used to confirm some of the identified relationships among components that interact in the real world, it should be emphasised that the proposed framework is qualitative in nature. It combines both existing theory and new empirical observations and identifies key requirements, design guidelines and principles that have to be followed when developing AR interfaces in the future. As such, its primary purpose is to provide an overarching frame of reference for the design process of AR browsers, especially when used by tourists in unfamiliar environments.

The framework captures knowledge that contributes to Human-Computer Interaction (HCI) as a sub-field of the wide Information Systems Design field. It can be used as a tool that helps researchers to analyse the present, but also in the pursuit of future knowledge. Within Human Computer Interaction, the framework sheds light on the underlying relationship among three components that interact when tourists use smartphone AR: (1) the people who roam around in an (2) unfamiliar physical environment and (3) technology as a tool that mediates how they experience such environments. Based on such interactions, the framework prescribes the properties of the digital layout that should be superimposed over physical objects in order to enhance the perception and understanding, or the learning of new environments.

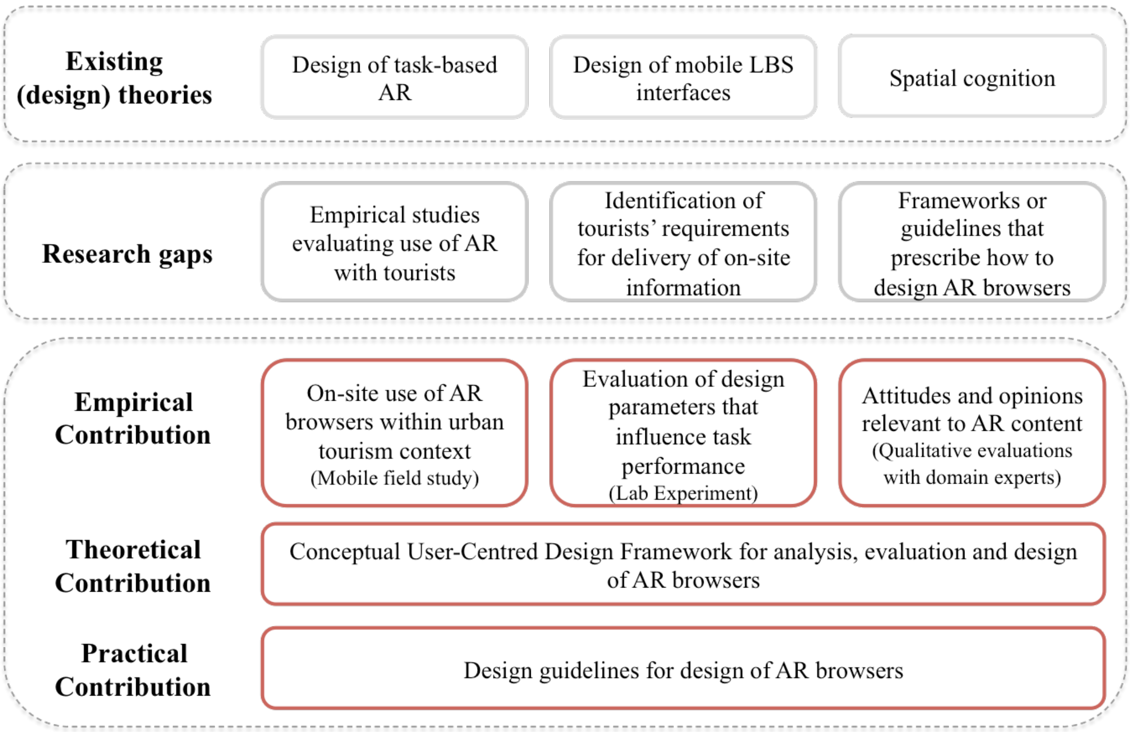
More importantly, however, the framework uncovers the latent but important influence of the physical environment on user requirements. The multi-disciplinary framework incorporates notions and concepts from Geo-Information Science and Environments Psychology and, as such, provides a new theoretical perspective for design and development, as well as research for AR. Unlike research which has been scattered in many (often implicit) theoretical and practical perspectives, the framework provides a clear and focused direction for AR used in large environments.

Unlike models and frameworks that try to explain the process of information acquisition prior to undertaking a trip, the framework provides a detailed description of the interactions that take place during on-site information acquisition by tourists. On a very high level, it captures the factors that trigger information search and that influence and determine how AR can change the experience with unfamiliar physical environments.

Practitioners in tourism should study the framework to deepen their understanding of the (potential) use of AR and how it can satisfy on-site information needs. HCI practitioners can use the framework to deepen their understanding of the underlying perceptual and cognitive phenomena that unravel when tourists make sense of unfamiliar environments through the use of AR.

The validity of the identified processes is grounded partially within existing theories and design knowledge within several relevant disciplines. In addition, this study aimed to generate a significant amount of empirical data that were used to validate and revise the major design propositions. Therefore, in line with the overall role and nature of Information Systems design theories, the framework captures new design knowledge and proposes how to design smartphone AR browsers through a series of new design guidelines (Figure 1.5).

**Figure 1.6. Contribution diagram**




Considering the unique and multi-disciplinary approach undertaken in this study and the scope of the obtained results, the thesis has smaller contributions relevant to new theoretical, empirical and methodological knowledge within the fields of Augmented Reality, Mobile Human-Computer Interaction and Geo-Information Systems design.

**Theoretical contributions:**

- A conceptual framework for analysis and design of smartphone AR browsers. The framework examines interaction with AR browsers and accommodates existing theories to explain the process of information acquisition in unfamiliar environments. The framework is of high value and relevance to researchers as it can be used to support the planning of experimental and user-based studies.

**Empirical contributions**

Research findings on:

- Work and embodied interaction with smartphone AR browsers in unfamiliar urban environments.
- Users' ability to carry out association of virtual AR annotations and physical entities in urban environments.
- The severity of problems that users experience with smartphone AR browsers in actual urban environments.
- Domain-experts' concerns regarding the design of smartphone AR browsers.

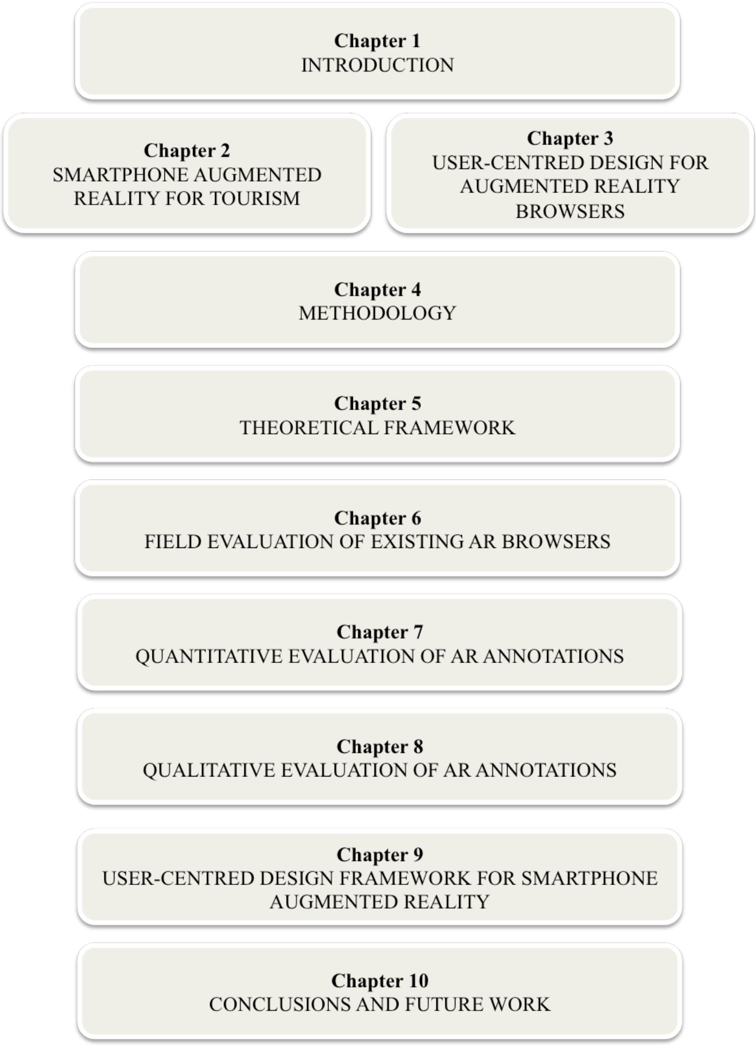
**Contribution to practice**

- Design guidelines and principles for more useful and usable design of smartphone AR browsers.

## 1.7. Structure of the thesis

Following the introduction to the study, described above, the thesis is divided into the following subsequent chapters (Figure 1.6):

**Figure 1.7. Structure of the thesis**




***Chapter 2 – Smartphone Augmented Reality for Tourism***

This chapter offers a review of state-of-the-art smartphone technologies relevant to tourism. It starts with description of the key enabling technologies and types of location-based services (mLBS). As a special type of mLBS, the key characteristics, similarities and differences of location-based Augmented Reality are then examined. Recent trends in ubiquitous and pervasive computing and the paradigm shift towards context-aware services is described. The chapter finishes with the key challenges that underpin the development of next-generation context-aware augmented reality services.

***Chapter 3 – User-Centred Design for Augmented Reality Browsers***

The chapter starts with definitions for usability and utility and moves on to explain the key factors that influence usability and utility of smartphone applications in general, and

AR browsers in particular. Description of User-Centred Design follows, which emphasizes the need to adopt the approach in order to identify meaningful design guidelines for AR browsers. The rest of the chapter reviews state-of-the art research relating to empirical evaluations of AR browsers, focusing on the specific type of virtual content delivered to users: AR annotations.

***Chapter 4 – Methodology***

Chapter 4 outlines the selected research paradigm and methodological approach. The chapter describes and justifies the selection of research methods from available empirical methods employed in human-computer interaction. The remainder of the chapter discusses how limitations are addressed by the research plan and methodological approach taken in this thesis.

***Chapter 5 – Theoretical Framework***

In response to the fragmented nature of current research that deals with design and development of AR browsers Chapter 5 introduces the unique approach adopted in this study, by describing the conceptual theoretical framework that underpins further empirical investigation of AR browsers used in tourism context. The framework incorporates existing empirical research and theoretical concepts in several disciplines related to the main phenomenon under study in this thesis, including geo-information science, environmental psychology, information science, and mobile human-computer interaction. The framework provides the conceptual and terminological basis for the empirical research presented in the thesis.

***Chapter 6 – Field Evaluation of Existing AR Browsers***

The chapter describes the procedures undertaken during the first empirical field-based evaluation of AR browsers, which investigated different aspects of actual use of such interfaces in unfamiliar urban environments. The key findings are then presented, which illustrate the major usability problems that tourists experience when they have to rely on AR browsers in large-scale environments. The findings emphasize the key role of perceived physical context and its influence on mobile interaction.

***Chapter 7 – Quantitative Evaluation of AR Annotations***

The chapter describes the second empirical evaluation, which adopted a laboratory-based experimental approach. The overarching goal of conducting further empirical testing was to validate the findings from the first empirical study and confirm the newly identified relationships that unravel during interaction with AR browsers. For this purpose, a laboratory-based experiment was conducted with 90 participants. The experiment was designed to compare task performance with three alternative AR annotation designs with both precise and imprecisely placed AR annotations. In order to generalise the results to different types of urban environments, different urban settings were used in three famous urban tourism destinations.

***Chapter 8 – Qualitative Evaluation of AR Annotations***

The chapter presents the procedures and findings from two qualitative evaluations of AR annotations. Both evaluations followed the design and procedures for pluralistic walkthrough evaluations and were carried out with domain expert users in Human-Computer Interaction, eTourism and Geo-Information Science. While the first empirical evaluation was carried out in controlled settings, the second was conducted in the field in an unfamiliar urban tourism context. The main goal was to obtain further feedback about content and issues related to relevance and usability of delivered information.

***Chapter 9 – User-Centred Design Framework for Smartphone Augmented Reality***

Chapter 9 starts with revisiting the major constructs and relationships in the proposed theoretical framework, developed in the beginning of the study. The chapter summarises and generalises the major findings from empirical work and describes the key user requirements that were elicited during this research project. It then deals with the implications from the presented work, describing elicited user requirements and suggesting design guidelines for smartphone AR.

***Chapter 10 – Conclusion and Future Work***

Chapter 10 returns to the aim and objectives of the research. It summarises the substantive empirical and theoretical contributions of the thesis. The text then provides formal evaluation of the work, concluding with opportunities for future research.

## 1.8. Chapter Summary

This chapter provided context for the research, highlighting briefly recent advance in AR research and emphasizing the need for a user-centred approach to design and development. Additionally, the aim and objectives driving the research were presented, along with a brief description of the adopted approach and the intended contribution of the thesis. The next chapter sets the scene for the research by introducing and describing the major themes that drive development of smartphone ISs. In particular, concepts, constructs, and recent developments within the field of Mobile Location-Based Services, smartphone Augmented Reality and Context-Based services are presented.

# CHAPTER 2

# SMARTPHONE AUGMENTED REALITY FOR TOURISM

**2.1. From Mainframes to Smartphone ISs**

Highlights key developments in Human-Computer Interaction

**2.2. Mobile Information Systems and Tourism**

Benefits of the smartphone to deliver information to tourists

**2.3. Mobile Location-Based Services for Tourism**

Enabling technologies and types of mLBSs

**2.4. Smartphone Augmented Reality Browsers**

Key definitions, enabling technologies and challenges

**2.5. Towards Context-Aware Smartphone AR**

Highlights the need to move beyond location-awareness towards more context-aware and adaptive mobile services

**2.6. Chapter Summary**

Synthesis of the key points in the chapter

## 2.1. From Mainframes to Smartphone Information Systems

In the past, computers were operated by trained programmers and technicians and placed mainly in specially equipped laboratories. The introduction of computers in corporate and government organisations (mainframes) between 1965-1980s allowed access of the general public to Information Systems (ISs), defined as “an integrated man/machine system for providing information to support the operation, management and decision-making functions in organizations” (Davis, 1974 cited in Zhang et al., 2004). The impact and importance of ISs grew significantly when personal computers became available during the 80s. Later, the emergence of the Internet provided a high-speed and low-cost way to connect and share information among thousands of distributed ISs.

Such developments were fundamentally important for the tourism industry, which has enormous potential for the use and adoption of new technologies (Buhalis and Law, 2008). Information Systems became central within the industry for both suppliers and customers of tourism and hospitality services. From the suppliers’ perspective, emerging new technologies transform the structure of competitiveness, allowing for lower costs and enhanced operational efficiency, new cost-effective and proficient marketing channels, and fast service failure recovery due to timely acquisition of relevant and updated information. From a consumer point of view, Information Systems allowed access to relevant information prior to a trip.

The tremendous technological convergence and interoperability in the 21st century led to the birth of more powerful mobile and wearable computers. A significant amount of research has recently been directed towards one of the most powerful mobile computers: the smartphone. The potential of the smartphone as a sophisticated information delivery channel is enormous (Krogstie et al., 2003; Fling, 2009), especially when it comes to tourism (Dickinson et al., 2013). This chapter captures state-of-the-art developments, which are fundamental when it comes to improving the design of smartphone AR used in tourism context. The chapter starts with a discussion of tourists and their characteristics aiming to revisit the need for swift and accurate information delivery during the on-site travel stage (Section 2.2). The chapter then discusses characteristics and recent trends in the swiftly developing field of mobile computing with focus on location-based services (Section 2.3) and smartphone Augmented Reality

(Section 2.4). Apart from benefits, the chapter discusses the limitations of such external visual representations and the need for more context-aware smartphone information systems (Section 2.5).

## 2.2. Mobile Information Systems and Tourism

### 2.2.1. Tourists and information needs

The United Nations World Tourism Organization (UNWTO, 2007) defines a tourist as "a traveller taking a *trip* to a main destination outside his/her *usual environment*, for less than a year, for any main purpose (business, leisure or other personal purpose) other than to be employed by a resident entity in the country or place visited". Tourists are often classified based on their reasons for travelling into three broad categories: 1) business and professional, 2) leisure and holiday, and 3) tourists travelling to visit friends and relatives. Irrespective of purpose of travelling, the definition emphasises an aspect related to tourists' characteristics, which makes the provision of information a critical necessity: trips are normally undertaken within an unfamiliar destination. In addition, travelling requires expenditure and the purchase of intangible services associated with high-risk, high-cost and high-involvement choices (Roehl and Fesenmaier, 1992; Költringer and Wöber, 2010).

In order to reduce the level of uncertainty and risks, tourists need to maximise their knowledge about a destination by acquiring as much information as possible (Fodness and Murray, 1997), often considering a wide variety of information sources. Information search and knowledge acquisition prior to a trip is often quite extensive and involves multiple information sources, especially when it comes to leisure-related travel which involves extensive expenditure (Schul and Crompton, 1983). Since holidays and leisure-related travel is considered to be a necessary part of a healthy lifestyle, most people will engage in the process of holiday-related information acquisition at least once a year.

Understanding information search behaviour is considered vital for both tourism scholars and practitioners (Fodness and Murray, 1997). Traditionally, information acquisition by tourists has been separated into three main stages: pre-trip, on-site (during trip) and post-trip (Steward and Vogt, 1999). Following traditions in consumer behaviour and marketing research, most studies have examined the use and influence of various information sources during the pre-trip stage (Gursoy and McCleary, 2004) when marketers and travel agents can influence the selection of a final destination, the booking of accommodation or the purchase of ancillary products. Conceptualizing travel decision-making, Fesenmaier and Jeng (2000) argue that information acquired in this stage is often used for making *core decisions*, which include the type of destination, the time of travel and accommodation.

Apart from acquiring information before arriving in an unfamiliar environment, tourists also engage in extensive on-site information search once they reach their destination. The acquired knowledge is used to make *secondary and en-route sub-decisions* (Fesenmaier and Jeng, 2000), which include selection of activities, where to eat or choosing which attractions to visit. The latter activity, also referred to as *sightseeing*, requires substantial amount of spatial (where), attribute (what) and temporal (when) data (Brown and Chalmers, 2003). In this context, on-site information acquisition is carried out not only as a risk reduction strategy, but also to maximise the quality of the trip and enhance the experience with the destination (Kah et al., 2011).

Unfortunately, on-site information acquisition remains an under-researched field. The need to investigate on-site information needs and information search behaviour has already been noted (Brown and Chalmers, 2003; Kah et al., 2011). This type of knowledge is considered essential not only in order to optimise the delivery of information, but also to improve marketing and promotion campaigns. Partially, the lack of knowledge could be explained by methodological difficulties connected with studying the needs of a wide and varied audience, as potentially anyone could be a tourist.

### 2.2.2. External tools that aid knowledge acquisition in tourism

The type of information people acquire, as well as the process that leads to obtaining knowledge about the environment, will heavily influence tourists' behaviour at the destination (Boulding et al., 2005). Both temporal and spatial information play a key role during mobility and are essential for a wide range of high- and low-level tasks, such as orientation, wayfinding and navigation. There are two main ways to acquire knowledge about unfamiliar environments (Siegel and White, 1975): 1) repeated physical exposure to the environment and 2) using external (visual) tools. When people are repeatedly exposed to a physical environment, they perceive, gather and store information about the places (locations) it consists of, and the routes between them.

With time, when enough information about individual locations and routes has been gathered, a person can develop a mental picture that represents that physical environment (Siegel and White, 1975). Later, tourists can refer to and use this accummulated information, often referred to as “internal” information search (Fodness and Murray, 1997). The problem is that tourists (especially first-time visitors) often do not have the time or resources to walk physically an unfamiliar environment. In addition, even if tourists have been exposed previously to a physical environment, travelling is an activity which is often intermittent and dispersed in time and this is why the obtained knowledge might no longer be available. However, such limitations can partly be overcome through the development and use of external tools that can enhance our cognitive abilities (Tversky, 2005).

Visual representations of information, such as text, diagrams, maps, web pages, graphics, instructions, and technical illustrations have many uses, because they allow us to learn, think and reason about places and times that are outside our immediate experience (Longley et al., 2010). Such representations are extremely powerful, because they make use of a large part of the brain devoted to visual sense, visual pattern finding and interpretation. If presented effectively, various spatial representations of information facilitate users to identify and localize objects; retrieve information regarding sizes, distances, directions, spatial relationships and patters (Kraak and Ormeling 2003). Therefore, such external tools can be used to optimize routes and mobility patterns within a destination.

Paper-based guidebooks, brochures, signposts and tourism maps are some of the most popular visual external tools that help tourists acquire knowledge about unfamiliar environments. Paper maps have been the dominant communication medium of geospatial information for centuries (Wood, 2003). Maps are especially useful to tourists, as they are able to capture and represent a large amount of spatial information about a specific area of interest within a single picture (Zipf, 2002). Guidebooks, on the other hand, provide quick access to (categorically or alphabetically ordered) historical, architectural and other thematic tourism-related information. Paper-based guidebooks and maps are extremely useful for tourists due to their high mobility. Their preparation and printing, however, take significant amount of time and often they might be out of date when finally available to the public. More importantly, maps and guidebooks are static representations of reality and, once printed, cannot be changed to satisfy the specific contextual needs of the user.

Digital computing vastly improved this situation, as electronic guidebooks and maps became available on desktop personal computers. A lot of effort was placed on developing computer-aided techniques for storing, analysing, processing and representing geospatial data, eventually resulting in the development of Geographic Information Systems (Longley et al., 2010). Additional advances in the late 1980s allowed the combination of maps with different visual and audio media (text, speech, images, animation). More importantly, the introduction of advanced graphics and the ability to interact with the map meant that users were now able to produce their own maps (Jones, 2013). Most mapping platforms and representations were still available only on desktop computers and, therefore, used mainly during the pre-travel information acquisition stage.

Increasingly more powerful handheld computers, emerging in the 1990s, combined the benefits of mobile and lightweight paper maps and guidebooks, coupled with the power of desktop multimedia cartography and Geographic Information Systems. Modern smartphone devices have more computing power than a 1980s PC, a high-resolution LCD screen and video cameras. No other single device combines the functionality of a standard Web browser, a games console, MP3 player, flashlight, TV or even a musical instrument. The smartphone is also the first truly personal, always-on, always-carried mass communication and information delivery medium with built-in payment capabilities (Fling, 2009).

Traditionally, both mobile and wearable computers, such as the wristwatch, video camera, GPS, and more recently the laptop, have had a fundamental role in supplementing the on-site experiences of tourists (Pearce, 2011). However, the smartphone is the only lightweight and affordable technology that combines the functionality of all of these devices (Oertel et al., 2002; Pearce, 2011). This is why tourism has been identified as the application area that can benefit the most from Mobile Information Systems (MobISs) delivered on smartphone devices (Umlauft et al., 2003).

### 2.2.3. Success Factors for Smartphone Information Systems for Tourism

As early as 1996, Long et al. (1996) envisioned mobile ISs that deliver more personal and relevant information to tourists. When the first mobile ISs first appeared on the market, their growing popularity was deemed as the long awaited "killer application" for tourism (Hamai, 2001). Despite such promise, a number of challenges and issues restricted their use. With time, it became apparent that information delivery through mobile ISs is more challenging, as it has to accommodate different hardware and form factors, a dynamic context, and special user behaviour and interaction (Dey, 2001; Gorlenko and Merrik, 2003; Krogstie et al., 2003; Looije et al., 2007; Matthews et al., 2009). This section discusses the key challenges or factors that need to be considered in order to make mobile ISs truly successful when it comes to tourists.

#### *2.2.3.1. Hardware and network constraints*

The first factor that has to be considered is the technology itself. As opposed to desktop computers, smartphones have a smaller screen, where only a limited amount of information can be presented at any one time (Krogstie et al., 2003). Therefore, there is a critical need to ensure that the delivered information is pertinent and useful for the tourist. Further, the limited colour ranges and resolution can impact the presentation of visual materials and the use of the mobile IS.

The smaller processing capabilities and limited battery life also pose constraints for the amount of operations that could be carried out on the device. Patchy network connectivity, or higher roaming charges, could also prevent tourists from accessing information on smartphones and degrade the overall experience with a mobile IS.

#### *2.2.3.2. Dynamic context of use*

Aside from technological limitations, many of which will undoubtedly improve over time, there are a number of issues related to the context in which mobile devices are used. Desktop computers are normally used indoors, in a stable and (often) predictable environment. In contrast, due to their high portability, the consumption of information on smartphone devices can unravel in many different circumstances that change constantly. This dynamic context of use, such as unpredictable and changing weather conditions, lightning level and noise, the presence of other people or devices, can influence how information is processed and used on the mobile device. The changing nature of such factors requires reconsidering traditional design strategies and methods, employed for desktop computers.

#### *2.2.3.3. Complexity of urban environments*

A number of typologies of tourism destinations exist. One fundamental classification, however, is natural (e.g. coastal, national park) and man-made (e.g. a city) destinations.

Man-made destinations can be further divided into rural and urban sub-types (Fletcher et al., 2013). Urban environments have been amongst the most significant tourist destinations (Edwards et al. 2008), attracting billions of visitors each year. One of the key activities urban tourists engage in is sightseeing, or visiting the various attractions within the urban destination. Acquiring information about such attractions (points of interest) heightens appreciation and engagement (Gursoy and McLeary, 2004) and is a key part of the overall experience with the destination. Apart from the increased geographical mobility of the population, exposure to unfamiliar urban areas is also becoming commonplace due to the growing rate of urbanisation processes. As a result, most people will need to regularly acquire new (geo)spatial knowledge and use it on-site, during their travel.

The world comprises of many physical objects, "revealing more detail the closer one looks, almost ad infinitum" (Longley et al. 2010, p.77). Densely built-up urban destinations are extremely complex, cluttered with many potential objects of interest that tourists might require information about. Apart from tangible points of interest, urban information search might be directed at intangible entities, connected with finding out what is special about places (attribute data) or whether something important is happening at the moment, or how physical entities have changed over time (temporal data). Presenting all of these on the small screen of the smartphone remains a challenge.

#### *2.2.3.4. Large amount of potentially relevant information*

As millions of networked sensors are embedded in physical devices to capture and stream data constantly, the information available about the physical world is increasing exponentially (Kitchin, 2014). Storing, processing, aggregating, analysing and using such large pools of data poses a number of challenges. A large amount of such information is related to tourism. This means that, in any densely built up environment, each object can be the source of substantial amount of information. Considering the small screen of the smartphone, however, browsing through such large amount of information might take considerable amount of effort and time, often not available to mobile users. This is especially the case with tourists, who are often time pressured, as they have only a limited amount of time at any destination.

#### *2.2.3.5. Limited interaction, attention and cognitive resources*

Another device-related limitation concerns the input capabilities of smartphones. The smaller size of mobile devices prevents the use of traditional desktop input methods, such as a keyboard or mouse. More imprecise input techniques are required, such as the use of a stylus, voice or thumb-based interaction (Krogstie et al., 2003).

**Table 2.1. Features of traditional desktop ISs and mobile ISs**

| Feature | Desktop PC | Mobile device |
|---|---|---|
| Task hierarchy | The main (cognitive) attention is on the computer task(s). | The main (cognitive) attention is on the primary task, not on the computer task(s). |
| Visual attention | The user can afford to direct her whole visual attention towards the interface/screen without interruptions or distractions. | The main visual attention is directed towards the real world, not towards the program. |
| Hand manipulation | The user can afford manipulation of the computer's input devices (e.g. keyboard, mouse, etc.) with both hands. | Manipulation of the device may be limited. |
| Mobility | The user is stationary, most of the time sitting in a comfortable position. | The user may be required to be highly mobile while operating with the IS. |

*After: Gorlenko and Merrick, 2003; Krogstie et al., 2003; Looije et al., 2007; Fling, 2009; Matthews et al., 2009*

Compared with desktop systems (Table 2.1), effective information delivery through mobile devices is challenging because it has to accommodate different forms of interaction, as well as a more varied and dynamically changing context of use (Dey, 2001; Gorlenko and Merrik, 2003; Krogstie et al., 2003; Looije et al., 2007; Matthews et al., 2009). Recently, various aspects of human interaction with the mobile device came under scrutiny (Table 2.1). In contrast to desktop computers, where attention is focused on the computer screen and the task at hand, mobile users are easily distracted by the environment where interaction takes place. Visual attention is dedicated to events that unravel in the real world, rather than the screen of the mobile computer. Human cognition is a very limited resource that is easily overloaded with information (Simon 1955). If not presented properly, the wide availability of content might hinder, rather than enhance decision-making.

#### *2.2.3.6. Key implications for mobile design and development*

Considering the challenges above, it becomes obvious why delivering information to urban tourists through smartphone information systems might be more challenging than expected. In order to be truly useful and usable, a mobile information system has to be responsive to the current information needs of the user, their goals, and the environment where interaction takes place. Ideally, the system will use such information to respond by adapting the content and functionality presented to the user. This property of mobile ISs, also called *context awareness,* is fundamental for improving the output from mobile interaction, but requires addressing a number of challenges, discussed later in this chapter (Section 2.6). One significant advantage of modern smartphone devices, especially when it comes to delivering relevant information to tourists, is the ability to provide location-specific information due to the availability of on-board sophisticated positioning sensors. The development of *mobile location-based services* was the first step towards more attentive and adaptive mobile information systems.

## 2.3. Mobile Location-Based Services for Tourism

### 2.3.1. Definitions and Enabling technologies

The term Mobile Location-Based Service (mLBS) appeared first in literature in the late 1990s (Raper et al., 2007). It was used to differentiate among information systems that use geospatial (positioning) information as filters for data query and presentation. Since then, this special type of information systems have been described through a number of definitions. Virranteus et al. (2001) defined mLBS as "services accessible with mobile devices through the mobile network and utilizing the ability to make use of the location of the terminals" (Virranteus et al., 2001, p. 66). Koeppel (2000) states that an mLBS is "any service or application that extends spatial information processing, or GIS capabilities, to end users via the Internet and/or wireless networks" (Keoppel, 2000, p. 2).

As illustrated by the definitions, most of the literature defines mLBSs as services that require *wireless* connection on *mobile* devices. It is important, however, to emphasize that Location-Based services can be accessed on desktop and laptop computers. MLBS are not specific to the smartphone and can be implemented on different types of mobile devices that can acquire and process positioning data, including tablets, multimedia phones, and smart watches. More recently, a number of efforts have been made to deliver mLBS functionality offline, or without wireless connectivity. This is why, it seems that the definition of Raper et al. (2007) captures the essence of mLBSs, without being too restrictive. The authors define mLBS as a special class of information systems that "deliver information depending on the location of the device and user" (Raper et al. 2007, p.5). Keeping this in mind, the scope of the current research has itself been limited to wireless Location-Based Services implemented on smartphone devices, which from hereon will be referred to as mobile Location-Based Services.

There are a number of technologies that had to mature in order to make location-based information delivery possible on smartphones. As a special type of location-based service, many of these technologies are essential for the implementation of AR browsers. This section provides an overview of the enabling technologies and technical advance related to mLBSs in general, while Section 2.4 discusses the special technical requirements relevant for AR browsers in particular.

#### *2.3.1.1. Processing unit and display*

A central processing unit (CPU) is necessary to carry out the arithmetical, logical and input/out operations with virtual content. This is the hardware component that changes most regularly and quickly within the industry. In 2014, high-end smartphone CPUs had powerful capabilities, and could run at speeds from 1.4 GHZ (e.g. iPhone 6) up to 2.8 GHZ (e.g. Snapdragon 810).

While the CPU is engaged with processing data, the smartphone display is where these data are presented visually to users. There are two key variables that influence the presentation of information: display size and display resolution. The characteristics of smartphone displays vary widely among manufacturers and even among models produced by the same company. In 2014, most smartphone display sizes ranged from 4 to 5.2 inches. Display resolution varied between 313 ppi (e.g. Moto X) to 432 ppi (e.g. Samsung Galaxy S5).

#### *2.3.1.2. Positioning of the mobile device*

Determining the location of the device that the user is carrying is essential for the operation of any mLBS. The approaches to determine the position of a mobile device can be generally divided into outdoor and indoor positioning methods. Outdoor positioning methods can further be divided into network-based (passive), handset-based (active) or hybrid.

Network-based positioning methods (WLAN, Cell-ID) use the transmitter base stations of mobile telecommunication networks. Locating a mobile device is achieved by measuring the signal travelling to and from a set of base stations. The signal measurements allow computing the direction and/or length of the individual radio path. The position of the mobile device is then determined using computational geometry (Brimicombe and Li, 2010). These methods require connection to server-side services and can work indoors if there is sufficient signal strength.

With device- or handset-based methods the mobile device determines its position based on signals it receives (Brimicombe and Li, 2010). The most widely popular service that operates on this principle is the American Global Positioning System (GPS), even though other systems are also available or under development, including Galileo (European Union), GLONASS (Russia), and Beidou (China). The basic principle behind GPS is trilateration based on distance measurements using satellites as reference points. This type of positioning does not require a network connection. A major problem with this approach, however, relates to overheads (the volume of data exchanged between a client and server) that affect the consumption of power and substantially decrease the smartphone's battery life.

The accuracy and consistency of the positioning data acquired through GPS is heavily dependent on a number of factors. Delay of signals due to atmospheric interference, multipath propagation (when the signal interacts with objects such as buildings or water bodies), multiple reflections and diffractions all cause inaccuracies and errors in the resulting data. The current accuracy of GPS sensors within smartphones varies significantly depending on the environment, but in open areas can reach 3-4 meters (Shaner, 2013). Considering movement of the user, the challenge for mLBS is to acquire the position of the device with sufficient accuracy over time. The purpose and type of LBS determine the level of accuracy and consistency that is acceptable and as the next section (2.4) discusses, Augmented Reality browsers require a substantially high level of both accuracy and consistency of positioning data.

##### *2.3.1.3. Storage and processing of geographic (geo-tagged) content*

Geographic Information Systems (GIS) are at the heart of mLBSs. GIS are special types of information systems that deal with the acquisition, integration, management, processing and visualisation of spatial data sets (Longley et al., 2010). GIS are often too heavy and currently their implementation on mobile devices is limited. However, Internet GIS (also Web-GIS and online GIS) addresses such difficulties as it makes GIS functionalities available to remote users via the Internet (Longley et al., 2010). Two main approaches are used: client-side and server-side. Internet GIS based on server-side means that the GIS software resides on the server and carries out the data processing and analysis. Client-side GIS allows users to download GIS functions and data and carry out data processing and analysis locally on the mobile device.

In the past, the feasibility of location-based services (and AR browsers) was mainly connected with the lack and availability of geo-tagged data (Langlotz et al., 2014). *Geotagging* is the process of assigning geospatial context information (geographical coordinates) to information resources (Yap et al., 2012). Today, geo-tagged content is also user-generated and is freely available to developers through popular repositories, such as GeoNames, CityGrid, Yelp, Zvents, Hoovers, Yahoo, and Trulia (Madden, 2011). The amount of geo-tagged data is increasing exponentially every day and is nearing our limit to process, store, transfer and deliver it to users in a way that is easy to understand and use.

#### *2.3.1.4. Data transfer and Network connectivity*

A range of standards and systems for wireless telecommunication allow fast and effective transfer of data to mobile devices. In the last decade, the capabilities of smartphones to receive and transfer data increased substantially with the implementation of third generation (3G) and fourth generation (4G) networks. Such networks offer higher data transmission rates, supporting fast mobile Internet, multimedia and video-conferencing applications (Brimicombe and Li, 2010). Fourth generation networks are expected to provide higher transmission speeds (100 Mbps to 1Gbps), larger capacity and high security. In parallel, developments such as the convergence of Wi-Fi and mobile networks (Hac, 2014), as well as plans for removing data roaming charges within specific geographic areas, will lead to the vision of using mobile phones at any time and anywhere.

### 2.3.2. Mobile Location Based Services and Tourism

Due to the huge potential of mLBS for tourism, the largest group of such services have been developed as mobile guides (Emmanoullidis et al., 2013). A mobile guide is "a portable, location-sensitive and information-rich digital guide to the user's surroundings" (Raper et al., 2011, p. 90). Considerable amount of research addresses the technical aspects and implementation of mobile tour guides (Kenteris et al., 2006;

Emmanoullidis et al., 2013). Among others, some of the most popular prototypical mobile guides include GUIDE (Cheverst et al., 1998), Hippie (Cheverst et al., 2001), and Lol@ (Umlauft et al., 2003). Location-based filtering of information is also available on widely popular commercial smartphone applications, such as Yelp (Yelp, 2014), Yell (Yell, 2014), and Foursquare (Foursquare, 2014), as well as the official mobile websites of tourism boards, such as Switzerland (MySwitzerland, 2014), or Estonia (VisitEstonia, 2014).

Smartphone location-based technologies have had a wide impact on the way people perceive and interact with physical space. On global scale, tourists are starting to consider travelling to destinations that they were not comfortable with before. On more local scales, the use of mLBSs engages tourists and results in longer distances travelled physically on-site (Michael and Michael, 2011). Koeppel (2000) recognized four primary functions of mLBSs for the mobile traveller:

- Localization of current position in space, persons, objects and places,
- Routing between objects and places,
- Search within a set proximity for objects and places,
- Information about travelling conditions, such as traffic-related data.

Mobile guides are often seen as the digital replacements of paper-based guidebooks and maps (Raper et al., 2011). The many differences between paper-based and interactive portable devices, however, has triggered a debate as to the type of functions and content, as well as the representation metaphors, that should be transferred to smartphone mobile guides (Raper et al., 2011). While a useful comparison, mobile location-based delivery of information has no analogue counterpart and the design and delivery of information through such services requires a new set of design methods and principles.

### 2.3.3. Limitations of Mobile Location Based Services

The emergence of location-based services created a lot of excitement in both academia and industry. Considered also the “killer app”, mLBSs were expected to “put the user in the center of rich and interactive world of spatial information…and facilitate new interaction techniques that enable the user to directly access and manipulate spatially-related information and services” (Fröhlich et al., 2008, p.251). General-purpose interfaces that disappear in the background, while being attentive to the user and allowing access to information at any time and place (also referred to as ubiquitous computing) have long been described in literature (Bush, 1945; Weiser, 1991). In 1999, Spohrer (1999) envisioned a global infrastructure that allowed the combination of virtual and physical worlds. Later, Fitzmaurice proposed “a world where electronic information will ultimately be everywhere” (Fitzmaurice, 1993, p.49).

While mLBS have taken us one step closer to achieving this vision, a number of limitations remain. Current smartphone devices limit access to information only within the scope of the smartphone screen. More importantly, virtual and physical worlds are separated spatially and users need to mentally integrate them in order to make sense of the delivered content. Most location-based service interfaces rely on the assumption that, as long as information is relevant to the current location, this integration would be achieved automatically within the mind of the user. However, empirical studies have shown that co-relating physical and virtual spaces might require substantial cognitive effort (Oulasvirta et al., 2009; Church et al., 2010; Kässi et al., 2014). At the very least, users need to constantly shift their gaze from physical to virtual space, in order to process information. This might require a huge physical effort, especially when users have to track changes in both physical and virtual worlds, which is the case, for example, during navigation and wayfinding. The separation of virtual and physical worlds has led a number of researchers to question the “any time, any place” tenet when it comes to ubiquitous delivery of information. From the many available technologies, Augmented Reality comes closest to achieving the vision of truly ubiquitous and pervasive computing.

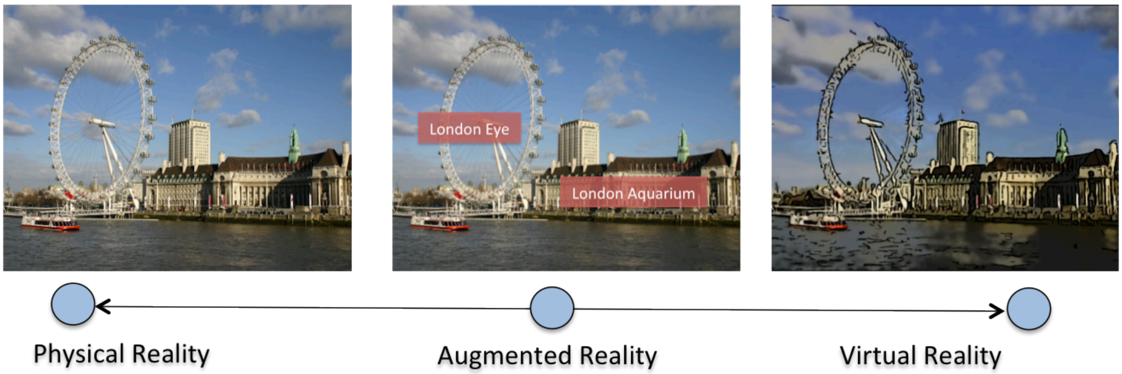
## 2.4. Smartphone Augmented Reality Browsers

### 2.4.1. The essence of AR

An Augmented Reality (AR) system enhances or augments the surroundings of its user in real-time with virtual (computer generated) information that seems to co-exists with the real world (Azuma et al. 2001). Augmentation of the physical environment can relate to any human sense (Höllerer and Feiner, 2004). However, visual representations of the environment play a pivotal role in supporting the activities of mobile users and this is why in this study the main focus will be on *visual* augmentation.

The evolution and development of visual AR is closely related to that of *virtual reality* (VR) technologies (Milgram et al. 1994). However, there is a difference between the two. As opposed to the completely synthetic virtual world of VR (Figure 2.1), in AR systems "a virtual world supplements the real world with additional information" (Feiner et al. 1997, p. 74). According to Milgram et al (1994), there are different ways that virtual information and the real world can be merged together, collectively termed *Mixed Reality*. The result can be placed along the Reality-Virtuality continuum (Figure 2.1). When virtual information is added to the real world, the result is *Augmented Reality*, whereas when real objects are added to virtual environments, the result is *Augmented Virtuality* (AV).

**Figure 2.1. The Reality-Virtuality continuum**




*Source: Milgram et al., 1994*

As opposed to VR and AV, AR has the potential to enhance the perception of reality in real-time attaching information to a specific place, because an AR system (Azuma et al., 2001):

- Combines real and virtual objects in a real environment.
- Runs interactively and in real (current) time.
- Registers and aligns real and virtual objects in three-dimensional space.

This widely accepted definition emphasises the fact that systems can only be considered Augmented Reality if they deliver information that is aligned with the actual physical environment of the user. Systems that overlay virtual information on top of pre-recorded videos (e.g. in sports) are often termed "pseudo" Augmented Reality (Langlotz et al., 2013).

### 2.4.1. Augmented Reality Browsers

Development of AR takes considerable time and resources. It is not a surprise then that in the past most AR systems were developed in domains that involve high risks, such as military operations (Julier and Rosenblum, 2000) and medicine (Fuchs et al., 1998). More recently, three-dimensional models and animations in smartphone AR has been used for marketing purposes (e.g. augmenting the pages of a magazine with video or 3D models), for games (e.g. virtual characters that appear on the street) or home shopping (e.g. placing virtual furniture in the living room) (Langlotz et al., 2014). The content delivered through such AR applications is often pre-prepared and developed for a specific type of context and use situation, making it less valid or irrelevant outside of the settings for which it was intended.

**Figure 2.2. Physical set up of the first mobile AR browser, The Touring Machine**

*Source: Feiner et al., 1997*

As discussed earlier, conceptual work has addressed the need and benefit of AR to provide highly relevant virtual information at any place and time (Fitzmaurice, 1993; Spohrer, 1999). AR that delivers general information about the environment has been at the core of concepts such as augmented memory (Spohrer, 1999) and augmented city (Matsuda, 2010). When applied to outdoor settings, such interfaces can be used within a number of application areas, ranging from gaming to tourism. Due to lack of infrastructure and technical limitations, this vision was not possible until the late 1990s when Feiner et al. (1997) developed the first mobile general-purpose AR interface, called the Touring Machine (Figure 2.2). The system delivered information about the surroundings of the user, but required expensive, bulky, obtrusive and heavy hardware. Since these early steps, AR has undergone enormous development and today the smartphone combines all of the necessary technology to augment the environment of the user with general-purpose content (van Krevelen and Poelman, 2010).

**Figure 2.3. Modern smartphones and other wearable devices combine all necessary technologies to augment the environment**

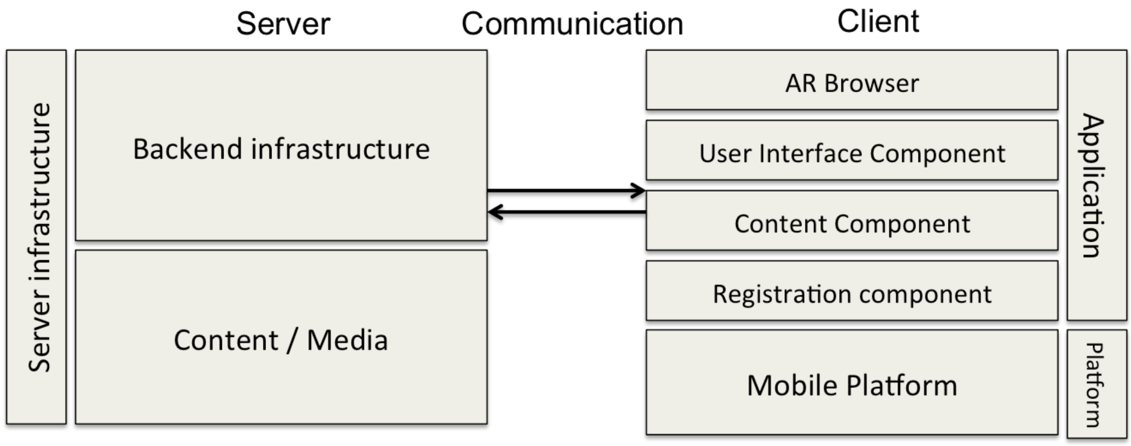

*Source: Wikitude, 2014*

The concept for general-purpose AR interfaces was further developed by Kooper and MacIntyre (2003) who implemented an AR system that mimics the functionality of a web browser, called the Real Wide Web Browser. More recently, such AR interfaces became popular as *Augmented Reality browsers* (Figure 2.3). Before being adopted in industry and academia, the term "AR browser" was first used by SPRXmobile when they presented the idea behind the AR application *Layar* (SPRXmobile, 2009). After the launch of Layar, AR browsers became one of the most popular commercial location-based smartphone applications (Langlotz et al. 2013).

### 2.4.2. Sate-of-the-art enabling technologies

Similarly to other types of location-based services (Section 2.3), smartphone AR browsers require a fast and powerful processing unit, high-resolution display, data storage and data transmission. There are, however, several differences in technical requirements, mainly pertaining to tracking the position of the device and registration of both physical world and virtual content. Figure 2.4 illustrates the common software architecture of AR browsers described by Langlotz et al. (2014). The *registration* component acquires the necessary positioning data. Once these data are captured, the *content* component initiates streaming of relevant data from the server, while the *user interface* component is responsible for the presentation of content on the screen of the smartphone device.

**Figure 2.4. Overview of the common architecture of an AR browser and its software components**




*Source: Langlotz et al., 2014*

It is important to examine the choices and state-of-the-art in each of these categories, as they would ultimately influence the user experience and usability of AR browsers.

#### *2.4.2.1. Processing unit and display*

This is a computational platform, necessary to process and generate the virtual content (e.g. images and text), process the tracking information and control the AR display. In 1997, AR systems required fast processing available only on mobile laptops (Feiner et al., 19997). However, modern smartphones have enough processing power to perform in the same way as those early mobile computers.

A display is also needed where physical and virtual objects are merged together and presented to the user. This is probably the most important part of any AR system and most development within the field has been directed at improvement of the form factor of both mobile and fixed AR displays. It is important to note that *mobile* AR displays can refer to three main types of form factors: *retinal* (lenses), *head-mounted* (HMDs), and *handheld* displays. Traditionally, head-mounted displays (HMDs) have received the most attention in research literature (van Krevelen and Poelman, 2010; Hua, 2014). With projects, such as Google Glass, HMDs still attract a lot of attention today. However, due to various technical challenges, most displays are still not feasible for the wide public and the smartphone remains the most popular augmentation device at present. Handheld displays include devices, such as Personal Digital Assistants (PDAs), tablets, feature and multimedia phones, and smartphones. In this study, *mobile Augmented Reality* refers to Augmented Reality implemented on and for *smartphone devices*.

#### *2.4.2.2. Data and Standards*

Similar to typical mLBSs, an AR browser requires access to geo-tagged content (Section 2.3). The storage format for AR is usually proprietary XML-based databases. Early AR browsers used the Keyhole Markup Language (KML), originally developed for 3D geo-browsers such as Google Earth. Later, however, it became clear that there is a need for an XML language that is specifically developed to address the needs of AR, making this an active research area. A number of other standards are currently being developed, including ARML (Augmented Reality Markup Language) and KARML (Keyhole Markup Augmented Reality Language) (Lechner, 2013). Work on development of AR standards is currently underway and is primarily connected with several ISO sub-commissions and the AR Standards community. Tourism and tourism-related use of AR is one of the use cases that the commission is currently working on.

#### *2.4.2.3. Tracking and Registration*

Just like other mLBSs, AR browsers require determining the position of the mobile device that the user is carrying (Section 2.3), together with the orientation of the user and the approximate height of the device. In addition, the device has to determine the position of the object that needs to be augmented in order to align virtual content precisely where it has to be. This process is generally referred to as *registration* and *alignment* (Langlotz et al., 2014). Once completed, the system also needs to track changes in the viewpoint of the user in order to keep seamless alignment between physical and virtual worlds, a process referred to as *tracking*. These processes are usually carried out by the *registration* component of an AR browser (Figure 2.4).

Broadly, the approaches for registration and tracking can be divided in two categories: *marker-based* and *marker-less* (Henrisson and Ollila, 2004). *Marker-based tracking* requires placing physical markers (e.g. QR codes, fiducial markers) in the environment that can be recognized by the system. The virtual content is then overlaid on top of these markers (Möhring et al., 2004). While extremely suitable for indoor scenarios, such an approach is less feasible for outdoor use of AR, since it: (1) requires instrumenting the whole world with physical markers; (2) is limited to the visibility of the markers; and (3) requires that the physical markers are scaled by distance in order to be recognized by the system (Henrisson and Ollila, 2004).

*Marker-less tracking* is considered more appropriate for settings in which users roam around in outdoor unprepared environments. There are three main approaches to marker-less augmentation: GPS-based, computer vision-based and hybrid. *Computer vision* algorithms work similarly to marker-based tracking, but they are able to recognize natural objects and features within the surrounding of the user (Henrisson and Ollila, 2004). They are, however, extremely resource-intensive. This is why most current smartphone AR browsers rely on GPS-based tracking. In essence, the data from the geomagnetic sensors (GPS, accelerometer, magnetometer) on board the mobile device is combined with the incoming data from the camera view in order to estimate the orientation and field of view of the user (Madden, 2011). Based on those parameters, data is extracted from a central database that contains geo-tagged (location-based) content and overlaid on top of the incoming live video feed.

#### *2.4.2.4. Output and Data Presentation*

In terms of output image, two approaches can be used: optical see-through and video see-through (Bimber and Raskar, 2005). In *optical see-through*, a virtual overlay is super-imposed over the real-world through the use of half-silvered mirrors. In *video see-through*, the virtual overlay is super-imposed on a real-time live video feed, acquired by a camera. Video see-through is the most widely used and implemented approach in current smartphone devices (Madden, 2011; Langlotz et al., 2014). The rear-facing camera of the smartphone is used to continuously capture and display the surroundings and simulate an experience similar to that if the device was transparent. Among many advantages, one benefit is that the incoming video (perceived as the view of the real-world) can be manipulated quite extensively.

#### *2.4.2.5. Input and Interaction*

Input technologies are used to enable the user to interact with virtual content. Traditionally, mobile AR systems made use of a mouse, or a stylus (e.g. Feiner et al. 1997). AR on mobile devices makes use of a stylus, touch- or voice-based interaction. In all cases, the user has to hold the device upright with an extended arm for prolonged periods of time, which can be very awkward and tiresome (Tokusho and Feiner, 2009). Holding the smartphone device with an extended arm also introduces the problem of tremors, as users usually have difficulties holding their hand still. As a result, perception and understanding might be challenged due to blurring effects.

#### *2.4.2.6. Data transfer and Network connectivity*

Network connectivity and protocols are necessary for acquiring virtual content. These do not differ from the standard mobile network protocols, described briefly in Section 2.3 and in more detail by Hac (2014).

### 2.4.3. AR content and AR annotations

AR browsers provide access to large amount of location-based data and deliver information about real physical objects through spatially registered virtual balloons (labels), called *AR annotations* (Wither et al., 2009). Having in mind that location-based content increases exponentially every day, it is not a suprise that AR annotations make up a great portion of all AR content (Wither et al., 2009). The use of the term "annotation", however, is not new and they have been used for centuries to personalise written text. Historically annotations have been known as *glosses*, or brief marginal notations in a text. Digital annotation methods attracted much attention since 1945 with the development of the Memex (Bush, 1945). Annotations are used in electronic publishing, on-line multimedia and learning systems, word-processing software and digital map production. Annotations aid learning and processing of information because they (Osviannikov et al., 1999): 1) are more easily accessible than a dictionary / encyclopedia; 2) direct the attention to specific words; 3) connect word forms to meanings with minimum interruption of the reading process; 4) force learners to read back and forth, triggering more lexical processing and retention; 5) can give multiple perspective on the same word/document.

While essential for AR, there is still no clear definition for AR annotations within the domain and the term is often used to refer to any type of AR content. Wither et al. (2009, p.680) define AR annotations as: "virtual information that describes in some way, and is registered to, an existing object". Maass and Döllner (2006, p.1) pose that annotations "represent textual or symbolic descriptions and provide explanatory or thematic information associated with spatial position". According to Wither et al. (2009) virtual information classifies as an AR annotation if it satisfies two criteria: 1) has a spatially dependent component (every annotation must be registered to a particular object) and 2) has a spatially independent component (there must be difference between the virtual content and what the user sees of the real world). Henrysson and Ollila (2004) pose that AR is particularly useful when there is a "close spatial relationship" between the physical object and the information to be displayed. When this relationship is weak, then the authors suggest using 2D maps, while when there is no spatial relationship, web pages and/or audio could be used. This study deals with AR annotations that satisfy the requirements of Wither et al. (2009). Therefore, other elements from the AR interface (Figure 2.5), such as the radar (abstract representation of surrounding annotations) is not referred to here as AR annotations.

**Figure 2.5. Annotations and Elements of an AR browser interface**


AR browser annotations vary widely in design parameters, but can comprise of the following elements (Figure 2.5):

- **pin**, also called a head, signifier or mark (Fedosov and Misslinger, 2014) – often this is a virtual pinmark or an icon which is overlaid on top of the POIs. The pin is the virtual part of an annotation that signals that there is virtual content related to a specific physical object.
- **leader line**, also called a connector line or pointer (Götzelmann et al., 2007) – the leader connects the pin (if there is one) with the body of the annotation.
- **annotation body** or bubble (Madden, 2011) – this is where virtual information about the POI is delivered. Often, the bubble is the only part of an annotation presented on the screen of the smartphone.

While no widely accepted taxonomy for AR annotations exists, a distinction is often made based on the content that the annotation body (bubble) contains into: textual, images, video and three-dimensional annotations. Textual annotations are the most widely used among AR browsers (Langlotz et al., 2014). In reality, however, most AR annotations are hybrid and combine different types of text and media. This wide availability of different types of information, combined with the specific characteristics of AR browsers, provides a unique opportunity to deliver useful and usable content to tourists.

### 2.4.4. Location-Based Augmented Reality Browsers and Tourism

Tourism has been identified as one of the most prominent areas for application of Augmented Reality technology (Höllerer and Feiner, 2004; van Krevelen and Poelman, 2010). As a result, a number of prototypical AR systems have been developed that augment the experience of tourists within museums (e.g. Choudary et al., 2009; Chang et al., 2014), cultural heritage and historical sites (e.g. Vlahakis et al., 2001). A number of projects and studies have examined the use of AR for auto or pedestrian (Walther-Franks and Malaka, 2008; Rehrl et al., 2014) navigation.

Many commercially available AR applications claim to support tourism-related activities, including the three most popular (in 2014) AR browsers: Layar (Layar, 2014), Wikitude (Wikitude, 2014) and junaio (Junaio, 2014). Recently, stand-alone smartphone applications also added AR view to provide content to their users. These include popular mobile travel guides, such as eTips (eTips, 2014), as well as general-purpose information providers, such as Yelp (Yelp, 2014) and Yell (Yell, 2014). In parallel, a number of destinations (e.g. Dublin, Tuscany, London, Amsterdam, Paris) have advertised the availability of AR in proprietary apps as part of a more memorable tourism experience. One of the first commercially available AR applications that was commissioned by a Destination Management Organisation (DMO) was Tuscany+ (Visit Tuscany, 2010).

Within academia, several projects and research studies have resulted with development of AR browsers for tourists. Addressing augmentation in indoor environments, Choudary et al. (2009) used an image-matching algorithm to augment museum artifacts. More recently, Seo et al. (2011) developed an application that is able to re-create past historical life at cultural heritage sites. Outdoor augmentation with tourism-related content has also been described in research literature. Marimon et al. (2009) developed an AR browser able to deliver 2D (images) and 3D (models) annotations for points of interest in the city of San Sebastian (Spain). Luley et al. (2011) reported the planned development of an outdoor AR browser for augmentation of rural destinations. Kim and Park (2011) describe the development of a smartphone AR application which delivers information to tourists visiting the National Palace Museum in Korea. Keil et al. (2011) developed a smartphone AR app that delivers different types of content (pictures, drawings and blueprints) for a key tourist attraction in the Darmstadt's Mathildenhoehe, the Olbrich House. More recently, Pereira et al. (2014) reported the development of a smartphone AR browser that provides information about POIs within the botanical garden of the University of Coimbra. The browser presented information about various plants in the garden.

### 2.4.5. Benefits of AR browsers for tourism

The swift technological advance, miniaturization of sensors and hardware and novel ways to deliver information has led to the development of the concept of *urban computing* (McFedries, 2014), where the city itself becomes the interface to information. The tourist is the "moving cursor", while the smartphone is used to "tap" into the information that this interface provides (McFedries, 2014, p.28). Instead of surfing webpages, tourists can now "surf" urban objects and entities in a digital urban environment.

As discussed earlier (Section 2.2), visual displays can be regarded as a direct extension of the human sensory and cognitive capabilities. Such external tools offload cognition and can facilitate users in processing information, reasoning and, ultimately, decision making. The way information is delivered also influences behavioural patterns. In this context, it is important to consider the specific characteristics of smartphone AR and how it could influence the overall experience with a destination and benefit tourists in their decision-making process. There are several specific benefits that need to be considered:

***Lower cognitive effort to find content***: Mobile location-based services allow users to retrieve information about specific points of interest. However, physical and virtual content exist in two different (physically separated) spaces. In theory, AR browsers solve this problem by merging the two spaces and, therefore, demanding lower physical and cognitive effort to find information about POIs. Ideally, the user can lift the display towards a specific POI and immediately see content overlaid on top of that physical object.

***Utilise the power of spatial indexicallity***: A concept related to AR is that of indexical representations (Kjeldskov and Paay, 2010). Indexical interfaces make use of physical space to provide information “just-in-place” or relevant only to a specific physical location or entity. The physical object and the information captured in an AR annotation make sense only when presented together. The concept is similar to presenting users at a train station with the train timetable, which would only make sense at that specific location and time. The main design implication is that the use of spatial indexicality can reduce the requirement for the amount of information delivered to users.

***Lower physical effort and increased safety***: AR has a huge potential for situations where shifting one’s focus from the physical world is detrimental, such as during navigation (Kjeldskov, 2003). Imagine a user who is trying to cross the street while looking at the smartphone display in order to keep track of the directions. In such situations, using AR is beneficial as the user does not have to shift back and forth his gaze and attention to verify the information.

***Providing information about non-visible features:*** Augmented Reality interfaces allow users to see and experience virtual information from places that might not be visible directly (e.g. occluded by other physical entities) in a more realistic manner. In parallel, tourists are presented with information about past or future times on-site. Annotations that combine rich media allow experiencing different types of content *in situ*. The use of 3D models, animations, interactive panoramas and other images provide on-site destination experiences that were not possible before.

***Direct and focus visual attention:*** AR browsers can be used to manipulate visual attention towards specific physical objects. This is especially valuable when tourists are unfamiliar with a destination and do not know where to look. Focusing attention on specific physical entities would increase the feeling of discovery and insight.

### 2.4.6. Affordances and mediating the tourist experience

In the context of seamless information acquisition, a tourist can learn about a large, unknown environment by interacting with a smartphone AR interface. The AR interface is then said to mediate the experience with the environment (Cheng and Tsai, 2013). In other words, an AR interface can be emotive and trigger experiences which otherwise would not unravel. The extent to which the experience can be mediated depends on the suit of affordances that support information acquisition through AR.

Despite its use and popularity in a number of design disciplines, the term *affordances* is still not well understood. Carrying out an extensive literature review, McGenere and Ho (2000) prove that the concept has been used in many different contexts and with many different meanings.

The term was first coined by James Gibson in 1979 to mean what the natural environment offers an animal or humans, either for good or for ill (Gibson, 1979, p. 56). Affordances are relationships that exist between the environment and an actor, they are “actionable properties of the world” (Norman, 1999, p. 38). A central question, then, is whether and to what extent information exists within the environment in order for humans and animals to perceive affordances. Gibson (1979) and later Gaver (1991) argued that affordances exist irrespective of whether humans perceive them or care about them.

Affordances are a relationship, part of nature (Gibson, 1979). They do not have to be visible, known or desirable and some of them have yet to be discovered. From this point of view, the physical environment provides affordances to tourists who might or might not be aware of them. Large unfamiliar environments can offer a number of affordances to tourists, which are difficult to recognise. This is where the role of various information technologies becomes essential, as such tools can communicate information that makes physical affordances more visible and easy to recognise. In addition, AR has a special role in the sense that it can visualise directly and communicate affordances to users.

The term *affordance* became popular in the HCI community in the mid 1990s because of Donald Norman. In his book, The Design of Everyday Things, Norman (1988) wrote that understanding how to operate a device has three major dimensions: conceptual models, constraints and affordances. As opposed to Gibson, Norman’s definition of affordances concerns opportunities for action. In other words, humans use affordances in order to determine the possible uses of a physical or digital object.

In the first sense and meaning of Gibson, affordances are connected with designing the utility of an object. In this sense, affordances are objectives. For instance, a number of studies have explored the educational affordances of mobile devices. Smartphone AR can provide perceived affordances through dynamic representations (Roschelle et al., 2003; Roschelle et al., 2007), which enhance what and how visitors perceive an urban environment.  In this context, it is important to consider both the affordance of smartphone devices and AR interfaces. An extensive review by Orr (2010) emphasizes three main affordances: (1) mobile devices as a representation tool, (2) mobile devices as a communication tool, and (3) limited learning vs. no learning at all.

Despite its infancy, the educational affordances of smartphone AR have also been examined recently. Cheng and Tsai (2013) discussed differences in education affordances of both location-based AR and image-based AR. They found that image-based AR supports development of spatial abilities, practical skills and conceptual understanding. On the other hand, location-based AR supported inquiry-based activities. The paper concludes with the need to investigate more closely the experience with AR, especially when it comes to cognitive load and motivations. They suggest basing further research on theories for spatial cognition and situated cognition, both considered in this study (Chapter 5).

In the second sense, affordances are subjective and are connected with designing the usability of an object (McGenere and Ho, 2000). This means that without clear and intuitive perceive affordances, AR will remain a technological curiosity, as it will be difficult for users to understand what benefit AR brings and how they can use such interfaces.

On one hand, the virtual content has to be designed so that it provides important visual cues that allow users to understand how they can use this virtual content. On the other, the physical objects around the designer have their own perceived affordances (Billinghurst et al., 2005). Billinghurst et al. (2005) follow the recommendations of Norman for designing in a way that allows good perceived affordances. These include: (1) the importance of making affordances visible, (2) giving feedback, and (3) providing constraints.

### 2.4.7. Challenges and Gaps

#### *2.4.7.1. Tracking and Registration*

Tracking and registration are two of the most challenging processes to implement successfully in smartphone AR and constitute a research area on their own. While there are many methods to determine the current position of the mobile device, state-of-the-art software and hardware can deliver only limited accuracy when it comes to large outdoor environments (Langlotz et al., 2014). Sensor-based approaches are mainly limited due to the cumulative error from incoming GPS, accelerometer and gyroscope data. Additionally, computer vision algorithms are still very inaccurate and resource intensive when it comes to large and unprepared outdoor environments. AR requires hyper-sensitive sensors and the tolerance for positioning errors is very small (Turunen et al., 2010). The lack of accurate registration and tracking of the mobile device results in lack of seamless integration of virtual and physical spaces (Figure 1.3). This problem is especially exacerbated when it comes to urban environments, as it may confuse users and lead to wrong decisions.

Research within urban AR for tourism has been directed exclusively at solving this problem. For instance, Marimon et al. (2009) made use of sensor-based and natural feature detection algorithms to develop a smartphone AR application for tourists as part of the MobiAR project. The main use cases concerned tourists that are willing to explore their surroundings, find interesting POIs and relevant information about them. The primary focus of the project was technical and aimed at proving the feasibility of combining both GPS-based and computer vision algorithms for registration and tracking. Later, delivery of social media content through smartphone AR in urban environments was explored by Turunen et al. (2010). The team investigated the feasibility of delivering social media (user-generated) content through AR browsers. An innovative feature of their experimental AR browser was the ability to overlay virtual annotations over moving targets, such as people. Their research emphasized the need for higher spatial accuracy when it comes to merging virtual content with moving targets.

The use of smartphone AR information systems has also been transferred to different types of destinations where tracking and registration poses additional challenges. The main aim of the project MARFT (Mobile Augmented Reality for Tourism) was to demonstrate the use of AR technology in rural tourism areas (Luley et al. 2011). While still conceptual, it is expected that the AR browser will provide tourists with cartographically correct AR annotations, overlaid on top of rural scenes. Using computer vision algorithms, Keil et al. (2011) developed an AR app that delivers original material to users for a key tourist attraction at the Darmstadt’s Mathildenhoehe

(Germany), the Olbrich House. The app overlays pictures, drawings and blueprints on top of the original house, which are interactive and could be used to access further content.

The studies described in this section placed special focus on improving tracking and registration, however, often concluding with imprecise and suboptimal results. In practice, it is questionable whether AR browsers can currently achieve absolute and immaculate tracking and registration (Livingston, 2013).

#### *2.4.7.2. Delivering useful content for AR browsers*

The amount of user-generated content and geotagged media increases exponentially every day, however, the density of available information is spatially unequal. Popular urban centres are cluttered with virtual annotations, while sub-urban or rural areas may lack interesting content (Langlotz et al., 2014). An additional concern is the availability of different types of data. Currently, AR databases rely heavily on textual content (Langlotz et al., 2014). While multimedia, such as videos, images, animations and 3D models have a significant potential to enhance tourists experiences, there is still lack of such content that is widely available to use within AR browsers. Most of the time, DMOs or other companies and organisations have to produce the content that should be available in an AR browser. This might require significant amount of time and resources.

#### *2.4.7.3. Delivering usable content through AR browsers*

An important additional challenge is the suitable presentation of content once it is available. Despite the huge availability of AR annotations in built-up urban destinations, several empirical studies have indicated that information delivered through AR browsers is difficult to understand and use. For instance, the main aim of the study by Olsson and Salo (2011) was to investigate whether expectations of early adopters of AR browsers are satisfied. Results showed that users of smartphone AR consider the content that is delivered inappropriate, irrelevant and excessive. Technical problems were also pointed out and included positional inaccuracies, software instability and bugs, and limited functionality. The results from the study emphasize the need to investigate what is the most suitable content for users and the way it should be delivered visually in different contexts. Selection of useful and relevant content, as well as suitable visual forms of representation for mobile AR require systems that are aware of the context in which they are used.

## 2.5. Towards Context-Aware Smartphone AR

Unlike desktop computers where traditional ISs operated, mobile devices are used in a variety of settings. Because users do not work with computers in isolation, a number of *internal* (e.g. goals, tasks, preferences) and *external* (e.g. lightning, noise, people) *situational factors* influence their work (Hackos and Redish, 1998; Bellotti and Edwards, 2001; Dourish, 2004). These factors have different names in literature, but are often referred to with the collective term *context* or *context of use* (Schilit and Theimer, 1994; Dey, 2001; Greenberg, 2001; Dourish, 2004).

Context has an important role in smartphone ISs design for two reasons. First, it is fundamentally important for a smartphone IS to fit within the context in which it is used (Hackos and Redish, 1998). Mobile ISs that do not fit in the context in which they are used may be cumbersome, annoying and difficult to use (Bellotti and Edwards, 2001; Dourish, 2004). Second, a change in context often influences the relevance and suitability of the information on the mobile screen (Fling, 2009). With mobile ISs, the change of context can be swift, sudden and dynamic, even throughout a single use session (e.g. a user coming out of a room onto the street).

The use of location as a contextual parameter to filter information and the development of location-based services provide an excellent example for these principles. Let us consider a user who is standing on Westminster Bridge. A mobile map which shows the current area, or an AR browser which displays information about the London Eye provide information about entities in the current location. They are both said to "fit within the context" to a certain extent. Both would be unsuitable if the user was located in front of the British Museum and was trying to obtain information about the opening times. The information has to change dynamically with changes of the location and orientation of the user. Indeed, tourists expect to have access to increasingly more intelligent services that are aware of much more than their location and adaptive to the current situation, their needs and requirements. Reviewing mobile tour guides, Baus et al. (2005, p.210) concluded that "in the future, mobile tour guides will have to take into account more and more situational factors in order to provide their users with a user-friendly experience".

A mobile IS that detects the change of context and triggers a consequent change in its behavior, content, or interface is called *context-aware* (CA). The process is often referred to as *context-based adaptation* (Schilit and Theimer, 1994; Dey, 2001). In the literature CA applications are termed also *reactive*, *responsive* and/or *adaptive* (Schilit and Theimer, 1994; Dey, 2001; Dourish, 2004). Some of the main benefits of context awareness include: (i) delivery of highly relevant and personalized information and services to the user, (ii) decreased demand for user interaction, (iii) simplified interface, (iv) automation of trivial tasks, and in some cases (v) increased perception of security and safety (Kjeldskov and Paay, 2010).

### 2.5.1. The representational approach to context

Mobile context-aware computing became an active research area in the mid 1990s when the first prototypes of context-aware systems emerged (Schilit and Theimer, 1994). For a decade, research in the field was mainly directed at establishing and listing all of the context parameters that a mobile system has to be sensitive to. Researchers believed that context is delineable and can be defined for a set of applications in advance. Apart from location, the range and nature of identified factors varied and included time, social situation, user identity, as well as environment factors, such as light, noise and weather (Schilit and Theimer, 1994; Feiner et al., 1997; Pascoe, 1998; Schmidt et al., 1999). Analysing previous research, Dourish (2004) called this type of approach the *Representational approach* to context and concluded that research was based on the following assumptions:

- *Context is a form of information* that can be known and hence encoded and represented in a software system.
- *Context is delineable* and can be defined for a set of applications in advance.
- *Context is stable* and, although it varies from application to application, it does not vary from instance to instance of an activity or event.
- *Context and activity are separable* and, hence, while the activity exists within a context, the context itself can be separated from the activity.

Following the Representational approach to context, a number of prototypical mobile applications have been developed in the last years, discussed in the next section.

### 2.5.2. Context-aware services for tourism

Vast progress has been made with respect to context-aware systems for travel and tourism. Because the selection and delivery of relevant information is key for tourists, location- and context-awareness have been discussed extensively in eTourism literature (Hinze and Buchanan, 2005; Kenteris et al., 2006; Höpken et al., 2010; Gavalas et al., 2014). In recent years, the number of prototypical context-aware mobile tourism applications grew rapidly. Most attention has been devoted to location-based services (Hinze and Voisard, 2003; O'Grady et al., 2005; Umlauft et al., 2003; Wiesenhofer et al., 2007; Raper et al., 2011). However, a number of other parameters have also been considered, such as user interests, time of day and season (Cheverst et al., 2000), age and user profile (O'Grady et al., 2005), available time, costs, travel history, walking speed, or opening hours (Hinze and Buchanan, 2005; Martin et al., 2011). Different types of mobile context-aware tourism prototypes have been extensively reviewed in literature (e.g. Hinze and Buchanan, 2005; Gavalas et al., 2014).

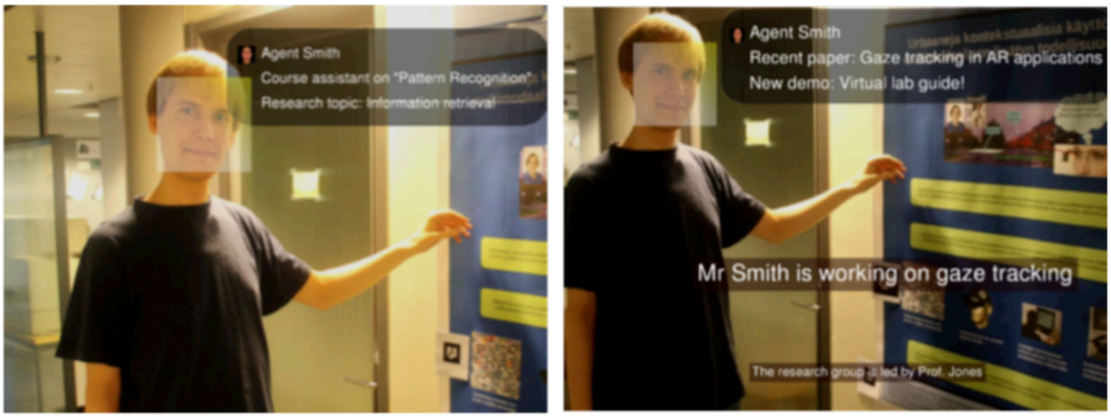
### 2.5.3. Context-aware mobile Augmented Reality

While for many types of mobile ISs context-awareness and adaptation are still optional, AR ISs on the contrary depends on being adaptive to the physical context in which they are used (Kjeldskov, 2003). At the very least, obtaining spatial information (location and orientation) is a key requirement for AR systems. The need for more adaptive content as part of context-aware AR systems has already been recognised (Kooper and MacIntyre, 2003; Langlotz et al., 2014).

Addressing this need, Bell et al. (2001) introduced the concept of *view management*. The idea was to adapt the layout of virtual information and its representation delivered through AR based on the observed actual characteristics of the physical scene. To show the feasibility of the concept, an HMD AR system was implemented which adapted the representation of annotations based on changes in visibility, size and position of physical objects (Bell et al., 2001). The authors argued that in the future, additional constraints (or context parameters) need to be added in order to make the system more usable and useful. Since then, a number of studies have investigated adapting AR to a variety of parameters. Most studies have considered special use cases and adaption for 3D models (e.g. Kalkofen et al., 2009), where a set of different context parameters are important (e.g. lightning conditions). Only a few studies, however, have investigated the feasibility of implementing context-aware AR browsers. For instance, Zhu and Owen (2008) considered adapting their AR shopping assistant to the user's preferences for shopping, location and the characteristics of the product that they are interested in. Ajanki et al. (2010) developed a CA AR system

(Figure 2.6) that was sensitive to objects in the immediate field of view of the user (determined through eye movement patterns) and the conversation at hand (speech detection algorithms). The system is able to detect the object of interest that the user is looking at (Figure 2.6, left). Then, it is able to adapt the provided information about that object, based on the conversation that is at hand (Figure 2.6, right).

**Figure 2.6. Context-based adaptation for Augmented Reality**




*Source: Ajanki et al., 2010*

Mendez (2010) considered the structure of the background that surrounds the computer-generated content, while Speiginer and MacIntyre (2014) used proximity to change dynamically the level of detail (LOD) for 3D augmentations. Other recommended context parameters include the field-of-view (Kjeldskov, 2003), the focus of attention on objects and people (Ajanki et al., 2010), visibility (Makita et al., 2009), lightning conditions and shadows (Papagiannakis et al., 2005), background textures of the surroundings (Jankowski et al., 2010), and their changing colours (Mendez, 2010).

There is only limited research that investigates the design and development of context-aware AR browsers in tourism context. The most relevant study pertaining to this area was described in Kourouthanassis et al. (2014). The main aim of the project was the development of an AR browser (CorfuAR) that delivered personalised content to tourists by automatically selecting and presenting content that matches tourists preferences. The filtering of content was based on three user profiles (thematic-based, entertainment-based and action-driven) adopted from the World Tourism Organisation tourists segmentation approach. A follow up user study, however, suggested that there was no difference in use or preferences between the personalised and non-personalised versions of the AR browser.

The main problem with development of context-aware AR is that it is extremely challenging to identify and measure context. Even if context is measured and captured, it is difficult to make inferences about its influence on interaction or intent. There is still an on-going debate what is the exact range and nature of the contextual parameters an AR system has to adapt to (Langlotz et al., 2014), which mimics the more wide debate about the relevance of context parameters in context-aware literature (Schmidt et al., 1999; Bellotti and Edwards, 2001; Greenberg, 2001; Dourish, 2004; Oulasvirta et al., 2005). The selection and combination of relevant context parameters is often determined on an *ad hoc* basis and as a proof-of-concept, rather than based on design principles and theories.

### 2.5.4. Towards a user-centred approach to context-awareness

Smartphone devices, and hence AR browsers, are used in a variety of physical environments, settings and circumstances. The powerful processors and variety of sensors of new smartphone devices can be leveraged to build smartphone applications which collect sensor data from the real world and use it to adapt to the context of use. However, the design and implementation of context-aware applications is not trivial and many challenges have to be addressed. The previous section revealed the multi-faceted nature of context and the many contextual parameters that have been identified as important for mobile interaction in general, and when it comes to delivering information to tourists through mobile ISs and Augmented Reality systems.

Recently, research has explored the limits to recognizing and labelling context. For example, simple activities of a person in a home environment can be recognized with about 80-85% accuracy (Intille et al., 2004). With the addition of many sensors to smartphones, contextual information can be sensed and recorded, however, a central question remains: “what to do with that information?” (Barnard et al., 2007, p. 83). This issue has come to a significant prominence within HCI research. For instance, Greenberg (2001, p. 23) argues that “although some contextual situations are fairly stable, discernible, and predictable, there are many others that are not. The result is that similar looking contextual situations may actually differ dramatically” in terms of the influence they have over the interaction with a mobile device. This is what led Barnard et al. (2007, p.83) to argue that “the domain of context-awareness is nearing a state where it is faced with an abundance of potentially relevant available data, but a deficit of knowledge of how to use it. Designers may assume that these contextual factors are important, and even intuitively design with them in mind, but what is missing is an understanding of how changes in context affect the user”. Kjeldskov and Paay (2010) pose that understanding the influences of both the physical environment and the human activities that unfold in that context is critical in order to move towards truly useful and usable adaptive context-aware smartphone applications. Hence, a different approach to recognizing and investigating context was needed.

While context-awareness and adaptation are directed at making information systems easier to use and more useful, many researchers agree that gathering more contextual information will not necessarily improve usability and help users meet their needs (Greenberg et al., 2001; Dourish, 2004; Christenen et al., 2006). The main implication from this shift of focus in viewing context is the need for an *empirical, user-centred design approach to understand mobile contexts* (Bellotti and Edwards, 2001; Greenberg, 2001; Oulasvirta et al., 2005; Kjeldskov and Paay, 2010). This is driven by the fact that revealing context cannot happen through theoretical reasoning only. As a dynamic, evolving and emerging property of action and interaction, context has to be studied empirically for individual types of applications.

## 2.6. Chapter Summary

This chapter discussed the need for on-site location-based information delivery in tourism and the potential of location-based interfaces (Section 2.2) to deliver more relevant content in tourism context (Section 2.3). Further review of recent developments in industry and academia revealed the characteristics of AR browsers (Section 2.4) and the benefits of this visualization paradigm to support tourists while on the go (Section 2.3). Despite the huge promise of AR browsers, there is a lack of research within the domain that has focused on tourists and tourism context of use. There are only several smartphone AR browser prototypes and projects dedicated to tourism-related functionality and content (Section 2.4.4). On a more general level, there are still a number of challenges and gaps related to design of AR browsers, related to technical, content and design issues (Section 2.4.6). Researchers have identified that the underlying problem for addressing such challenges is the lack of user research and understanding of user requirements in actual context of use. The latter is fundamentally important for moving towards the development of more usable and useful context-aware mobile information systems (Section 2.5). In summary, the lack of empirical investigations and knowledge regarding user requirements in actual context of use has stalled the development of context-aware AR browsers in general and within the domain in particular. While the context debate continues, a number of authors have recognized the need for an empirical, user-centred approach to design of context-aware mobile information systems (Section 2.5.4). The next chapter explores the key context of use factors that influence mobile usability and utility and how User-Centred Design can be used to investigate and improve the design of current or next-generation Augmented Reality browsers, which is also the key focus of this study.

# CHAPTER 3

# USER-CENTRED DESIGN FOR AUGMENTED REALITY BROWSERS

**3.1. The Smartphone as a Catalyst of Change**

Explains why the smartphone is a catalyst for change in Information Systems Design Theory

**3.2. Usability and Utility of Information Systems**

Key definitions and ways to measure usability and utility

**3.3. Mobile Usability and Context of Use**

Emphasizes the role of context and the need to adopt a user-centred design to mobile IS development

**3.4. IS Design Theory and Mobile User-Centred Design**

Key tenets of IS design theory generation and UCD

**3.5. Gaps in Existing Design Knowledge for AR Browsers**

Critical analysis of existing research that highlights the need for the research described in the thesis

**3.6. Chapter Summary**

Synthesis of the key points in the chapter

## 3.1. The smartphone as a catalyst of change

The adoption of desktop ISs by tourists, as well as the wide public, brought to surface many issues relating to their overall design (Grudin, 2012). Graphical user interfaces, functionality and presentation of information had to be suitable for an audience with no special computer background and technical skills. Designing ISs in a way that makes them easier to use became a central research topic of the newly emerging field of Human-Computer Interaction (HCI) (e.g. Norman, 1983). Early studies investigated different user interface designs, paradigms and metaphors and how they influenced task performance and acceptance by end users (Grudin, 2012). Consequently, the importance and influence of HCI grew in parallel to the increasingly fundamental role of technology in modern society.

The tremendous technological developments, convergence and interoperability in the 21st century led to the birth of more powerful mobile and wearable computers. This is when the field of Mobile Human-Computer Interaction (Mobile HCI) was born, with the main purpose to study and address the specific interaction aspects with such devices (Krogstie et al., 2003; Kjeldskov and Graham, 2003; Schleicher et al., 2014; Kjeldskov and Skov, 2014). A significant amount of research within the field of Mobile HCI is directed towards one of the most powerful mobile computers: the smartphone.

The early days of smartphones mimicked the dawn of the Internet era, where technological advance was prioritised over the ease of use of websites. Difficult to use or unintuitive mobile ISs waste their users' time, cause frustration and annoyance, prevent users from completing their tasks and discourage further interaction with the product (Bevan and MacLeod, 1994; Abras et al., 2004). Many empirical studies suggest that smartphone applications and websites are still difficult to use and understand (Nielsen and Budiu, 2013). This is why a lot of attention has been placed recently on ensuring qualities such as *utility* (expressed as functions and content that people really need) and *usability* (expressed as ease of use, efficiency, effectiveness, learnability, memorability and satisfaction) of mobile information systems. Both utility and usability are essential and will ultimately determine the acceptance and success of mobile information systems. The first part of this chapter (Section 3.2) introduces both concepts and examines the key factors that influence them when it comes to mobile information systems and location-based services (Section 3.3).

Since the introduction of personal computers to the mass market, a number of guidelines and heuristics have been proposed that aim to improve the usability and utility of interactive digital products. Popular user interface guidelines include Schneiderman's "Golden rules for Interface Design" (Schneiderman et al., 2013) or Nielsen's 10 usability heuristics (Nielsen, 1993). Design knowledge, expressed as guidelines, heuristics and checklists, is often accumulated after extensive empirical research. The main problem is that such detailed and specific guidelines remain pertinent mainly to desktop user interfaces. Designing usable and useful mobile information systems is not trivial. The small size of the smartphone display, limited input and dynamic context of use call for new user interface design principles and guidelines (Gorlenko and Merrick, 2003; Fling, 2009) captured in design knowledge and theories. The second part of this chapter (Section 3.4) looks at the overall process of Information Systems Design theory generation. A number of studies have shown that the key to designing successful context-aware mobile applications is to break away from the traditional way of thinking about computing and to place users in the centre of all design activities. Designing with users and context of use in mind is the key concept that underpins a User-Centred Design approach. After presenting the key principles of UCD, the final part of this chapter (Section 3.5) describes the very limited number of projects that have adopted a UCD approach to Augmented Reality used in urban tourism context. Adopting a UCD methodology is especially important in view of the lack of design theories and guidelines when it comes to smartphone Augmented Reality browsers.

## 3.2. Usability and utility of Information Systems

### 3.2.1. Defining usability, utility and user experience

Usability and utility are fundamental qualities of products when it comes to supporting users with achieving their tasks and goals. Despite this, for many years, utility, as well as usability "remained a fuzzy concept, which has been difficult to evaluate and impossible to measure" (Bevan and MacLeod, 1994, p.132). Both terms have been used broadly in literature often referred to with the umbrella term *usefulness*, which is defined as the ability of a product to satisfy its users' needs and goals (Grudin, 1992; Nielsen, 1993). Intrinsically connected to overall usefulness is the concept of utility. It concerns whether the "functionality of the system in principle can do what is needed"

(Nielsen, 1993, p. 25). Software that provides the right functionality for its users is considered useful. When it comes to information systems, utility concerns the delivery of relevant information (Wilson, 1992) to its intended users. A useful information system maximises the relevance (the match between the provided information and the information need of the user) of information. In this context, relevance is defined as the match that exists between an information source and an information need as seen by its inquirer (Wilson, 1992; Mizzaro, 1997). The precise focus of design activities that ensure or assess utility can vary among domains and will depend on the product being designed. Ensuring high utility, however, does not imply that the system is easy to use (Grudin, 1992) and this is why the overall usability of the product has to be considered.

Shackel (1981; 2009) was one of the first authors to propose an operational definition for *usability*, which later became accepted in both academia and industry. In his widely cited paper, Shackel (1981, p.24) defined usability as "the capability in human functional terms to be used easily and effectively by the specified range of users, given specified training and user support, to fulfil the specified range of tasks, within the specified range of environmental scenarios". This became the basis for the first internationally recognised definition that focused on operationalization of the term, introduced by the International Standards Organisation through ISO13407, later renamed to ISO9241, "Ergonomics of human-system Interaction". Part 210 of the standard, called "Human-Centred Design for Interactive Systems" defines usability as: *"The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use*" (ISO9241-210, 2010).

The term *usability* is often confused with that of *User eXperience* (UX). However, there is difference between the two. UX is defined as "*a person's perception and responses that result from the use or anticipated use of a product, system or service*" (ISO 9241-210, 2010). The ISO standard suggests that UX "*includes all the user's emotions, beliefs, preferences, perceptions, physical and psychological responses, behaviours and accomplishments that occur before, during and after use"*. Hence, some authors argue that UX is similar to the concept of *satisfaction* in usability (Bevan, 2009). The definition explains why UX and usability are often used interchangeably. There are two distinctive objectives common for both user experience and usability studies: 1) optimising human performance; 2) optimising user satisfaction with achieving goals (Bevan, 2009).

This is also why both usability and UX studies make use of the same methods available as part of a user-centred design methodology. Nonetheless, there is a subtle difference among usability and UX studies, as the accent within a user-cented design lifecycle is often placed on different aspects of interactive systems (Table 3.1). More often, however, both within academia and industry, UX is used as an umbrella term that incorporates usability (e.g. Bevan, 2009).

**Table 3.1. Comparison between usability and user experience**

| Usability studies goals | User experience studies goals |
|---|---|
| • Designing for and evaluating overall effectiveness and efficiency.<br>• Designing for and evaluating user comfort and satisfaction.<br>• Designing to make the product easy to use, and evaluating the product in order to identify and fix usability problems.<br>• When relevant, the temporal aspect leads to a concern for learnability. | • Understanding and designing the user's experience with a product: the way in which people interact with a product over time: what they do and why.<br>• Maximising the achievement of the hedonic goals of stimulation, identification and evocation and associated emotional responses. |

*After: Bevan, 2009*

### 3.2.2. Measuring usability and utility

A set of *observable* and *quantifiable* metrics is needed in order to evaluate the two different, but at the same time related aspects of information systems: their utility and usability. Variations on aspects and measures abound in literature and will depend on the adopted definition for usefulness. For instance, Nielsen (1993) suggested learnability, memorability, efficiency, errors and satisfaction as main criteria that determine the usability of software. Later Shackel (2009) proposed effectiveness, learnability, flexibility and attitude as key usability attributes of ISs. The most widely used set of attributes that measure usability are the ones proposed by the ISO9241-210 standard: effectiveness, efficiency and satisfaction (Table 3.2).

The nature of the product, domain and the tasks of the users will determine which criteria are most suitable to use (Shackel, 2009). For instance, consider the design of complex software that supports expert tasks, such as drawing and computer-aided design. Due to the variety of functions, buttons and menus that such software packages combine, it is important to test how fast users will learn to work with the product (learnability) and whether they will retain this knowledge over time (memorability). In comparison, smartphone information systems in tourism should take minimal amount of time to learn. Because they are used spontaneously and intermittently (over large periods of time), the interface should be intuitive immediately, without requiring users to make a conscious effort to learn how buttons and menus work. In such cases, effectiveness (e.g. success rate) and efficiency (e.g. time) can be used to evaluate the extent to which users can learn quickly how to work with the interface.

More recently, Harrison et al. (2013) reviewed existing usability models and argued that cognitive load (Table 3.2) is an important aspect to consider when evaluating interaction with mobile applications. The authors pose that scales, such as the NASA Task Load Index (TLX) (Hart and Staveland, 1988), are essential for mobile studies where users are often forced to work in  dynamic and changing settings.

**Table 3.2. Metrics for measuring usability aspects of products**

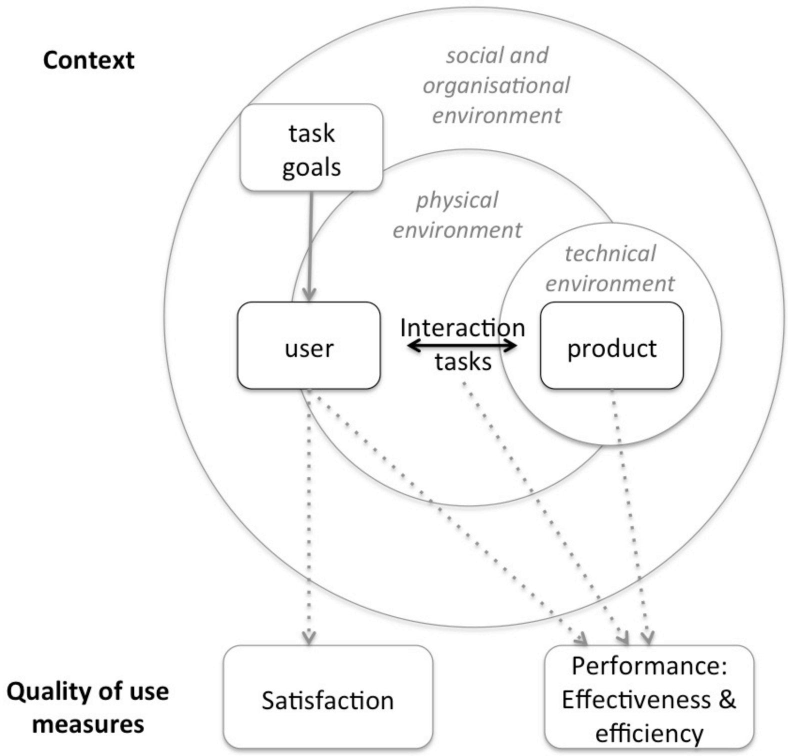
| Criterion | Definition | Metrics |
|---|---|---|
| **Effectiveness** | Accuracy and completeness with which users achieve specified goals. | Success rate<br>Number of errors |
| **Efficiency** | Resources expended in relation to the accuracy with which users achieve goals. | Time on task<br>Mental effort rating scales<br>Physical effort rating scales |
| **Satisfaction** | Freedom from discomfort and attitudes to the use of the product | Attitude rating scales |
| **Cognitive load** | The amount of cognitive processing required by the user to use the application | Subjective workload ranking |

Unlike usability metrics, no consensus exists on a standard set of criteria that can be used for evaluating the utility of information systems. Partially, this is because the measure of utility is heavily dependent on the domain for which the information system is developed, as well as the information seeking context and tasks. This is why often usefulness of information is treated as an extension of the concept of *relevance* (Tsakonas and Papatheodorou, 2006). Relevance of information denotes the match which exists between delivered content and information needs as seen by the inquirer (Wilson, 1973). Empirical studies within Information Science have suggested that relevance is a multi-faceted concept, determined by a number of aspects and contextual parameters (Mizzaro, 1997; Case, 2012). Empirical research in tourism confirms the situational aspects of relevance, and have shown that trustworthiness, the perceived value of content, timeliness, and the degree of difficulty in understanding information content are only some of the major determinants of information utility for tourists (Fesenmaier and Vogt, 2008).

## 3.3. Mobile usability and context of use

Situation factors (the user, the tasks and the environment) are so important for usability that "changing any relevant aspect of the context of use may change the usability of the product" (Bevan and MacLeaod, 1994, p. 138). From the point of view of mobile information systems, changes in context influence the relevance (utility) of information delivered to users (Chua et al., 2011). Situational factors also impact performance (Bevan, 1995) and how easy it is for users to understand information, which is visually displayed on the screen of the smartphone device (Figure 3.1) (Bevan, 1995; Krogstie et al., 2003; Fling, 2009). Therefore, it is critical to consider the range of context parameters and how they influence mobile interaction, perception and use of information.

**Figure 3.1. Context of use frames mobile interaction**




*Source: Bevan, 1995*

According to the ISO definition, usability and utility are closely related to context and can be measured only when there is an interaction between a product (on a smartphone) and a specified user (Figure 3.1) (ISO 9241-210, 2010). Within Mobile HCI, there is a general agreement that context of use (CoU) frames, surrounds, defines and ultimately influences the interaction between users and mobile computers (Bevan, 1995; review in Jumisko-Pyykkö and Vainio, 2010). Understanding context of use characteristics is important for designers, as well as usability researchers alike.

### 3.3.1. Context of Use models

There are a number of models, definitions and frameworks that have been proposed to describe and explain CoU. Most definitions echo the representational approach to context (Section 2.5), as researchers try to capture and list the context of use variables that will influence usability of mobile information systems (Dey et al., 2001; Bradley and Dunlop, 2005). For instance, the widely adopted ISO9241-210 standard defines context as “the users, goals, tasks, equipment (hardware, software and materials), and the physical and social environments in which a product is used”. The standard sets the following definitions:

- User: is the person that interacts with the product.
- Goal: is the intended outcome of the interaction.
- Task: comprises of the activities undertaken to achieve a goal.

Defining the nature and influence of CoU and its implications for design has been the goal of more extensive research (Winters and Price, 2004; Bradley and Dunlop, 2005; Bradley, 2005; Jumisko-Pyykkö and Vainio, 2010) and is out of scope for this study. However, CoU is an important concept that influences usability. Reviewing available research on context of use in Human-Computer Interaction and mobile HCI, Jumisko-Pyykkö and Vainion (2010) proposed a new context of use model. The model lists categories of context parameters that might influence the interaction of users and mobile ISs. It excludes users and activities, as something that happens “in context”, which is in line with a user-centred approach to context (Section 3.3.3). There are five main context of use categories, captured in the CoU model (Jumisko-Pyykkö and Vainio, 2010). In essence, the authors argue that interaction of users and mobile information systems might be influenced by properties of the surrounding physical, temporal, task, social and technical context.

### 3.3.2. Context of use parameters

Adopting the approach and model developed by Jumisko-Pyykkö and Vainio (2010), this section reviews empirical and theoretical research that has contributed to understand better how context influences the use of mobile information systems, and in particular augmented reality.

#### *3.3.2.1. Physical context*

Physical context describes the apparent features of the situation in which human-computer interaction takes place, including spatial location, functional place and space, sensed environmental attributes, movements and mobility, and artefacts present.

In the past, research in mobile computing and mobile Information Systems for tourism has focused exclusively on *location* (Section 2.5.2). While location-awareness is a concrete step towards context-based adaptation, it is only one of the elements of physical context that influences mobile interaction. The design and development of mobile location-based services, such as map-based interfaces, and especially mobile Augmented Reality, has revealed the importance of studying space in more detail (Sarjakoski and Nivala, 2005). Aspects such as the function of space (e.g. city zone, home, office), perception of space and other material characteristics of location have been highlighted as relevant (Sarjakoski and Nivala, 2005).

Literature has discussed the granularity of obtaining physical coordinates with respect to mobile HCI. Apart from physical coordinates in space, *orientation* of the user is considered an important aspect of mobile interaction with mLBSs (Cheverst et al., 2000; Sarakjoski and Nivala, 2005). Orientation is also an extremely important aspect of interaction with mobile AR (Kjeldskov, 2003) and has to be acquired in order to deliver more relevant information to users.

In addition, *environmental attributes*, such as weather conditions, lightning and noise level are also considered important physical aspects that influence usability (Barnard et al., 2007). For instance, lightning level could affect the legibility of mobile maps and this is why the map colours and background illumination should be adapted to such environmental factors (Sarakjoski and Nivala, 2005). Lightning level is also very important for Augmented Reality applications, as the lack of light (e.g. use by night), or very bright sunshine (Herbst et al., 2008) could hinder legibility.

Finally, *present artefacts* represent physical objects surroundings the human-computer interaction (Kjeldskov and Paay, 2010). In tourism context, artefacts are mainly nearby attractions and points of interest (Cheverst et al., 2000; Tan et al., 2009). Surrounding landscape and physical configuration of the environment has been considered important for both mobile maps (Sarjakoski and Nivala, 2005) and mobile Augmented Reality (Kjeldskov, 2003). In fact, Kjeldskov (2003) argues that the actual physical space in view of the user is the most fundamental contextual parameter triggering adaptation in mobile AR. Different contextual detail would be necessary, for instance, for outdoor city exploration AR application, a museum exhibit or browsing information for products in the store.

Looking at research on context-aware systems confirms that the predominant view in the field of AR is that context is simply the surroundings of the user that can be augmented with information (Bell et al., 2001; Kjeldskov, 2003; Bell et al., 2005). In this sense, a number of additional parameters that describe physical space, such as structure, visibility, proximity, priority, background textures, and empty space have been identified as important to interaction with mobile AR (Bell et al., 2001; Kjeldskov, 2003; Bell et al., 2005; Makita et al., 2009; Kruijff et al., 2010).

*Proximity* to important points of interest can be used to filter out information on AR displays (Bell et al., 2001). The main idea is to determine whether objects are visible or not based on their distance from the location of the user. This approach is quite limited and, instead, ***visibility*** (or the occlusion relationships between objects) has been proposed to determine the features that are currently visible from the position of the user (Kruiff et al., 2010). These approaches are limited as the importance of a physical object is not simply a function of distance and visibility (Bell et al., 2005). In such cases, more complex *priority rule-based* approaches have been adopted, where the combination of visibility, location, as well as the role of the object with respect to the current task of the user is determined (Bell et al., 2005).

*Background textures* (Jankowski et al., 2010) and *availability of empty or low-priority space* (Feiner et al., 1997; Bell et al., 2001; Bell et al., 2005) have been used to adapt the presentation of content in AR. In addition, the structure of the physical environment, expressed as depth ordering, scene distortions, clutter, object relationships, surfaces and object segmentation (Kruiff et al., 2010) are also important context factors that could influence how users perceive, interpret and use information delivered by AR systems.

#### *3.3.2.2. Temporal context*

Temporal context describes the user's interaction with the mobile computer in relation to time in multiple ways, such as duration, time of day or year, the situation before and after use, action in relation to time and synchronism. When it comes to tourism, relevant temporal aspects might also include the latest happening events, duration of stay at a current destination, last visited date, number of repeated visits to a destination, and acceptable waiting time (Tan et al., 2009).

In Mobile HCI, *duration* illustrates the length of the use session, which often depends on the task of the user and the surrounding environment. For instance, preferred time to consume mobile video varies between 35-40 minutes, depending on the duration of a given situation (e.g. waiting time) (O'Hara et al., 2007). In tourism context, duration might also refer to the time period allocated within a destination (Tan et al., 2009). A quite different approach to duration has been applied in context-aware AR implemented on head-mounted displays. Considering physical context and the multiple physical objects that could be augmented with information, *duration of gaze* has been used to determine the potential objects of interest that the user would like to acquire information about (Ajanki et al., 2010).

The *time of day, week and year* indicates relative periods of user interaction in relation to time. In the context of everyday use of mobile devices, this parameter has been used to describe the peaks of user interaction in relation to time of the week (Halvey et al, 2006). Time of day, season and year are also contextual parameters that have been used in mobile tourism ISs mainly to adapt the type of information delivered to tourists (Cheverst et al., 2002; Hinze and Buchanan, 2006). Time of day might also influence physiological information needs, such as the need for locating food venues.

*Before and after use* emphasizes the need to study actions that are carried out prior or post use session. For instance, after using a mobile tour guide, tourists might want to extend the experience of a certain event by taking digital souvenirs (Kaassinen, 2005). In urban tourism context, this aspect might cover, for instance, already booked or selected tourism services (Höpken et al., 2010).

*Action in relation to time* highlights the temporal tensions of actions (Tamminen et al., 2004). For instance, the perceived usability of a mobile information system will be influenced if tourists are in a hurry or are waiting in line to visit an attraction.

Temporal tensions might also influence the speed of walking and available time to stay at a point of interest (Kramer et al., 2006) and should be used to adapt the interface of the mobile IS. *Synchronism* describes the status of interaction in relation to communication with people (Jumisko-Pyykkö and Vainio, 2010). For instance, calling to book a reservation in a restaurant is a synchronous two-way communication, while texting the reservation details is an asynchronous activity.

Even though not included in the original CoU model, *context history* is another parameter that has been identified as important in tourism literature (Sarakjoski and Nivala, 2005; Hinze and Buchanan, 2005; Hopken et al., 2010). Already visited points of interest, time since the POI was last visited, route history, or number of repeated visits to a destination might all influence the information needs of tourists (Sarakjoski and Nivala, 2005; Tan et al., 2009; Hopken et al., 2010).

#### *3.3.2.3. Task context*

Task context refers to demands of the situation on the attention of the user and captures relevant aspects, such as multitasking, interruptions and task domain (Jumisko-Pyykkö and Vainio, 2010). *Multitasking* describes the necessary multiple parallel tasks that users need to carry out and which compete for cognitive resources. For instance, parallel tasks to the interaction with the mobile device might be walking, sidestepping, and planning routes (Oulasvirta et al., 2005). *Interruptions* are events that break the attention of the user temporarily. In mobile HCI, interruptions can be caused by technical problems (e.g. patchy network connectivity), social (e.g. people interrupting the use session) or physical (e.g. lightning level changes) context (Kaassinen, 2005).

The *task domain* represents the macro level of task context. Jumisko-Pyyko and Vainio (2010) divide task domains into two main categories: goal-oriented (work) and action-oriented (entertainment). The main difference between the two is the aspects of user interaction that have to be measured. Goal-oriented task domains prioritise performance and this is why efficiency and effectiveness are both important aspects of interaction. For the action-oriented task domains, such as gaming, mobile video or music consumption, the action itself is the goal of interaction (Jumisko-Pyykkö and Vainio, 2010). According to this classification, this study is concerned with goal-oriented task domains, as work-related applications, guides and navigational assistants are examples of highly goal-oriented tasks.

#### *3.3.2.4. Technical and information context*

Early research in Mobile Human-Computer Interaction focused on the hardware and network constraints of mobile devices compared to desktop computers and how such characteristics influence mobile usability (Table 3.3). These included the limited screen size, display resolution and colours, processing power, storage space and patchy network connectivity. Desktop systems require heavy computations and were originally designed for large screens, which means that they do not scale well to mobile devices.

**Table 3.3. Comparison of desktop and smartphone devices**

| Feature | Desktop PC | Mobile device |
|---|---|---|
| CPU and storage capacity | 4-6 processors x 2-3GHz (CPU) and 320GB storage capacity. | Between 1.3 - 2.7GHz (CPU) and 32GB storage capacity. |
| Screen size | Varies. Standard between 13-21". | Varies. Normally around 3.5" – 4.7". |
| Network connectivity | Constant, does not change. | Varies and depends on the location of the user. May be interrupted frequently due to the mobile nature of the user. |
| Network transfer rate | Depends on the ISP, but does not vary significantly. | Varies, depending on the network connectivity. |
| Power | Unlimited when connected to a power supply. | Short battery life. |

Jumisko-Pyykkö and Vainio (2010) describe technical and information context as the relation of other relevant systems and services including devices, applications and networks, their interoperability, information artefacts or access, and mixed reality to the user's interaction with the mobile device.

Key hardware considerations in AR are mainly related to the type of mobile device used to deliver augmented content. The influence of type of display is threefold: duration of use sessions, interaction and field of view (see overview in van Krevelen and Poelman, 2010). First, the type of display influences the way users will interact with the device and, in turn, with augmented content. For instance, input in smartphone devices is carried out through touch-based interaction, while mobile glasses and head-mounted displays require gesture-based interaction and/or voice input. Second, the type of display also influences the field of view (FOV) of the user, or the extend of the observable world (Kruijff et al., 2010). Smartphone devices have limited FOV, constricted to the viewing parameters of the smartphone camera. Third, the duration of a use session is also different. For instance, head-mounted displays can provide “always-on” and continuous augmentation, while smartphone devices require that the device is taken out and lifted vertically towards the object of augmentation.

#### *3.3.2.5. Social context*

Social context describes the other persons present, their characteristics and roles, the interpersonal interactions and the surrounding culture that influences the users’ interaction with a mobile computer. Jumisko-Pyykkö and Vainio (2010) classify the physical and virtually present persons during interaction into self, group and organization. The presence of other people might influence the way users interact with mobile devices, depending on their status relative to the user (e.g. familiar or unfamiliar). Unfamiliar people that are present during interaction trigger the need for users to create “private spaces” during use of mobile devices (Tamminen et al., 2004).

When it comes to mobile LBSs, other people’s characteristics, information needs and their roles might also influence interaction (Brown and Chalmers, 2003; Paay et al., 2009; Paay and Kjeldskov, 2010) and is especially important in tourism, as leisure tourists rarely travel alone. For instance, the presence of travel companions might influence and trigger collaborative use of information sources towards a shared understanding of space (Paay et al., 2009). In addition, the need for sharing visits with distant people also influences design of mobile ISs (Brown and Chalmers, 2003).

Culture denotes the macro level of social context in terms of values, routines norms and attitudes (Jumisko-Pyykkö and Vainio, 2010) in the place where the mobile IS is used. While the importance of cultural context has been discussed when it comes to tourism mobile ISs (e.g. Raptis et al., 2005), its influence on design has not been studied so far.

#### *3.3.2.6. Tourists characteristics, knowledge and abilities*

Potentially anyone can be a tourist, which means designing for users with widely different cultural backgrounds, education, expertise, knowledge, skills and abilities. Design of mobile information systems for tourism is highly challenging, as they have to satisfy the needs of a huge and varied audience. There is a large body of literature that examines the characteristics of tourists from a management, cultural, environmental and sustainable point of view. Research has been focused on tourists information seeking behaviour, mainly prior to their arrival at a destination (e.g. Gursoy and McCleary, 2004). More recently, a number of papers have tried to address this lack of research, focusing on tourist characteristics and behaviour (Brown and Chalmers, 2003), tourists experiences (Pine and Gilmore, 1995) and the role of mobile technologies within the overall tourist experience (Neuhofer et al., 2013). While not considered as a context parameter in the original CoU model by Jumisko-Pyykkö and Vainio (2010), the specific characteristics of tourists as users of information will ultimately influence the usability and perceived utility of AR browsers. In terms of characteristics, tourism and HCI studies have emphasized the role of demographics, user interests, preferences, cognitive and physical abilities and already acquired knowledge and experience (Poslad et al., 2001; Cheverst et al., 2002; Srakjoski and Nivala, 2005; Hinze and Buchanan, 2005; Höpken et al., 2010) when it comes to usability and utility of mobile context-aware applications.

*Demographic aspects* important to design of mobile ISs in tourism include age and nationality. Age influences mainly the type of information and its representation on the screen of the mobile device. For instance, symbols on mobile maps should be more simple and entertaining for younger users (Srakjoski and Nivala, 2005). Nationality has been considered in the context of preferred language of use.

*Interests and preferences* are two of the most commonly used contextual parameters in the design of context-aware mobile tourism applications (Hinze and Buchanan, 2005). Often, the main implication is that information is categorised according to the topic it refers to (e.g. history, architecture, shopping). Different ways have been explored to capture and infer interests and preferences automatically, for example, through already visited locations and feedback for visited points of interest (Poslad et al., 2001; Hinze and Buchanan, 2005) or user profiles (Umlauft et al., 2003). User profiles that capture interests and preferences have also been considered when it comes to context-aware AR applications (Seo et al., 2011). There is still no uniform way to automatically determine the interests and preferences of tourists, and this is why most often mobile ISs rely on manual input from the user.

*Perceptual, cognitive and physical abilities*, such as memory, learning, problem-solving, and decision-making have also been identified as important parameters, especially when it comes to using visual displays, such as mobile maps and routing services (Sarakjoski and Nivala, 2005). Perception is one parameter that is fundamentally important for AR systems (Kruijff et al., 2010). Considering the many factors that influence perception in AR, Kruijff et al. (2010) conclude that "perceptually correct augmentation remains a crucial challenge" for designers of AR systems (Kruijff et al., 2010, p. 3). Perceptual issues "relate to problems that arise while observing and interpreting information" either from the combination of the generated virtual world and physical space, or the real world only (Kruijff et al., 2010, p. 3).

Likewise, already acquired knowledge or familiarity with a specific destination is an important fact that could influence the usability and utility of ISs (Davies et al., 2010). The influence of familiarity has not been studied in detail (Davies et al., 2010) when it comes to perceived utility and usability of mobile location-based services or Augmented Reality.

##### *3.3.2.7. Travel context*

In addition to user characteristics, the purpose (e.g. business versus leisure), itinerary and logistics (travel modes) behind an individual trip have also been proposed as important aspects of mobile context in tourism (Höpken et al., 2010). While not used explicitly in context-aware tour guides, the type of trip could potentially influence time availability of tourists (with business travellers having less time for sightseeing). In turn, temporal context could be used to adapt the length of a proposed tour or itinerary. Available time, together with the opening hours of attractions were used in m-To Guide (Kamar, 2003) to push information to tourists about currently open attractions in their vicinity. However, tourists might still want to visit a point of interest outside of opening hours.

Proposing the TILES contextual framework, Tan et al. (2009) identified a set of 42 contextual parameters that they consider important when it comes to delivering information to tourists during the on-site information acquisition stage. The parameters were divided in 5 major categories (Temporal, Identity, Location, Environment, Social). They included events around the year, duration of stay, preferred language, number of repeated visits, carvings, acceptable wait time, travelling speed and, among many others, traffic and road conditions. The authors propose that such context parameters should be considered when it comes to designing mobile tourism applications. While comprehensive, the framework does not propose (or discuss) which context parameters are relevant in different situations. Lack of such discussion leaves the question as to whether the use of such contextual information would improve mobile interaction.

### 3.3.3. Determining relevant context parameters

As Section 2.5.4 discussed, predicting and listing the number of relevant contextual parameters prior to design is a very difficult task and connected with what Dourish (2004) calls the *representational approach* to context. A number of researchers have argued the limitations of the representational approach (e.g. Greenberg et al., 2001; Dourish, 2004), mainly due to the fact that gathering more contextual information will not necessarily improve usability and help users meet their needs (Christenen et al., 2006). Dourish (2004) proposed an alternative to the representational approach, which he called the *Interactional approach* to context. In his highly cited paper Dourish argues that "context isn't [viewed as] something that describes a setting; it's something that people do. It is an achievement, rather than an observation, an outcome rather than a premise" (Dourish, 2004, p.6). The author goes on to pose that:

- *Contextuality is a relational property* – it is not simply the case that something is or is not context; rather, it may or may not be contextually relevant to some particular activity.
- *The scope of contextual features is defined dynamically* – rather than considering that context can be delineated and defined in advance.
- *Context is occasioned property* – it is particular to each occasion of activity or action. It is relevant to particular settings, particular instances of action, and particular parties to that action.
- *Context arises from the activity* – it is not "just there" but rather is actively produced, maintained and enacted in the course of an activity.

The main implication from this shift of focus in viewing context is the need for an *empirical, user-centred design approach to understand mobile contexts* (Bellotti and Edwards, 2001; Greenberg, 2001). This is driven by the fact that revealing context cannot happen through theoretical reasoning only. As a dynamic, evolving and emerging property of action and interaction, context has to be studied empirically for individual types of applications.

A number of researchers support this point of view (Bellotti and Edwards, 2001; Greenberg, 2001; Dourish, 2004; Tamminen et al., 2004; Oulasvirta et al., 2005; Kjeldskov and Paay, 2010). It seems that the presented approaches (Representational and Interactional) adopt two separate and opposing standpoints, as most of the time they are applied in isolation. However, Dourish (2001, p. 232) argued that, "these are in fact two aspects of the same broad program", while Oulasvirta et al. (2005, p.198) added that "both camps benefit from considering the alternative approach and a middle ground can be found". While the Representational approach to context facilitates the fast development of context-aware applications, there is a critical need for a more empirical, user-centred approach to studying context for novel mobile applications. At present, there is a lack of studies undertaking the latter approach within the application area of eTourism, therefore, an empirical, user-centred approach to context was adopted for the purpose of this study.

## 3.4. IS Design Theory Development and Mobile User-Centred Design

Information Systems Design is a multidisciplinary research area, involving experts and researchers from fields such as computer science, management, software engineering, databases and scientific visualization. Within such fields, the definitions of the term design abound (Dix et al., 1998; Carroll, 2000). However, there are two broad general views: (1) design as the *process* of creating an information system, expressed as recommendations for implementing specific algorithms and techniques and (2) design as the *result* of that process, expressed as the qualities and characteristics that an information system should possess.

Design (both as properties and process) can relate to different aspects of an IS, such as (1) the logical user interface (e.g. information architecture), (2) the physical user interface (hardware components), or (3) the graphical user interface (e.g. layout) (Heo et al., 2009). In all of these cases, the core that drives a design process and the qualities and properties that ISs should possess, is captured through the collective term *design knowledge*. Design knowledge is accumulated and described in *design theory* (Gregor and Jones, 2007; Gregor and Hevner, 2013). The main focus of this study is to make a theoretical contribution through generating design knowledge expressed as the qualities and characteristics that Augmented Reality browsers should possess in order to meet user requirements. Therefore, it is essential that the general process of generating design knowledge and contributing to the formulation of design theories is reviewed. This is the purpose of the following sections.

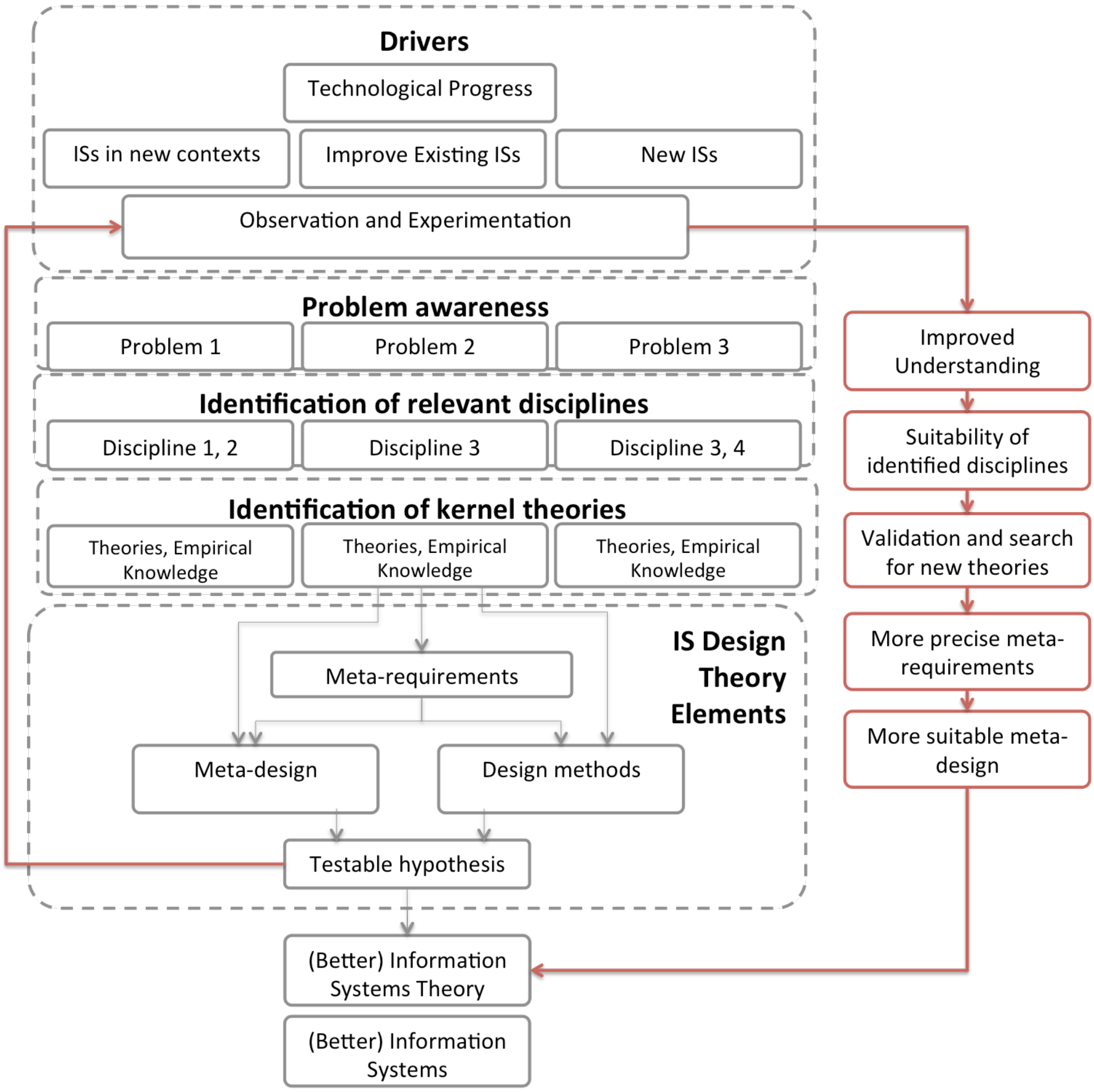
### 3.4.1. Information Systems Design Theory Generation

The lack of "native theories" within the Information Systems domain have been predominant for years (Weber, 2003). More than ten years ago, Weber (2003, p. iii) wrote that this is not surprising, as *"we have a reputation in using and adapting theories developed in other disciplines. Little wonder, that we see few high-quality theory papers in our discipline, in spite of the significant insights that such papers can provide about information system-related phenomena"*. This has remained the predominant view among IS researchers who still believe that the nature of research within the field requires to borrow theories from other disciplines, rather than create its own (Straub, 2012). A number of researchers, however, have argued that the development of Information Systems Design Theories (ISDTs) is the ultimate purpose of ISs design science research (Walls et al., 1992; Hevner et al., 2004; Gregory and Mautermann, 2014).

According to Walls et al. (1992), an ISDT is a *prescriptive theory* that guides the process of ISs creation (Walls et al., 1992). A number of authors have extended the original definition and process of developing ISDTs (Markus et al., 2002; Hevner et al., 2004; Gregor, 2006; Gregor and Jones, 2007; Kuechler and Vaishnavi, 2008; Arazy et al., 2010). In parallel, several ISDTs have been proposed in the ISs literature (e.g. Azary et al., 2010). In essence, a design theory encompasses knowledge represented through conjectures, models, frameworks and design principles (Gregor and Jones, 2007) that *prescribe* what qualities a specific class of ISs should possess in order to achieve certain *goals*. This characteristic distinguishes design theories from descriptive and predictive theories as well as from routine design practice. For instance, the Cognitive Information Processing theory is a descriptive theory, which, among other things, says that new information enters short-term memory before it enters long-term memory (Matlin, 2013). It does not say, however, how to facilitate learning through an IS. In comparison, an ISDT for e-Learning uses the Information Processing theory to prescribe what should be the qualities of that IS so that information enters long-term memory faster so that students can learn better.

The *general process of ISs design theory generation*, illustrated in Figure 3.2, was first described by Walls et al. (1992), and later refined by Markus et al. (2002), Hevner et al. (2004), Gregor and Johnes (2007) and Arazy et al. (2010). The process is initiated by *problem awareness*. Common *drivers* that trigger problems or opportunities for design include the introduction of traditional ISs into new contexts of use (e.g. word processing on the smartphone), technological progress that enables the improvement of existing ISs (e.g. new algorithms for information retrieval), or the achievement of tasks traditionally performed without the use of ISs. All of these developments result in the need to design artifacts in areas where "existing theory is often insufficient" (Hevner et al., 2004, p. 76), and this is where the opportunity arises for "IS design research to make a significant contribution…[by addressing] fundamental problems in the productive application of information technology" (Hevner et al., 2004, p. 76-77).

**Figure 3.2. The process of ISDT development**



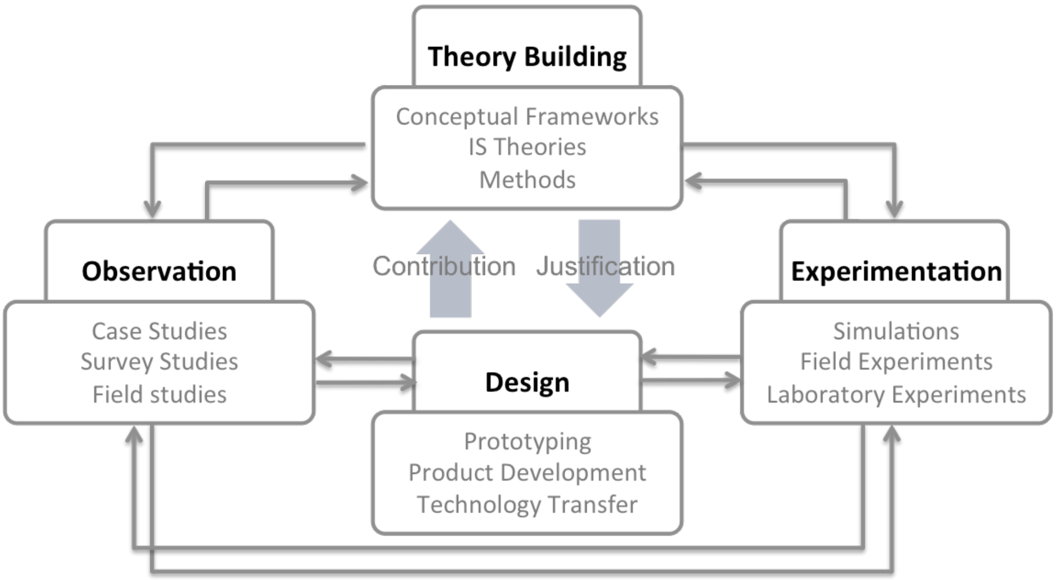

*After: Walls et al. (1992); Markus et al. (2002); Hevner et al. (2004); Gregor and Jones (2007); Arazy et al. (2010)*

In order to tackle such problems and develop new theories for ISs design, a researcher first identifies existing knowledge in *relevant research disciplines* that could help in understanding the problem (Walls et al., 1992; Hevner et al., 2004). The goal of the researcher is to identify *kernel theories* (Walls et al., 1992) or the most relevant existing empirical research, concepts, frameworks, models, instruments, and constructs that would help in addressing the problem (Hevner et al., 2004; Arazy et al., 2010). Kernel theories (and especially prior empirical research) are valuable because they allow the researcher to understand the context of use of the IS, and the prerequisites it should satisfy to achieve its purpose. The latter are also referred to as *meta-requirements* (Walls et al., 1992). Against this background, the following step is concerned with proposing a *meta-design*, namely by hypothesizing about the set of qualities and functionality that could satisfy the meta-requirements. After meta-design is identified, the researcher has to formulate *hypotheses* about the proposed design. These are tested through *experiments and observation,* where the purpose is to prove that the proposed qualities and functionality of the IS resolve the initially identified problem.

The term "hypotheses" can trigger associations with quantitative experiments and implies that a design theory is often formulated as a quantitative model (as a result of quantitative experiments). Within Information Systems and HCI, a model is an abstract conceptualization of a process. Examples include motor-behaviour models of HCI (McKenzie, 2003) or the GOMS model (Stuart et al., 1983). However, a review of existing design theories (Hevner et al., 2004; Gregor, 2009; Arazy et al., 2010) suggests that they can be formulated and are often expressed as a series of qualitative statements, coupled together in a meaningful way so that relationships among them are recognized. An example is the design theory developed by Arazy et al., (2010), where qualitative statements were used to derive hypotheses, which were then tested through quantitative research (administered through an online survey).

In the early stages of a new type of IS development, or when there are significant changes of the environment where an IS is used, experiments and observations play a fundamental role for design theory generation (Nunamaker and Chen, 1991; Markus et al., 2002; Gregor, 2009). This is because collected empirical data allow more accurate understanding of the problem, and thus result in more precise definitions of relevant research, meta-requirements and meta-design. This is also illustrated in Figure 3.2 (red loop). The development of an ISDT is, hence, an iterative process. The desired outcome from each iteration is *better IS design theory* (Markus et al., 2002).

**Figure 3.3. The role of user-based studies in ISDT generation**




*After: Nunamaker and Chen, 1991*

The seminal paper by Nunamaker and Chen (1991) emphasizes the fundamental need for empirical research and its role for generating design knowledge (Figure 3.3). When developed, an IS changes the experiences and needs of people. This is why a new IS serves both "as a proof-of-concept for the fundamental research and provides an artifact that becomes the focus of expanded and continuing research" (Nunamaker and Chen, 1991, p.92). Further empirical research is then needed to generate new theories and explanations. This is also the fundamental core concept behind User-Centred Design.

### 3.4.2. User-Centred Design Approach

Originally called *usability engineering* (Nielsen, 1993), today the practice, philosophy and methodology of designing usable products is widely referred to as User-Centred Design. While first applied to design of desktop software and websites, User-Centred Design has also been recognised as the most widely used and relevant methodology to address problems with development of context-aware mobile ISs. As the previous section discussed, the need for adopting user-centred approach is mainly driven by the need to observe mobile interaction in actual context of use.

Successful design for the mobile medium requires a dedicated attention to the user, tasks, goals, needs, and the changing context during the use of a mobile product: concepts which underpin the essence of User-Centred Design (UCD). UCD is one of the major concepts that emerged from early HCI research, describing an approach (and methodology) to design in which the end-users of a product shape out its final outlook (Abras et al. 2004). UCD is both a philosophy and a framework for product design and development that was initially introduced in 1986 by Norman and Draper and comprised a set of rules for designing more usable and useful information systems (Norman and Draper 1986). Since then, a number of authors have contributed to the initial theory constructs (Abras et al., 2004; Hacklay and Nivala, 2011; Cooper and Reimann, 2014) leading to the recognition that today UCD is "one of the guiding principles for designing usable technologies" (Hacklay and Nivala 2011, p.91). In 1999, UCD was officially recognized by the International Standards Organization as an international best practice for design through the introduction of ISO 13407:1999 "Human-centred design processes for interactive systems". The standard was later revised and released as ISO 9241-210 in 2010 (ISO 9241-210, 2010). Recently, UCD has also been widely recognized as the most effective practice for designing effective mobile user experiences (Garrett, 2011).

The key result of both IS theory generation and UCD is design knowledge for better information systems through the identification of requirements, design qualities and methods. Indeed, designing useful and usable information systems requires a thorough analysis of user requirements (Byrd et al., 1992). Requirements analysis "involves end users and systems analysts interacting in an effort to recognize and specify the data and information needed to develop an information system" (Byrd et al., 1992, p. 117). The main focus of UCD is to discover early in the design process what are the requirements and needs of users and how context (or changes thereof) influences these requirements.

In the past, the role of user requirements within the overall *System Development Life Cycle* (SDLC) was often undermined and they were only elicited in the beginning of the design cycle (Ahmed and Cox, 2014). After the failure of many (often expensive) information systems, this model was heavily criticized. Many participatory and user involvement methods have since been developed that aim to understand, capture, analyse and elicit user requirements (Byrd et al., 1992). These principles lie at the heart of user-centred design (Hackos and Redish, 1998; Abras et al., 2004; Hacklay and Nivala, 2010): 1) early focus on users, tasks and environments, 2) active involvement of users throughout design, and 3) iterative design. The main goal is to place users in the center of design from the product's planning stages, to its implementation and testing.

A UCD lifecycle undergoes iteratively several key stages: (1) context of use analysis, (2) requirements specification, (3) design, and (4) evaluation. The process is similar to design knowledge generation, described in the previous section, where the key aim is to identify user requirements and propose design methods and qualities that the IS should possess. However, UCD breaks from the traditional linear approach of IS development. The main difference is that the design of an IS can go through each stage several times prior to implementation, as each consequent iteration provides more information and knowledge relevant to previous stages.

### 3.4.3. User-Centred approach applied to design of AR

While in the mid 1990s an increasing interest in ISDTs could be noted (Walls et al., 2004), nowadays they are still relatively sparse in the literature, and are mainly limited to prescribe the design of ISs for well-understood organizational processes (e.g. Hevner et al., 2004; Markus et al., 2002), or web-based ISs (Arazy et al., 2010). Gradually, with the growing importance of Human-Computer Interaction, the focus of design research shifted towards individual, rather than organizational use of information systems. This also poses challenges for IS theory generation, as information systems are used all over the world, in a variety of contexts and by users with different demographics, education, technical experience and background. Design knowledge within HCI is, therefore, expressed as design principles, heuristics and checklists based on kernel (relevant) theories and empirical observations. One of the most tangible results from UCD is the compilation of design guidelines, heuristics and checklists based on extensive empirical user research. When generalized enough, the common expectation is that such guidelines can be used to design useful and usable ISs, which will be used by different types of users and in different situations.

There are a number of User Interface (UI) design guidelines (e.g. Schneiderman's Golden Rules of Interface Design) and heuristics (e.g. Nielsen's 10 Usability Heuristics) available to designers. The specific characteristics of the mobile medium transformed traditional design principles and guidelines regarding the architecture, visual appearance and behavior (functionality) of mobile applications (Fling, 2009). The field of mobile interaction design has been characterized by swift and dynamic development in the last couple of years. Design guidelines and best practices for mobile products are still inconsistently defined and categorized (Allen and Chudley, 2013) and the need for further empirical observations in various contexts of use has been noted.

Coming up with design principles generated through kernel theories and user-based studies is especially important when it comes to new user interface metaphors or visualization paradigms, such as Augmented Reality. While increasing in number, empirical studies within the field of AR are still rare (Livingston, 2013; Tiefenbacher et al., 2014). In 2005, Swan II and Gabbard (2005) examined 1104 papers on AR and found that only 38% address human-centred design issues. Stunningly, despite the recognized need for involving users in design activities, only 2% of the studies described a formal user-based study. The authors emphasized the need to ground new design for AR interfaces on empirical observations and findings from actual user studies. This is especially important for those technologies that fundamentally alter the way humans perceive the world. As the following section describes, the lack of design guidelines and frameworks for design of AR used in tourism context can be explained by the scarce empirical research conducted in such contexts.

## 3.5. Gaps in Existing Design Knowledge for AR browsers

When smartphone AR applications first appeared on the market, their popularity grew exponentially for a very short time (Madden, 2011). This was evidenced not only by the increasing number of available applications, but also by the huge amount of downloads. Both are predicted to increase in the future, with an estimated 200 million users by 2018 (Juniper Research, 2014). A number of studies and research projects have since then focused on implementation feasibility and technical advance connected with the development of location-based AR browsers (e.g. Tokusho and Feiner, 2009; Geiger et al., 2014). When the initial excitement wore off, it became evident that early impressions and expectations of users are, to a large extent, negative (Olsson et al., 2009). As discussed earlier, the two major drivers that trigger problems and low perceived utility and usability include the introduction of ISs to new contexts of use, or development of new ISs. Indeed, commercial AR browsers are used in many different contexts of use, both indoors and outdoors (Olsson et al., 2009) and were not originally (or specifically) developed for use by tourists. All of these developments result in the need to design artifacts in areas where "existing theory is often insufficient" (Hevner et al., 2004, p. 76). Design knowledge and theory is still very limited, especially when it comes to location-based AR browsers (Parker and Tomitsch, 2014). This is where the opportunity arises for "IS design research to make a significant contribution…[by addressing] fundamental problems in the productive application of information technology" (Hevner et al., 2004, p. 76-77). Hence, there is a critical need for new design knowledge related to AR browsers, especially when it comes to tourism. There are, however, a number of studies that provide a useful basis for further design knowledge generation.

### 3.5.1. Empirical studies and user requirements for AR

The process of designing AR browsers is not trivial, as designers have to make a number of decisions with respect to available design parameters (e.g. annotation colours, fonts, backgrounds). This task might be very difficult, considering the fact that the design space for AR in general is difficult to understand (Sandor and Klinker, 2009). This is why empirical studies, examining the effect and impact of various design parameters in different contexts of use are key to ensuring that the resulting system is usable and useful.

Traditionally, AR interfaces have been studied in the military (Julier and Rosenblum, 2000) and medicine (Fuchs et al., 1998) domains. Such studies are often directed at examining the use of an AR system in a specific situation and for a specific task (Kalkofen et al., 2009). This strand of research is mainly concerned with visualization of 3D graphics, their appearance and perception by users. Key user requirements that have been identified relate to providing virtual information in such a way that users perceive it as part of the real world.

There have been very few user evaluations of AR annotations in outdoor settings. The main concern that empirical studies have addressed is the contrast between the physical environment and the virtual annotations. Leykin and Tuceryan (2004) developed an algorithm that changes the layout of the virtual annotation depending on the texture of the physical background in order to improve legibility. Their empirical results show that background textures affect readability only if the contrast between the virtual and the physical world is low. Gabbard et al. (2007) examined the effect of illuminance and text drawing styles on a text identification task. In their experiment they used an optical see-through AR system. During the text identification task, they varied the background against which the text was superimposed. In total, six background textures were used, commonly found in urban outdoor environments: pavement, granite, red brick, sidewalk, foliage and sky. The results of their study suggest that a billboard and green text performed best against all textured backgrounds.

Building upon this work, Jankowski et al. (2010) carried out an experiment where they varied the text drawing style, image polarity, and background (physical world) to determine their effect on legibility for a reading task. Their results confirmed that the billboard style supports the fastest and most accurate performance, while background texture and image polarity had little to no effect.

Authors have also discussed the key user requirements that have to be fulfilled when it comes to AR browsers. One of the key user requirements for AR annotations is that they are easy to read at all times (Bell et al., 2001). In order to be legible, the interface should provide annotations that do not overlap (Azuma and Furmanski, 2003). Considering the changing nature of background textures, annotations should adapt their layout in order to maintain legibility (Gabbard et al., 2007; Jankowski et al., 2010). Annotations should also be big enough so that users can read and process their content (Bell et al., 2001). A usable AR interface should also prevent excessive movement of virtual content and maintain frame consistency (Thanedar and Höllerer, 2004).

The second most important user requirement that has to be satisfied is to ensure *unambiguous association,* or that users are able to associate each virtual annotation with its corresponding physical entity (Bell et al., 2001; Azuma and Furmansky, 2003; Bell et al., 2005; Grasset et al., 2013). This process has been called *co-referential relationship* (Hartmann et al. 2005) and is also a requirement for map-based interfaces where it is referred to as *referential mapping* (Oulasvirta et al., 2009). In order for this process to be successful, a key requirement for AR is that the interface places annotations precisely on top or nearby the object of reference.

Literature identifies that the *precise placement* (alignment) of virtual and physical worlds solves the problem of referential mapping (Bell et al. 2001). Substantial work has been carried out with respect to precise placement of AR annotations on AR video-see through HMD systems (e.g. Azuma and Furmanski, 2003; Ishiguro and Rekimoto 2011), and more recently, for smartphone devices used in urban environments (Grasset et al. 2013). Such AR annotation placement algorithms and strategies draw heavily from Cartography (Bell et al. 2001) where the precise placement of labels on a map is critical for its utility (Imhof 1975; Christensen et al. 1992). Surprisingly, there are no studies that investigate the effect of placement on association of virtual and physical spaces empirically. Thus, it is still questionable whether placement is the most important design parameter that will ultimately determine the usability and utility of an AR browser. The situation is the same when it comes to empirical observations that confirm the effect of different design variables on usability and utility, or discuss what design decisions need to be made to improve the presentation of information.

In line with the historical development of the field, user requirements for AR browsers have been identified mainly with respect to providing information on head-mounted displays. While a useful starting point, such requirements may no longer apply to design of smartphone AR browsers. For instance, while precise placement of annotations has been considered fundamental for HMD AR, the study of Turunen et al. (2010) indicate that it may no longer apply, or is considered less relevant when it comes to delivering information through smartphone AR browsers (discussed in the next section). In addition, most user requirements have not been elicited based on empirical user studies, but extracted from literature that discusses label design and placement in cartography (e.g. Imholf et al., 1975) or annotation of Information Rich Virtual Environments, virtual graphics and 3D interfaces (Hartmann et a., 2005).

**Table 3.4. Studies, discussing parameters that are important for design of AR browsers**

| Reference | Year | Parameters identified as important | User-based study | Outdoor | Smart phone | Results |
|---|---|---|---|---|---|---|
| Azuma and Furmanski | 2003 | Overlap, placement | YES | NO | NO | Association can be achieved even if placement is suboptimal. |
| Leykin and Tuceryan | 2004 | Font size, font colour, contrast | YES | NO | NO | Legibility is only affected when text contrast is low. |
| Gabbard et al. | 2007 | Drawing style | NO | NO | NO | No empirical results. |
| Kim et al. | 2009 | Label size, colour, transparency, hybrid | Informal | NO | NO | Hybrid approach aids legibility |
| Wither et al. | 2009 | Location complexity<br>Location movement<br>Semantic relevance<br>Content complexity<br>Interactivity<br>Permanence | NO | NO | NO | No empirical results.<br>Taxonomy for analysis of AR annotations. |
| Jankowksi et al. | 2010 | Annotation style, font colour | YES | YES | NO | Billboard style most suitable for different backgrounds. |
| Choi et al. | 2010 | Grouping of labels based on distance to physical object | YES | YES | YES | Automatic grouping performs better. |
| Ganapathy et al. | 2011 | Density, accuracy, delay | YES | YES | YES | 7 annotations, up to 3 sec. delay. |

There are many aspects of a visual display that ultimately influence the usability of a product, as discussed in Section 3.3.2. The empirical studies described above emphasize the need to examine the impact of various design parameters on the utility and usability of AR annotations, delivered through AR browsers. Many other factors, however, have been proposed (Table 3.4). Their influence on the usability and utility of delivered content remains to be investigated further.

The problem is that most of the time only a few parallels are made between AR interfaces and mobile location-based services as tools for knowledge acquisition. As Table 3.4 illustrates, most empirical studies documented in literature that addressed directly the usability and utility of AR annotations were carried out indoors, or with head mounted displays.

### 3.5.2. Empirical studies with smartphone AR browsers

Despite their popularity, recent evidence suggests that the usability and perceived utility of AR browsers is very low. For instance, Olsson and Salo (2011) collected data from 90 early adopters of AR browsers through an online survey. Results showed that participants consider many aspects of current AR browsers problematic, including: content, technical and functional aspects, user interfaces and social aspects. AR content was deemed of poor quality, largely inappropriate, irrelevant and excessive. Technical problems included imprecise placement of annotations, software instability and bugs. The limited functionality of AR browsers and lack of social features were also criticized. While touching upon use of AR browsers in unfamiliar settings, this study considered mainly everyday situations and activities in urban areas. In addition, a significant limitation of the study was the lack of actual observation of users in real context of use.

An early study comparing access to POI information through different interface modalities was carried out by Fröhlich et al. (2006). The authors developed and compared empirically four different low-fidelity prototypes that allow access to information about POIs: (1) pointing gesture, (2) map, (3) radar and (4) Augmented Reality. The AR prototype received negative feedback from users and low subjective ranking compared to the map and pointing paradigm. The results from the study indicated that users preferred the map view for accessing information about remote (non-visible) POIs. A later study that investigated the use of a functioning prototype of an AR browser was documented by Turunen et al. (2010). During the study, ten users were instructed to walk across a park and use a prototypical AR browser in order to find information that annotates another user, walking towards them. Results show that users tolerated displaced annotations, but provided negative feedback related to the movement of virtual content within the display.

Similar results were obtained by Ganapathy et al. (2011). They carried out a field study, where 12 participants were asked to stand at a specific location (Portland waterfront, Oregon, USA) and carry out several tasks with an AR browser prototype (e.g. select the annotation for the Embassy Suits Hotel). Users were then interviewed about their experience. The results show that the participants preferred up to 7 annotations on the screen and were willing to tolerate up to 3 seconds delay for labels to appear on the screen. In terms of association, users indicated that they could tolerate up to 3.5 mm offset between the AR annotation and the physical object. The most interesting items that users wanted to acquire information about were visible objects, points of interest and restaurants. Addresses and public building names received a lower score. User feedback also addressed the way virtual annotations were represented on the screen. Users were dissatisfied with the simple blue dot that annotated physical objects and suggested that virtual AR annotations should be visualised with a wider variety of icons. These studies highlight the need to re-examine current user requirements for AR displays and study empirically how context of use influences usability.

An important aspect of investigating usability is that it can only be measured with representative users, and is only meaningful when evaluated in the context in which the product is used (Section 3.3). In the context of tourism, Toh et al. (2010) explored tourists' needs and requirements for smartphone AR. A set of contextual interviews and a field study resulted in a list of unmet needs of tourists. These included the need for translation of signs, more simple user interfaces, providing augmented photos of places that tourists could not visit, and effective navigation. One of the main conclusions for further research was the need to find appropriate ways to enrich tourists' experience through better content.

Probably the most relevant study investigating the use of AR browsers in urban tourism context was documented by Linaza et al. (2012). The authors carried out a field study with 15 users in the city of San Sebastian. The participants were asked to work with an AR browser for 1-1.3 hours and then interviewed about their experience. Overall, users placed very high importance "to the quality and quantity of the multimedia content available at each POI" (Linaza et al., 2012, p. 268). Results, however, show that test subjects found it difficult to understand the provided information.

In a more recent study, Lao and Humphreys (2014) examined the use of AR browsers (Layar) in everyday settings. They conducted interviews with 12 early adopters of AR. The study reported that participants used AR to enhance their interpretation and relationship with space. Informants reported that AR allowed for a heightened awareness of space, even when not using the technology and that AR can motivate people to scrutinize ordinary places. Another recent user-based study revealed that AR visualizations facilitate interaction with the environment (Cabral et al., 2014). In a study comparing the use of maps and AR, participants that worked with the AR interface interacted much more with physical space, compared to map users, whose attention was primarily directed towards the smartphone device (Cabral et al., 2014). Such findings emphasize the suitability of using AR for tourism.

Providing relevant information about POIs in the island of Corfu was supported by an AR browser (CorfuAR) developed to explore how such technology impacts emotions and user experience (Kourouthanassis et al., 2014). Results from empirical research indicated that the use of AR in tourism context was mainly associated with positive emotions. The filtering of content within the AR browser was based on three user profiles (thematic-based, entertainment-based and action-driven) adopted from the World Tourism Organisation tourists segmentation approach. A follow up user study, however, suggested that there was no difference in use or preferences between the personalised and non-personalised versions of the AR browser. One of the key conclusions from the study was the need for coming up with more usable and useful AR interfaces that minimise cognitive load.

Given the widely recognized need and importance of UCD in the AR field (Section 3.4.3), as well as the wide popularity of AR browsers, one would expect that there is a well-established body of literature that assess and documents their usability and utility. However, empirical studies that investigate the usability and utility of AR browsers have been very limited and far too little attention has been paid to travel and tourism contexts. As a result, designers are often forced to make design decisions blindly and without knowledge of how design parameters would influence usability and utility of AR browsers when used in actual context of use. Until now, there has been only limited attempts to identify and elicit user requirements when it comes to design of smartphone Augmented Reality browsers used in urban tourism context. Identifying user requirements for AR browsers and proposing new design principles for AR are also two of the objectives of this study.

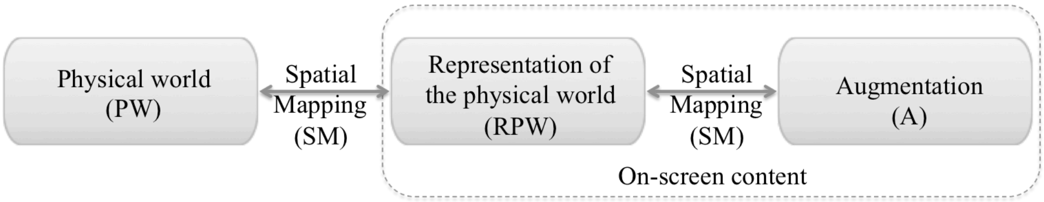
### 3.5.3. Existing design frameworks and guidelines for AR

A number of design principles and patterns for smartphone UIs have been proposed and described in literature (Fling, 2009; Banga and Weinhold, 2014; Neil, 2014). While trying to achieve generalisation, such guidelines rarely apply to design of user interfaces for different domains. Existing frameworks, models and design principles still lack specific directions that address the design of AR browsers. Indeed, while there are a number of commercial design solutions for AR, there is still a lack of rigorous research that classifies available visualization styles and techniques (Parker and Tomitsch, 2014).

Recently, Parker and Tomitsch (2014) proposed a classification of mobile AR visualizations based on the type of task that users will perform with such information: overview, zoom, filter, details-on-demand, relate, history, and extract. Their review found that fishbowl overview (a collection of points surrounding the user) together with 2D maps are the most common type of visualization in AR apps. They concluded that AR browsers should provide "customized filtering of information, finding related information and saving" (Parker and Tomitsch, 2014, p. 231). While the paper claims to provide directions for design of more usable AR browsers, the study makes recommendations out of context, not considering actual user needs. It is questionable whether and when users need filtering of information or saving content, as well as how these could be implemented in a useful and usable manner.

Until recently, AR was mainly the focus of study exclusively within the Computer Science and Computer Graphics domains. Few parallels are made between AR interfaces and mobile location-based services. This is also reflected in existing frameworks and models. For instance, Hansen (2006) developed an annotation taxonomy that tries to explain the different types of relations between annotations and reference objects. Alzahrani et al. (2011) developed a formal model for *physical annotations* (annotations about one or more physical entities), which consists of three main categories: 1) the annotation, 2) the physical entity, 3) the link between the annotation and the annotated target. Wither et al. (2009) proposed a framework that discusses several dimensions for AR annotations: location complexity, location movement, semantic relevance, content complexity, interactivity and annotation permanence.

**Figure 3.4. The framework proposed by Vincent et al. for analysis and design of AR browsers.**




*Source: Vincent et al., 2012*

Vincent et al. (2012) propose that an AR interface is composed of two layers (Figure 3.4): 1) the representation of the physical world, and 2) the digital augmentation with three mappings between them. The framework proposes two spatial mappings that describe how the AR interface relates to the physical space. The first spatial mapping describes the coupling between the physical world and the representation of the physical world on the screen of the handheld device. The second spatial mapping describes the position of the AR annotations in relation to the representation of the physical world within the AR interface. Both spatial mappings can have one of three properties: 1) conformal (absolute mapping), 2) relaxed and 3) none. While extremely useful for analysis of AR interfaces, the framework does not consider user characteristics. It is also limited because it only considers placement and position of annotations on the smartphone screen. Hence, other design parameters are excluded from analysis.

The overview of existing design and research frameworks for AR reveals the lack of incorporating and considering the user within design activities. The provided literature review emphasizes there is still lack of design theories and design knowledge, expressed as frameworks and captured in design principles and guidelines. This has led to development that is mainly focused on the technical aspects of AR, rather than developing and designing more usable and useful interfaces. This study builds on top of existing AR design theory. However, it is also directed at generating new insights regarding user requirements and how they can be satisfied in AR interfaces. By incorporating existing theories and new empirical data, this thesis proposes a new user-centred design framework that contributes to the general theory of designing AR information systems and their interfaces.

## 3.6. Chapter Summary

This chapter set to provide justification for the research in this study by examining current gaps in design knowledge related to smartphone AR used in tourism context. With the aim to improve usability and utility (Section 3.2), researchers need to investigate and elicit user requirements in actual context of use in order to determine the role and influence of contextual parameters on the use of mobile ISs (Section 3.3). To this end, User-Centred Design (Section 3.4) approach provides an overall mind-set and a framework for placing the user in the centre of design and structuring empirical work.

An overview of existing empirical work within the domain of AR (Section 3.5.1) and tourism (3.5.2) revealed that there are several important gaps that need to be addressed further. First, most empirical work within AR has focused exclusively on problems with legibility of AR annotations. Issues related to the utility of delivered content or other problems that users might experience in urban environments have remained unaddressed. Second, while there are several empirical studies that address the needs of tourists, most have placed accent on technical development and feasibility, rather than rigorous evaluation in actual context of use. It is evident that further work with representative users (tourists rather than urban residents), environments (urban settings, rather than natural environments), and tasks (obtaining information about the surroundings, rather than navigation) needs to be carried out. On a more general level, this lack of empirical work serves to explain the scarce research that proposes design guidelines or design frameworks (Section 3.5.3) that could be used to improve or evaluate the design of existing AR browsers. Overall, the presented analysis yielded justification for further theoretical and empirical work that will contribute to the development of design guidelines for more usable and useful AR browsers. The next chapter describes the UCD methodology chosen to guide the research activities.

# CHAPTER 4

# METHODOLOGY

**4.1. Introduction**

Describes the purpose of the chapter

**4.2. Research paradigm**

Justifies the adopted research paradigm

**4.3. Research Approach and Design**

Presents the adopted UCD approach and the scope and goals of mixed methods research designs

**4.4. Research Methods**

Describes the methods selected in each research stage

**4.5. Risks and Validity of Research Findings**

Reviews high-level risks and the measure undertaken to ensure the credibility of the results

**4.6. Chapter Summary**

Synthesis of the key points in the chapter

## 4.1. Introduction

In Chapter 2 it was established that there are still a number of challenges connected with the design of AR browsers, mainly driven by lack of knowledge regarding user requirements and needs. Chapter 3 then described User-Center Design (UCD) as an approach to generating design knowledge that is based on actual user needs. Identifying user requirements, problems and needs is critical in order to propose meaningful guidelines for design of smartphone Augmented Reality Browsers in tourism. Therefore, UCD was selected as the overarching methodology for this research. This offered particular value to achieving the aim of this research project, as it provided a framework and a structured approach that guided empirical data collection, and thereafter its analysis and translation into user requirements and design guidelines. This chapter begins by describing the adopted paradigm (Section 4.2) and overall research approach (Section 4.2) for this project. The specific data collection and analysis methods are further detailed and justified in Section 4.4. The chapter then concludes with a discussion of the judging criteria for the research and the strategies adopted to ensure the internal and external validity of the findings.

## 4.2. Research Paradigm - Pragmatism

The design of any product involves making a range of explicit and implicit assumptions reflected in the adopted epistemological and ontological stance of the designer (Hirschheim and Klein, 1989; Smith, 1997; Gregor and Jones, 2007). In Information Systems design research these assumptions affect the system under development and relate to the users (organization), the task at hand and what is expected from the designer (Hirschheim and Klein, 1989; Smith, 1997). Despite the need to make such assumptions clear, epistemological and ontological questions have received attention only recently in IS literature (Gregor and Jones, 2007). As a result, the field boasts "a rich tapestry of diverse research methods, paradigms, and approaches that are multidisciplinary and multi-national" in nature (Becker and Bjorn, 2007, p. 198).

A useful classification of epistemological and ontological stances that has been discussed most intensely in IS literature is the one developed by Burrell and Morgan in 1979 (Smith, 1997; Avison and Fitzgerald, 2002). The authors describe two epistemologies: subjectivist and objectivist. While the subjectivist position denies the natural sciences approach to study the social world, the essence of the objectivist position is to "apply models and methods derived from the natural sciences to the study of human affairs. The objectivist treats the social world as if it were the natural world" (Burrell and Morgan, 1979, p. 7). The same authors distinguished between integrationist (order) and coercionist (conflict) ontological views. The former emphasizes the order and consensus in society, while the latter stresses change and conflict (Burrell and Morgan, 1979).

Hirschheim and Klein (1989) adapted the proposed epistemological and ontological stances from Burrell and Morgan (1979) and mapped them on a continuum to represent the four different paradigms in Information Systems Development (Table 4.1).

**Table 4.1. Some of the most influential paradigms in Information Systems research**

| Paradigm | Systems development | Elements used in defining IS |
|---|---|---|
| **Functionalism** | Proceeds from without, by application of formal concepts | People, hardware, software, rules as physical or formal |
| **Social Relativism** | Proceeds from within, by improving subjective understanding and cultural sensitivity through adapting to social change | Subjectivity of meanings, symbolic structures affecting evolution of sense making |
| **Radical Structuralism** | Proceeds from without, by raising ideological conscience and consciousness | People, hardware, software, rules as physical or formal, objective |
| **Neohumanism** | Proceeds from within, by improving human understanding and the rationality of human action | People, hardware, software, rules as physical or formal; subjectivity of meanings |
| **Pragmatism** | From practice, by reason and actions that change existence. | Actions, knowledge, artefacts |

*Source: Hirschheim and Klein, 1989; Goldkuhl, 2011*

More recently, however, researchers have started to investigate the role of other paradigms into IS design and development. Within this area of research, it has been identified that pragmatism has played a crucial role in IS design research, even though practitioners and researchers within the field “seldom explicitly ground their research in the pragmatist research paradigm” (Goldkuhl, 2011, p. 141). An extensive review of design literature showed that such lack of explicit discussion is common not only to IS research, but also to other design fields, such as, among others, architecture, urban planning and design (Melles, 2008). One of the main reasons for this, Melles (2008) argued, is that each design discipline brings unique affiliations, views and practices with it.

Melles (2008, p.90) poses that pragmatism is not only intrinsic to design, but provides “a robust epistemological methodological terrain for design research”. According to Dewey (1929), the essence of pragmatism is action and change and the relationship between them. The central ideas behind Dewey’s pragmatism were recently reviewed by Biesta (2010). The paper discusses in detail the main constructs and ideas behind Deweyan pragmatism: (1) knowledge is derived from experience that supports action, (2) knowledge is concerned with the relationship between actions and consequences, (3) everyone’s experience (and knowledge) is equally real, and (4) action is situation-dependent. Such an account of the fundamental principles of pragmatism is not only compatible with, but also underpins the nature of design (Melles, 2008) and design approaches, such as user-centered design, collaborative design or interaction design. Having the primary aim to contribute to Information Systems Design theory, it was considered appropriate to base the research philosophy on the emerging paradigms within the field. Apart from being widely spread within the IS design discipline, pragmatism is also aligned with the main philosophy of User-Centred Design.

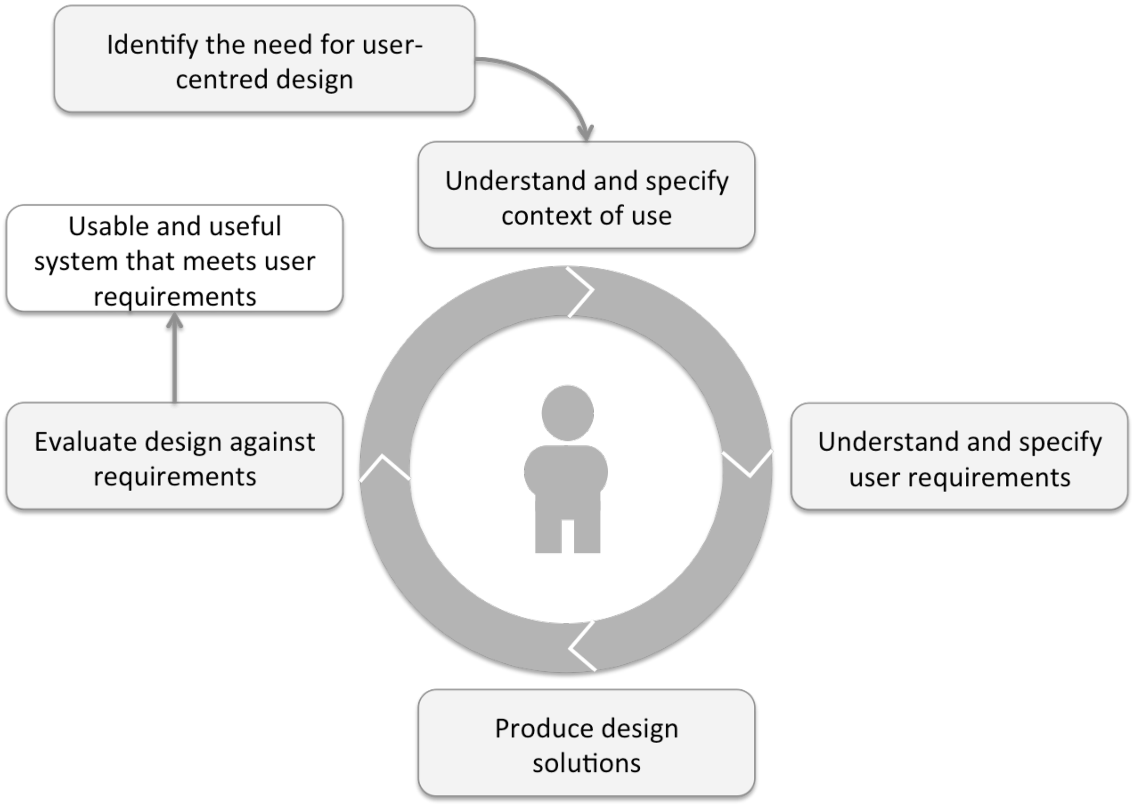
## 4.3. Research Approach and Design

### 4.3.1. High-level UCD Approach for AR

Design methodologies aim to provide designers and developers with a mental framework that organizes available techniques in different design stages. A number of comprehensive models are available to researchers for implementing a UCD methodology, complete with detailed discussions of possible activities for data capture, analysis, modeling and representation (Mayhew, 1999; ISO9241-210, 2011; Cooper and Reimann, 2014). The developers of these models, however, also acknowledge that it is not always feasible to employ every activity and technique, due to time, budget and other resource constraints (Mayhew, 1999).

A UCD lifecycle generally undergoes iteratively several key stages: (1) context of use analysis, (2) requirements elicitation, (3) design, and (4) evaluation (Figure 4.1) (ISO 9241-210, 2010).

**Figure 4.1. The User-Centred Design lifecycle**




*After: ISO 9241-210, 2010*

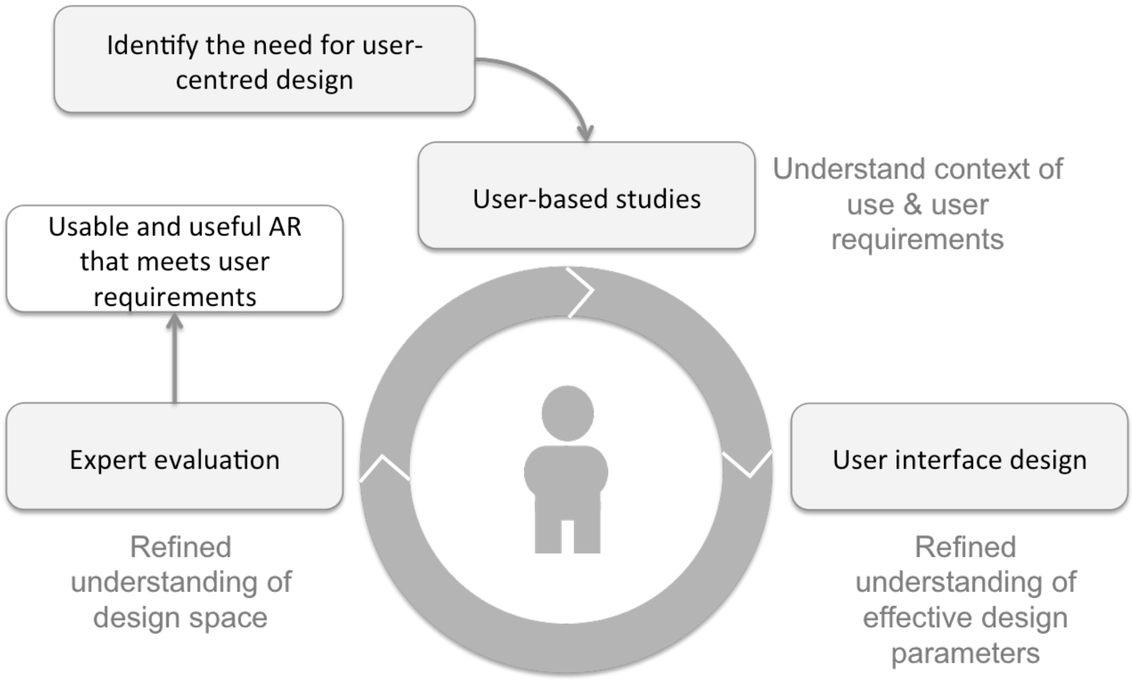
UCD normally starts with a thorough description and understanding of the ***context of use*** in which the interactive system is currently used or in which it will be used. Analysis is directed at identifying the most relevant context of use parameters (Section 3.3.2) that could potentially influence the use, usability and utility of the current or future system (ISO 9241-210, 2010). These include the current tasks of users, their characteristics, the environment or what users know and how they know it (Crandall et al. 2006). There are a number of data collection methods that allow deeper understanding of current or future contexts of use, including surveys, interviews, focus groups, but also observations and testing of how users use current products (Mayhew, 1999). When it comes to eliciting requirements for mobile devices, several reviews of mobile HCI methods (Kjeldskov and Graham, 2003; Kjeldskov and Paay, 2012) emphasized the need for user-based studies carried out on the field in actual context of use.

The primary goal of data collection in context of use studies is to understand user needs and what problems users experience that the system has to solve (Hacklay and Nivala, 2010). User needs and problems are captured and expressed as ***user requirements*** towards the future (or current) system. Notations and models have been developed to specify the elicited user requirements. Typically, usability experts use a type of graphical modeling language and a set of notations that reveal identified user behavior, the application's behavior and the physical settings where the system is or will be used (Mayhew, 1999). The activities involved in data gathering and requirements elicitation are closely related and iterative in nature. This is why often in literature they are treated as one single stage of a research project, often termed simply Requirements Analysis (Mayhew, 1999; Abras et al., 2004; Hacklay and Nivala, 2010).

Requirements analysis elicitation for Augmented Reality interfaces poses a range of unique challenges. In part, this is due to the relative youth of the field (Livingston, 2013), but also to the inherent challenge to design interfaces that combine physical and virtual spaces. Following an extensive review of AR literature (Swan II and Gabbard, 2005), Gabbard and Swan II (2008) argued that the novelty of AR interfaces, as well as the lack of design guidelines and heuristics, emphasizes the need for user-centred approach to design and development. In their widely cited paper, the authors proposed a UCD approach for AR interfaces that was later applied successfully in various domains (Gabbard et al., 2007; Linaza et al., 2012; Olsson et al., 2013). The approach is especially successful in new domains and for emerging technologies where there is lack of context of use understanding, design guidelines or heuristics. As discussed in Chapter 3, this is also the case with AR browsers used in urban tourism settings. This is why this approach was found suitable for the needs of the study (Figure 4.2).

The design lifecycle proposed by Gabbard and Swan II (2008) starts with ***user-based studies*** (Figure 4.2), which are "critical for driving design activities, usability, and discovery early in an emerging technology's development (such as AR). As the technological field evolves, lessons learned from conducting user-based studies are not only critical for the usability of a particular application but provide value to the field as a whole in terms of insight into a part of the user interface design space" (Gabbard and Swan II, 2008, p. 514). Observation of user performance is critical to understand the impact of different design parameters and what combinations support optimal user performance under various conditions.

**Figure 4.2. Adopted UCD approach in the thesis**




After requirements have been elicited, they need to be translated into specific ***design*** solutions (Figure 4.2). According to Hackos and Redish (1998) the process of translating user requirements to specific designs in order to arrive to an optimal design is a creative and open-ended process where many decisions have to be undertaken. Thus, the goal is to generate as much diversity as possible. Designers have a tremendous freedom in shaping up the design of a system according to their own understanding of the design space, the characteristics of the task at hand, or the users (Kling, 1977). However, design should not turn into a series of subjective choices based on personal preference, but rather be a tangible representation of product goals (Watzman, 2012).

Good design is an activity that reveals multiple solutions to a problem (Watzman, 2012). Therefore, it is a highly accepted practice to develop several versions of the design, or design alternatives (Cooper and Reimann, 2014). Developing multiple design solutions was also recommended by Gabbard and Swan (2008). When it comes to novel AR interfaces, user interface design activities help designers to "explore the design space prior to investing time in system development and, moreover, can explore a number of candidate designs quickly and easily" (Gabbard and Swan II, 2008, p. 515).

Each design alternative represents a specific *hypothesis* that can be directed at understanding whether: (1) the selected design elements will solve the problem, (2) whether the selected design elements fit the context of work, (3) whether the developed conceptual work model is accurate enough. Iteratively throughout the design process, the ***evaluation*** of developed design alternatives (Figure 4.2) plays a key role, as it provides further insight for designers (Mayhew, 1999). When the design alternatives have been created, Gabbard and Swan II (2008) recommend that these designs are evaluated through expert evaluations, or user-based studies. In cases where initial understanding of the design space has been achieved, evaluation of the design alternatives can be carried out through experimental user-based studies. With time, the accumulation of empirical findings in context of use studies and evaluation of designs provides a collection of informal design guidelines and metaphors available to designers and researchers (Gabbard and Swan II, 2008).

### 4.3.2. Mixed Methods Research Approach

The basic premise of scientific research is to gain knowledge using a structured and systematic approach. Involving the study of people, research within HCI draws heavily from social research within the realm of social sciences, such as anthropology, psychology and sociology (Grudin, 2012). There are three main approaches to undertaking social research, *qualitative, quantitative*, and *mixed methods* mainly distinguished by the particular data collection and data analysis techniques they employ.

Originating within the natural sciences, ***quantitative research*** involves the objective collection of data in order to test a theory or hypothesis (Creswell, 2013). Measurements are collected through predetermined instruments and are systematic, producing precise, quantitative information about reality (Kumar, 2014). The main goal is to quantify the extent of variation in a phenomenon. Following trends within psychology, early HCI research was based on standard quantitative human performance measurements (Lazar et al., 2010). Such measures are still relevant and widely used today. They are based on a task-centric model and include time for completion of tasks and number of errors. The model is based on the assumption that the usage of computers can be broken down into specific tasks, which can be measured in a discrete way (Lazar et al., 2010). Apart from being less time consuming (as opposed to qualitative data collection and analysis), quantitative data collection enables identifying causal relationships in a structured and systematic way as it provides precise (numerical) data. On the other hand, many of the phenomena that require the use of computers involve more complex measures, such as satisfaction, enjoyment, fun and aesthetics. More importantly for mobile devices, such a quantitative approach may prevent the researcher from capturing the influence of context or understanding why performance evolves in a certain way. When these are important a qualitative research approach is recommended.

***Qualitative research*** strategies and methods were developed within the social sciences to aid the study of phenomena and their meaning from the point of view of the participants involved (Creswell, 2013). A qualitative approach follows an "open, flexible and unstructured approach to enquiry" (Kumar, 2014, p. 14). One of the key elements of qualitative research is to observe participants' behaviour during an activity. The collected data then sheds light on how people carry out activities and why they engage in specific strategies. Becoming more common in the field of HCI, qualitative research represents an exploratory approach to studying why users engage in various use strategies and involves interpretive approaches of analysis such as ethnography, case study and phenomenology (Lazar et al. 2010). A qualitative research approach has a number of benefits for improving the design of ISs: it allows revealing user perceptions, experiences and feelings about interfaces; enables researchers to understand better the influence of context of use; allows collecting rich and explanatory data, which might not be anticipated by the researcher (Lazar et al., 2010). Qualitative research is, however, time consuming. More importantly, most qualitative research approaches are often critiqued for being too subjective and undermine the importance of performance when it comes to design of computer interfaces.

***Mixed methods*** is a third relatively new methodology that combines both qualitative and quantitative research techniques. Originating in the 1980s, it was developed to reduce the respective limitations and biases of quantitative and qualitative approaches, and also enables the result of each technique to inform subsequent actions (Creswell, 2014). Although differences of opinion still exist (Crewsell and Plano Clark, 2011), a generally accepted definition was proposed by Johnson et al. (2007). In their highly cited paper, the authors define mixed methods as "*the type of research in which a researcher or team of researchers combines elements of qualitative and quantitative research approaches (e.g. use of qualitative and quantitative viewpoints, data collection, analysis, inference techniques) for the purposes of breadth and depth of understanding and corroboration*" (Johnson et al., 2007, p. 123).

The procedures adopted in mixed methods research can be informed by a theory or a framework and are incorporated into a specific mixed methods study design. Within HCI and design research, both quantitative and qualitative data are frequently collected and analysed (Lazar et al., 2010). This is also how UCD operates, with the combination of various techniques enabling the analysis of both qualitative and quantitative data (Gould and Lewis, 1985; Abras et al., 2004). Ensuring meaningful and comprehensive results during UCD for smartphone AR and context-aware applications also favours the use of mixed, multi-method, multi-data and multi-analysis methodological approach (e.g. Oulasvirta et al., 2009; Ajanki et al., 2010).

When adopting a mixed methods approach, researchers need to select and specify the research design of the study that guide data collection, analysis and interpretation of the data (Creswell and Plano Clark, 2011). There have been a number of typologies proposed in literature that describe the various alternative mixed methods designs (Teddlie and Tashakkori, 2010; Creswell and Plano Clark, 2011; Creswell, 2013). The most widely cited classification was proposed by Creswell and Plano Clark (2011) where the authors distinguish between three basic (convergent parallel, explanatory sequential, exploratory sequential) and three advanced (embedded, transformative, and multiphase) designs. Studies can either rely on the described designs (typology-based approach) or consider and interrelate multiple components from them (dynamic approach) (Creswell and Plano Clark, 2011). The goals and scope of the current study were considered accordingly, with the decision ultimately made to approach the UCD research methodology from a mixed methods perspective, involving the collection of both qualitative and quantitative data. This is because identifying problems and user requirements for more effective AR information systems requires gathering both quantitative performance data, but also qualitative behavioural and attitude observations.

### 4.3.3. Pragmatic Interpretivist Approach

As with other disciplines, which are an amalgamation of other scientific fields, Information Systems and Human-Computer Interaction both lack "explicit discussion of their underlying epistemological commitments" (Harrison et al., 2007, p.1), or a rigorous body of work that describes research paradigms and worldviews. On one hand, this is because of the influence of a broad range of disciplines, each with its own worldviews. On the other, both HCI and IS deal with what Simon (1969) called "the sciences of the artificial", or building artefacts. The sciences of the artificial have their roots within engineering, design, human factors, but also within psychology, and the social sciences. Researchers within those fields adopt and use a paradigm that is similar to the field they have experience with, without discussing underlying worldviews or epistemological commitments (Harrison et al, 2007).

Recently, researchers started discussing the various research paradigms and their use within Computer Science and Information Systems (Villiers, 2005), as well as within HCI (Goldkuhl, 2011). Understanding HCI history is largely about understanding a series of paradigm shifts. Following the scientific tradition in other disciplines and fields, early Information Systems research was grounded in the positivist tradition. The same was the case with HCI, which followed mainly trends and methodology within Psychology (Harrison et al., 2007).

Equated with the scientific method, the positivist paradigm holds that knowledge is absolute and objective. Positivist research findings are usually collected through quantitative research methods. In positivist research, the researcher is independent from the study. However, when it comes to human behaviour, multiple interpretations can exist and more recently, researchers focused on conducting interpretivist research. Interpretivism aims to find new interpretations or underlying meanings and adheres to the ontological assumptions of multiple realities. Originating within the social sciences, interpretivist research is now becoming more widely accepted within Information Systems (Roode, 2003) and Human Computer Interaction (Lazar et al., 2011). Interpetivism results with subjective findings, which may differ among researchers. It is, nonetheless, appropriate and valuable view for studies that examine complex behaviour and phenomena.

Trauth et al. (2001, p.7) states that "interpretivism is the lens most frequently influencing the choice of qualitative methods". However, Goldkuhl (2011) argued that this in not necessarily the case, especially when it comes to IS research. Reviewing existing literature, Goldkuhl (2011) emphasises the appropriateness of pragmatism and its effectiveness as a paradigm for IS research. Since pragmatism is concerned with action and change, the author argues that this is especially suitable worldview for IS research, which introduces new artefacts in the world and studies their influence on organisational and individual behaviour.

The process of acquiring information within a multi-faceted and layered urban environment is complex and contextually dependent on multiple factors. It concerns interpretation and meaning making within a complex physical world through a smartphone device that mediates the experience. Therefore, this study adopts a pragmatic interpretivist approach to understand tourists' requirements for design of more useful and usable AR browsers. The study explores the interaction between the tourist, their context and the AR interface, which is assessed by parallel and sequential analysis of qualitative and quantitative data. Qualitative research is carried out as part of a pragmatism research paradigm (Goldkuhl, 2011). This approach is particularly suitable when it comes to HCI research and, in particular, new design theories for information systems, as it relies on "action, intervention and constructive knowledge" (Goldkuhl, 2011, p.1). As a result, the final framework and the outcome of the thesis are based mainly on qualitative findings. This means a lower degree of certainty, which is natural for qualitative (interpretative) research.

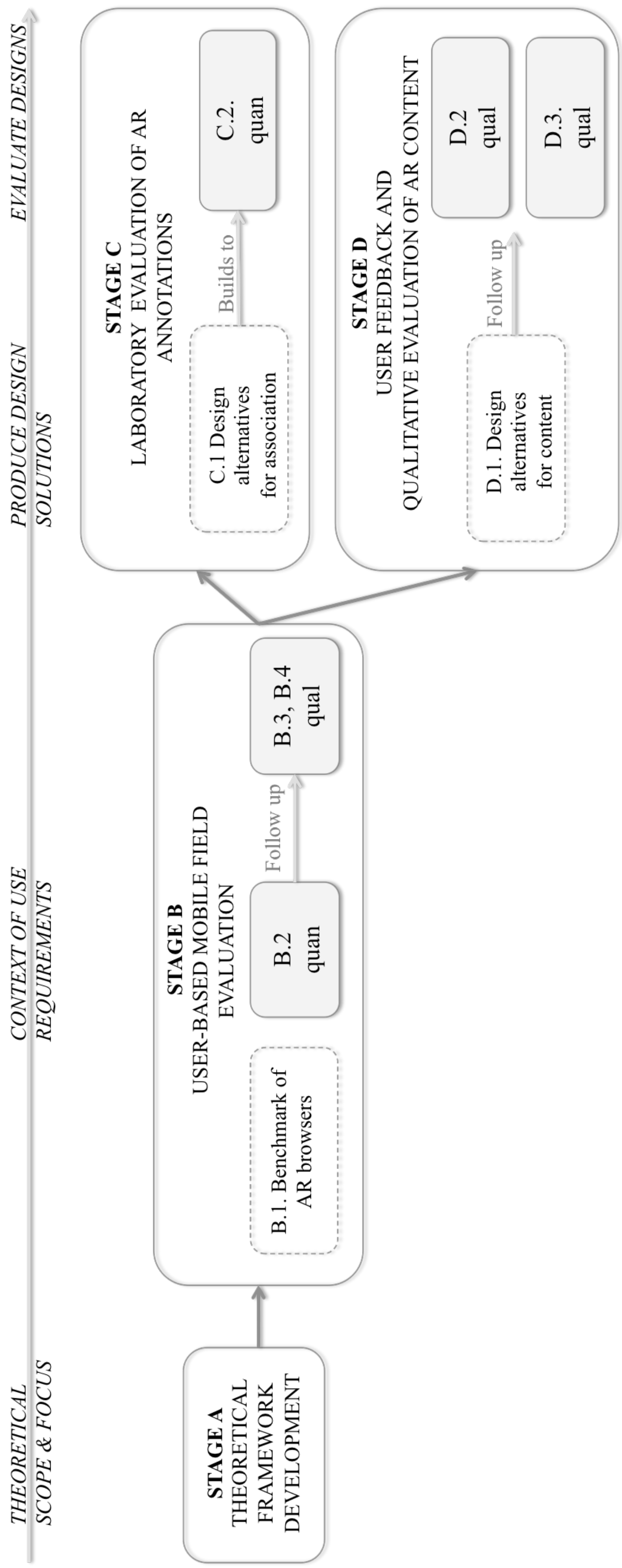
## 4.4. Research Methods

There is still a lack of design knowledge expressed as design guidelines and principles that prescribe how to implement usable and useful AR browsers for tourists in unfamiliar urban environments. Partially, this is due to the fact that design and development of AR interfaces is mainly advanced in disciplines outside of tourism and geo-information science where focus is placed on AR as a gadget that changes visual perception, rather than a tool that facilitates information acquisition about large-scale unfamiliar environments. This has also led to the lack of empirical evaluations of AR annotations that elicit on-site tourists' requirements for more efficient and effective AR browsers. To address such gaps, and following a mixed methods user-centred design approach, this research project employed a series of data gathering, analysis and modeling techniques, broadly separated into five stages (Table 4.2).

In considering the selection of a mixed methods design strategy, several factors were considered, following the recommendations of Maxwell and Loomis (2003): the study's purposes, conceptual framework, research questions, methods, and validity considerations. The literature review (Chapter 2 and Chapter 3) facilitated identifying and formulating research questions and revealed the lack of a conceptual or theoretical framework that prescribes or captures guidelines and principles for design of usable and useful AR browsers. The development of such a framework is closely related to identifying the most relevant kernel theories (Section 3.4.1) that will guide the design of the future system. Theoretical frameworks are commonly used to guide and direct empirical (mixed methods) research (Corbin and Straus, 2008; Creswell, 2014). This is why the first stage of this study comprised the design and development of a theoretical framework, described in Chapter 5,that further helped to focus the study and select appropriate research design. Apart from identifying relevant constructs explored further in the study, the framework is presented as a visual model that captures the key relationships between such constructs.

The framework was also consulted when selecting and adopting the mixed methods design for the study. Ultimately a dynamic approach was adopted, which uses elements of convergent, exploratory and explanatory sequential designs (Figure 4.3). A convergent parallel design uses concurrent timing to collect, analyze and interpret quantitative and qualitative data (Creswell and Plano Clark, 2011). This approach places equal priority on both methods. Collecting both qualitative and quantitative data in the initial phase of the study (Requirements Analysis) was considered highly beneficial and in line with the identified objectives and research questions (Chapter 1). On one hand, quantitative data allowed objective measurement of user performance with AR interfaces. On the other, qualitative data enabled revealing users' reasoning, cognitive patterns and strategies. Analysis of quantitative (quan B.2, Figure 4.3) and qualitative (qual B.3, B.4, Figure 4.3) data were kept separate. Unlike the more traditional convergent parallel design, however, in this study qualitative data were used to shed light and gain insights on the obtained quantitative results.

**Figure. 4.3. Mixed methods strategy adopted for the study**



The next part of the study incorporated elements of an exploratory sequential design. In essence, a second quantitative study was conducted (quan C.2, Figure 4.3) where several hypotheses about design of AR browser annotations were tested. The main aim of the quantitative study was to test and generalize the initial findings, as recommended in Creswell and Plano Clark (2011). On one hand, the purpose was to confirm observations and insights obtained through qualitative data and analysis in the first part of the research project. In addition, the quantitative results were also used to obtain further insights on user requirements and effective design of AR browsers. While the quantitative study was directed at user performance and usability (effectiveness and efficiency), two additional qualitative studies (qual D.2 and qual D.3, Figure 4.3) were conducted as a follow-up. The main aim was to explore further in depth issues related to satisfaction and utility (content) of AR browsers. Each of the five research stages is further described below.

### 4.4.1. Stage A – Theoretical Framework Development

The design of AR browsers, or any information system, is a complex research problem that could be explored through the lenses of a range of theoretical perspectives. Being interdisciplinary in nature, there exist a number of theories, frameworks and models within HCI and ISs that could be used to drive such a research project. The main aim of this study is to advance IS design theory that prescribes the design of more usable and useful AR browsers. Following the general process of design theory generation (Walls et al., 1992; Hevner et al., 2004; Arazy et al., 2010), the first step was the identification of kernel theories, or the most relevant empirical research, concepts, frameworks, models and constructs that could help in understanding the design space for AR browsers used in urban tourism context.

**Table 4.2. Research stages and methods in the thesis**

| Stage | Method | Research approach | Why | Where / With whom | How / Analysis |
|---|---|---|---|---|---|
| **A** | ***Theoretical Framework Development*** | | | | |
| **A.1.** | Literature review in several relevant domains; identification and synthesis of kernel theories | --- | **Needed for:** theoretical basis of the study; focus of data collection activities; identifying major constructs and relationships | Research laboratory | Reading HCI, eTourism, Geo-Information Science and Environmental Psychology literature |
| **B** | ***Mobile Field Based Evaluation of AR Browsers*** | | | | |
| **B.1** | Comparative evaluation and benchmarking of AR browsers | --- | **Needed for:** classification of AR browsers' design parameters **In order**: to select representative designs for data collection | **Where**: Bournemouth, UK London, UK | Identification of similarities and differences in design of annotations |
| **B.2** | Observation of task performance | **Quantitative** | **Needed for:** systematic identification of problems and influence of design on task performance **In order:** Identify important design parameters; | **Where**: Bournemouth city centre, UK **With whom**: 14 participants | Quantitative Analysis through general linear models and statistical procedures |
| **B.3** | Think-aloud protocol | **Qualitative** | **Needed for:** Identification of cognitive behavior and reasoning about designs **In order**: understand why tourists experience problems; what strategies they adopt when using AR browsers | | Protocol Analysis (Exploratory Sequential Data Analysis) |
| **B.4** | Contextual interviews | **Qualitative** | **Needed for:** User feedback regarding design; **In order**: Understand attitudes, opinions and preferences towards designs | | Qualitative Analysis |

**Table 4.2. Research stages and methods in the thesis (Continued)**

| Stage | Method | Research approach | Why | Where / With whom | How / Analysis |
|---|---|---|---|---|---|
| **C** | ***Laboratory Evaluation of AR Annotations*** | | | | |
| **C.1** | Design Concepts and Artifacts | --- | **Needed for:** Translating requirements into specific designs in order to test research hypotheses | Research laboratory | Preparing several design alternatives implemented as mock-ups |
| **C.2** | Lab-based experiment | **Quantitative** | **Needed for:** Testing the strength of identified relationships between design and user performance **In order to**: identify key design parameters that improve use of AR browsers | **Where**: Bournemouth University, UK **With whom**: 90 participants | Quantitative Analysis through general linear models and statistical procedures |
| **D** | ***User Feedback and Qualitative Evaluation of AR Content*** | | | | |
| **D.1.** | Design Concepts and Artifacts | --- | **Needed for:** Translating requirements into specific designs in order to test research hypotheses | Research laboratory | Preparing several design alternatives implemented as mock-ups |
| **D.2.** | Pluralistic walkthrough with domain experts | **Qualitative** | **Needed for:** Identifying additional concerns about design of AR browsers **In order to**: Understand thoroughly user requirements | **Where**: Innsbruck, Austria **With whom**: 9 participants | Qualitative Analysis |
| **D.3.** | Field-based Pluralistic walkthrough | **Qualitative** | **Needed for:** Obtaining feedback regarding content in AR browsers **In order to**: Understand user requirements towards AR content | **Where**: Paris, France **With whom**: 10 participants | Qualitative Analysis |
| **E** | ***Conceptual User-Centred Design Framework for AR Browsers*** | | | | |
| **E.1.** | Data Integration | --- | **Needed for**: Integration of data from all research phases **In order**: To enable triangulation and holistic understanding | Research laboratory | Conceptual Design Framework development |

The process started with review of literature within several domains, including tourism, eTourism, geo-information science, environmental psychology, information science and geospatial cognition. The development for the framework was necessary for two main reasons. First, as opposed to previous research within AR (Chapter 3), this study adopts an innovative perspective towards AR browsers as tools for geospatial knowledge acquisition for large-scale urban environments. As such, the framework facilitated identification of the most important constructs and relationships that determine and influence this process. From this perspective, and building on previous research within AR and geo-information science, it was used to deconstruct the AR interface and identify potentially important design parameters that could influence the usability and utility of AR browsers.

Second, the use of AR browsers in tourism context can be influenced by a number of factors (Section 3.3). Eliciting user requirements, then, would require extensive and long data collection in many different settings. Apart from its resource-intensive nature, such evaluation might be flawed, as many different factors could influence the results, posing a threat to the internal validity of the data. In order to focus data acquisition, the theoretical framework was necessary to identify the major potential contextual parameters that could influence interaction with AR browsers in tourism context.

### 4.4.2. Stage B – Mobile Field Based Evaluation of AR Browsers

Following the recommendations for adopting a UCD approach to design of augmented reality described in Section 4.3.1 (Gabbard and Swan II, 2008), the second stage of the research project started with the preparation of a user-based study, aiming to explore and elicit user requirements towards AR browsers. In this stage, a choice had to be made among alternative user-based data collection methods.

#### *4.4.2.1. Data Collection Methods*

There are a number of traditional approaches (e.g. Nielsen, 1993; Mayhew, 1999; Kjeldskov and Graham, 2003; Lazar et al., 2010; Allen and Chudley, 2013) and innovative methods (e.g. Bowser et al., 2013) used to obtain empirical data during requirements elicitation for mobile products. These can be broadly separated into two categories: *observational* and *inquiry methods* (Karat, 1997).

***Inquiry methods***, such as questionnaires, interviews and focus groups, are based on asking current or potential users about their tasks, characteristics, environments, and/or their opinions, attitudes and perceptions with regard to a current or future product (Lazar et al., 2010). Questionnaires allow the fast collection of data from geographically distributed populations and are used within mobile HCI to gather information about current problems with mobile products, or eliciting user requirements and needs (Kjeldskov and Graham, 2003). Interviews and focus groups allow obtaining direct feedback from users and can provide deep insight into user tasks, their current context, problems that they experience or current and future needs (Lazar et al., 2010). Inquiry methods, however, suffer from several major limitations. All three of the methods can lead to problems with recall, since data collection is separate from the task and context under consideration. Users might not remember or be able to recall what they actually do, use or what information they need, and this is why user requirements and needs might be biased or limited. This is especially the case when it comes to novel interfaces and interaction metaphors, such as Augmented Reality.

In such research context, ***observations*** of actual use and interaction with a product are more beneficial (Kumar, 2014). User observation involves users accomplishing specific representative activities with a product. Its main advantages reside on the richness of the gathered data and the unbiased perspective of the workflow. Observational research in HCI often provides the richest insights with regard to how people use technology, how different design features or the context surrounding interaction with an information system influence behavior (Lazar et al., 2010). Improving the usability and utility of AR browsers requires a deep understanding of the problems users experience in actual use settings, as well as the key factors that influence such use. This is why observing tourists and their performance with AR was considered suitable for the purposes of this study.

Within HCI, observational studies have typically been carried out in a research laboratory (Mayhew, 1999; Lazar et al., 2010; Kjeldskov and Skov, 2014). Due to the specific nature of mobile interaction, arguments have been made for carrying out experiments on the field, in actual context of use, especially when it comes to mobile interfaces and Augmented Reality (Kjeldskov and Graham, 2003; Oulasvita et al., 2009; Oulasvirta, 2012; Kjeldskov and Skov, 2014). One of the primary concerns, among others, is that "phenomena critical to usability might be difficult or impossible to study in the lab" (Oulasvirta, 2012, p. 60). The primary benefit of this approach is that field studies can reveal how users interact with different environments (Oulasvirta, 2012; Kjeldskov and Skov, 2014), which is especially important considering the fact that physical context is part of the AR browser's interface.

One of the most common types of user observations are ***usability studies*** (Lazar et al., 2010). Usability testing involves representative users attempting representative tasks with a prototype or a product in representative environments (Spool et al., 2008). This method is increasingly being adopted to understand how people work with hand-held devices, such as smartphones. Usability testing is directed towards finding user interface flaws that cause problems for users and challenges that they experience with an interface (Spool et al., 2008). Recently, a number of authors have adopted and contributed to the development of mobile usability testing methods (Rosenbaum and Kantner, 2007; Oulasvirta et al., 2009; Oulasvirta, 2012; Kjeldskov and Skov, 2014). In line with the identified research objectives, a mobile usability study was also considered suitable, as it allowed gaining deeper insight on the problems users experience with AR browsers, as well as understanding the contextual parameters that influence knowledge acquisition through such applications in large-scale urban environments.

When observation is directed at understanding how users work with a specific product, people can be asked to verbalise their thinking during or while they interact with an interface. This method is also known as ***think-aloud protocol*** (van Someren et al., 1994). It is one of the most widely used methods in usability studies (Lazar et al., 2010) and has been recognized as the most valid approach to obtain complete data on cognitive processes (Ericsson and Simon, 1993). In this study, think-aloud involved the analysis of recorded verbal and action protocols (van Someren et al., 1994) that resulted from asking participants to voice their thoughts when executing several preliminary identified tasks with AR browsers.

Apart from understanding the participants' reasoning strategies, it was also considered necessary to obtain direct feedback from users regarding their attitudes, opinions and perception towards AR browsers. To this end, field-based studies often contain a form of interview, referred to as ***contextual inquiry*** or *situated interview* (Rosenbaum and Kantner, 2007). Contextual inquiry (CI) is a qualitative requirements gathering and analysis method, originally adapted from the fields of psychology, anthropology and sociology (Holtzblatt et al., 2005). Contextual inquiries are focused primarily on the context of use of products, as participants are asked to explain how and why they do things (Holtzblatt et al., 2005; Rosenbaum and Kantner, 2007). This approach allows obtaining additional shared understanding of what is happening and uncover the meaning of actions, which aids the interpretation of interaction in a later stage. Data derived in this manner overcomes the drawback of discovering tacit knowledge about unconscious and habitual work practices (Holtzblatt et al., 2005). This type of UCD technique was therefore particularly suitable for revealing tourists requirements towards AR browsers in context and was selected in order to augment the data collection in this study.

Field-based user studies can be carried out with research prototypes or existing (commercial) products. The evaluation of existing systems is highly recommended when there are already developed applications available. The use of (several) available systems can shed light on the impact of various design parameters on user performance and problems (Spool et al., 2008). Considering the availability of AR browsers and their wide popularity among consumers, it was considered appropriate to focus the first empirical data collection in this study by using already existing commercial AR browsers. In 2012, when the design of the study was being prepared, there were more than 500 commercial AR applications on the various smartphone app stores. In order to select representative designs for the empirical evaluation, it was first necessary to obtain a thorough understanding of available designs through an overview of existing solutions, how they work, their advantages and problems. Formally, this analysis was carried out in the form of a ***comparative evaluation and benchmarking*** (Allen and Chudley, 2013), where the primary aim was to classify existing design and select the most representative sample for follow up evaluation. In essence, the researcher selected and tested how 23 AR browser applications work at two locations (Bournemouth city centre and London). In total, 1500 screen shots were collected and further compared in order to detect general design patterns and differences among the AR browsers. This allowed the selection of the final 4 browsers that were used during the mobile field evaluation with representative users.

#### *4.4.2.2. Analysis, Modeling and Data Representation*

Analysis of data obtained through usability testing are similar to other types of mixed methods research, as it involves both quantitative and qualitative analytical procedures (Lazar et al., 2010). ***Quantitative data analysis***, such as the time it takes participants to complete tasks, was carried out using descriptive statistics and General Linear Models, with the type of AR interface as a between-subject independent variable. Descriptive statistics (mean and range) (Lazar et al., 2010; Field, 20013) were used as a preliminary indication of user performance with the various AR annotation designs. Inferential statistics (Field, 2013) were then used to explore differences in objective (interval) measures (ANOVA) and subjective (ranking) measures (Kruskall-Wallis).

Afterwards, ***qualitative analysis*** was carried out. The main purpose was to gain understanding of user tasks, as well as the emerging cognitive patterns and reasoning behind them. This is also the primary goal of ***Cognitive Task Analysis*** (Chipman et al., 2000), which provides systematic procedures for understanding cognitive processes behind tasks. More than 60 different approaches to cognitive task data analysis and representation have been developed (Crandall et al., 2006). When it comes to analysis of activities that unfold in time and data obtained through think-aloud protocols, Sanderson and Fischer (1994) recommend using ***exploratory sequential data analysis (ESDA)***. ESDA examines activities as they unfold sequentially in time. The difference in analysis is that it emphasizes preserving the integrity of events as they occur naturally with time (Sanderson and Fischer, 1994). Many ESDA techniques have been developed or ported from psychology, social sciences and other fields. Considering usability studies and think-aloud analysis, however, ***protocol analysis*** is common among usability practitioners (van Someren et al., 1994). These techniques rely primarily on ***qualitative*** analytical strategies (Creswell, 2012), where content is first transcribed, coded and then thematically organized to uncover meaning and patterns. This was also the approach adopted for analyzing behavioural data captured within the video/audio recordings from the mobile field evaluation.

Qualitative data analysis approach was also used with regard to the obtained user feedback and data during the contextual inquiries. The analysis consisted of three primary stages (Lazar et al., 2010; Creswell, 2012): (1) identification of the major components or themes (coding), (2) making connections among the themes (axial coding), (3) integrating the data around a central theme in order to understand the phenomenon under study (selective coding).

Both qualitative and quantitative data were then used to ***model user behaviour***. All models in HCI use some form of a graphical language to represent existing knowledge about work (Beyer and Holtzblatt, 1998). Work models are common to all UCD methodologies (Mayhew, 1999), as they provide an external representation of work and can be used to share and communicate knowledge among teams. They also serve as a tool for the designer to check whether he/she is not forgetting some aspect that will cause their design to fail. Several different types of models were considered (Beyer and Holtzblatt, 1998), including flow, sequence, physical, and cultural. ***Task sequence models*** were considered most suitable, as they focus on description of an individual sequence of work, illustrating in detail how a user accomplishes a goal in a specific work instance (Beyer and Holtzblatt, 1998). Even though sequence models resemble standard task analysis models (e.g. flow charts, procedural work models), they are more than simple representations of observed behavior. Sequence models include trigger actions and focus on *why* the user is doing what they are doing, incorporating cognitive elements, such as reasoning and problem solving behavoir (Holtzblatt et al., 2005). Individually, sequence models are suited to represent activities in fixed and mobile settings.

Once a task sequence was developed for each user and for each task situation, the models were used to recognize patterns among users and identify the common structure of work. This is the essence of what Beyer and Holzblatt (1998) call the ***consolidation process***. In essence, the developed models were merged, excluding fine detail and using only high-level patterns. This final consolidated model captured the “single statement of the practice that must be supported, improved, replaced or obviated” (Beyer and Holtzblatt, 1998, p.140). This step was important, as the resulting consolidated model developed in this study provided a first approximation of a general framework that can be used for design of AR browsers.

#### *4.4.2.3. Development of Design Concepts and Artifacts*

After data analysis, user requirements were translated into specific ***design concepts*** and artefacts. Concept development is an iterative process of “re-evaluating and combining existing concepts, visualizing them and validating them with users” (Nieminen et al. 2004, p.227). Good design is an activity that reveals multiple solutions to a problem (Watzman, 2002). Therefore, it is a highly accepted practice to develop several design concepts also called ***design alternatives*** (Allen and Chudley, 2013). The main goal is to generate as much diversity as possible to explore the benefits and drawbacks of each design. Designers have a tremendous freedom in shaping up the design of a system according to their own understanding of the design space, the characteristics of the task at hand, or the users (Kling, 1977). However, design should not turn into a series of subjective choices based on personal preference”, but rather be a tangible representation of product goals (Watzman 2002). In the end of this stage it is often good to have a specific artefact that can be evaluated with users (Mayhew, 1999; Lazar et al., 2010).

Together with user requirements and goals, several visualisation techniques were considered that allow transforming design concepts into specific design artefacts: scenarios, storyboards, 2D/3D models, mock-ups, low-fidelity (paper) or high-fidelity (functional) prototypes. During the design process of a certain software program, product mock-ups allow designers to test their ideas and concepts (Beyer and Holtzblatt, 1998; Mayhew, 1999). At the end of Stage B, the development of such design artefacts was necessary for two reasons. On one hand, such artefacts represent specific hypotheses about design and could be used to test observed and identified relationships and patterns during the field based evaluation of AR browsers. To this end, design artefacts in the form of AR mock-ups were used in a follow-up laboratory experiment (Stage C). On the other, such design artefacts can be evaluated by users or domain experts (Gabbard and Swan II, 2008), and allow understanding further user requirements (Stage D).

### 4.4.3. Stage C – Laboratory Evaluation of AR Annotations

The field-based evaluation (Stage B) provided rich insights as to how and why tourists work with AR browsers and the influence of the physical environment on the process. Subsequent consolidation of the task sequence models, developed as part of the analysis procedures, revealed the common determinants for errors and problems that users experienced when trying to obtain knowledge about unfamiliar urban environments. Based on this acquired understanding, several hypotheses were formulated regarding the major implications for design of AR annotations. The hypotheses were translated into specific design alternatives, implemented as design mock-ups. In order to increase the generalizability of results and confirm the hypotheses, the mock-ups were tested with 90 representative users in a lab-based experiment. ***Experimental research*** enables the identification of causal relationships among variables. Influenced by the predominant paradigm in psychology, experimental research has been applied widely to HCI (Oulasvirta, 2009). The adopted procedures followed standard experimental design used within HCI (Lazar et al., 2010). Even though the study aimed at simulating actual use of AR browsers in different contexts of use, this approach allowed to control the experimental settings in order to increase the generalizability of the findings (Shadish et al., 2002). Due to their increased power in identifying causal relationships, an experimental set up was also considered suitable in order to test the identified hypotheses.

### 4.4.4. Stage D – Qualitative Evaluation of Alternative AR Designs

Having obtained a deeper understanding of factors that influence usability of AR annotations (Stage C), it was considered necessary to obtain a holistic perspective on the determinants of their utility (e.g. relevance and usefulness of content). To this end, elicited user requirements towards content (Stage B) were translated into several design concepts and tangible design mock-ups. Several alternative approaches were then considered in order to assess the developed design alternatives. Iteratively throughout a user-centred design lifecycle, ***usability inspections*** help to evaluate new design concepts, ideas and alternatives (Nielsen and Mack, 1994). The main benefit of the method is that it allows a thorough and wide evaluation of design artefacts and revealing a large set of potential usability problems (Cockton et al., 2009). Such activities help designers and researchers to understand better what are the most effective design parameters for a specific interface. Usability inspections are especially useful when it comes to design of novel Augmented Reality interfaces and are a highly recommended activity (Gabbard and Swan II, 2008) that ultimately contributes to the development of new design guidelines for such interfaces.

Originally introduced by Nielsen and Mack (1994), there are a number of usability inspection methods available to researchers (Nielsen and Mack, 1994; Preece et al., 2002; Cockton et al., 2009), including heuristic walkthrough, cognitive walkthrough and pluralistic walkthrough. Heuristic evaluation and cognitive walkthroughs are two of the most commonly used methods. During heuristic evaluation, an evaluator uses a small set of design principles and rules of thumb (Nielsen and Mack, 1994) to confirm usability problems and possible design violations. This is one of the most commonly adopted usability inspection methods (Cockton et al., 2009). The main problem is that the evaluation relies on available design guidelines and know-how that are yet to be developed for AR browsers. Alternatively, during a cognitive walkthrough (Lewis and Wharton, 1997), analysts evaluate an interface by stepping through each task that users will carry out and discuss potential problems.

Cognitive walkthroughs can be carried out by individual evaluators, or in a group session as a *pluralistic cognitive walkthrough* (Bias, 1994). A pluralistic walkthrough allows gathering feedback from multiple perspectives in a single session. This approach was selected as the method enabled obtaining a more holistic understanding on potential usability and utility problems in various settings and scenarios. In the final stage of this research project, two pluralistic cognitive walkthrough evaluations were carried out. The first pluralistic evaluation followed a traditional set up where: (1) domain experts were presented with each AR annotation design alternative; (2) asked to assume the role of the user; (3) allowed time to write down their comments and concerns; and (4) asked to discuss each design alternative as a group. Confirming that novices could be trained successfully to carry out pluralistic walkthroughs (John and Packer, 1995), Folstad et al. (2012) recommended that domain experts or work-experts are included in the evaluation. Therefore, the first pluralistic walkthrough involved experts from several domains, including Tourism, eTourism and Augmented Reality.

Because of the strong connection between virtual AR annotations and physical environments, the second qualitative pluralistic walkthrough evaluation was carried out on the field. In essence, two groups of 5 evaluators walked a pre-determined route and discussed consecutively several design alternatives of AR browser annotations. The main advantage of the adopted approach was that evaluation and discussions were carried out not only with representative participants (experts in HCI and Geo-Information Science), but also in representative settings (unfamiliar urban tourism context). The evaluation, therefore, captured problems with content and impressions grounded not only in expert knowledge and opinions, but actual context of use.

### 4.4.5. Stage E – Conceptual User-Centred Design Framework for AR Browsers

The main aim of the final stage of this study was to bring together the empirical findings in one consolidated framework by expanding and building upon the developed theoretical framework (Stage A). While the final results (Stage E) are presented in the end of the thesis, this process was iterative and carried out throughout the study. It started with incorporating theory in the developed consolidated cognitive task model (Stage B) in order to explain and reason about identified relationships. The process follows the main steps of development of psychological models (van Someren et al., 1994) and conceptual design frameworks (Beyer and Holtzblatt, 1998) where the results from task analyses, theory, empirical knowledge and literature are brought together to prescribe design guidelines for user interfaces.

The developed conceptual framework (Chapter 9) aimed to explicitly state "the meaning of the psychological theory in the context of the task" (van Sommeren, 1994; p. 55). It specifically related properties of user goals (information acquisition about unfamiliar urban environments) and tasks (e.g. association) with existing knowledge on cognitive processes. As a result, the framework can be used to generate hypotheses about user behavior during mobile interaction and, thus, optimize the design of smartphone AR browsers.

## 4.5. Risks and Validity of Research Findings

The research approach adopted in this study is acknowledged to pose a number of risks. In addition to the challenge to satisfy both quantitative and qualitative validity criteria, there is a set of additional concerns that mixed methods approach faces. These are discussed in the following section, together with the adopted strategies to minimise their impact.

### 4.5.1. Validity Criteria in Quantitative Research

There are several threats to validity (Shadish et al., 2002; Creswell, 2014) that could influence the outcome and inferences made based on quantitative data collection and analysis during the study. In order to adopt consistent and effective strategies for limiting their effect, a validity typology by Shadish et al. (2002) was used (Table 4.3). Defining validity as "the approximate truth of an inference" (Shadish et al., 2002, p. 34), the authors describe strategies that could be adopted to ensure the statistical conclusion validity, internal validity, construct validity and external validity of a study. The strategies that were adopted in this study in order to deal with such threads to validity are also described in Table 4.3 and discussed further.

***Statistical conclusion validity*** refers to the "validity of inferences about the connection between treatment and outcome" (Shadish et al., 2002, p. 42). Threats to validity (Table 4.3) are common to quantitative experimental studies and include low statistical power, violated assumptions about statistical tests and measurement errors. Several strategies were considered and adopted in order to ensure that the conclusions about relationships among variables are correct. First, care was taken that sample sizes and data collection in-between and between subjects was suitable for each study. Second, a cut off point of 0.05 in significance level was used for all statistical tests. Care was taken that measurement errors (e.g. duration in task time) are reduced by using triangulation between obtained data (video/audio recordings) and sophisticated software packages (IBM SPSS and nVIVO) that allow more accuracy in analysis. Corrupted audio/video recordings were excluded from analysis.

**Table 4.3. Description of specific threads to validity and strategies to negotiate their effects**

| Validity | Threat | Description | Strategy |
|---|---|---|---|
| Statistical Construct Validity | Low statistical power | Incorrect conclusion about relationship between variables | Adopt standard statistical tests |
| | Violated assumptions of statistical tests | Violation of statistical tests assumptions | |
| | Unreliability of measures | Measurement errors | Triangulation of video/audio data collection |
| Internal Validity | Ambiguous temporal precedence | Lack of clarity which variable occurred first | Measurement prior, during and after task completion |
| | Selection | Differences in respondent characteristics | Random selection of participants and assignment to conditions |
| | Maturation | Naturally occurring changes over time | Short-term mobile field study |
| Construct Validity | Inadequate explication of constructs | Failure to adequately explicate a construct | Adequate definition of concepts and constructs |
| | Mono-operation bias | Operationalization underrepresents the construct of interest | Adopt standard constructs from HCI |
| | Mono-method bias | When all operationalizations use the same method | Use multiple measures for each construct |
| External Validity | Interaction of the causal relationship (CR) with units | An effect with certain units might not hold with other units | Increase variation of profile characteristics |
| | Interaction setting and treatment | An effect found in one context might not hold with other | Increase variation of settings |
| | Interaction of the CR with settings | Interaction of history and treatment | Claims restricted to current time period |

*After: Shadish et al., 2002; Creswell, 2014*

***Internal validity*** refers to drawing correct inferences about the causal relationship between variables and results (Creswell, 2014). Threats to internal validity include ambiguous temporal precedence, history, selection and maturation (Shadish et al., 2002). Several strategies were adopted (Table 4.3) to minimise their effect where relevant. In order to prevent *ambiguous temporal precedence*, care was taken that measurement is carried out prior, during and after task completion. In all empirical studies, data recording protocols were developed that reflected when measurements are taken. The effect of *selection* was counter-balanced by randomly assigning participants to conditions. Finally, several strategies were considered in order to address *maturation* effects, expressed as naturally occurring changes over time. In order to prevent learning, it was considered suitable to limit the duration of all empirical studies to no more than 1-1:30 hours. Randomization of tasks, locations and applications (e.g. see Oulasvirta et al., 2009) was implemented during the laboratory experiment in order to prevent carryover and learning effects.

***Construct validity*** refers to the quality of the conceptualization and operationalization of the relevant concept (Shadish et al., 2002). In order to increase construct validity during the empirical evaluation, Shadish et al. (2002) suggest the following strategies, which were also adopted during empirical data collection and analysis: adequate definition of constructs, use of standard metrics (time, task accuracy) for determining usability of information systems, and the use of multiple measures (time, task accuracy, certainty, difficulty) for each construct (e.g. usability).

***External validity*** refers to the generalizability of research findings to various settings, populations and treatments (Shadish et al., 2002; Creswell, 2014). Several strategies, described in Table 4.3, were adopted where possible in order to ensure the external validity of the findings. The key concern was to reflect on whether the characteristics of the participants, settings and time periods when evaluation was conducted could influence the size and direction of effects and inferences. During the field and laboratory evaluations, it was considered necessary to increase the variation in participants' profile characteristics (e.g. age, background, experience with smartphones and AR) and test settings (e.g. urban type of terrain, physical structures, distance to POIs).

### 4.5.2. Validity Criteria in Qualitative Research

Where qualitative techniques for data analysis are applied, an important consideration to any research methodology is to ensure the *qualitative validity* and *reliability* of the research design (Creswell, 2014). In order to ensure the credibility of research findings, Creswell (2014) suggests using at least one of a total of 8 different strategies: triangulation, member checking, rich and thick description, clarification of research bias, presentation of negative/discrepant information, prolonged time spent in the field, peer debriefing, and external audit. With the current study relying on qualitative data analysis techniques, several approaches were adopted in order to ensure the qualitative validity and reliability of research findings:

- **Triangulation** – use of different data sources to build coherent justification for findings. In this sense, qualitative data analysis relied on establishing themes and claims from several sources of information from participants.
- **Peer debriefing** – this strategy was used in order to increase the accuracy of the results. All findings from the qualitative studies were reviewed by peers who asked questions and validated the drawn conclusions.

## 4.6. Chapter Summary

This chapter described pragmatism (Section 4.2) and mixed methods research (Section 4.3) adopted as part of a user-centred design approach to achieve the aim and objectives in this study. With the main aim to contribute to Information Systems Design theory and generate design knowledge relevant to improving usability and utility of AR browsers in urban tourism context, the chapter then presented the stages of this research and the selected methods for empirical data collection and analysis. This research project consisted of 5 stages (2 theoretical and 3 empirical). Starting with theoretical framework development (Stage A), this thesis progressed with the implementation of four primary data collection studies: (Stage B) mobile field-based evaluation of existing AR browsers, (Stage C) a laboratory evaluation of AR annotations, (Stage D.2) a pluralistic walkthrough with domain experts, and (Stage D.3) field-based qualitative evaluation of alternative AR annotations. These activities were followed by the development of a conceptual user-centred design framework (Stage E) that combined existing theory and the empirical findings from this project. Finally, the chapter described the risks and limitations connected with both quantitative and qualitative research, together with the adopted strategies to minimise their impact.

# CHAPTER 5

# THEORETICAL FRAMEWORK

**5.1. Theoretical Framework Development**

Presents the approach undertaken to develop the theoretical framework

**5.2. Information Acquisition in Urban Tourism Context**

Reviews how people acquire knowledge about space

**5.3. Information Needs and Cognitive Gaps**

Examines the triggers that lead to information needs and information search behaviour in unfamiliar environments

**5.4. AR Browsers and Landmark Knowledge Acquisition**

A unique approach to AR browsers is undertaken, and their role as external tools that facilitate information acquisition is identified

**5.5. Design and Context Parameters Influencing Usability**

Provides the scope for further research by examining the key context of use factors and design parameters that influence usability

**5.6. Chapter Summary**

Synthesis of the key points in the chapter

## 5.1. Theoretical Framework Development

A review of the literature revealed several conceptual (Chapter 2) and empirical (Chapter 3) shortcomings of existing user-centred approaches to design of AR browsers used in urban tourism destinations. The scarce research that investigates usability of AR browsers has led to a lack of understanding of user requirements and design knowledge, expressed as guidelines for more useful and usable AR browsers. There is a strong need to place the user in the centre of design and explore the role of AR browsers as tools for (geo)spatial knowledge acquisition.

Following the main principles of User-Centred Design (Abras et al., 2004), early user involvement through user-based studies is critical when design is directed at novel interfaces, such as AR (Gabbard and Swan II, 2008). However, on-site empirical evaluations with smartphone applications in tourism can be influenced by a number of factors (Section 3.3). Therefore, in line with the general process of Information Systems theory generation (Section 3.4), it was considered critical to first identify the key constructs and factors (and relationships among them) that play an important role for design of AR browsers in urban tourism context. Adopting a unique approach to AR interfaces as tools that facilitate (geo)spatial knowledge acquisition, key concepts, constructs and previous empirical findings were gathered in a preliminary theoretical framework (Stage A, Table 4.2), described in this chapter. Apart from setting the boundaries for research, the framework was used to guide further empirical data collection. The development of the framework was guided by three main principles (Antunes and André, 2006):

(1) identify the key constructs and elements, relevant to interaction of tourists with AR browsers in urban tourism destinations.

(2) be open for exploring and interpreting human factors in (geospatial) knowledge acquisition through AR browsers, thus requiring relatively open-ended constructs and abstracts elements.

(3) link the elements and constructs in a purposeful and meaningful way.

The main focus of attention within this study is on improving design of smartphone AR for tourists, who roam around in unfamiliar urban environments. Therefore, it is critical to investigate how tourists acquire and store (geospatial) information in general, and more specifically, about large-scale urban environments. Investigating how AR browsers mediate information acquisition in this way helps in identifying the role of such tools, as well as their benefits and drawbacks in comparison to other available smartphone interfaces that deliver location-based information. Due to the multi-disciplinary nature of the examined phenomenon, existing empirical and theoretical knowledge within several relevant disciplines (Information Science, Geo-Information Science, Environmental Psychology, Cognitive Psychology, Tourism) was brought together. Empirical observations served to identify further kernel theories, confirm the identified processes and add new ones. Building upon this sound theoretical foundation, the framework was later revised by incorporating the findings from all consequent empirical studies. This led to the development of a conceptual user-centred design framework for design and evaluation of AR browsers (Chapter 9).

## 5.2. Information Acquisition in Urban Tourism Context

In order to optimise their on-site visit and enhance their experience with a destination, tourists require access to substantial amount of spatial and attribute data (Brown and Chalmers, 2003), or uncategorised and raw facts about their surroundings. Mobile Location-Based interfaces are intended to represent these data in a structured, organised and clear way. Only in this way, when the provided data makes sense in the context in which it is delivered, it becomes information (Longley et al., 2010). Information is then stored in working memory and used to build an internal representation of reality, called a ***mental model*** (Johnson-Laird and Byrne, 1991). Such internal representations of reality are used to recognise and interpret new data, and formulate conclusions about the world. The construction of mental models depends on cognitive abilities and skills for understanding, imagining and navigating through physical spaces and influences the way location-based interfaces are used (Davies et al., 2010). In turn, people's behaviour with mLBSs is heavily influenced by the image they form about their physical and non-physical environment.

### 5.2.1. Spatial and Attribute Information Acquisition

In order to understand how smartphone Augmented Reality browsers can aid tourists, it is important to look at how spatial information is acquired, processed, stored and used. These aspects are part of a vast research area that studies how people perceive and understand (physical) space, called ***spatial cognition*** (Hart and Moore, 1973). A lot of research has been carried out and reviewing all aspects of spatial cognition is out of the scope of this study. This is why this section only looks at aspects that have important practical implications for the design of AR browsers:

- How people learn and store knowledge about geographic space
- The role of external visual displays for spatial knowledge acquisition
- How acquired spatial knowledge influences the understanding of space

Kuipers (1982) coined the term ***spatial knowledge acquisition*** to mean the process in which people retain information about space that they can use at a later stage. Information that is processed, organized and stored in long-term memory is captured in a ***spatial cognitive map*** (Tolman, 1948). When directly exposed to an unfamiliar environment, the user starts storing knowledge about it through the process of ***environmental mapping***. Downs and Stea (1973, p.9) note that this is the "process composed of a series of psychological transformations by which an individual acquires, codes, stores, recalls and decodes information about the relative locations and attributes of phenomena in his everyday spatial environment". The cognitive map is built up of three types of knowledge (Siegel and White 1975; Stern and Leiser 1988):

- ***Declarative / landmark knowledge*** – knowledge about discrete objects and entities, also called landmarks.
- ***Procedural / route knowledge*** - knowledge about paths and routes between landmarks.
- ***Configuration / survey*** knowledge - spatially organized knowledge of locations and routes.

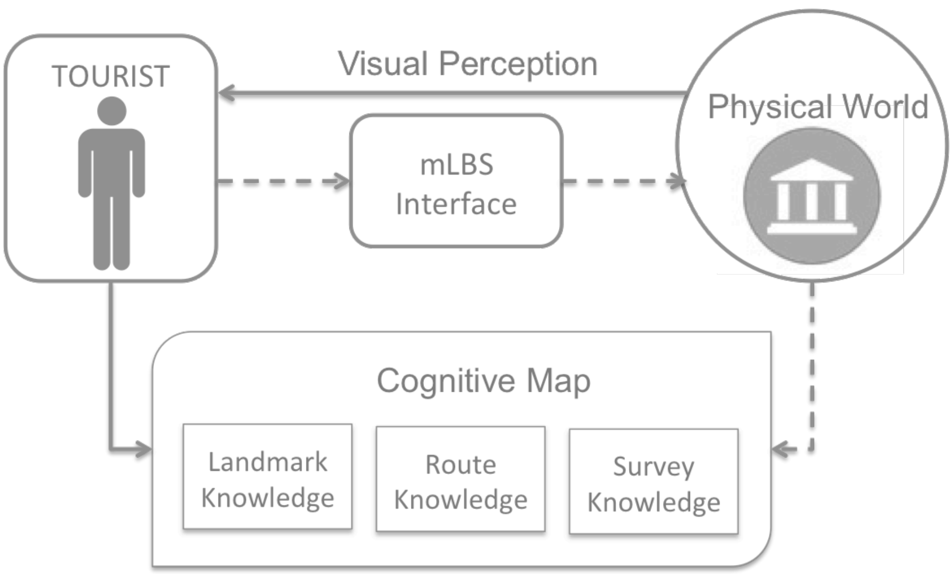
Environmental mapping is an iterative process that develops with time and is associated with “repeated exposure” to a physical environment (Lobben 2004). The construction of cognitive maps starts with knowledge acquisition about landmarks, used as anchor points to organize spatial information (Hart and Moore, 1973; Siegel and White, 1975). This process is mainly influenced by visual perception, or the ability to recognise and interpret visual sensory stimuli. After repeated exposure, humans are able to connect mentally individual landmarks (route knowledge) and organise this knowledge in a map-like allocentric representation of the environment (survey knowledge) (Hart and Moore, 1973).

Apart from learning and memorising their position in space, tourists need to acquire a lot of additional facts about physical entities, captured in attribute data. The common vision behind location-based interfaces is that they should communicate information that extends beyond what is visible around users (Fröhlich et al., 2008) and that otherwise would remain unnoticed. Apart from learning *where* physical entities are (geospatial data), tourists require access to additional facts about them, such as the date when they were built, their name, special characteristics or history. Typically, this information is contained and communicated through ***attribute data*** (Longley et al., 2010). Attribute data can be appended to any object, and are often stored in a database together with geo-spatial information.

### 5.2.2. Communicating Information Through mLBSs

Both attribute and spatial information are essential for tourists, as they aid and influence navigation, decision-making and the on-site experience with a destination. Apart from being time-consuming and resource-intensive, the acquisition of information through physical exposure requires significant cognitive effort. For tourists repeated visits to a place might not be possible due to time pressure and limited resources, and even if they occur, they are often separated in time. In addition, physical presence does not presuppose ready access to attribute information, such as the history, name, or other special characteristics of points of interest within the surrounding environment. This is why external tools, such as location-based interfaces (Figure 5.1), that help tourists obtain this information play a critical role within their on-site visit, as they facilitate attribute and spatial information acquisition.

**Figure 5.1. In addition to physical exposure, visual external tools can help tourists obtain valuable spatial and attribute information about their surroundings**

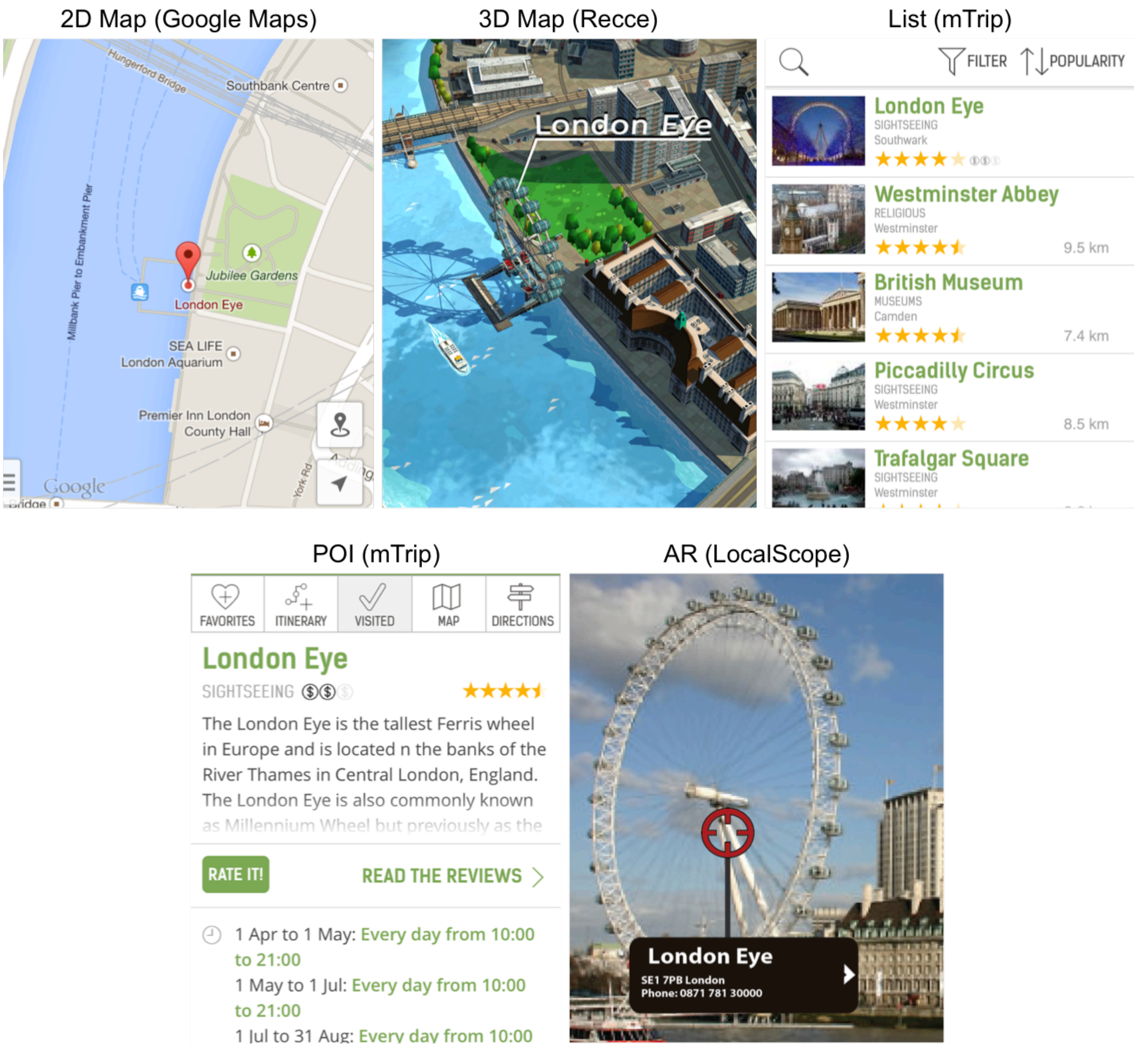


There are a number of visual representations of data that could aid spatial and attribute information acquisition about unfamiliar tourism environments. These include two-dimensional (2D) and three-dimensional (3D) maps, list, individual point of interest (POI) interfaces and Augmented Reality (AR) (Figure 5.2). Each visual representation captures specific aspects of reality and employs its own system of signs or "language" for communicating spatial and attribute information. Empirical evidence (Oulasvirta et al., 2009; Kraak and Ormeling, 2010; Baldauf et al., 2011) shows that each interface supports different types of search and information acquisition.

***Maps*** remain the most powerful means for visual representation of spatial knowledge (Kraak and Ormeling, 2010) and smartphone map-based interfaces are the most popular type of mLBS (Meng, 2008). A map is a simplified and abstract visual representation of reality (Longley et al., 2010). Maps provide an overview of larger territories at different scales, and therefore allow fast acquisition of survey knowledge (Fremlin and Robinson, 1998). However, this depends on users being able to interpret and match the signs captured on a map. This process might require significant cognitive and physical effort (Levine, 1982; Oulasvirta et al., 2009) when users have to quickly and efficiently align a map with the landmarks in the surrounding environment.

***Three-dimensional representations*** of reality (3D maps) "involve volumetric instead of flat representation of space, realistic instead of symbolic representation of objects, more variable views that are directional and bound to a first-person perspective, more degrees of freedom in movement, and dynamically changing object details" (Oulasvirta et al. 2009, p.303). Empirical research indicates that 3D mobile maps allow faster recognition of objects (landmarks) in the surrounding environment (Oulasvirta et al., 2009). They enhance performance when it comes to route knowledge acquisition and navigation. However, 3D mobile maps are significantly more difficult to control and navigate and may lead to disorientation (Oulasvirta et al., 2009). More importantly, while they provide a good overview of a larger territory, it is very difficult to overlay additional information about POIs due to their photo-realistic (and visually complex) nature.

**Figure 5.2. Different types of mobile Location-Based Interfaces**




Declarative (landmark) knowledge is critical for informed decision-making (Davies et al., 2010) and is especially important in tourism. Mobile 2D and 3D maps facilitate landmark knowledge acquisition by helping tourists identify and learn *where* points of interest are in space in relation to their current position. Apart from knowing where landmarks are, tourists also search for further (historical, architectural, etc.) information about points of interest. ***List-based*** interfaces organize and display information about POIs based on some ranking criteria. Typically smartphone mLBSs lists are accessed through a “Near me” button, which triggers filtering of information based on the current location of the user. Nearby POIs are located at the top of the list, while far off locations are further down. From lists, users can also access information about individual POIs. ***Individual POI*** interfaces convey more detailed descriptions of spatial features (e.g. buildings) and non-spatial, temporal or other type of attribute data (e.g. history of a building; architectural style, etc.). Lists and POI interfaces are simple and easy to understand when it comes to acquiring landmark knowledge. However, they are limited in the amount of contextual detail that they can convey.

The main problem with on-site access of information through map or list-based interfaces is that users have to constantly shift their gaze between virtual and physical worlds in order to relate the information on the screen with their physical surroundings. The overall format, completeness and way of presenting this type of information might not be essential for navigation, but affects the overall experience and engagement of tourists with a destination (Gursoy and McCleary, 2004). Considering the characteristics of Augmented Reality, it seems that the display will be most suitable to enhance and contribute to landmark knowledge acquisition. Such declarative knowledge does affect navigation or wayfinding, however, in this study we consider static, rather than mobile (moving) users.

## 5.3. Information Needs and Cognitive Gaps

An AR browser is, in essence, an information system. Its use, utility and usability will ultimately depend on the information needs and behaviour of the user. Information needs and behaviour have been the focus of study of Information Science since the late 1940s. In spite of this, theoretical progress in the area has been slow, due to the complexity and the many factors involved in information search behaviour (Wilson 2006). Major theories and definitions within Information Science are, nonetheless, highly relevant to this study. As Raper et al. (2007, p.25) argue "Information science is concerned with issues of information need, management and retrieval, all of which are taken to extremes in LBS. Information design requires an understanding of information needs as information overload readily sets in if needs are not carefully considered."

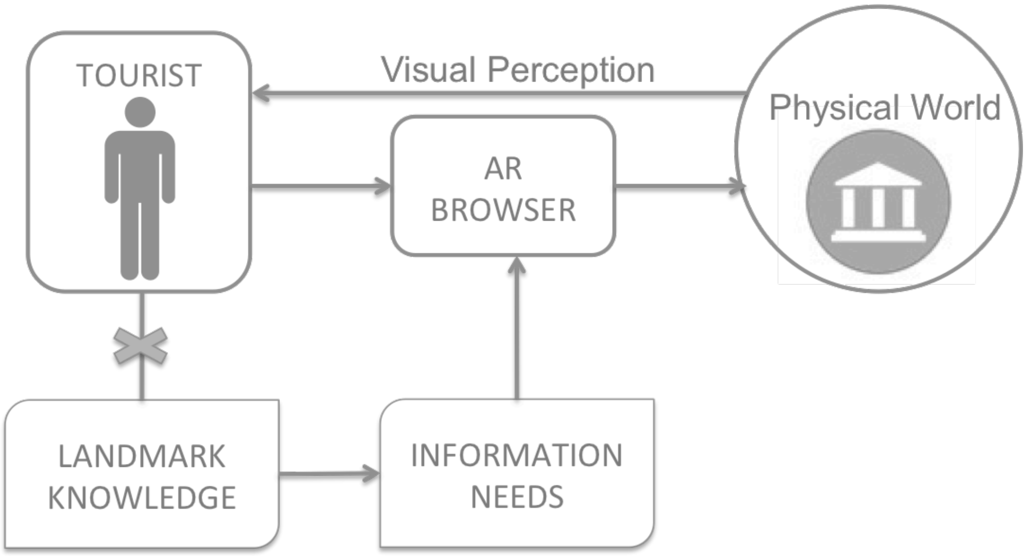
Information science defines ***information needs*** as "consciously identified gaps in the knowledge available to an actor…[that] may lead to information seeking [behavior] and formulation of requests for information" (Ingwersen and Jarvelin 2005, p.20). Proposing the sense-making metaphor, Dervin (1984) theorized that humans move through time and space until they reach a *cognitive gap*, where an information need is perceived. Such gaps must be bridged through the acquisition of new information until the perception no longer exists (Krikelas 1983; Wilson 2006). Therefore, interaction with a mobile location-based interface results from the recognition of some lack of knowledge, perceived by the user. When the user consciously recognizes that need it is often expressed as a question (query).

The three (interrelated) ***drivers*** that trigger information-seeking behavior are physiological (e.g. need for food), affective (e.g. need for attainment) and cognitive needs (e.g. need to learn) (Wilson, 2006). Cognitive needs underpin the process of *informal learning*. Tough (1979) defined informal learning as "a major, highly deliberate effort to gain certain knowledge and skill (or to change in some other way)". According to Livingston (2001) the process comprises of "all forms of intentional or tacit learning in which we engage either individually or collectively without direct reliance on a teacher or externally organized curricula" (Livingston, 2001, p. 2).

Prototypical mobile technologies often address scenarios related to tourists because this emphasizes the need to move around an unfamiliar environment and the desire to learn about the world around us. Information needs in such situations are driven mainly by cognitive needs. From a spatial cognitive point of view, the experienced cognitive gaps will trigger information seeking behavior related to acquiring more knowledge (informal learning) about individual landmarks, the paths between them (route knowledge), or the overall configuration of the environment (survey knowledge).

During the exploration of an unfamiliar environment, tourists' attention will first be directed at salient objects (landmarks) that will serve as reference points for organising additional geospatial knowledge (Hart and Moore, 1973). Landmarks are defined as distinctive spatial features (due to their colour, shape, semantic value), that are used by individuals to organize information about large-scale environments (Hart and Moore, 1973; Bluestein and Acredolo, 1979).

**Figure 5.3. Cognitive gaps related to lack of knowledge about physical landmarks trigger interaction with the AR browser**




The acquisition of landmark knowledge is critical for decision-making, and influences to a large extent the learning of new environments (Evans, 1980). In tourism literature, such individual physical entities (landmarks) are often called Points of Interest (POIs). A POI is an individual location, building, monument or other physical entity that is interesting and important from a tourist point of view. The main implication is that a perceived cognitive gap for such landmarks (or points of interest) will trigger an information need, expressed as search queries (Figure 5.3). The tourist will then interact with the smartphone AR browser in order to find answers to such queries.

## 5.4. AR Browsers and Landmark Information Acquisition

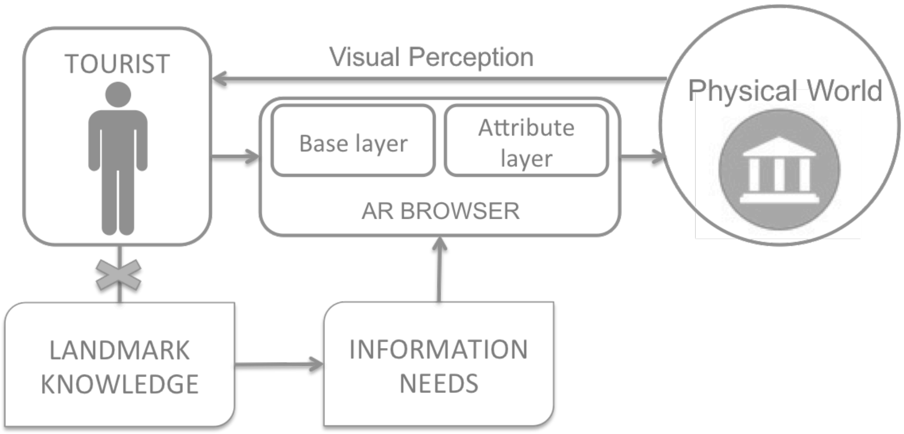
The ultimate goal of mLBSIs is to communicate information about spatially distributed phenomena (Raper et al., 2011) within the surroundings of the user. In order to be useful and usable, tourists have to match (mentally relate or associate) the information presented on the screen of the smartphone with physical space. All mobile location-based service interfaces depict (virtually) in some form geographical space and/or the attributes of objects and events within that space. This information is captured in two layers (Figure 5.4):

- Base layer – This layer is also called *the representation of the physical world* (Vincent et al., 2012). It captures and communicates information about physical entities within space (e.g. roads, houses, buildings, monuments).

- Attribute layer – This layer is also referred to as *augmentation* (Vincent et al., 2012). It captures and communicates additional information about the represented physical entities (e.g. name, age, architecture).

The following sections examine both of these elements and the (cognitive) processes that users need to carry out with each in order to make sense of the AR interface.

**Figure 5.4. Composing elements of the AR browser interface.**

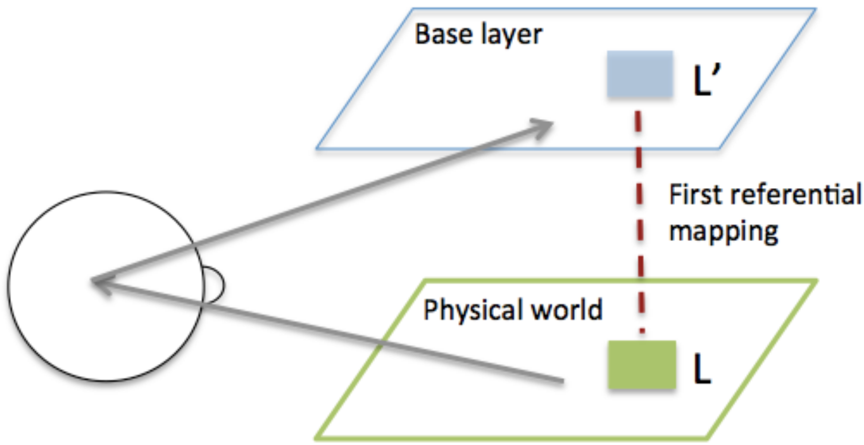



### 5.4.1. The base layer and first referential mapping

The ***base layer*** encompasses the components that represent the physical surroundings around the tourist. This representation allows the user to understand where physical entities are located within the surroundings. The most obvious difference among mLBSIs is the abstraction level of the base layer, or the level of detail that is used to represent the physical world. Map-based interfaces encode physical entities through point, line and area features (Longley et al., 2010). *Virtual environments*, or 3D maps often make use of volumetric photorealistic or non-photorealistic computer generated models to represent real world entities. Typical AR browsers have the lowest level of abstraction, as the base layer consists of an unaltered video feed that captures the visible surroundings of the user.

In order to understand a spatial visual display, the user must first understand the basic ***referential relationship*** between the real world and the representation of the real world on that display (Lobben, 2004). According to Levine (1982), for this to happen the user needs to perform *structure matching*, or relate specific visual cues in the base layer to their relating visual cues in the environment (Figure 5.5) until the two spaces (virtual and real) overlap in the mind of the user.

**Figure 5.5. First referential mapping**




*After: Oulasvirta et al., 2008*

As a result of this process, the user can then associate each perceived physical landmark (L) with only one target represented on the base layer (L’). This process is also called *projection, superimposition* (Bluestein and Acredolo, 1979), or ***referential mapping*** (Oulasvirta et al., 2009). As the next section will discuss, the referential mapping between the physical world and the virtual representation of the physical world is not the only referential mapping that users have to make. To prevent ambiguity, here onwards this process will be referred as *first referential mapping*.

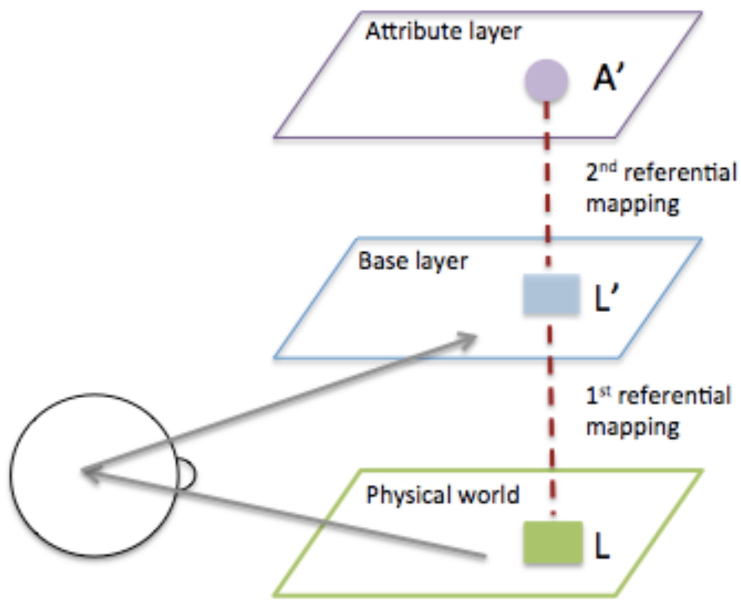
The first referential mapping has been studied exclusively within the domain of spatial cognition and Cartography. Most research has addressed referential mapping in the context of 2D maps which require that at least two pairs of points (e.g. L1 – L1’ and L2 – L2’) are matched visually with their physical counterparts before the user can superimpose the virtual and physical spaces (Levine, 1982; Bluestein and Acredolo, 1979). Whereas the mapping process when using 2D maps as visual communication tools have received a lot of attention, other types of mLBS interfaces have not. As a consequence, there is very scarce knowledge on how referential mapping is achieved through different mLBSIs. However, studies with 3D maps suggest that a lower abstraction level of the base layer decreases the cognitive demands on the user (Oulasvirta et al., 2009). Since most AR browsers use unaltered visual representation of the environment, it seems reasonable to assume that users will be able to match immediately the virtual video feed with the physical world.

### 5.4.2. Attribute layer and second referential mapping

The ***attribute layer*** is superimposed on the base layer and contains additional information associated with individual spatial entities. An attribute can be used to *identify* a place or an entity (e.g. name, street address, social security number), *quantify* a characteristic (e.g. temperature, age), or help classify the entity into a category (e.g. class of land use, type of building, function). The range of attributes that can be represented within a single interface is very large. The type of information and the way it is represented depends on the type of base layer and the purpose of the mLBSI. However, ***textual (labels)*** are the most common form of attributes, represented on all types of mLBSIs, including 2D maps (Meng, 2008), 3D maps, virtual environments (Hartmann et al. 2005) and AR browsers (Wither et al., 2009).

Textual labels have been studied extensively in a number of disciplines, such as Cartography, Scientific Visualistion, and Information Rich Virtual Environments. In the context of mobile maps, once the first referential mapping is carried out (Section 5.4.1), the user then has to match (associate) the textual label with only one (virtual) entity represented on the base layer of the map (Figure 5.6). This process of association is fundamental for any type of mLBSI and is here onwards referred to as ***second referential mapping***.

**Figure 5.6. Second referential mapping**



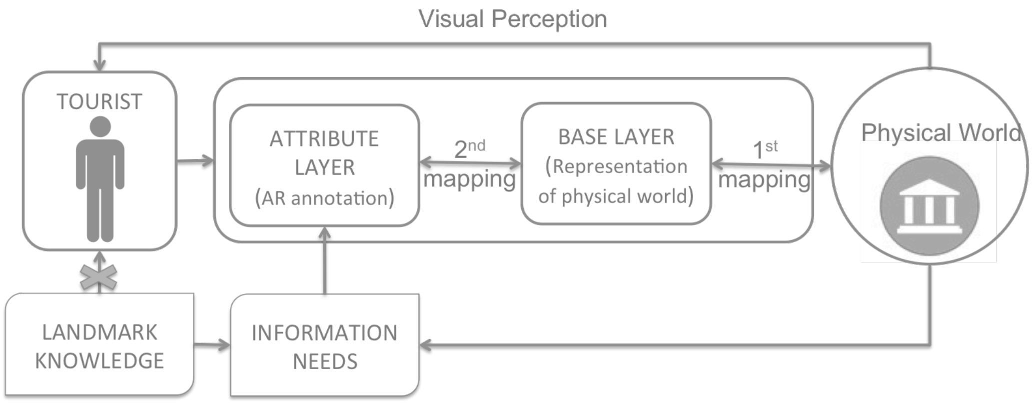

*After: Imhof 1975; Christensen et al. 1992; Dijk et al. 2002*

The first referential mapping is often achieved through ***visual coupling***, or matching visual cues from both physical and virtual spaces. Typically, the second referential mapping on maps, 3D virtual environments and other mLBSIs is achieved through ***spatial coupling***, or precisely placing the textual label on top (internal textual labels), or nearby (external textual labels) of the representation of the physical object on the base layer (Imhof, 1975; Hartmann et al., 2005). Design and development of AR browser annotations has drawn mainly from research within cartography (Feiner et al., 1997; Bell et al., 2001; Azuma and Furmanski, 2003; Leykin and Tuceryan, 2004; Bell et al., 2005; Grasset et al., 2012). Following principles for label placement on maps, it is assumed that users will be able to carry out the second referential mapping as long as the virtual AR annotation is precisely placed on top of its reference physical object. Association through placement is also supported by the Feature Integration Theory (Treisman and Gelade, 1980), which states that features of objects are known to belong together based on the fact that they co-occur at the same position in space. Therefore, as long as the virtual labels occur at the same position as the physical objects, the user will mentally associate the two as belonging together. Following these principles, it seems that users will experience problems when using AR browsers due to large positioning errors and imprecise placement of annotations.

## 5.5. Design and Context Parameters Influencing Usability

AR browsers are a very special type of information system, mimicking the properties, functionality and content of standard *web-based browsers* and *location-based services*. In order to be useful and usable, AR browsers have to be designed so that users can interpret the information on the screen and match it to the real-world entities around them. In effect, this means that three processes have to be carried out effectively and efficiently and will influence the usability of AR browsers: (1) first referential mapping, (2) second referential mapping, and (3) answering the questions that triggered the search for information (Figure 5.7). It is important to understand what are the key parameters, both of the system, as well as the context where activity takes place, that will influence these three processes.

**Figure 5.7. The initial conceptual model for investigating design and context parameters that influence the usability of AR browsers**




## 5.5.1. Design parameters

AR interfaces can be decomposed into two separate layers: the base layer (representation of the physical world), and the attribute layer (the virtual AR annotations). In order to be useful and usable, users have to be able to effectively and efficiently carry out the first and second referential mappings in order to understand the relationship between the screen of the smartphone and the physical world.

### *5.5.1.1. Base layer*

The first (and only) paper to discuss the different types of relationships between the physical environment and the base layer in the context of AR was published by Vincent et al. (2012). The authors pose that there are three types of relationships (called *spatial mappings*) between the actual physical world and the base layer: conformal, relaxed and none. AR systems that use the incoming real-time video feed are said to have a conformal (one-on-one) spatial mapping. Relaxed spatial mapping is achieved if, for example, the user can freeze the current frame and edit it. According to the authors, the third category comprises of systems that do not provide spatial mapping between the physical world and its representation on the screen of the phone. However, revisiting the definition for AR (Chapter 2), it is questionable whether such systems can be called Augmented Reality. The paper does not discuss, however, the design parameters that can be manipulated in order for users to understand and be able to work with an AR browser depending on the type of spatial mapping it supports.

Outside of AR, a number of studies within cartography and Geo-Information Science have discussed the design parameters that could influence the first referential mapping. Orientation and ego-centric alignment (Levine, 1982; Oulasvirta et al., 2009) or the ability to orient the display to match the orientation of features within the physical environment is one such parameter.

#### *5.5.1.2. Attribute layer*

Current research within AR emphasises the role of the precise position of annotations for ensuring the second referential mapping (Section 3.5). Label placement is often defined by the angle and radius away from the reference object (Azuma and Furmanski, 2003). Most work within AR has been based on the assumption that badly positioned labels may negate the benefit of augmenting the display, as the user cannot associate labels with their reference object (Azuma and Furmanski, 2003; Grasset et al., 2012). Examining more closely the results from such studies, however, it seems that task performance is mainly influenced by label overlap, rather than label placement (Azuma and Furmanski, 2003). AR annotation research draws from cartography and label placement for maps where association of attribute and base layers is influenced mainly by (im)precise label placement. There is a general lack of research that investigates the effect of, for example, different angles, and distances from the reference object on association.

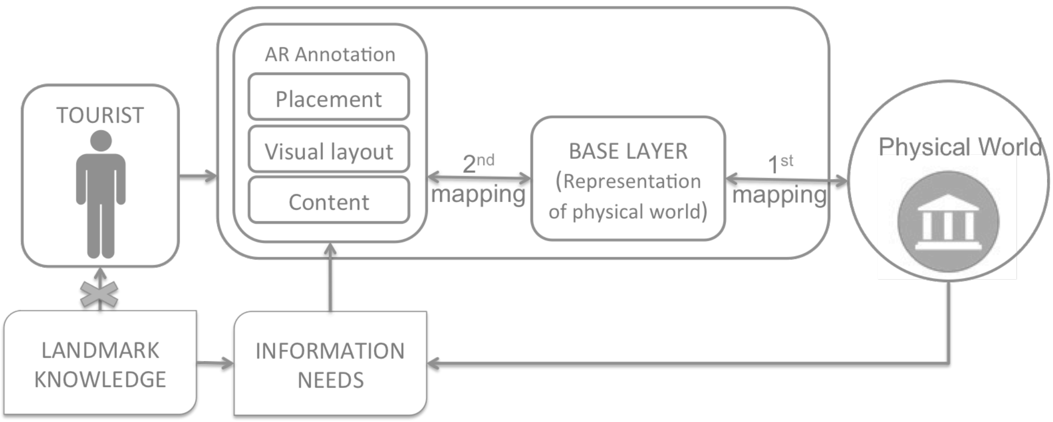
Likewise, research that discusses other design parameters and their effect on association has been very limited so far. Studies that examine the influence of design parameters on usability have addressed only requirements towards legibility (e.g. Gabbard et al., 2007; Jankowski et al., 2010). Empirical work suggests that parameters, such as the colour of the annotation's background or text influence legibility (Jankowski et al., 2010). Adopting a more broad approach to design, there are several relevant disciplines where design of virtual annotations has been discussed and empirically evaluated. In particular, studies within Information Rich Virtual Environments (Bowman et al., 2003), virtual 3D graphics (Hartmann et al., 2005), virtual environments (Maass and Döllner, 2007) or videos (Thanedar and Höllerer, 2004) have been concerned with aesthetic and functional design of virtual annotations. While it is questionable to what extent the findings and discussions from such studies apply to AR browsers used in urban tourism context, they provide a more complete picture of the overall design space for virtual annotations (Table 5.1). Apart from annotation background colour, a number of other design parameters, such as background opacity, structure of content, font size and frame colours could influence work with AR browsers.

**Table 5.1. Design parameters for AR browsers**

| Annotation element | Design parameters |
|---|---|
| **Annotation Placement** | Fixed / Dynamic<br>Internal / External<br>Object / Global |
| **Annotation Body** | Size<br>Colour<br>Transparency |
| **Annotation Frame** | Weight / Thickness<br>Colour<br>Transparency |
| **Annotation Content** | Structure<br>Type (text, image, 3D model, animation)<br>Amount |
| **Annotation Leader line** | Curvature<br>Thickness<br>Length<br>Colour<br>Transparency |
| **Annotation Screen Layout** | Spatial<br>Linear |

Considering the broad design space, design parameters can be classified in three broad categories: placement, visual layout and (type of) content. Such design variables are important to have in mind as they could potentially influence the overall usability and utility of AR browsers (Figure 5.8).

**Figure 5.8. Design space for AR browsers**

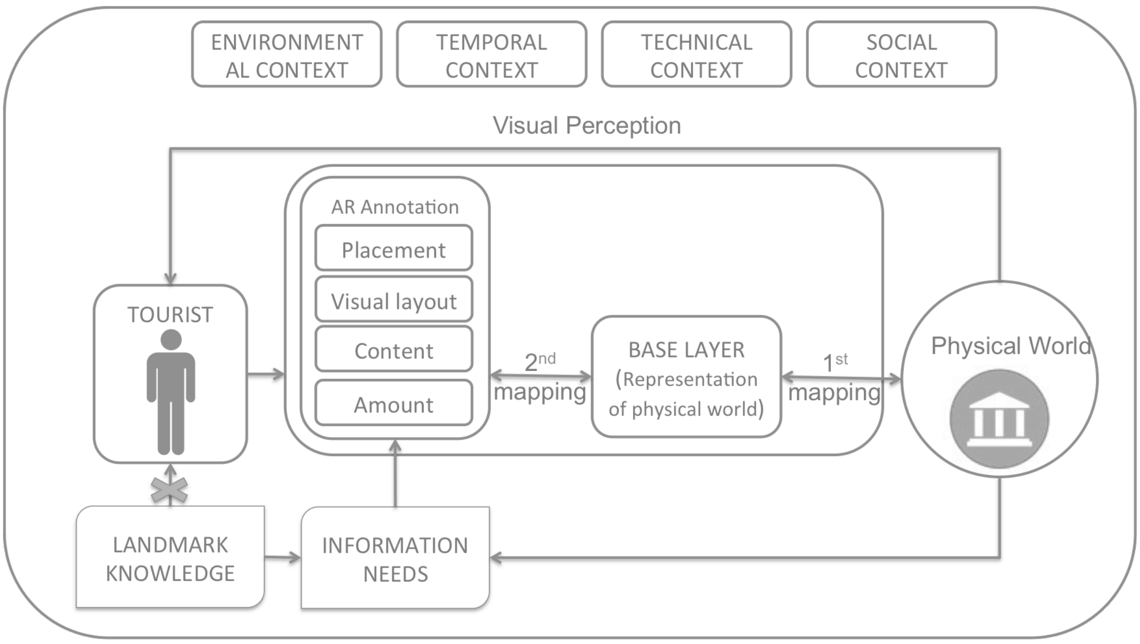



## 5.5.2. Context parameters

Section 3.3 discussed the fundamental role of context of use and how different parameters influence work with mobile ISs. Environmental context factors, such as lightning level, or bright sunshine, could influence and hinder visual perception and legibility (Herbst et al., 2008). Temporal factors, such as the duration of a visit or a use session could contribute to lack or presence of high time pressure. Similarly, the time of day and season could influence the perceived utility and relevance of information (Tan et al., 2009). Technical context, such as the type of mobile device used, could in turn impact the colour range and resolution of the display, influencing the legibility of delivered information. Considering the constructs that underpin mobile interaction with an AR browser (Figure 5.9) in urban destinations, variations in two main context categories seem to be most influential on work with AR browsers: physical and user context.

There are a number of aspects of ***physical context*** that could potentially influence interaction with AR browsers (Section 3.3.2.1). The visibility of physical objects, textures and backgrounds, and urban structure all seem to be important and influence interaction with AR interfaces (Feiner et al., 1997; Bell et al., 2001; Jankowski et al., 2010). More importantly, such context parameters also form part of the user interface, which makes them a critical (and unique) element of interaction with AR. Indeed, research on the influence of the physical world has been identified as a key necessity for developing and designing AR applications (Kjeldskov, 2003). There have been limited attempts to address this need so far, even though several workshops and conference call for papers clearly identify the need for research that investigate the role of "reality" within the design of mobile AR applications (e.g. Workshop on Designing Mobile Augmented Reality at MobileHCI, 2013).

**Figure 5.9. Context of use and AR browsers**



Potentially anyone can be a tourist, which means designing for users with widely different cultural backgrounds, education, expertise, knowledge, skills and abilities. While not considered as a context parameter in the original CoU model by Jumisko-Pyykkö and Vainio (2010), the specific characteristics of tourists as users of information will ultimately influence the usability and perceived utility of AR browsers. In terms of characteristics, Section 3.3.2.6 reviewed some of the studies that have emphasized the role of demographics, user interests, preferences, cognitive and physical abilities and already acquired knowledge and experience when it comes to usability and utility of mobile context-aware applications.

When it comes to user context, consumer research and tourism literature identify ***familiarity*** with a destination as the most important factor that influences information needs and behaviour (Gursoy and McCleary, 2004), but fails to discuss the essence and meaning of the term, as well as how familiarity can be measured. Likewise, a key assumption, which remains unexplored in geo-information literature (Davies et al., 2010), is that *familiarity* is a key factor that will influence the use of location-based interfaces. Already acquired landmark knowledge is expected to influence the amount of information and detail that is needed on mobile location-based interfaces (Zipf, 2002). Indeed, past studies in tourism indicate that when faced with a problem or information need, tourists first use information stored in memory (*internal search*), gathered from personal experiences with the destination or a similar one (Gursoy and McCleary, 2004). The type of information that tourists have stored in long-term memory would affect how they perceive a destination, as well as the type of (additional) information they will look for during their on-site stage of travel (Gursoy, 2003).

When prior accumulated knowledge is insufficient, tourists will search for external sources of information. The extent to which tourist engage in external search for information will depend mainly on the cognitive and physical effort that tourists have to allocate (Gursoy and McLeary, 2004). This is why it is paramount that information sources are easy and intuitive to use. On-site attribute information acquisition by tourists through traditional (e.g. guidebooks) and new (mobile devices) channels has received relatively less attention, compared to other travel stages (Section 2.2).

A similar situation arises when it comes to domains, such as environmental psychology where there is no commonly accepted definition for *unfamiliar* environment. As a result, alternative terms abound in literature, including *new, novel* and *nomadic* (Magllwo et al., 1996; Blajenkova et al., 2005). In such domains, familiarity is often measured through the number of times a tourist has visited a destination or the time spent researching and learning about a place from other information sources. As discussed in Section 5.2, knowledge acquisition about landmarks can be achieved directly, through physical exposure, as well as indirectly, through visual displays or other information sources. Therefore, in this study the term "unfamiliar" is used to specify that: (i) the user has had no previous physical exposure to that environment; (ii) the user has no previous specific (landmark) knowledge of that environment.

## 5.6. Chapter Summary

This chapter described the theoretical framework that captures the key constructs and relationships underpinning mobile interaction with AR browsers in urban tourism context. The chapter first explored the different types (spatial and attribute) of information that tourists need access to and how this is stored in memory (Section 5.2). In line with the set out objectives of this study, the chapter then explored the characteristics of various mLBSs interfaces and their role in on-site knowledge acquisition. Particular attention was devoted to AR browsers and their potential to provide fast acquisition of spatial and attribute information about discrete POIs or landmarks (Section 5.4). Adopting this point of view, the AR interface was further deconstructed to two inter-connected (base and attribute) layers. The framework also captured how information needs are triggered (Section 5.3) and discussed the various design and context of use parameters that could influence the usability and utility of AR browsers when used in urban tourism settings (Section 5.5). Apart from setting the boundaries for research, the framework was used to guide further empirical data collection, described in the next three chapters.

# CHAPTER 6

# FIELD EVALUATION OF EXISTING AR BROWSERS

**6.1. Study Goals and Questions**

Introduces the specific research questions that guided the field study

**6.2. Exploratory Analysis of Commercial AR Browsers**

Describes the procedure undertaken to select the AR browsers evaluated in this study

**6.3. Method and Procedure**

Describes in detail the design of the mobile field evaluation procedures

**6.4. Analysis**

Describes the undertaken qualitative and quantitative analysis

**6.5. Findings**

Reviews the major findings from the study

**6.6. Task Sequences and Consolidated Model**

Describes task analysis and consolidation of the findings

**6.7. Discussion**

The findings are explained and discussed with respect to relevant literature

**6.8. Chapter Summary and Conclusion**

Synthesis of the key points in the chapter

## 6.1. Study Goals and Questions

By adopting a User-Centred Design (UCD) approach, the main aim of this thesis was to generate design knowledge relevant to effectively support on-site information acquisition in unfamiliar urban environments. As discussed earlier (Chapter 4), user-based evaluations are critical for design of AR as they allow designers to understand better user requirements and the design parameters that could be used to satisfy them. Within the framework of ISDT generation, user-based studies allow better understanding of the problem that the IS has to solve, but also the problems that users experience with it. The obtained understanding, in turn, allows researchers to identify more appropriate and relevant kernel theories that could be used during the design process and to extract design guidelines for new ISs.

After the theoretical framework was developed, the next stage (Stage B, Table 4.2) aimed to observe in real time how tourists make use of AR browsers in unfamiliar urban environments. In order to focus the investigation, it was considered necessary to find answers to the following specific questions:

**Question 1**: To what extend the content delivered through current AR browsers can satisfy the location-based information needs of tourists?

**Question 2**: To what extent tourists are able to carry out association tasks through AR annotations within unfamiliar physical environments?

**Question 3**: Which elements of context of use influence the association process?

**Question 4**: What are the most significant problems that tourists experience that could prevent effective and efficient work with AR browsers?

In essence, the mobile field study simulated a real situation in which tourists were asked to identify interesting and/or important points of interest (POIs) in an unfamiliar environment and find out more about key tourist attractions by using four, different in design, commercial AR browsers. The procedures, analysis and results described in this chapter have been published in Yovcheva et al. (2014), Yovcheva et al. (2013a), Yovcheva et al. (2013b).

## 6.2. Exploratory Analysis of Commercial AR Browsers

In 2012, when the design of this study was being prepared, there were more than 500 commercial AR applications on the various smartphone apps stores. A large amount of those apps were (and still remain in 2014) AR games. The number of AR browsers was also considerable. Carrying out an empirical study with a large number of smartphone AR apps is neither feasible, nor necessary, as the variability of designs could lead to biased results. To this end, an exploratory evaluation and classification (Stage B.1, Table 4.2) was carried out prior to commencing the field study in order to classify AR browsers in several major categories and select the most representative applications for further evaluation.

The method adopted in this study is broadly based on the principles of comparative evaluation and benchmarking (Allen and Chudley, 20013), but has been modified following the procedures proposed by Pakanen et al. (2011). Competitor benchmarking involves comparing and contrasting digital products in the same or different industry domains (Allen and Chudley, 2013). Apart from evaluation of their weaknesses and strengths, the main benefit of this method was that it allows designers to spot similarities or patterns in product design, also called *design patterns*. The main aim here was to document both design patterns and differences among various AR browsers.

Pakanen et al. (2011) propose a variation of the competitive benchmark approach. In their study the authors wanted to improve the design of elements used to display location-based information in smartphone social media applications (Pakanen et al., 2011). In order to explore the design space for AR browser applications and identify the potential problems for users, the study followed a similar approach: (1) identify and browse available applications, (2) collect screenshot images from available commercial applications, and (3) group elements which have similar appearance/functionality. This activity aimed to document in more detail the similarities and differences among AR annotations used in current AR browsers. Available AR browsers were classified in several categories, depending on how they support tourism-specific tasks. The main goal was to make an informed decision regarding the AR browsers that were included in further actual observation with users.

### 6.2.1. Preliminary selection of AR browsers

After undertaking a preliminary inspection, it became evident that a large amount of commercial AR applications are available. Twenty-three applications (Appendix 1), that satisfy several preliminary criteria, were examined:

- Provide visual augmentation of the environment through annotations in AR view, excluding AR games and audio guides.
- Deliver content for the territory of the UK, related to urban leisure experiences, excluding dedicated AR applications that are specific for other regions.
- Are available for iPhone smartphones and can be downloaded/purchased from the Apps store, since the development of such applications is guided, to some extent, by general user interface guidelines.
- Use a marker-less, GPS-based approach to track, register and align virtual and physical objects.

### 6.2.2. Data collection and analysis

In order to be able to compare the design of the selected AR browsers, screenshot images were first collected at two specific locations: the City Square in Bournemouth (UK) and Westminster Bridge in London (UK). Data collection was carried out by the researcher in four consecutive days and took over 70 hours. In order to focus data collection, specific tasks were carried out with each AR browser. Selection of representative tasks for data collection was based on common tasks with mLBS (e.g. recognize, identify, and locate) (Jakobsson, 2003; Reichenbacher 2004) and mobile tourism services (e,g, find restaurants / museums) (e.g. Church and Smith, 2009; Hinze et al., 2010). A single general scenario was defined, described in Figure 6.1. The selected scenario was used to emphasize specific aspects of interaction and define explicitly the scope of study (bold).

**Figure 6.1. A description of the basic scenario**

> ***Basic Scenario***
>
> *"Jane is a tourist who has just arrived in the city (Bournemouth / London). It is her first time in the city (**unfamiliar environment)**. Jane does not know anything about the city (**no prior knowledge**), but wants to explore it and learn more about it (**knowledge acquisition**). She goes to the tourist information desk and decides to follow the route they advise her to take. While on the route, Jane will use a smartphone AR application to obtain information about her surroundings" (**outdoor use of AR**).*

Five use cases within this scenario were further defined. Each of them became the basis for one of the five tasks carried out during data collection with each AR browser:

- (Task 1) Locate / identify a point of interest (museum) in the immediate surroundings.
- (Task 2) Locate / identify one particular type of attraction (Russel-Cotes Art gallery for Bournemouth; the National Gallery for London).
- (Task 3) Obtain more information about the selected POI.
- (Task 4) Obtain directions towards the selected POI.
- (Task 5) Explore different types of available information within the immediate (visible) surroundings without specific pre-defined criteria.

During data collection, the researcher opened each application and carried out the identified tasks, making screenshots whenever there was a change in the interface or data representation.

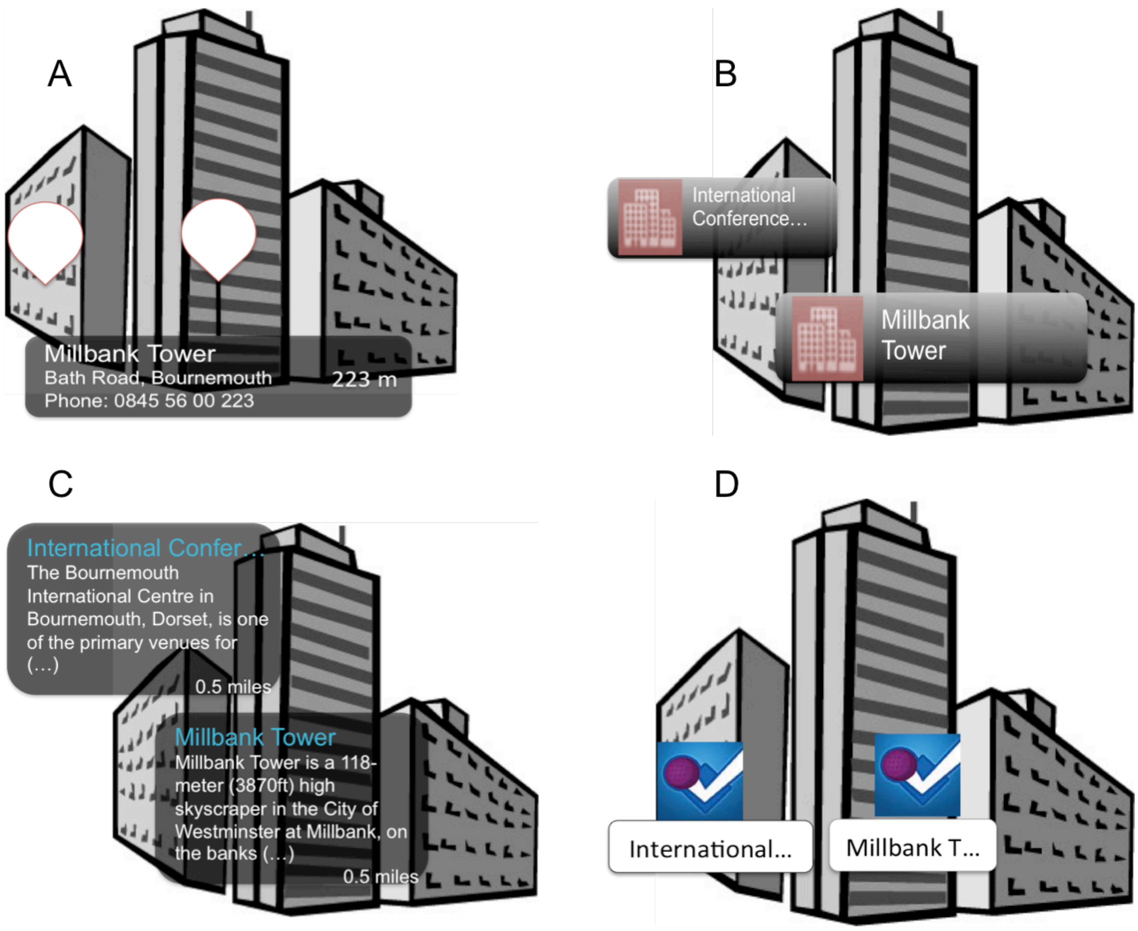
### 6.2.3. Classification and selection of AR browsers

In total, 1500 screen captures were compared, taking under consideration the previously identified criteria. After the screenshots were collected, the applications were grouped according to their high-level characteristics. The final classification and selection of AR browsers was based on the criteria, identified within the theoretical framework (Figure 6.2):

1. Visual layout and the use of graphical elements (pointer, single annotation, symbol, expanded view) used to annotate the environment;

2. The type of content within the annotation (e.g. name, address, description);
3. Amount of AR annotations;
4. Amount of information within AR annotations.

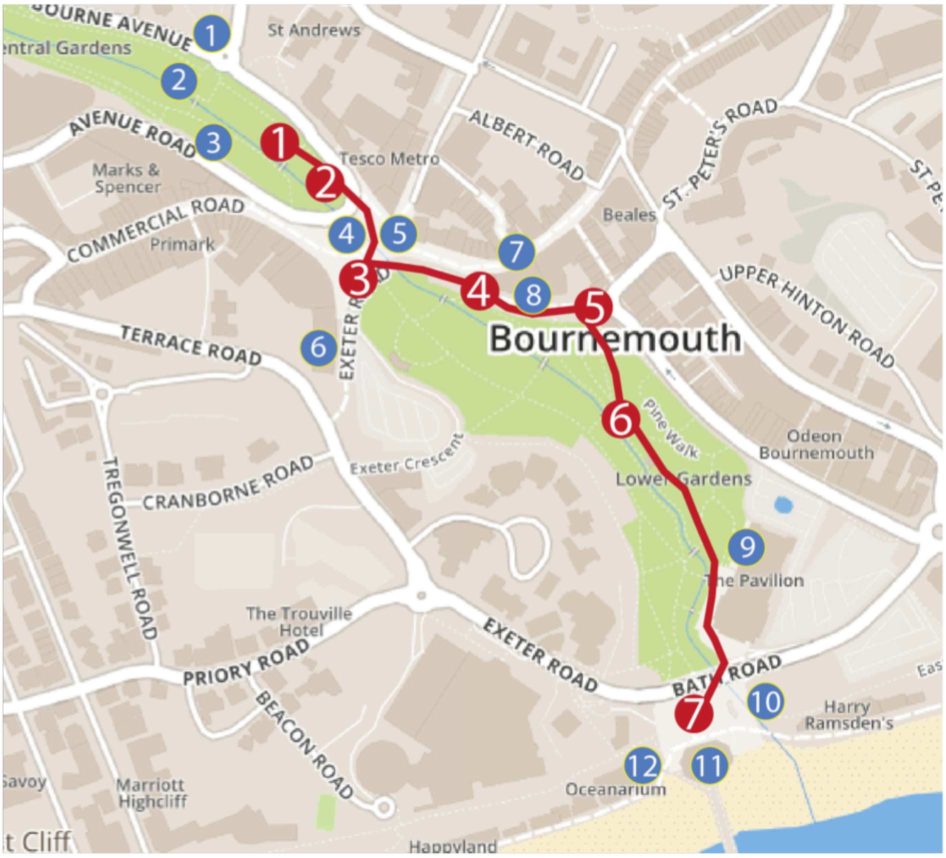
**Figure 6.2. Selected commercial AR browsers: A) LocalScope, B) Junaio, C) AcrossAir and D) Wikitude**




## 6.3. Method and Procedure

After the four browsers were selected, a number of decisions had to be made regarding the procedures and techniques adopted for the mobile field-based evaluation (Stage B.2, B.3 and B.4, Table 4.2). The next sections describe the route, locations, participants and materials used during the study, as well as the pilot that was conducted prior to commencing the empirical data collection.

**Figure 6.3. A map of the central part of Bournemouth with the selected locations and targets for the field study**





1. The Town Hall (TH)


2. War Memorial (WM)


3. Central Gardens (CG)


4. Camera Obscura (CO)


5. V-Club (VC)


6. Bournemouth Square (S)


7. The Arcade (A)


8. Information Centre (IC)


9. The Pavillion (P)


10. Waterfront Building (WB)


11. Amusement Arcade (AA)


12. The Oceanarium (O)

### 6.3.1. Route, locations and targets

Data gathering was conducted in the centre of Bournemouth (UK) on a predefined route, which: (i) was a representative tourist route, recommended by the tourist information centre; (ii) included different types of urban environment (park, typical urban, coastal urban); (iii) included different types of target objects (see below); (iv) provided different visibility (full, partial) to target objects.

Along the route, 40 potential points of interest were considered. Twelve different in function and structure points of interest were finally selected (Figure 6.3), visible from 7 key locations in the city. Seven of the targets were either stand-alone structures with easily distinguishable contour (e.g. Town Hall, War Memorial), while the rest were surrounded by other (similar or different) physical entities (e.g. V-club, the Arcade).

### 6.3.2. Preliminary testing of AR browsers

The selected AR browsers (Junaio, Wikitude, AcrossAir and LocalScope) were tested extensively prior to the study on three separate occasions at different times of the day. Screen captures and videos were collected on the indicated locations and with the selected targets. The GPS error was also measured with a Garmin GPS eTrex receiver. The positions for the stops were then adjusted so that there is no more than 3-4m positioning error (the minimum for the route). Where the annotation for the target was not precisely overlaid on top of that object, or within the boundaries of its representation on the screen, this was noted down (Table 6.1). The number of the annotations appearing on the screen was also recorded and ranged between 2 and 13 in all conditions.

**Table 6.1. Targets where the AR annotation appears to be placed on top of the target (v) or displaced (x) and number of annotations on the screen**

| Target | WM | TH | CG | CO | VC | S | A | IC | P | AA | O | WB |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Position** | x | v | v | x | v | v | v | x | v | x | x | x |
| **Number** | 3-6 | 4-7 | 3-5 | 9-13 | 2 | 6-8 | 2-4 | 4 | 3-5 | 3-6 | 2 | 6-8 |

### 6.3.3. Participants

The field study had to be executed with a specific type of user representing a visitor to an unfamiliar city using a smartphone AR application (Table 6.2). Suitable test persons included people with broad characteristics and technical abilities, but had to be unfamiliar with the city centre of Bournemouth. Approaching visitors, however, is challenging due to the time pressure they experience when on holiday.

**Table 6.2. Profile for the participants in the field study and assigned group**

| Test person | Sex | Age | Country of origin | Background / Expertise | Dm. Hand | Corrected vision | Smartphone experience | AR exp. |
|---|---|---|---|---|---|---|---|---|
| **JUNAIO** | | | | | | | | |
| **TP1** | M | 55 | United Kingdom | Transport Management | Right | NO | NO | NO |
| **TP2** | M | 39 | New Zealand | Tourism | Right | YES | YES | NO |
| **TP3** | F | 38 | China | Marketing | Right | YES | YES | NO |
| **TP4** | F | 19 | Bulgaria | Communication and Media | Right | YES | YES | NO |
| **TP5** | F | 29 | Germany | eTourism | Right | YES | YES | YES |
| **LOCALSCOPE** | | | | | | | | |
| **TP6** | M | 61 | United Kingdom | Psychology | Right | YES | YES | NO |
| **TP7** | M | 19 | United Kingdom | Archeology | Right | NO | YES | YES |
| **TP8** | F | 22 | Indonesia | Tourism and Hospitality | Right | YES | YES | NO |
| **WIKITUDE** | | | | | | | | |
| **TP9** | F | 28 | Poland | Psychology | Right | YES | YES | NO |
| **TP10** | F | 24 | Austria | Marketing and Design | Right | NO | YES | NO |
| **TP11** | F | 33 | Spain | Social work | Right | NO | YES | NO |
| **ACROSSAIR** | | | | | | | | |
| **TP12** | F | 24 | Spain | Education | Right | YES | YES | NO |
| **TP13** | M | 19 | Bulgaria | Computer Animation | Left | NO | NO | NO |
| **TP14** | F | 26 | Italy | Economics | Right | YES | YES | NO |

Prior to the field study, an invitation was prepared which indicated that the participant should be unfamiliar with Bournemouth but have strong desire to learn about the city (Appendix 2). It also gave further details about the duration of the experiment. The invitation was then distributed in two versions: digitally as a PDF and as a hard copy printed on an A5 sheet. The invitation targeted the newly arriving international and national students (added to their welcome package, distributed personally at welcome events, through facebook and twitter) to Bournemouth University, but also visitors to

Bournemouth in the indicated period. Therefore, the invitation was also distributed to several travel blogs and websites (CoachSurding.com).

In total, more than twenty visitors to Bournemouth were willing to participate in the study. Due to time and weather constraints not all of them could participate in the study. Fourteen participants (9 female, 5 male) completed the study, all of whom were relatively unfamiliar with the city centre of Bournemouth. Three of the participants were from the United Kingdom, while the rest had different nationalities (New Zealand, Poland, Spain, Bulgaria, Germany, Austria, Indonesia, Macao and Italy). Their mean age was 31.3 (range 19-61). Thirteen of the test subjects were right handed, 60% had corrected vision at the time of the test (eyeglasses and contact lenses). Two of the participants had no experience with smartphones prior to the test. None of the participants used AR regularly, but two of the test subjects reported trying similar types of applications prior to the test. Each participant was given £10 at the end of the study.

Each participant was assigned randomly to work with one of four commercial AR browsers: 5 participants worked with Junaio (Junaio, 2014), 3 with LocalScope (Cynapse, 2014), 3 with Wikitude (Wikitude, 2014) and, finally, 3 with AcrossAir (AcrossAir, 2014).

### 6.3.4. Evaluation tasks

The participants carried out four types of tasks with the allocated AR browser: match, reverse, reverse overview and decision, as described in Table 6.3.

**Table 6.3. Tasks and criteria for task completion**

| Task Name | Task definition |
|---|---|
| Match | The goal was to find an annotation about a specific physical target. This is similar to the pointing paradigm adopted for the usability testing of maps, where a subject is asked to find a feature on the map about a specific object of interest (Ottoson, 1987). |
| Reverse | The reverse version of the pointing paradigm (Sholl, 1987), where the subject is asked to find a corresponding physical target for a specific annotation. |
| Reverse Overviev | The instructions were to look at the annotations on the screen and match (associate) as many annotations as possible with actual physical targets in the surroundings. |
| Decision | The subject was asked to select a specific object of interest (e.g. café, attraction) where he/she would like to go next. |

Prior to executing each task the researcher took the device from the hands of the test person (TP) and re-loaded the content of the application. This was necessary because during the initial pilot study it was observed that there are problems with the automatic re-load of the content of AR browsers with the change of position of the user.

### 6.3.5. Procedure

The field study utilized the general high-level planning recommended by Jumisko-Pyykö and Utriainen (2011). Each experiment was carried out with one representative user for a period of 1.5-2 hours. It comprised of several main stages, also described in more detail in Table 6.4: (1) pre-test phase, (2) test phase, (3) post-test phase. The data was enhanced by contextual inquiries (Beyer and Holtzblatt, 1998), post-test interviews and additional background information collected through questionnaires.

**Table 6.4. Procedures for conducting the mobile field study**

| Phase | Sub-phase | Activities |
|---|---|---|
| **PRE-TEST** | INTRODUCTION AND PRACTICE | The researcher explains the purpose of the experiment. The participant signs a consent form and is asked to practice each task twice. |
| **TEST** | LOCATION 1 (1-7) | Information needs questions (the participant is asked to point out the physical objects/features they are interested in and to formulate questions connected with the location/specific targets) |
| | TARGET 1 (1-12) | Pre-task familiarity assessment questions -> Matching task and think aloud-> Rate certainty and difficulty -> Post-task discussion |
| | TARGET n… | (Same as above)<br>Information about objects of interest |
| | CONTEXTUAL INTERVIEW | Discussion about the experience, feedback, comments, suggestions |
| | LOCATION N… | (Same as above) |
| **POST-TEST** | POST-TEST PROCEDURE | Annotation design feedback<br>Design exercise<br>Background questionnaire, Santa Barbara Sense of Direction scale |

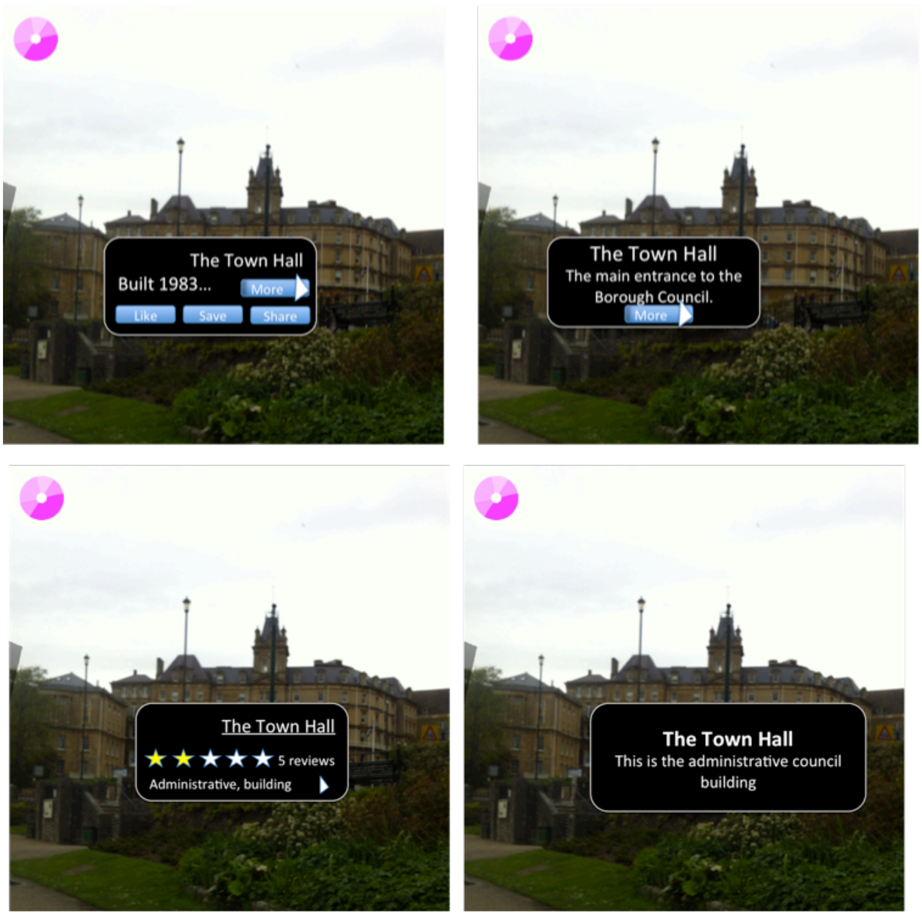
During the pre-test phase the researcher informed each TP that: their participation is voluntary and anonymous; they can quit at any time without providing further explanation; evaluation is directed towards the interface, rather than their own skills, knowledge and abilities. Each TP signed a consent form (Appendix 3). The researcher then explained the goal of the experiment and offered instruction on how to operate the smartphone. Each TP practiced all of the test tasks and thinking aloud twice before the start of the test. The participant was then taken to the first test location, instructed where to stand and in which direction to turn.

During the test phase, each TP was taken to all of the 7 locations consecutively and asked: (i) to identify the physical objects of interest; (ii) to formulate questions; (iii) to carry out matching tasks for 12 key points of interest in the city centre; (iv) to think aloud; (v) to discuss the results of each task. After the TP carried out all tasks at the location, the researcher provided information about each object of interest, simulating a real guide tour around the city. The contextual interview was carried out while walking towards the next location.

The study entered the final post-test phase when all of the tasks were completed. Each participant was provided a hot beverage at a food venue and asked to:

**Fill in a background questionnaire:** The questionnaire captured key demographic characteristics, such as age, experience with smartphones and AR (Appendix 4).

**Figure 6.4. Design alternatives for AR annotations with different content shown to participants after the field study.**



**Evaluate alternative AR annotations:** Afterwards, participants were asked to comment on the design of several annotations (Figure 6.4) with respect to scenarios of use that they were provided. The scenarios are outlined below:

- Learn: You are walking around in a new city and decide you want to learn something about this building…
- Eat: You are walking around in a new city and decide that you are hungry…
- Do: You are walking around in a new city and decide that you want to do something…

Each participant was asked to comment on the type of information that they would need and select one of the annotations which would satisfy their information needs best.

### 6.3.6. Contextual Inquiry

The contextual inquiries were carried out when users finished with all of tasks at a test site and started moving towards the next location (Table 6.4). During the CI, a protocol (interview guide) was used that formed part of the experimental protocol. Following the recommendations in the literature, the questions included in the protocol were open-ended and high-level in nature (Beyer and Holtzblatt, 1998; Rosenbaum and Kantner, 2007). The interview guide included very general questions about the information users obtained (*What are your impressions of the delivered information?*) and needed further (*Did the information answer your questions? What other information would you like to have access to now?*) in the current settings.

### 6.3.7. Presentation of stimuli materials

Each AR browser was viewed on an iOS iPhone 4 smartphone. The device has a 3.5" multi-touch display with 640x960 pixels resolution, 5MP camera, 30fps video, 1GHz Cortex-A8 CPU, HSDPA 3G network connection.

### 6.3.8. Equipment

The data was collected through a mobile field-testing mini-camera system, developed by Delikostidis and van Elzakker (2009) and made available for this research by the University of Twente, ITC Faculty. The system consists of two pairs of audio transceivers, three mini video cameras, a laptop, a handheld video recorder and two pairs of video transceivers. The input from all of the devices is synchronized at real-time and recorded through a video quad processor. The system allowed capturing data relating to (1) the user, (2) the environment, (3) the interaction that is taking place between the user and the device.

During the field study the equipment was fitted in a backpack carried by the test subject (Figure 6.5). The backpack was connected through a cable to the researcher's display. Whilst the experiment took place, the researcher stayed at approximately 1.5-2 meters behind the TP.

**Figure 6.5. The mobile field testing set-up**

### 6.3.9. Protocols

In order to make sure that all procedures were followed for each test subject in the same manner, several protocols were prepared and carried out. These included:

1. Preparation protocol – describing all of the activities that are necessary before meeting the test subject, including making sure that all materials are ready, the batteries are charged, etc.

2. Introduction protocol – this described the words for the participant that introduced the experiment and the general procedure.
3. Practice protocol – these were activities and tasks during the practice.
4. Experimental protocol – these were the activities and tasks during the test.
5. Post-experiment protocol – these were the activities that were carried out at the debriefing session

Before each experiment the researcher made sure that all necessary documents are printed and ready, including all protocols and the consent form. Following the preparation protocol, the researcher also made sure that the SD card is empty, all batteries are charged and ready to use. Before each experiment the equipment was connected and tested in the research laboratory to make sure that everything is working properly. The researcher also made sure that there is an extra backpack and a bottle of water ready for each participant. This was necessary because participants were required to walk and speak at the same time. As a backup, the researcher also carried an additional smartphone battery, audio recorder, mobile camera and video glasses in case there are technical problems with the equipment.

### 6.3.10. Pilot testing

Due to the complexity involved in designing and carrying out a mobile field study, it was critical to conduct pilot testing before actual data collection commenced. Pilot studies can uncover a number of problems connected with the technical, physical or logical set up of a study (Lazar et al., 2010).

In order to provide valuable information, pilot testing was carried out with 3 users prior to obtaining the equipment for the field study and 5 users in the actual context of use (along the predefined route). The main objectives were to: (i) make sure that the field study does not take more than 1.30 h.; (ii) task instructions and questions are clear and unambiguous; (iii) all of the data that is necessary for addressing the research questions is obtained; (iv) there are no major technical problems with the equipment or obtained audio / video recordings. The obtained data was used solely for the purpose of identifying major problems regarding the set up of the field study and was not used during the final evaluation and analysis of commercial AR browsers.

The data from the pilot was transcribed and reviewed. It became obvious that some of the questions might be ambiguous. For instance, the question “After having looked at the screen now, describe in your own words what information is available around you” was confusing and was removed in consequence.

In addition, it was also observed that the viewing direction of the first mini camera on the hat of the user provides limited information as it is directed straight down. This set up is necessary for experiments where participants work with maps and/or other interfaces that require users to hold the device horizontally. However, in this study the participants lifted the device vertically in order to preview information and this is why it was considered suitable to rotate the camera at 90 degrees.

## 6.4. Analysis

The resulting video and audio data was a rich information source of quantitative and qualitative nature. All of the resulting data files were stored digitally and backed up on several external HDs. The files were renamed according to the TP numbers and the type of material. For instance, the data for TP3 includes: one document file named TP3_background_questionnaire.doc, two video/audio recordings named as TP3_video_1.avi and TP3_video_2.avi and one audio recording named as TP3_Final_interview.mp3. In this way it was easy to organize the data and perform searching and analysis activities faster.

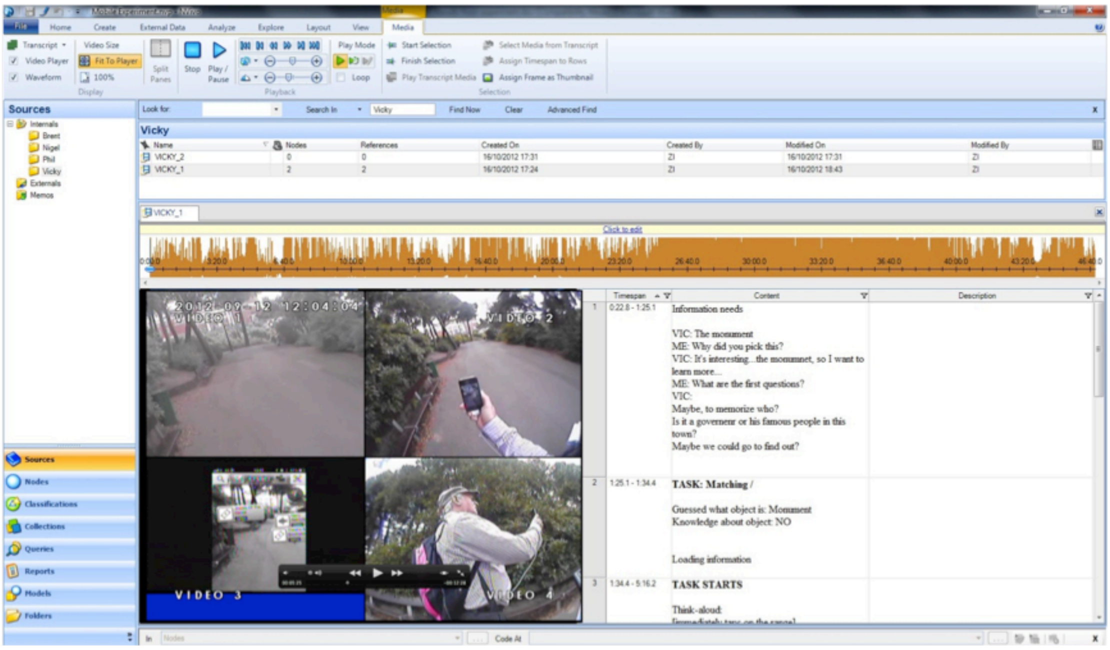
The total size of video/audio files and documents collected was 120GB. An example of the video recordings of the experiment, in the form of a screenshot, is shown in Figure 6.6. The presented video frame shows test person 1 (TP1) standing on Bournemouth Square in front of one of the selected targets for the experiment (V-Club), using the Junaio AR browser. The four different video signals, which were captured synchronously and in real-time, are shown together during the video playback. Each video signal has an identification name on it, such as “ENVI” (the environment in the viewpoint of the TP captured by the front camera on his/her hat), or “DISP” (the screen capture of the mobile device).

**Figure 6.6. A screenshot from the acquired video recordings**


The first step was the preparation of the data, which included verbatim transcription from the four video signals. The analysis was quantitative and qualitative, as the data captured the thinking-aloud and the post-task discussions, as well as the behaviours and reactions of the TPs to different events. At the same time, quantitative information, such as the time it took to complete the tasks, were also extracted.

For the transcription and qualitative analysis, QSR's nVivo, version 10.2 (QSR International, 2014) software package was used. This software program was found most suitable as it allows textual transcription and annotation to be directly linked to the corresponding video/audio material (Figure 6.7), i.e. time stamping of video, audio and textual materials. The analysis included a total of 168 matching tasks, 14 reverse tasks, 14 overview tasks and 14 decision tasks, altogether 14.5 hours (869.13 minutes) of recording. The developed taxonomy integrated coding categories and individual codes for the observed behavior (embodied interaction) with the device, observed (usability) problems and issues plus the obtained feedback (comments) during the think-aloud, post-task discussions and contextual interviews. Care was taken that the time code of each event refers as closely as possible to the start of the event. A degree of precision of 0.1s was deemed to be the maximum, yet enough for the purpose of the study. Word-for-word transcriptions of concurrent and post trial verbal protocols were examined together with the integrated videos.

Figure 6.7. Screen capture from nVivo



After transcription and time stamping, quantitative analysis was carried out with the IBM SPSS, version 19 (IBM, 2014) software package. General Linear Models were used, with the type of interface (Junaio/Wikitude/AcrossAir/LocalScope) as a between-subject, and the type of object as a within-subject, independent variable. Finally, qualitative analysis was carried out, followed by the development of task sequence models, which were later consolidated into one task model.

## 6.5. Findings

Despite being unfamiliar with AR, all tourists had a very positive attitude towards the technology and liked the idea of augmenting their environment with information. The technology was expected to be especially useful in unfamiliar urban spaces. After the practice phase, all tourists individually expressed the opinion that AR is easier to use than other information sources, such as guidebooks and maps, because it *"picks on the things that are around" (TP9)* and *"there is a cut off in the distance and you don't get stuff that's on the other side of town…but literally within eyeshot" (TP7)*. Shorter time periods to find relevant information was considered one of the biggest advantages of AR browsers, *"…because without having a map and reading in another place it can still show you something interesting…otherwise it usually takes a lot of time"(TP4).*

However, as the field study progressed, the participants voiced a number of negative comments and expressed their dissatisfaction with the provided (or lack of) content and specific elements of the interface:

In addition to this, results from the matching task indicate that the content of AR annotations is far from optimal. The following sections discuss first the quantitative and qualitative results for the matching task, then the results with respect to the information needs task, followed by the observed general usability issues.

### 6.5.1. Association of AR Annotations and physical entities

Association is the process of (mentally) being able to relate a virtual annotation to only one physical entity. If users are unable to carry out this process, the LBS interface and, in the case of this study, the interface of an AR browser, becomes difficult or impossible to use. During the field study, the participants carried out 168 association (matching) tasks. The following section describes findings from quantitative (performance) and qualitative (observations of body movement, interaction with device and content, feedback, thinking-aloud) data.

#### *6.5.1.1. Performance measures*

Objective (success, time) and subjective (difficulty and certainty) performance measures were collected for each matching task. A matching task was considered successful if the participant could identify the correct annotation (from all other annotations displayed on the screen of the smartphone) for a specific target object. The means for success rate, time, certainty and difficulty for each of the four AR browser applications can be found in Table 6.5.

**Table 6.5. Mean performance results for individual browsers**

| AR Browser | Mean success rate (%) | Mean time (sec) | Mean certainty (1-5) | Mean difficulty (1-5) |
|---|---|---|---|---|
| **Junaio** | 65.5 | 31.7 | 4.05 | 2.02 |
| **Wikitude** | 50 | 33 | 4.47 | 2.00 |
| **LocalScope** | 50 | 24.3 | 4.33 | 1.82 |
| **AcrossAir** | 52.4 | 35.9 | 4.53 | 1.46 |

The average success rate for the four AR browsers was similar (50%-60%). Test subjects using LocalScope and Wikitude had the lowest percentage of successfully completed tasks. The results from a one-way between-subjects ANOVA showed that there is no significant effect of interface ($F(1,3) = 2.646$, $p=0.06$) on task success.

The overall mean completion time for all 168 matching tasks was 31.9 seconds. The group using AcrossAir had the highest average time, while the group using LocalScope was the fastest. The results from a one-way between-subjects ANOVA showed that the effect of interface type on task completion was non-significant, $F(1,3) = 0.842$, $p = 0.842$.

Subjects working with AcrossAir experienced the highest level of certainty and lowest difficulty. The results from a Kruskal-Wallis test showed that there is no statistically significant difference in reported certainty across the four interfaces ($Z=-2.267$, $p=0.23$). The same was the case with experienced difficulty ($Z=-1.755$, $p=0.79$).

There was, however, difference in the performance when participants tried to match virtual AR annotations with different types of physical targets (Table 6.6). The participants were highly successful (100%) during the matching task for the Oceanarium and the Tourist Information Centre. Test subjects made the most errors when carrying out the matching task for the Waterfront building (14%). There was a significant effect of object type on task success ($F(1,11) = 8.443$, $p = 0.04$).

**Table 6.6. Performance measures results for individual target objects**

| Target | Mean success rate (%) | Mean time (sec) | Mean certainty (1-5) | Mean difficulty (1-5) |
|---|---|---|---|---|
| **War Memorial** | 92.8 | 59.7 | 4.4 | 2 |
| **Town Hall** | 85.7 | 18.9 | 4.4 | 1.7 |
| **Central Gardens** | 78 | 31.3 | 4.8 | 1.4 |
| **Camera Obscura** | 21 | 41.5 | 4.8 | 1.9 |
| **V-Club** | 46 | 29.2 | 4.3 | 2.1 |
| **The Square** | 57 | 39.1 | 3.9 | 2.4 |
| **The Arcade** | 57 | 41.5 | 3.9 | 2.3 |
| **Information Centre** | 100 | 10.7 | 5 | 1 |
| **Pavilion** | 85.7 | 41.5 | 3.6 | 1.9 |
| **Amusement Arcade** | 31 | 26.1 | 4.8 | 1.6 |
| **Oceanarium** | 100 | 5.2 | 4.6 | 1 |
| **Waterfront building** | 14 | 37.7 | 3.2 | 2.5 |

TPs completed the matching faster when they had to associate the annotation for the Oceanarium (5.2 s) and the Tourist Information Centre (10.7 s) with their reference targets. On average, the participants were the slowest when performing the matching task for the War Memorial. The results from a one-way between-subjects ANOVA showed that the effect of object type on task completion time was significant, $F(1,11) = 3.774$, $p = 0.008$.

Users reported the highest certainty and lowest difficulty when matching virtual annotations with the Oceanarium and the Tourist Information centre targets. The task for the Waterfront building resulted in the lowest experienced certainty and highest difficulty (Table 6.6). The results from a Kruskal-Wallis test showed that there is a statistically significant difference in reported certainty across the twelve objects ($Z=-4.402$, $p=0.000$). The difference in reported difficulty across the four interfaces was also statistically significant ($Z=-3.955$, $p=0.000$).

The number of the annotations influenced mainly the time for each task. During all matching tasks the test subject first scanned through all of the available annotations on the screen and related each element within each individual annotation (name, distance, symbol) with the target object in order to make a conclusion.

#### 6.5.1.2. Factors that influence association

Prior to each of the 168 matching tasks, the participant first verbalized their assumption about the nature of the target object. The visual cues (e.g. height, textures) that the TPs used to infer the (non)visual characteristics (e.g. function) of the target object influenced how participants interacted with the smartphone screen and their conclusion in the matching task.

After making an assumption about the target object, each participant examined the smartphone screen and the available AR annotations. They then compared the visual characteristics of the target object and the inferred non-visual characteristics of the building (e.g. its function) with the elements and information contained within each virtual AR annotation. As Table 6.7 illustrates, the participants used the keywords (e.g. club, arcade, centre, café) contained within the name of the point of interest to make conclusions about association. The Junaio and Wikitude groups used the provided symbols in the annotations rarely, as they had difficulties understanding what they refer to due to their abstract nature. Distance was one additional parameter that created confusion, as participants were unable to judge correctly what is the actual distance to a target object.

**Table 6.7. Used visual cues, perceived name and function for each target object and elements within the AR annotations used to infer association**

| Target object | Used visual cues | Perceived name | Perceived function | Number of TPs | Used annotation element(s) |
|---|---|---|---|---|---|
| Oceanarium | Name on building | Oceanarium | Aquarium, zoo for fish | 14 | Name |
| Tourist Information Centre | Sign on building | Information Centre, desk, office | Tourist information centre | 14 | Keywords in name |
| Central Gardens | N/A | Gardens, park | Gardens, park | 6 | Keywords in name |
| The Square | N/A | Square | Square | 6 | Keywords in name |
| The Arcade | Name on building | The Arcade | Shopping mall | 10 | Name |
| | Contour | No name | Various | 2 | None |
| Amusement Arcade | Name(s) on building | Aruba | Pub, bar | 11 | Name |
| | | Amusement arcade | Entertainment, video games | 2 | Keywords in name |
| Camera Obscura | Name on building | Obscura café | Café, restaurant | 11 | Name |
| | Object on building | Camera Obscura | Rotating camera for entertainment | 1 | None |
| Pavilion | Pre-existing knowledge | Pavilion | Theatre, dance, performance | 2 | Name, symbol |
| | Architecture | N/A | Concert venue, theatre, pavilion | 8 | Symbol |
| | Architecture, size | N/A | Gala casino | 2 | Keywords in name |
| Town Hall | Pre-existing knowledge | Town Hall | Borough council | 1 | Name |
| | Architecture | N/A | Various | 8 | Name |
| | Architecture | N/A | Hotel, office | 3 | None |
| V-club | Contour, Architecture | N/A | Church | 11 | None |
| | Posters at front | N/A | Disco club | 1 | Keyword in name |
| War Memorial | Contour | N/A | Monument, war memorial | 11 | None |
| Waterfront building | Architecture | N/A | N/A | 1 | None |
| | Architecture, size, colour, materials | N/A | Shopping, offices, bar, club, hotel, industrial, pavilion | 10 | Various |

Tourists were most successful and required less time to match annotations with their target object if at least one visual cue of the physical object matched the content of the annotation. Such was the case for the Oceanarium matching task, where the name on the physical object matched the name of the annotation. As a result of this, the task was 100% successful and required on average only 5 seconds to complete (Figure 6.8).

**Figure 6.8. A view of the smartphone screen when matching annotations for the Oceanarium**

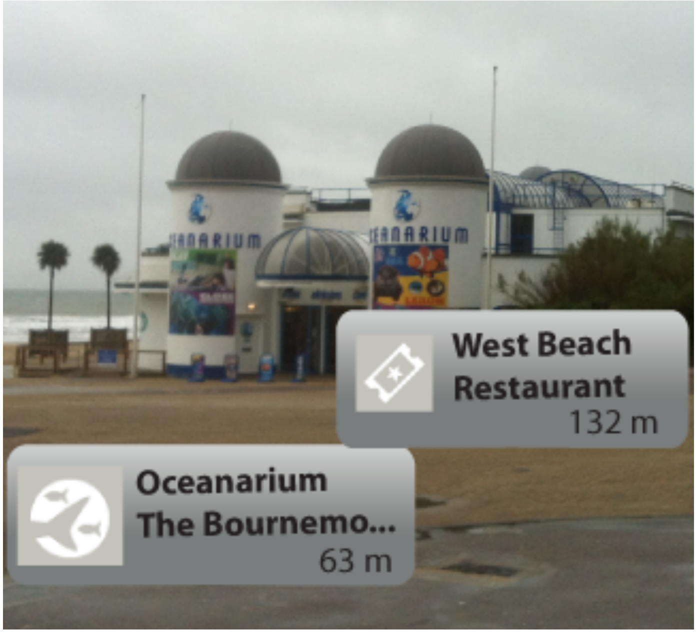



**Figure 6.9. A view of the smartphone screen when matching annotations for the Waterfront building**


When there was at least one match between the visual characteristics of the target object and the content of the annotation, participants were successful with relating the two, despite the imprecise position of the annotation. For instance, both the annotations for the Oceanarium and the Information Centre were displayed either lower on the screen (Information Centre) or to the right (Oceanarium) of their physical target object. This required that the participant scans the virtual annotation space before making a conclusion. The position of the AR annotations influenced success only when the virtual annotation appeared in a position which exceeded the 45 degrees lateral angle (see also

Section 6.5.1.3). In such cases, the subject automatically assumed either that there is no information about the target object or made errors in his/her conclusion.

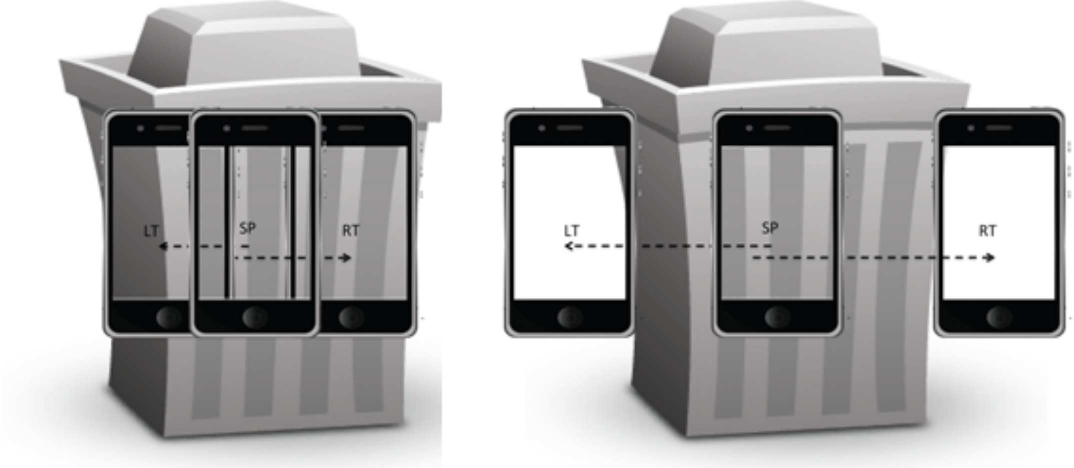
Association of annotations with physical targets was especially difficult and resulted in many errors when none of the visual characteristics of the target building matched the content of the annotation. When test subjects had to rely on other physical or structural properties of the target, they made errors, took longer time and reported higher difficulty and lower certainty. For instance, only one of the participants inferred correctly that the AR browser does not provide content about the Waterfront building. All other participants made errors in concluding that different annotations on the screen match the target object (Figure 6.9). The task required 35 seconds on average to complete. Likewise, participants experienced difficulties when the target annotation was superimposed visually over several targets with similar characteristics.

#### *6.5.1.3. Embodied interaction*

In the beginning of each matching task, each participant was asked to turn in the direction of the target. When the task started, the subject pointed the smartphone directly towards the centre of the target. If the annotation was not within view, or the participant was uncertain whether they have found the right information, they started moving the device laterally, either slightly to the left or to the right in a way that the physical representation of the target was kept within the borders of the screen.

Movement of the device differed also with respect to the distance to the target object. The lateral movement of the device was minimized when the target object was further away from the user and its representation could fit within the viewport. The lateral movement was bigger when the object was closer and its representation could not fit within the viewport of the smartphone (Figure 6.10).

**Figure 6.10. Strategy for lateral movement of the smartphone device when matching annotations with distant (left) and close (right) target objects.**




In all matching tasks for discrete objects, the lateral angle did not exceed 45$^{o}$ from the starting point in either direction. If the annotation was not within these boundaries, the participant concluded that there is no information about the target.

The matching tasks for the Central Gardens and the Square differed slightly as users pointed the device towards the ground when they first started the task. When they could not find the required information, they started moving the device laterally towards the left or right. When this strategy did not yield results, users raised the display and started following the same strategy as for discrete objects. In all cases, the process was accompanied by comments of confusion and frustration.

Further differences in hand and body movement amongst the browsers stemmed from the differences in design. For instance, LocalScope users had to stand very still in order to limit the movement of the directional pointer on the screen. Wikitude users were required to tap on all annotations appearing within the viewport in order to obtain information about physical targets. AcrossAir users had to swipe the display in order to view the annotations for targets further away from their current position. In all cases, the interaction with the display did not impact on the generally adopted reasoning strategy when associating annotations with their target, described in the previous sections.

### 6.5.2. Augmenting the right objects in the cityscape

Tourists information needs can be expressed as interest in particular points of interest (physical entities) and specific questions about those physical entities (Section 5.3). An

AR browser is expected to provide information about the points of interest that the user is already interested in. Therefore, on arrival at the first two test locations, the first set of questions asked was towards eliciting the location-based information needs of the participants and the triggers for those needs.

#### 6.5.2.1. *Selected points of interest and user-generated questions*

The results show that there was no uniform preference for specific physical objects or buildings in the surroundings. The participants identified a set of 16 different local objects of interest, including the central gardens, St. Stephen's church and the town hall. Figure 6.11 illustrates the identified objects for location 2 (The Square), together with the number of participants that selected each POI.

**Figure 6.11. Identified objects of interest at location 2 (the Square), together with number of participants that selected the POI (blue pins), starting position (red pin) and orientation (red triangle)**


The reasons for selecting specific objects of interest varied but all depended on the perceived qualities and characteristics of the physical objects/buildings in the surroundings. The participants' choice for objects of interest was mainly influenced by different visual cues in the environment. The architecture of the building, the contour, as well as the name on display were the most often visual cues participants referred to when selecting POIs. For instance, the building of the V-club drew attention immediately due to its imposing height and architecture when compared to the other buildings visible from the location on the Square.

Apart from attracting attention, the participants used visual cues to infer the non-visual attributes of the buildings (e.g. their function, significance and importance). This process, and the conclusions that each participant made, impacted the selection of POIs. For instance, when considering V-club, all participants inferred that the structure is a church. However, only 6 out of 14 TPs wanted to know more about the building. The inference of function influenced the presence or lack of interest in the building. For instance, TP9 pointed our that she has a general interest in churches, while TP10 mentioned that she would not like further information as there are many churches in her home town.

Participants could identify the function of the business most often by reading the name displayed on the entity. Using pre-knowledge, most TPs could infer correctly that Debenhams is a department store or the Moon in the Square is a pub. Such cafes and restaurants did not seem to attract further attention or trigger information needs. However, having just arrived in the UK, TP13 pointed that the names are unfamiliar and wanted to know what these places are:

> *"The first thing is this building…Debenhams…I don't know what it is…it's just when I see the signs on the buildings, like the Moon in the Square and I think it's really interesting and maybe I would go to see what it is…" (TP13).*

Inferred non-visual properties also influenced the types of questions users formulated. For instance, the questions of all 6 participants who inferred that V-Club is a church were very similar in nature. In another example, most TPs were not interested to know more about the Camera Obscura, as they inferred incorrectly that the building is only a café/restaurant. However, TP1 noticed that the building contains a rotating camera, inferred that this is an entertainment venue, and wanted to know how much it costs to get inside. During the same task, TP4 used the contour of the building to infer that it is a symbol of the city and is important for the surroundings. She wanted to know what is the function of the building and why it is important for the residents. Using the specific architecture of the building, several participants assumed that the town hall is a hotel and wanted to know when the building was built. Their questions were similar in nature and type. The same situation resulted in a different set of questions when TPs identified the administrative function of the object using the architecture and the size of the building as cues.

#### 6.5.2.2. Expectations and dissatisfaction with current augmented objects

The inferred non-visual properties of POIs influenced the expectations of users with respect to the content they will find on the AR browser interface. Once an object was identified as important and/or interesting, the tourist expected to find an annotation about that object. On several occasions, the participants could not find the annotation which referred to the object they have identified as important. This happened either because the annotation was missing, or the participant could not associate the available annotations with the physical entity. In ten of the test sessions, the inability to find the annotation for a specific object that the user identified as interesting resulted in confusion, especially when the TP expected that the object is important for the history and/or present of the city:

> *TP4: "I would say I can't find it…which is weird because it's the symbol of the city, it should be there" (TP4).*

In such cases, the participants took additional time to go through most of the annotations on the screen and reason about their content. When this happened, the participant reported that working with the application is more difficult than they have expected:

> *"It was difficult…because it's one of the most important points in Bournemouth so there must be some information" (TP5).*

Throughout the field study, the participants commented that they are unsatisfied with the type of objects the AR annotations refer to. All of the twelve test persons expected to find more information about specific tourist attractions in the city. The wide availability of content about local restaurants, cafes and shops was considered less valuable for the current situation:

> *"I think it's really commercial…it's just about shops and restaurants, but not really explaining the background of the city" (TP5).*

The lack of content for tourist-specific attractions also influenced the overall impression of quality of the provided content:

> *"I don't know whether that's complete information…if it's purely advertising then I am not sure about the quality of the information…if this is the case I'd be worried what's missing" (TP1).*

The tourists considered monuments, historic buildings and other tourist attractions (such as the Bournemouth balloon), as some of the objects that need to be augmented with content. The need to filter out content was brought up by six participants and was considered a critical functionality for AR browsers:

> *"I want it to filter out and only show attractions and at the moment I can't see, or I can't work out anything that I am seeing if it is an attraction" (TP1).*

The lack of content for POIs that were considered important from a tourist point of view led to confusion. However, this situation seemed less problematic if the TP considered the target of less cultural or social importance. For instance, lack of information for target 4 was less problematic for users who considered that this is the Obscura Café, rather than users who identified the object as an important POI. In a number of cases, the TP concluded that the object is less important and/or interesting than they have originally anticipated:

> *"Maybe because it is not that much important and [that is why] there was no bubble" (TP12).*

### 6.5.3. Providing the right content

#### *6.5.3.1. Lack of relevant content*

Table 6.8 describes the specific questions that the participants formulated during the field study. The table also illustrates how the visual characteristics of a physical entity, as well as its inferred non-visual properties, influenced the specific questions that the participants formulated.

All of the participants expected that the content in the AR annotations will answer their specific object-based questions. The type of content within the AR annotations was criticised when this did not happen. For instance, the use of an address within the annotations was considered irrelevant as a type of information that users wanted to access when looking at a specific POI. The provided content was considered superficial and non-informative. All TPs expressed the need for information that explains "*what something is*". In this case, categorical symbols (e.g. building) were considered less useful. More importantly, the participants commented on the redundancy of information:

> *"I can see everything…so I don't need the app, I would delete it immediately" (TP10).*

**Table 6.8. Used visual cues and questions that expressed an information need**

| Visual cues (triggers) | Inferred non-visual attributes | Point of interest | Questions |
|---|---|---|---|
| Architecture | Function: castle | The V-club | Is this a church or a castle? |
| Architecture | Function: church | The V-club | What is the name of the church?<br>What type of church it is?<br>How old it is?<br>When was it built?<br>How is it called?<br>Is there anything special inside? |
| Architecture | Function: hotel | Town Hall | What is this?<br>When was it built? |
| Architecture | Function: administrative building | Town Hall | No specific questions |
| Name on building | Function: café | Camera Obscura | No questions |
| Name on building | Significance, novelty | Debenhams<br>Moon in the Square | What is Debenhams?<br>What is the Moon in the square? |
| Contour and previous knowledge | Significance, uniqueness | Camera Obscura -> symbol of the city<br><br>Camera Obscura -> interesting building | What is this?<br>What was it before?<br>What was it used for?<br>What is special about the Obscura café? |
| Contour | Uniqueness | The Bournemouth Eye | What is the name of the balloon?<br>How long has it been there? |
| Contour | Function: attraction | The Bournemouth Eye | How far up does it go?<br>How much does it cost?<br>What is its purpose? |
| Contour | Function: monument | War Memorial | Who does it commemorate? |
| Surface / texture / colours | Age | V-club<br>War Memorial | How old is this?<br>When was it built? |

In the same context, many TPs agreed that there is a need for *"more specific information*" *(TP7)*, as they would not like to *"know just that"* (TP8), referring simply to the names of the entities around. In many cases, the name of the object was considered of less value as unfamiliar names meant little for tourists who just arrived in the UK.

#### *6.5.3.2. Influence of content on perception of space and POIs*

Throughout the study it was observed that the provided content influenced the perception of space and specific objects within the surroundings. In all experiments, the visual clutter on the display (availability of more than six annotations) was interpreted as a signal for the importance and centrality of a place: "*There are many bubbles together, so it must be many things around there" (TP9)*. In five of the test sessions, the TPs used the amount of annotations as an indication of where they should go next:

> *"From here it says that's where all the action is…so that's where I want to be…" (TP6).*

Available content also influenced perception towards specific objects. Points of interest were discarded when the TPs could not understand their name, what the symbol stands for or the description of the object. For instance, TP10 assumed that one of the most important conference and concert venues, the Bournemouth International Centre is *"maybe something not interesting for a tourist" (TP10)*. Four of the TPs identified the need for further description of the POIs around them in order to make a decision:

> *"Not really appeal to me until they've got some description" (TP3).*

Images and photos of the actual POI within the AR annotation were considered critical, especially when the annotated entity was not visible from the current location.

Apart from perception of urban space, content also influence decision-making during the study. During the decision task, test subjects were asked to examine the AR annotations and use their content to select one specific POI that they would like to visit. The final decisions are described in Table 6.9, together with the reasons for selecting a specific POI. For eight of the TPs, the virtual content did not provide useful visual cues and they discarded all of the annotations:

> *"Yeah, from here it doesn't actually look that intriguing…it just looks like there's a church and offices behind…so there's not really any clue as to whether the shops carry on around the corner…" (TP6).*

Only six of the participants made a decision based on the available annotations. In all cases, a point of interest was selected because its referent annotation "stood out" from the rest of the content. For example, TP2 and TP3 selected the Delice de Champs annotation because it contained a unique symbol (the Eiffel Tower). Both participants assumed that the annotation refers to a special type of food venue, where French food is served. For TP4 and TP13, the selection was based on the name in the annotation which was considered "something interesting".

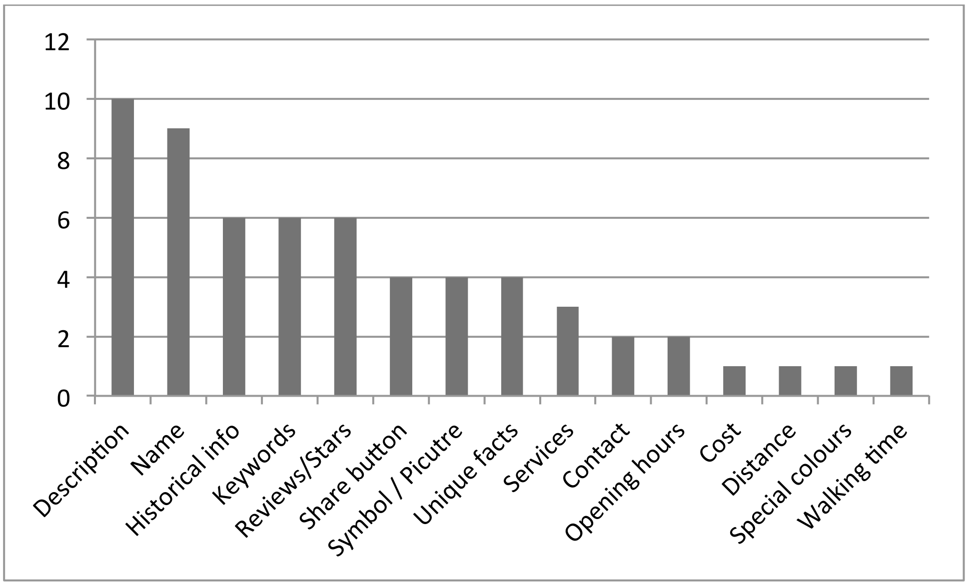
**Table 6.9. Decision task results**

| TP | Selected virtual annotation | Reasons | Assumption what object is |
|---|---|---|---|
| **TP1** | NO | Nothing looks interesting | N/A |
| **TP2** | Delice de Champs | Interesting and unique symbol (Eiffel Tower) | Patisserie, food |
| **TP3** | Delice des Champs | Interesting and unique symbol (Eifel Tower) | Food |
| **TP4** | Waterstones | Unfamiliar name, interesting and unfamiliar symbol | Entertainment, attraction |
| **TP5** | Past times | Interesting symbol | Entertainment |
| **TP6** | NO | Nothing looks interesting | N/A |
| **TP7** | NO | Nothing looks interesting | N/A |
| **TP8** | Horseshoe Common | Unfamiliar, interesting name | Shop, entertainment |
| **TP9** | NO | Nothing looks interesting | N/A |
| **TP10** | NO | Nothing looks interesting | N/A |
| **TP11** | NO | Nothing looks interesting | N/A |
| **TP12** | NO | Nothing looks interesting | N/A |
| **TP13** | Oceanarium | Interesting name | Oceanarium |
| **TP14** | NO | Nothing looks interesting | N/A |

#### *6.5.3.3. User preferences for content and information assets*

Apart from feedback obtained during the study, preferences for content were also examined during the final design exercise where users had to compose their own annotation. All participants specified that the provided information that they would find interesting would vary depending on the type of the building and the object being augmented. Preferences varied and different participants came up with different elements and structure for the AR annotations. However, they all contained more or less the same elements and pieces of information (Figure 6.12). The figure shows that there was a clear preference for including the *description* of objects of interest. However, the range of elements that participants wanted to have access to varied and included the walking time, cost to enter, special colours, the services that are offered, symbols and pictures.

**Figure 6.12. Preferred type of content in AR virtual annotations for tourism-specific objects when visiting an unfamiliar environment.**




## 6.5.4. General usability issues

There were no significant problems with the legibility of content. The observed usability problems were separated in four different categories and each problem was assigned a severity ranking (Table 6.10).

**Table 6.10. Usability problems experienced during the field study**

| Problem | Application | | | |
|---|---|---|---|---|
| | Junaio | LocalScope | Wikitude | AcrossAir |
| **1. Overlap of annotations** | 3 | 0 | 1 | 0 |
| **2. The size of the annotation was too small** | 2 | 0 | 1 | 0 |
| **3. Distance-based filtering of annotations led to errors** | 1 | Not used | 1 | Not used |
| **4. The linear layout of the annotations led to confusion** | 0 | 1 | 1 | 0 |
| **5. The movement of annotations led to confusion** | 0 | 1 | 3 | 3 |
| **6. The application did not load properly** | 3 | 3 | 3 | 3 |
| Legend: 0 - No problem<br>1- Catastrophic problem, preventing the successful completion of a task<br>2- Major problem, resulting in increased time, reported difficulty and/or lower certainty<br>3- Minor problem, resulting in minor confusion, negative comments or overall attitude | | | | |

Overlap of annotations was most problematic for Wikitude users, as it resulted in participants not being able to access the content of certain AR annotations. Likewise, the size of the annotations was mainly a problem for Wikitude users as they were forced to tap several times on an annotation before being able to access its content. The size of annotations in Junaio led to difficulties with reading the content related to targets which are far from the current position of the user.

Apart from overlap, a linear layout led to confusion, as it provided no cues regarding where targets are positioned in the surroundings. As a result of this, users often had to go through all of the content and use the distance indicator to make conclusions, which often led to errors and increased time. The movement of annotations (e.g. due to hand tremors) was most problematic with LocalScope. This is because the white virtual pinpoints appear too close on the screen and the directional pointer switches very fast from one annotation to another. For this reason, the participants using LocalScope tried standing very still in order to use the application. However, this strategy often did not solve the problem. The (sudden) movement of annotations also led to negative feedback for the Wikitude and AcrossAir groups. Loading times and crashes were very common with all of the selected applications.

Only three TPs (TP2, TP4 and TP10) used distance-based filtering of annotations. The main goal of the test subject when using the function was to reduce the amount of annotations on the screen of the smartphone. In all cases the distance parameters were set incorrectly as users either over- or under-estimated the distance to a target object. the When this happened, the participant could recognise that they have made an error:

> *"Probably I would take the information as a guidance, it's probably not perfect, but my ability to judge distance is not perfect either"(TP2)*

In all cases the use of the distance-based filtering led to TPs expressing confusion or annoyance.

## 6.6. Task Sequences and Consolidated Model

After qualitative and quantitative analysis of the data, work continued with task sequence modelling. The first step was to create work sequences for each of the 168 matching tasks. For each of the instances, an abstract step was defined that "states the work done in each of the instances independently of the specifics of that instance" (Beyer and Holtzblatt, 1998, p. 173). To facilitate understanding of the process, the instances and abstract steps were also represented graphically. Each task fell within one of 21 different models. Table 6.11 shows two of the developed sequences for TP1 and TP6 when carrying out matching task 2 (the Town Hall).

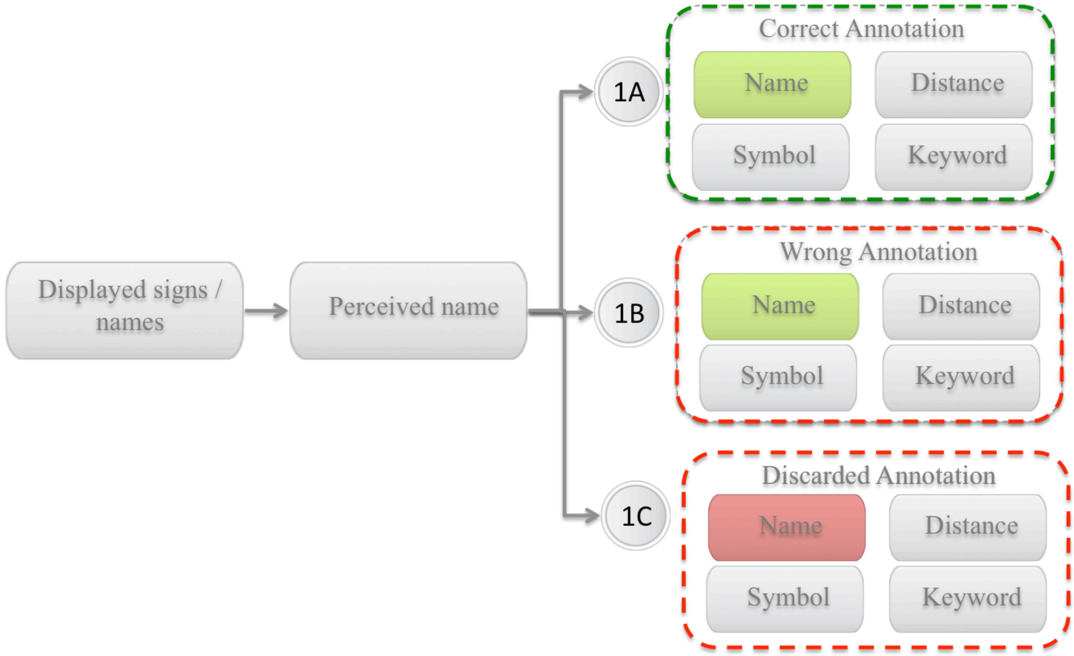
**Table 6.11. Work sequence for TP1 and TP6**

| TP1 (Junaio) | TP6 (LocalScope) |
|---|---|
| Assumption prior to task: Civic building<br><br>User raises display directly towards the building<br><br>User reads names within the annotations<br><br>User selects the annotation for the Town Hall<br><br>User makes conclusion based on the name of the building | Assumption prior to task: Hotel<br><br>User raises the display towards the building<br><br>User reads names within the annotations<br><br>User selects the annotation for the Citizen's Advice Bureau<br><br>User moves the display<br><br>User considers available names in other annotations<br><br>User corrects their answer to Town Hall<br><br>User makes wrong conclusion based on the names of the buildings |

After several iterations and further analysis, it was possible to generalize the sequences to 2 different models (strategies) that exhibited common properties. Both of the models are described below and visualized graphically on Figure 6.13 and Figure 6.14.

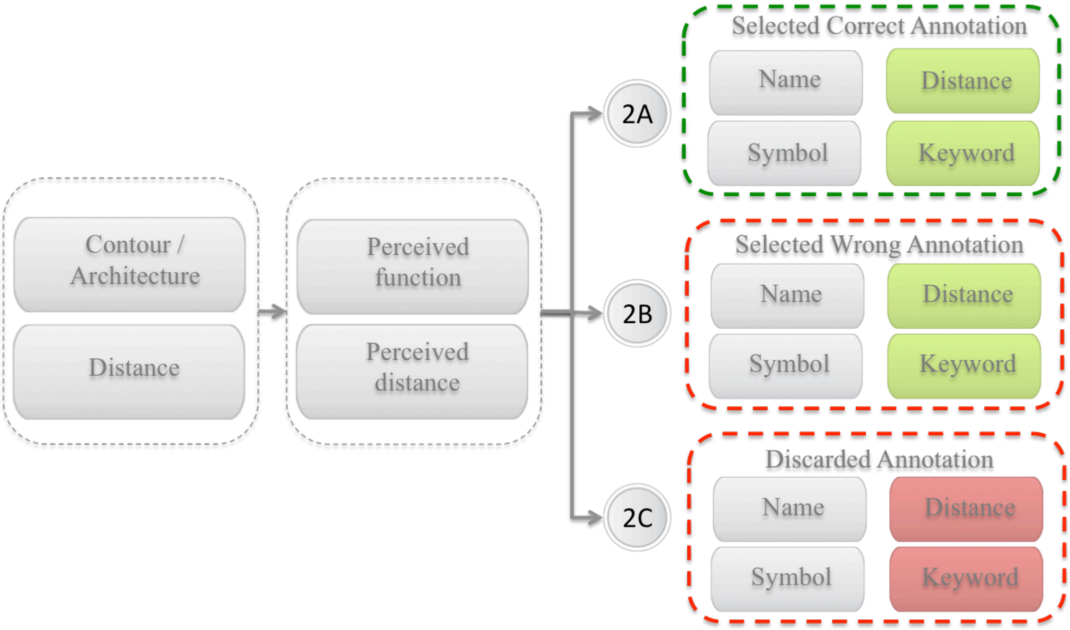
**Strategy 1 – Direct Visual Match** (Figure 6.13) – This strategy for matching physical objects with virtual AR annotations was observed when there is a name or a sign displayed on top of the target object. The displayed name influenced the assumption of the test person and they assigned it to signify the whole target object (perceived name). The test persons then scanned the AR annotations on the screen, looking for a match by using the name contained in each annotation. This strategy had three outcomes. The task was successful (case 1A) when the actual name of the target (Oceanarium) matched the perceived name by the test subject (e.g. Oceanarium) and the name contained within the annotation (Oceanarium). However, in some cases the perceived name of the target object (Obscura café) did not match the actual name of the building (Camera Obscura) and such tasks led to the test person either selecting the wrong annotation (case 1B) or discarding all annotations altogether (case 1C).

**Figure 6.13. Matching annotations with real physical targets based on displayed signs or names on top of the object with three possible outcomes: 1A) selecting the right annotation, 1B) selecting a wrong annotation, 1C) discarding all annotations on the screen.**




**Strategy 2 – Indirect Visual Match** (Figure 6.14) – In cases where the target object did not have a sign or a name, test persons used other visual cues, such as the architecture of the building, its contour, texture materials, and/or size to infer its non-visual attributes, such as its function. Since the test persons did not know in advance what the name of the target object is, the primary elements used in the AR annotations were the keyword (if available) and the distance that is displayed within the annotation. If the perceived function of the target object (e.g. disco club) matched a keyword in the AR annotation (e.g. V-club), test subjects were successful in finding the right annotation (case 2A). If, however, the perceived function of the building (e.g. church) did not match the content of the AR annotation (e.g. V-club, disco symbol), the test person either selected the wrong annotation (2B) or discarded the annotations altogether (2C).

**Figure 6.14. Matching annotations with real physical targets based on other visual cues with three possible outcomes: 2A) selecting the right annotation, 2B) selecting a wrong annotation, 2C) discarding all annotations on the screen.**



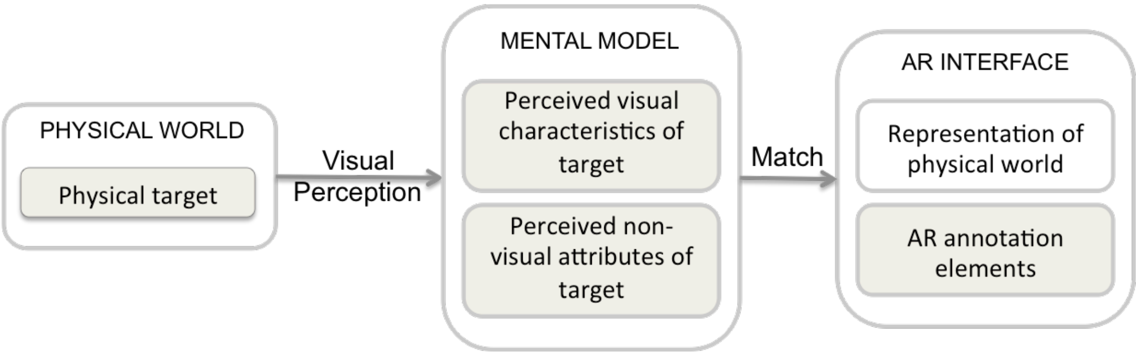

The developed graphical models suggested that there are several common tasks (Table 6.12) that users go through: (1) *use of specific environmental cues* which help them to (2) *reason and make conclusions about the (physical) target object*, after which they make (3) *use of the AR interface elements to reach a conclusion about association*.

After each sequence step was described in detail, the sequences were merged in a high-level consolidated task sequence model (Figure 6.15). Consolidated models are invaluable in revealing hidden task structures common to a wider user population (Beyer and Holtzblatt, 1998). They can be used to aid design because they show what aspects of the task can be supported through the current system, and what needs to be re-designed in the future. As tools for design, consolidated models also reveal problems and are extremely valuable for identifying tacit knowledge. A consolidated sequence model was necessary in order to reveal the structure of the matching task that is common to all observed instances. The final step of the analysis, therefore, involved merging the developed task sequence models in one consolidated task sequence structure.

**Table 6.12. Identifying common tasks in work sequences**

| Activity | Intent | Abstract step |
|---|---|---|
| **Use of environmental cues** | Determine useful visual cues<br>Identify non-visible attributes | Select visual cues for comparison with interface elements<br>Select visual cues that can be used to determine non-visual attributes |
| **Reasoning about target** | Identify non-visible attributes<br>Determine probable cues to match with interface elements | Make a conclusion about the non-visible attributes of target |
| **Use of annotation elements** | Associate each element with identified environmental cues | Use of interface elements to match selected visual cues |

**Figure 6.15. A graphical illustration of activities undertaken when using AR browsers, captured in the consolidated sequence model**




## 6.7. Discussion: The key role of visual salience and urban legibility

The results from the study show that participants were able to understand and work with AR browsers even without prior experience with AR or smartphones. The initial reactions of the participants with respect to AR confirm that this visualisation paradigm can create positive first impression and is still considered innovative and interesting by tourists. The study also revealed problems when associating physical targets and virtual content. Building on the developed theoretical framework (Chapter 5), the empirical evaluation revealed that tourists rely on perceived visible and non-visible characteristics of physical targets in order to match them with their corresponding virtual annotation. Consequent cognitive task modelling through task sequence models and consolidation shows that there are two main strategies that users adopt, depending on the physical target. This section summarizes and discusses these observations and the results from the study.

#### 6.7.1.1. Association of physical targets and virtual content

The overall performance of participants when associating annotations with their reference physical object was similar within the four groups in terms of time (31.9 seconds) and success (60%). While not ideal, these measures indicate that AR has a huge potential when it comes to knowledge acquisition for POIs in the immediate visible surroundings. Summarising the results for association, it was observed that performance with AR annotations was mainly influenced by two properties of physical objects: ***visual characteristics*** and ***legibility***.

When users first initiated the association process, they used the visual characteristics of the target object and employed different ***visual cues*** to relate the virtual content with its physical reference. Association was faster in the cases where there was a direct match between the perceived visible characteristics of the physical object (e.g. colour, physical name on display) and the elements of the virtual annotation (name, colours, pictures).

In the absence of a direct visual match, users tried to match the elements of the annotation with the ***inferred non-visible attributes*** of the building. This worked in situations where users could infer correctly at least one of the non-visible attributes of an object (e.g. the function of a building) and match it with an element within the AR annotation (e.g. the symbol for the Information Centre). The matching process failed when the visual characteristics of the building, or its ***inferred non-visual attributes***, did not correspond to information within the AR annotations.

Empirical research within urban architecture, planning and design has shown that people use the visual characteristics (appearance) of physical entities to make inferences about urban environments (Craik and Appleyard 1980; Nasar et al., 2005). This process is often referred to as ***legibility*** of urban environments. *Legibility* is the degree to which it would be easy for a person to infer the non-visual properties of an object, building, or a place. A highly legible environment is easily learned and remembered. This is a fundamental concept in urban geography, urban planning and architecture. A legible city facilitates its residents to find their way, “find a friend’s house or a policeman or a button store” (Lynch, 1960, p. 4). The ultimate goal of architects and urban designers, then, is to design buildings in a way that will instigate the appropriate perception (and behavior) for the building/place (Sullivan, 1918).

Researchers have shown that visual features, such as size, form, clarity, contour, colour, and dominance serve as useful probabilistic cues for determining the non-visual attributes of places, objects, and buildings (Craik and Appleyard, 1980). Most of the time, this process is unconscious and happens automatically. Non-visual attributes include the social status, cultural importance and function of buildings and structures. Non-visual attributes that have been studied empirically include the social status of residential homes (Lynch 1960; Royse 1969; Duncan 1973), and the cultural importance (Nasar 1989) or function of a building (Nasar et al., 2005).

In this study it was expected that difference in design would result in difference in thinking and reasoning. However, in all 168 matching task cases, the users used a similar strategy / reasoning to carry out the tasks. One possible explanation for this is that the information items did not differ amongst the four browsers, as they included: symbol/name/distance (Junaio), symbol/name/distance/keywords (Wikitude), description/name/distance (AcrossAir), name/address/distance (LocalScope). Users used those information assets to make conclusions about the association of an annotation and the target object.

The use of non-visual cues and their influence on the use of the AR interface also suggests that acquired landmark knowledge will influence significantly the association process. As users acquire more landmark knowledge, are able to rely only on visual cues (objects that they have seen before) and use acquired knowledge (e.g. name of the building) to match the annotation and the physical target. This observation is extremely interesting to investigate further in the context of other types of visual displays, as the relationship between familiarity and the use of geospatial technologies remains unexplained (Davies et al., 2010).

##### *6.7.1.2. Influence of perceived characteristics on information needs*

Both the visible and non-visible properties of physical entities play an important role and influence the (lack of) triggers for an information need, the expectations regarding availability of content, and the information users look for when they interact with the AR display. The visible and non-visible properties of buildings and physical entities are used to make inferences regarding POIs. Both of these properties influence the perception of the user in terms of whether certain POIs are interesting and/or important to know about and, hence, trigger or hinder the formulation of information needs. For instance, despite being visually attracted to the structure and architecture of the V-Club, users disregarded the POIs as the inferred non-visual properties did not lead them to conclude that it is a unique and/or important to know about. Legibility also influences the questions (queries) that users formulate regarding the entities around them. For instance, the questions that users posed during the field study differed depending on the assumed function of the POI (church vs disco club).

Despite differences in terms of use of visual cues, it is clear that the visible characteristics of the environment play a significant role and determine which landmarks will attract the attention of the user. Indeed, geo-information science literature defines landmarks as entities that "stand out", in comparison to adjacent items, because they are ***visually salient*** (Hirtle and Jonides, 1985). Until recently, the underlying notion in environmental psychology and geo-information science was that *visual salience* is an intrinsic property of specific physical objects (Appleyard, 1969; Hart and Moore, 1973). Properties such as significant height, complex/different shape and bright exterior were all considered to set apart objects and make them more memorable than others (Hart and Moore, 1973; Presson and Montello, 1988). While previously researchers were trying to define and list the key characteristics and properties of entities that define visual salience, Raubal and Winter (2002) note that the term can be used for "any of the elements of the city". Recently, Caduff and Timpf (2008) argued that such notions are wrong as salience is rather "a unique property of the trilateral relation between the ***feature*** itself, the ***surrounding environment***, and the ***observer's*** point of view, both cognitively and physically" (Caduff and Timpf, 2008; p.250). Therefore, both the characteristics of the user, the visual characteristics of the physical entity, as well as the surrounding ***context*** will determine whether an entity is perceived as a landmark.

The main implication is that visual salience (rather than distance-based proximity) will direct the attention of the user to specific entities, objects or elements. Knowledge of what visible characteristics people rely on to learn an environment is not sufficient, however, to explain why users formulate different questions when attracted visually to the same landmark. Another property of the physical environment is also important in order to explain this observation. Apart from the visible characteristics of physical entities, their non-visible properties also play an important role in attracting attention, and later recall. For instance, Harrison and Howard (1972) found that recall was related to components of location (actual physical location) and components of meaning (including economic, political, social, religious, ethnic, historical, and functional considerations). Reviewing empirical literature, Chalmers and Knight (1985) concluded that apart from distinctiveness (contrast with surroundings) and visibility, the functional or cultural significance of buildings and objects play a major role in the selection of landmarks. It is, therefore, important to understand how users draw conclusions about such non-visual attributes of entities within urban environments as they will influence the use of a mLBS interface.

## 6.8. Chapter Summary

This chapter described the first empirical evaluation in this study, focusing on the actual use of AR browsers in urban tourism context. A field-based evaluation was carried out in order to observe actual context of use and user strategies, behaviour, reasoning and problems when they work with AR browsers. Specific measures and decisions had to be undertaken during the design of the mobile evaluation, with careful selection of routes, stops, targets, tasks and questions (Section 6.3). Section 6.4. described the qualitative and quantitative analysis strategies employed to understand mobile interaction with AR browsers. The results from the field evaluation (Section 6.4) indicated that there are several critical problems that tourists experience when trying to find information about their surroundings through AR browser (Section 6.5). These included the lack of relevant content and support for effective association of virtual annotations and physical targets. The findings contributed to a better understanding of the overall context of use of AR browsers Furthermore, observation of user strategies to relate and superimpose virtual and physical spaces provided rich and useful results that were later validated through a laboratory-based experiment, described in the next chapter.

# CHAPTER 7

# QUANTITATIVE EVALUATION OF AR ANNOTATIONS

**7.1. Study Introduction and Goals**

Introduces the major hypotheses that guided the experiment

**7.2. Method and Procedure**

Describes in detail the design of the experiment, the preparation of stimuli materials and dependent variables

**7.3. Findings**

Reviews the major findings

**7.4. Discussion**

Revisits the hypotheses stated in the beginning of the experiment

**7.5. Chapter Summary**

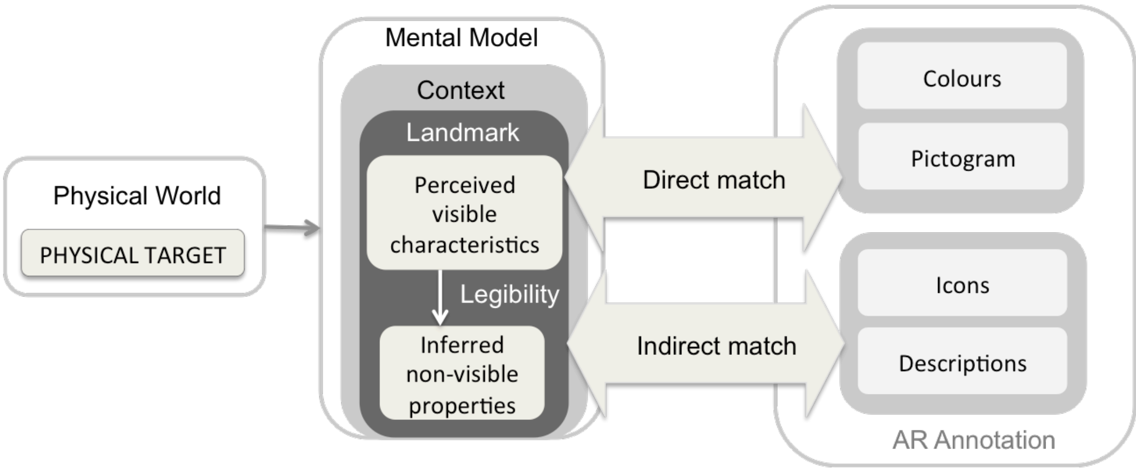
Synthesis of the key points in the chapter

## 7.1. Study Introduction and Goals

The main aim of this study is to investigate and improve the overall usability and utility of AR browsers when used in urban tourism context. To ensure this, tourists have to be able to successfully and efficiently associate virtual annotations with the physical targets they relate to. Literature has identified and discussed the importance of *spatial coupling* of virtual and physical objects (Chapter 5). This means that previous development and design has been based on the assumption that the position (or placement) of the AR annotation near or on top of its physical counterpart is enough to allow for successful match between the two. The findings from the first empirical study with commercial AR browsers and tourists (Chapter 6) indicate that association of physical targets and virtual AR annotations in complex urban environments depends on whether users are able to match the perceived visual or non-visual characteristics of the target with (at least one of) the visual elements of the annotation. From here onwards, this process will be referred to as ***visual coupling***, in order to distinguish between it and spatial coupling (using the position of the annotation to carry out association).

The findings from the mobile field study, illustrated in the consolidated task sequence model (Section 6.6) were further re-examined and used to expand the developed theoretical framework (Chapter 5). The findings suggest that there are two sub-processes that underpin visual coupling: direct and indirect visual coupling. In ***direct visual coupling*** users are able to match the visible characteristics of the physical target with one of the visible characteristics of the AR annotation. In ***indirect visual coupling*** users rely on inferred non-visual attributes of the physical target (e.g. its function) and match that with the elements (colour, symbols, text) of the AR annotation (Figure 7.1).

**Figure 7.1. Visual Coupling model of association: Direct and indirect visual match for association of annotations**




Visual coupling means that the perceived visible characteristics of the physical target, as well as its inferred non-visible attributes are both essential contextual factors that need to be considered when designing AR annotations. The model illustrated in Figure 7.1 allows to analyse different strategies and alternative designs for AR annotations. For instance, let us assume that the user is interested to learn more about the target object on Figure 7.2. In this case, at least in the mind of the user, the building becomes the physical target object and all surroundings become the (background) context. In order to acquire information about that target the user will first try to match the perceived visual cues (visible characteristics) with at least one element (e.g. colours, symbols/pictogram/icons) within the AR annotation. In a situation where the user is unfamiliar with the name of the building (the target object) or lacks knowledge about the context, she will be unable to match successfully the annotation with its target (Figure 7.2, A). However, if the AR annotation contains an element, e.g. an icon or a pictogram with a clear contour, that matches directly one element of the target object, then association will be successful (Figure 7.2, B).

A different scenario allows illustrating how indirect matching works. In situations where the user is able to correctly infer the function of the building, the addition of a keyword or a cartographic symbol (Figure 7.3) is assumed to allow successful association. The process, however, depends on the user being able to successfully determine that the building functions as a church and interpret the symbol (Figure 7.3, A). In this situation, association will be unsuccessful when the elements of the annotation suggest that the target object is a disco club (Figure 7.3, B), while the user has determined that the function of the building is a church.

**Figure 7.2. Design alternatives for AR annotations that rely on direct visual match**

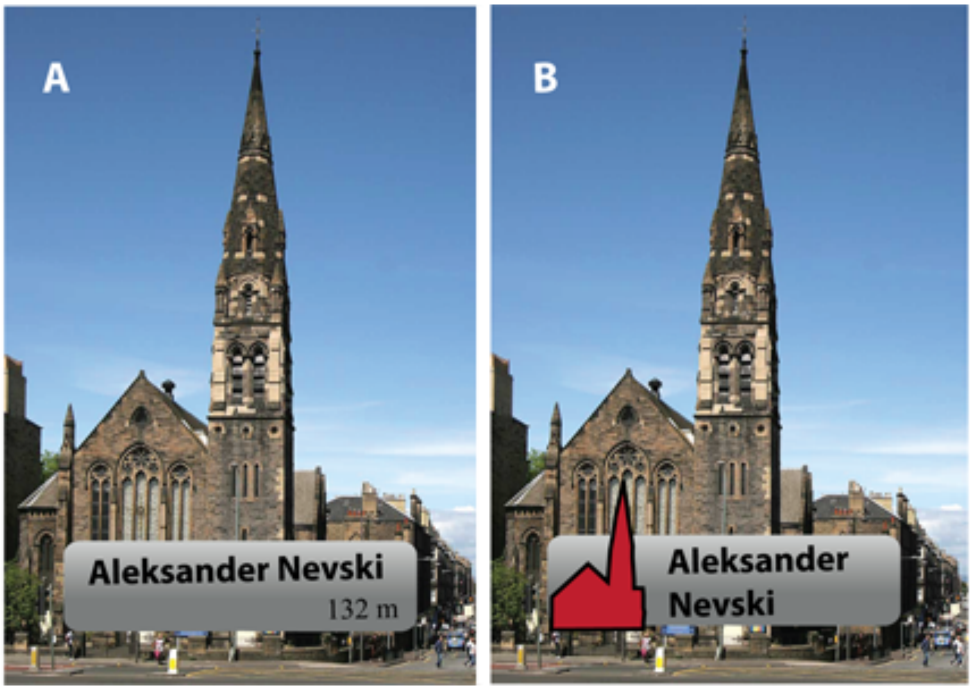
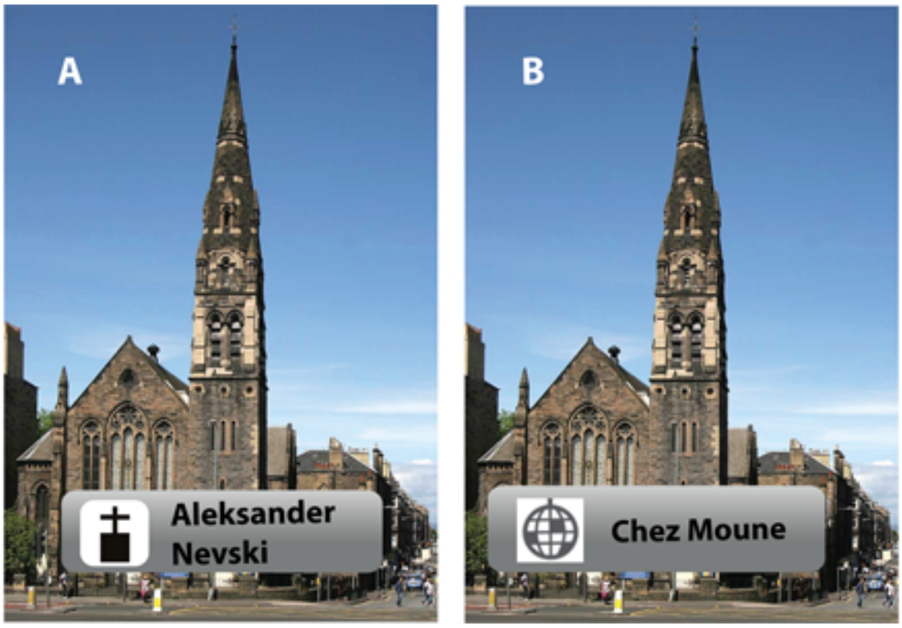
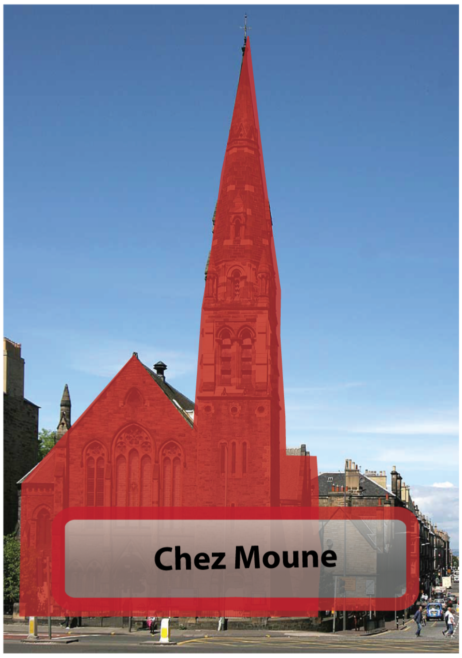



**Figure 7.3. Design alternaitves for AR annotations that rely on indirect visual match**




One of the ways to achieve a direct visual match, for instance, is by using a small picture within the AR annotation. As a true and photorealistic representation of the actual target object, at least in theory, users will be able to match the visual cues in the picture (contour, shape, textures) with the perceived visual characteristics of the target (contour, shape, textures) (e.g. Elias and Paelke, 2008). However, it is also important to consider the various factors that would influence the direct visual match, such as rotation of the picture, lightning conditions and weather. For example, if the viewing angle of the picture and the actual viewing angle of the user do not match it would be more difficult for users to relate the two spaces. Other factors, such as weather conditions, the size of the image, and lightning could influence the way users perceive the physical target and carry out the direct visual match of the two spaces.

**Figure 7.4. The colour-coding technique**




A way to address such challenges is through the use of colour-coding, which involves a semi-transparent overlay on top of the target which matches the colour of the annotation (Figure 7.4). In theory, the user will be able to match directly the two colours, even if the annotation is not precisely placed over the physical target. Colour-coding does not require interpretation and does not depend on the orientation of the user. The problem, then, is that the physical objects within urban environments often do not have one predominant colour, but various textures that are difficult to simulate. In addition, changing the background colour of the annotation could impact on the legibility of the text inside the annotation. AR displays, however, provide a suitable alternative to ensure a direct visual match by manipulating the representation of the physical object on the smartphone display (i.e. the base layer). In order to maintain legibility, the frame of the annotation, rather than its background, could be used to match the perceived colour of the physical target.

The model illustrated in Figure 7.1 can help analysing whether and which alternative designs will be more successful then others. In order to provide guidelines for design of AR browsers, the identified relationships between performance and direct and indirect visual coupling have to be confirmed and examined further. To this end, a laboratory experiment was carried out with 90 participants. Apart from validating the findings from the mobile field study, the overarching aim of the experiment was to examine the effect of direct and indirect visual coupling on user performance.

## 7.2. Method and Procedure

The experiment (Stage C, Table 4.2) was conducted in laboratory settings, following principles of experimental design in HCI (Lazar et al., 2010). While it would have been beneficial to test the AR annotations in actual context of use, a laboratory environment allowed to test more locations and targets while still controlling for external and confounding variables. At the same time, a laboratory environment was preferred as taking test subjects to different cities and locations would have been very resource intensive.

### 7.2.1. Hypotheses

One of the key implications from the mobile field study was that when tourists worked with commercial AR browsers, they were more successful when they relied on a direct match between the visual appearance of the annotation or its content to the perceived visual characteristics of the target object. Theoretically, task performance will improve if there is a direct match between the visual characteristics of the target object and the AR annotation, compared to using only a keyword where association relies on indirect match. Therefore, the hypothesis that the experiment addressed is:

**H1: Task performance will improve when there is a direct visual match between the graphical variables of the annotation and the representation of the target object, compared to the use of pictograms where users have to rely on mental rotation.**

During the field study (Chapter 6) it was observed that placement influences the speed with which targets and virtual content are related. Since the direct visual match relies on observable characteristics of both physical target and the AR annotation, the placement of the AR annotation would not influence the association process, and this is why a second hypothesis is:

**H2: When placement is imprecise, task performance will improve when there is a direct visual match between the graphical variables of the annotation and the representation of the target object, compared to the use of pictograms where users have to rely on mental rotation.**

A large number of AR browsers both within academia and industry make use of textual description (Langlotz et al., 2014). The use of names, keywords and descriptions is useful if users are able to match them with at least one characteristic of the observed target object. In cases where there are no physically visible names on the target object, users have to rely on inferred non-visual attributes of the building (indirect match) in order to relate the physical and virtual spaces. This process is ultimately reliant on knowledge and experience that tourists might lack, especially in unfamiliar environments. A large number of textual AR annotations are often combined with directional pointers (leader lines). Directional pointers would be useful when there are many physical objects clustered together, as then the annotation would not cover or overlap with these target objects. Despite their usefulness, however, keywords rely on indirect matching, and therefore, would perform worse than colour-coded designs:

**H3: Task performance will improve when there is a direct visual match between the graphical variables of the annotation and the representation of the target object, compared to the use of keywords, where users have to rely on the inferred non-visual attributes of the target object.**

In order to test the difference in performance between the colour-coded approach and the use of keywords, an additional hypothesis was that:

**H4: When placement is imprecise, task performance will improve when there is a direct visual match between the graphical variables of the annotation and the representation of the target object, compared to the use of keywords, where users have to rely on the inferred non-visual attributes of the target object.**

In theory, the use of photorealistic 3D pictograms would improve the association process, as users would be able to match the overall appearance (shape, contours, textures) of the symbol with the visual characteristics of the building. A number of empirical studies have confirmed that pictorial symbols have the advantage to be recognized easily, because no interpretation process is necessary (Elias and Paelke 2008). From a perceptual point of view, it is sufficient to match the represented symbol to the observable visible patterns in the environment. This process, however, depends on the detail included in the pictorgram (Bruyas et al., 1998). Therefore, the effectiveness of symbols would depend on how they are represented within the AR annotation. More complex 3D pictograms take up a lot of screen space and in order to include them within the AR annotations, their size has to be reduced. This means that users might have problems in associating the symbol with the physical target.

One of the most common elements used within AR annotations is a categorical symbol. The visual coupling model (Figure 7.1) suggests that the associative power of such elements is week and would work only if two conditions are met: the annotated target stands out from its context and has visual characteristics that could be used by the user to determine its function. Substituting categorical symbols with a pictogram symbol would be useful as then users will be able to carry out a direct (visual) match between the base and attribute layers and associate the target with its annotation successfully. It is proposed here that association of annotations and targets requires at least one direct visual match. This process will be faster than using indirect match with keywords:

**H5: Task performance will improve when there is at least one visual match between the graphical variables of the annotation and the representation of the target object, compared to the use of keywords.**

Users would be able to match pictograms with their targets faster, than when they rely only on a keyword, especially when the annotations are not directly superimposed on the target. Therefore, it is hypothesised that:

**H6: When placement is imprecise, task performance will improve when there is at least one visual match between the graphical variables of the annotation and the representation of the target object, compared to the use of keywords.**

The examined processes concern only the immediate surroundings of the user, and association between virtual annotations and visible physical targets. It would have been possible to test the association of annotations with non-visible targets, but this would have made the experiment longer, and required additional tasks and a substantially different set up (where maps and actual virtual environments are possible used). While acquisition of non-visible targets is important, it is a secondary requirement for AR browsers. Therefore, it was decided that only visible targets are included during the experiment.

### 7.2.2. Design, conditions and dependent variables

The structure of an experiment is typically determined by the number of independent variables. Experiments with one independent variable have a basic one-level design, while more independent variables require a factorial designs (Lazar et al., 2010). Once this is determined, a second choice concerns the conditions to which each participant will be exposed. In *between-group design* each participant is exposed to only one experimental condition. The main advantage is that users do not learn from different task conditions. When the experiment involves a smaller target participant pool or tasks that are less susceptible to learning a *within-group design* is more appropriate. Then, each participant is exposed to multiple experimental conditions (Lazar et al., 2010). A *split-plot factorial design* combines the benefits of both between- and within-subject designs where one set of independent variables are examined through a between-group approach, while other variables are investigated through a within-group approach (Lazar et al., 2010). The main aim of the laboratory experiment for this study was to investigate the effect of different designs on user performance. Since task performance (time, errors, certainty and difficulty) can be influenced by users first using one design, it was considered suitable that this variable is investigated through a between-subject approach. In addition, the experiment aimed to test whether there are any differences in performance when users start the matching process with the physical target in mind (matching task) or the AR annotation (reverse task). It was also interesting to examine the effect that imprecisely placed annotations have on task performance. In order to test the effect of task and placement, therefore, the experiment adopted a within-subject approach. Therefore, the experiment had a split-plot (3x2x2) design, with *AR annotation design* as a between-subject and type of *task* and *position* of annotations as within-subject variables (Table 7.1).

**Table 7.1. Independent variables in the laboratory experiment**

| Independent variable | No of variables | Type of variable |
| --- | --- | --- |
| **Task** | 2 | Match, Reverse |
| **Design** | 3 | Symbol (Pictogram/Abstract), Pointer, Colour-coded |
| **Placement** | 2 | Precise, Imprecise |

Three different design alternatives for AR annotations were developed to test the formulated hypotheses.

***Design 1 - Pictogram*** **(P)** – incorporated a photo-realistic 3D model (pictogram), mimicking the contour, shape and textures of the target object (Figure 7.5).

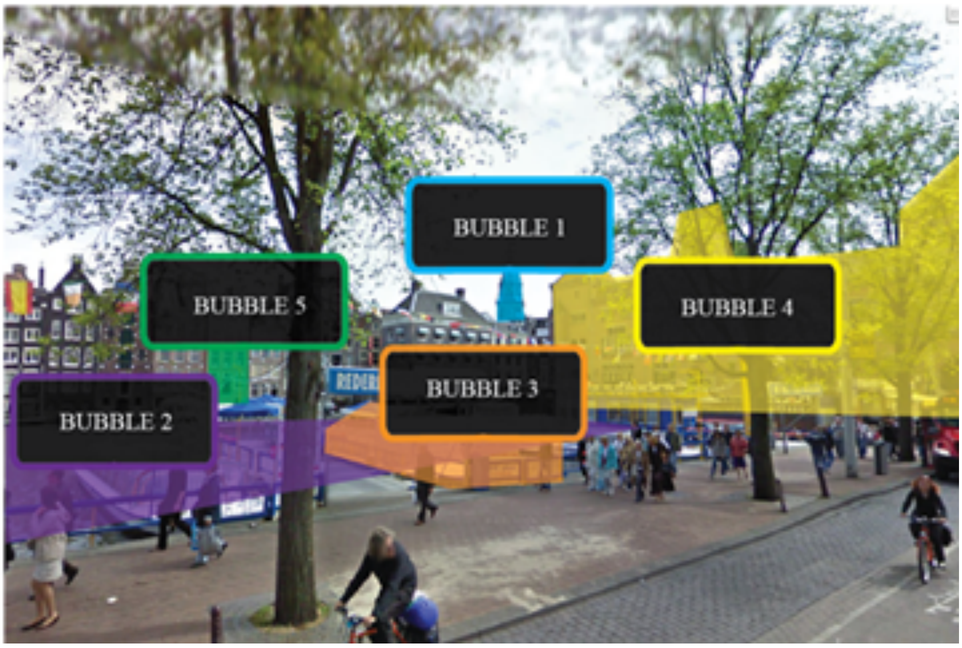
**Figure 7.5. Design alternative 1 (P) contains a pictogram representing the target object**


***Design 2 - Keyword*** **(K)** – incorporated a red crosshair directional pointer and a keyword. Since association based on inference of function is not always helpful, it was expected that this design will perform worse than Design 1 (H5), especially when the position of the annotations is imprecise (Figure 7.6) (H6).

**Figure7.6. Design alternative 2 (K) contains keywords that describe the target object**

***Design 3 – Colour-Coded*** **(CC)** – change of perception of the target object and association with the reference annotation is expected to perform the best, compared to Design 1 (H1, H2) and Design 2 (H3, H4) (Figure 7.7).

**Figure 7.7. Design alternative 3 (CC) uses a colour-coded highlighting model overlaid on top of the target object, which matches the frame colour of the corresponding annotation**




Apart from direct and indirect visual coupling (H1, H3, and H5), the experiment set out to investigate the effect of spatial coupling (the placement of annotations) on the predicted effects (H2, H4, and H6). Therefore, each design was tested with absolute placement and incorrectly placed annotations. The dependent variables measured during the experiment included: task time, task success (errors), certainty and difficulty.

### 7.2.3. Selected targets

One key implication from the field study (Chapter 6) was that performance of association will be influenced significantly if users are familiar with the physical objects and have acquired knowledge about them in advance. In such situations, the participants will rely on non-visual cues (e.g. name, function) to match the AR annotation with its physical target. Therefore, selecting popular scenes and tourist attractions (e.g. the Eifel Tower) could influence task performance during the experiment. However, since the AR annotations contained only numbers (and not names) and generic keywords (for Design 2, e.g. "building"), it was considered that any location and target could be included in the experiment, as long as no physically visible names are present on the physical targets.

More than ten popular urban destinations were considered during the preparation of the experiment, including Amsterdam, Rome, London, New York, Los Angelis, Delhi and Beijng. Three popular urban tourist destinations were selected: London, Berlin and Amsterdam. Ultimately, the selection of the locations depended on the availability of street view data, as this was required in order to simulate realistic scenarios during the experiment. Another factor that influenced the selection of locations and targets was the availability of 3D models that were later reduced in size and used as pictograms for Design 1.

Within the three main urban destinations, several different in structure and composition urban environments were selected, including: narrow and wide streets; roundabouts; squares; and canals. The selected locations included variable in contour and textures buildings, but also uniform in nature cityscape (Appendix 5). The selected targets were different in type, contour, shape, textures and colours buildings. They were positioned at variable distances from the current position of the user and included fully visible and partially visible structures. Whole targets (church, museum), as well as parts of buildings (shop, café, construction cranes) were selected for the experiment. The targets also varied in terms of length and height, textures, and colours.

### 7.2.4. Procedure for implementing the experimental mock-ups

Each of the design alternatives was implemented as an interactive digital mock-up (augmented photos), using the Axure RP Pro (Axure, 2014) package. During this procedure, specific measures were taken so that all variables for the three designs, apart from the selected manipulations (independent variables) remain constant.

First, screen captures were made of the selected location in Google Maps Street View. The AR annotations were then prepared in vector format with Adobe Illustrator and superimposed on the photos. In order to prevent the effect of various design variables, each of the AR annotations had the same specifications (Table 7.2).

The annotations were positioned directly above the visual centre of each target for the “precise placement” condition, and moved 150 points downwards for the “imprecise placement” condition. The positions of the annotations were identical among all of the three designs. In order to isolate the effect of pre-knowledge, all additional information, such as descriptions and keywords (apart from Design 2) was removed. Instead, and to make identification of the annotations easier to report, each annotation contained only a numbered label (e.g. Bubble 1, Bubble 2). The order of the numbers in which the annotations appear on the screen were positioned randomly throughout the screen to prevent learning.

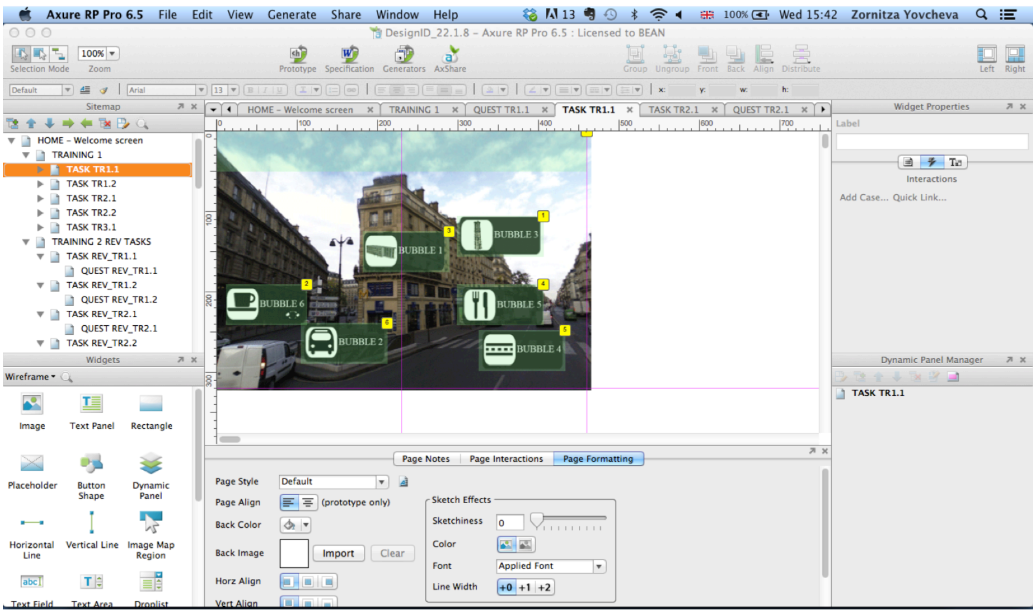
**Table 7.2. Characteristics of the interactive low-fidelity AR prototypes**

| Variable | Values |
|---|---|
| **Size of annotation bounding box** | 95x45 pixels |
| **Font** | Time New Romans, capital letters |
| **Font size** | 12 pt. |
| **Background colour** | Black |
| **Opacity** | 94% |
| **Frame colour** | No frame colour (except D3) |

The pictograms for the landmark design were extracted from GoogleSketchUp and exported as Collada models, as the format allows preview in 3D. The orientation was adjusted so that each landmark is viewed from the same angle. Custom symbols were also added (designed with Adobe Illustrator), following the international ISO convention for design of symbols for public information systems (ISO28564-1, 2010).

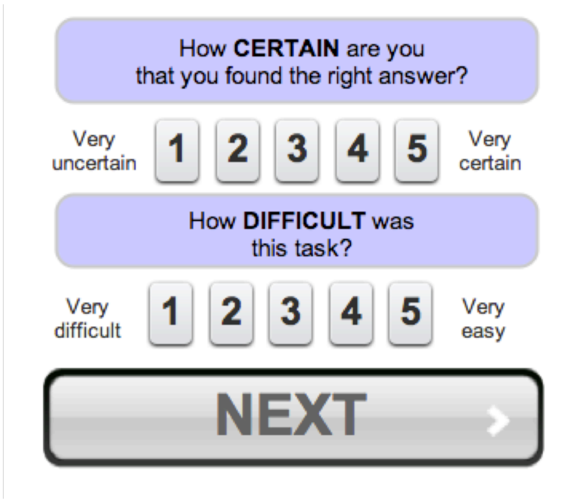
For the colour-coded design the visible targets were overlaid with a semi-transparent (60% opacity) layout, mimicking the outline of the target. The colour for the frame of their corresponding annotation was the same as the selected layout for each target. Care was taken the colours are neutral (green, yellow, purple) and do not distract or attract attention to individual targets, using the same hue and saturation.

**Figure 7.8. Screenshot of one of the prototypes, implemented with Axure RP Pro**



The prepared designs were imported in Axure RP Pro (Axure, 2014). Each mock-up consisted of 67 screens (Figure 7.8). A welcome and introduction screens were added. Axure RP Pro allows adding interactivity to the prepared augmented photos. The AR annotations acted as hyperlinks that jump to a subsequent page that required the user to assess the certainty of their answer and the difficulty of the task (Figure 7.9). Participants could not go back to already completed tasks.

**Figure 7.9. Certainty and difficulty screen**




### 7.2.5. Tasks

Two types of tasks, identical to the ones used in the mobile field study, were used:

*(M) Matching task* – similar to the first experiment where the reverse of the pointing paradigm was adopted. The TP had to look at the screen of the computer, where the target object was indicated (surrounded by a red rectangle), and then find the annotation on the screen of the smartphone which is associated with this target object.

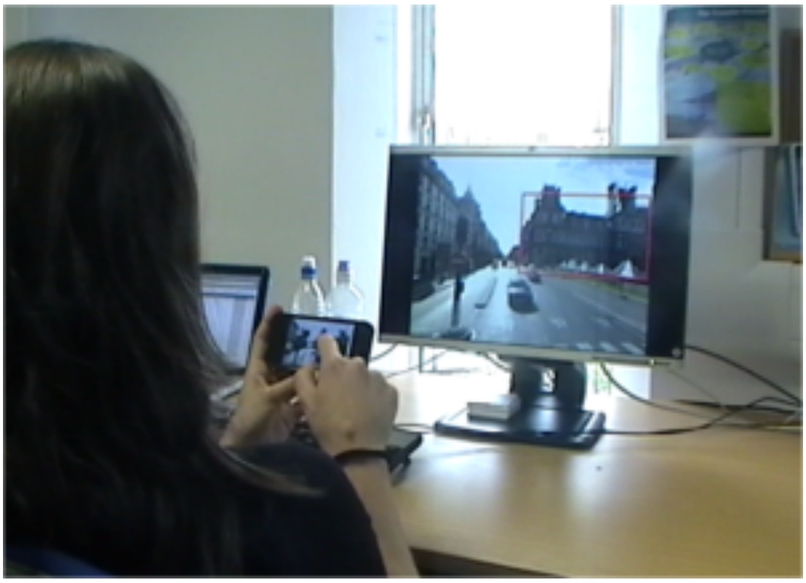
*(R) Reverse task* – similar to the pointing paradigm, the TP had to find a specific bubble on the screen of the smartphone (its number was indicated in the upper left corner of the screen), and then point to the screen of the computer and indicate which target object the annotation refers to.

### 7.2.6. Presentation of stimuli materials

Each AR mock-up was viewed on an iOS iPhone 4 smartphone. The device has a 3.5" multi-touch display with 640x960 pixels resolution, 5MP camera, 30fps video, 1GHz

Cortex-A8 CPU, HSDPA 3G network. The original picture, representing the target objects and the environment, was presented on a stationary computer (Figure 7.10).

**Figure 7.10. Laboratory set up**



During the experiment, care was taken that the researcher stays slightly behind the participant, so that they are not within their immediate field of view. In this way the participant would not be disturbed and would focus on the tasks at hand when the experiment starts. The author of the thesis operated the computer through a remote control.

### 7.2.7. Procedure

The length of each experiment varied depending on the user, but took no less than 20 and no longer than 35 minutes. Upon arrival, each participant was given a short introduction and asked to sign a consent form (Appendix 6). The introduction explained how AR interfaces work and the purpose of the experiment. Participants were asked to imagine that they are a tourist in a new city and that their goal would be to find out whether there are AR annotations (referred to as bubbles) for specific targets. Additional explanations were provided about the AR annotations: they have a name and a number; that they can refer to a whole building, or parts of a building. However, the participants were not given additional explanations about specific designs (e.g. keyword, colour-coding) and the differences among the designs. During the introduction each participant was instructed that:

- There are two types of tasks and that they will practice each task first.
- Their participation is voluntary, anonymous and they could quit at any time.
- They will be timed so they need to carry out the tasks as quickly and as accurately as they can.
- When the test starts, the role of the researcher is only to observe and record the data. They should imagine that the researcher is not present in the room.

After the introduction, the participant was provided enough time to practice the tasks. Each participant carried out 6 practice trials. The test started only after the researcher made sure that there are no additional questions and that each participant understood correctly the tasks, how to work with the smartphone device and how to rate the experienced certainty and difficulty after each task. The tasks for the test were randmomized to isolate carry over and learning effects.

### 7.2.8. Participants

Ninety test subjects (students and lecturers at Bournemouth University) were recruited and randomly assigned to one of the three groups. The average age was 31.14 (range 18-66). From those, 49 were female. Most of the participants used smartphones every day and have never used AR prior to the test. Table 7.3 describes further the characteristics of the test subjects.

**Table 7.3. Characteristics of the participants (n=90) in the laboratory experiment**

| Characteristic | Level | Frequency (%) |
|---|---|---|
| **Gender** | Male | 46 |
| | Female | 54 |
| **Age** | 18-25 | 25 |
| | 26-30 | 34 |
| | 31-40 | 30 |
| | >41 | 11 |
| **Use of Smartphones** | Every day | 65 |
| | Every week | 20 |
| | Less than once a week | 10 |
| | Never | 5 |
| **Use of AR** | Every day | 0 |
| | Every week | 0 |
| | Less than once a week | 4 |
| | Never | 96 |

### 7.2.9. Data capture and recording

Due to problems with video and on-screen capture, the time for each participant was measured through a stop-watch. A special form was developed (Appendix 7) in order to make sure that all of the data is collected during the experiment. The researcher recorded all of the data while the participant was carrying out their tasks.

### 7.2.10. Pilot study

A pilot was carried out with 5 participants. The trial sessions were recorded through two video cameras, capturing how each participant carried out the tasks from two different angles. Several different alternatives for capturing the screen of the smartphone were considered. Unfortunately, the device does not allow this unless the iPhone is tampered with and special software is installed. An alternative was sought by capturing the screen through the mobile application UX Recorder. The application was developed for user testing and acts similarly to a web browser. The pilot study resulted with insights regarding the timing, wording of instructions and the feasibility of the experiment:

- Recording with UX Recorder is processor-intensive and significantly slows down the experience, as the user had to wait for individuals pages of the mock-ups to load;
- The screens for certainty and difficulty had to be adjusted, as users tried tapping on the numbers;

The results from the trial informed the final set up for the larger study. It was decided that the experiment will not be recorded through video/audio capture. This was also not necessary given the quantitative nature of the experiment and the additional time that video/audio recording analysis would take. Instead, additional measures were taken so that all of the time data is recorded during the procedure.

## 7.3. Findings

The collected raw data were prepared and analyzed using IBM's SPSS analysis package. In total, 7200 measures (30 TPs x 20 tasks x 3 groups) were collected, 1800 for each of the dependent variables. The following sections describe the results with respect to errors, time on task, certainty and difficulty.

### 7.3.1. Task accuracy and errors

An error was recorded when the participant could not match the right physical target for the indicated virtual annotation (matching task) or could not relate the right virtual annotation with its physical target (reverse matching task). The total number of errors for all 1800 observations was 367 (20.4%). Table 7.4 shows the observed errors in each of the conditions.

**Table 7.4. Total number of errors in the three experimental conditions**

| Position of annotations | Design 1 Pictogram | Design 2 Keyword | Design 3 Colour-coded |
|---|---|---|---|
| **Precise** | 75 | 36 | 20 |
| **Imprecise** | 65 | 132 | 39 |
| **All tasks** | 140 | 168 | 59 |

As expected, in terms of absolute numbers, Design 2 performed the worst, with a total of 168 errors for all conditions. The total number of errors with Design 1 was smaller (140). The group using Design 3 made the least number of errors (59). A one-way between-subjects ANOVA was conducted to examine the effect of design on errors. The results show that there was a significant effect of design on task accuracy ($F_{2,87}$ = 24.924, p=0.000).

Post-hoc comparisons for all conditions with precise placement were carried out using the Tukey HSD test (Table 7.5). The results indicated that the difference in number of errors between Design 3 (20 errors) and Design 1 (75 errors) is significant (p=0.00). No significant difference was found between task accuracy with Design 3 (20 errors) and Design 2 (36). Finally, Design 2 (20 errors) outperformed Design 1 (75 errors) and significantly improved task accuracy for all tasks with precise placement (p=0.00).

**Table 7.5. Post-hoc comparison results**

| Hypotheses | D3 $\neq$ D1 | D3 $\neq$ D2 | D1 $\neq$ D2 |
|---|---|---|---|
| **Precise placement** | 0.00 | 0.2 | 0.00 |
| **Imprecise placement** | 0.06 | 0.00 | 0.00 |

The results from post-hoc comparisons with the Tukey HSD test for all conditions with imprecise placement are also illustrated in Table 7.5. The results show that there was no significant difference between Design 3 (65 errors) and Design 1 (39 errors) in terms of task accuracy (p=0.06). As expected, the results show that Design 3 (65 errors) significantly reduced the number of errors, compared to Design 2 (132 errors). The same was the case when comparing task accuracy performance with Design 1 and Design 2. As expected, Design 1 improved task accuracy performance significantly (p=0.00).

No significant effect was found of placement of annotations on task accuracy for Design 1 (t=1.069, p=0.294) and Design 3 (t=2.102, p=0.154). There was, however, a significant effect of placement on task accuracy for Design 2 (t=-9.401, p=0.000).

### 7.3.2. Response time (time-on-task)

The mean time for task completion for all 1800 tasks was 6.67 seconds. The results from a one-way ANOVA indicated that the time-on-task for the three designs was significantly different ($F_{2,87}=17.443$, $p < 0.0001$).

The mean time for task completion was the lowest for the group using Design 3 (5.25 sec) (Table 7.6). As expected, the two other designs required more time for completing the tasks. However, Design 2 (mean time = 6.5 sec) performed better than Design 1 (mean time = 8.25 sec).

**Table 7.6. Mean time for task completion in the three experimental conditions (in sec.)**

| Position of annotations | Design 1 Pictogram | Design 2 Keyword | Design 3 Colour-coded |
|---|---|---|---|
| **Precise** | 8.35 | 5.64 | 5.22 |
| **Imprecise** | 8.16 | 7.26 | 5.27 |
| **All tasks** | 8.25 | 6.5 | 5.25 |

The results from post-hoc comparisons with Tukey HSD are illustrated in Table 7.7. As expected Design 3 outperformed significantly Design 1 in both the precise (p=0.00) and imprecise (p=0.00) placement conditions. When annotations were precisely placed on top of their reference object, the mean times for Design 3 and Design 2 did not differ significantly. This was not the case when annotations were misplaced and performance with Design 3 (mean time = 5.27 sec) was significantly better than the mean time for Design 2 (mean time = 7.26 sec).

**Table 7.7. Post-hoc comparison results**

| Hypotheses | D3 ≠ D1 | D3 ≠ D2 | D1 ≠ D2 |
|---|---|---|---|
| **Precise** | 0.00 | 0.7 | 0.00 |
| **Imprecise** | 0.00 | 0.00 | 0.2 |

When annotations were precisely placed on top of the target object, there was a significant difference in mean task times between Design 1 and Design 2. The expectation here was that Design 1 will perform better than Design 2, which was not the case. The mean time for Design 2 (5.64 sec) was significantly lower than Design 1 (8.35 sec), with $p=0.00$. When annotations were displaced, however, no significant difference was found among the two designs.

Comparing the effect of placement, Table 7.7 indicates that the mean times for Design 1 and Design 3 were similar, irrespective of the position of the annotations. The results from a paired samples t-test indicate that placement had a significant effect on task completion time only for Design 2 ($t=-6,881$, $p=0.000$).

### 7.3.3. Certainty

The average certainty for all tasks was 4.00 (1-5), which indicates that most of the time users felt certain that they provided the right answer. As expected, Design 3 was associated with the highest certainty (4.4) (Table 7.8). Reported certainty was lower for Design 1 (3.86) and the lowest for Design 2 (3.75). The results from the Kruskal-Wallis test indicated that the means for all tasks were significantly different, $H(2)=19.885$, $p<0.001$.

**Table 7.8. Average certainty (1-5) in the three experimental conditions**

| Position of annotations | Design 1 Pictogram | Design 2 Keyword | Design 3 Colour-coded |
|---|---|---|---|
| **Precise** | 3.77 | 4.13 | 4.4 |
| **Imprecise** | 3.97 | 3.38 | 4.4 |
| **All tasks** | 3.87 | 3.75 | 4.4 |

Post-hoc comparisons were carried out with the Man-Whitney test with Bonferroni correction ($p=0.0167$) (Table 7.9). The results for precisely placed annotations show that Design 3 outperformed Design 1 ($U=182.5$, $Z=-3.96$, $p<0.001$) and Design 2 ($U=284.0$, $Z=-2.471$, $p=0.013$). No significant difference was found, however, in reported certainty between Design 1 and Design 2 ($U=311$, $Z=-2.05$, $p=0.17$).

**Table 7.9. Post-hoc comparison results**

| Hypotheses | D3 ≠ D1 | D3 ≠ D2 | D1 ≠ D2 |
|---|---|---|---|
| **Precise** | 0.000 | 0.013 | 0.17 |
| **Imprecise** | 0.000 | 0.000 | 0.008 |

The results were similar for the condition with imprecisely placed annotations, where significant differences were found between the conditions. As expected, the reported certainty for Design 3 were significantly higher when compared to Design 1 (U=227, Z=-3.29, p=0.001) and Design 2 (U=139, Z=-4.606, p<0.001). The difference in reported certainty means between Design 1 and Design 2 was also significant (U=270, Z=-2.66, p=0.008).

The results from Wilcoxon Signed-Ranked test indicated that there is significant effect of placement of annotations on certainty for both Design 1 (Z=-2.267, p=0.23) and Design 2 (Z=-4.402, p=0.000). When annotations were precisely placed, users felt more certain using Design 2 than Design 1. The reverse situation was observed when the annotations were imprecisely placed on top of objects, as users felt more certain in their answers with Design 1, as opposed to Design 2. No effect was found of placement on reported certainty for Design 3.

### 7.3.4. Difficulty

The mean reported difficulty was 4.03, which indicates that users found the tasks rather easy than difficult. Table 7.10 shows the mean difficulty scores for all designs. Users found Design 3 the easiest to work with (4.28). Unexpectedly, Design 2 (3.91) was found easier to work with when compared with Design 1 (3.89). The mean differences among the designs for all tasks were significant (H2=7.282, p=0.026).

**Table 7.10. Average difficulty**

| Position of annotations | Design 1 Pictogram | Design 2 Keyword | Design 3 Colour-coded |
|---|---|---|---|
| **Precise** | 3.8 | 4.17 | 4.24 |
| **Imprecise** | 3.97 | 3.65 | 4.31 |
| **All tasks** | 3.89 | 3.91 | 4.28 |

Table 7.11 shows the significance levels when comparing the three designs and the reported mean difficulty ranks. As expected, when annotations were precisely placed on top of their reference target Design 3 (4.24) outperformed Design 1 (3.8) significantly

(U=277, Z=-2.55, p=0.011). The situation was different, however, when it comes to comparisons among the other designs. No difference was found between the mean difficulty scores for Design 3 and Design 2 (U=419.0, Z=-0.46, p=0.64), or Design 1 and Design 2 (U=311, Z=-2.05, p=0.04).

**Table 7.11. Post-hoc comparison results**

| Hypotheses | D3 ≠ D1 | D3 ≠ D2 | D1 ≠ D2 |
|---|---|---|---|
| **Precise** | 0.011 | 0.64 | 0.04 |
| **Imprecise** | 0.011 | 0.007 | 0.321 |

When annotations were not precisely placed on top of their reference target, Design 3 outperformed both Design 1 (U=278, Z=-2.54, p=0.011) and Design 2 (U=269, Z=-2.68, p=0.007). No significant difference was found between Design 1 and Design 2 (U=383, Z=-0.993, p=0.321).

The results from Wilcoxon Signed-Ranked test indicated that there is significant effect of placement of annotations on certainty for both Design 1 (Z=-2.128, p=0.033) and Design 2 (Z=-3.955, p=0.000). No effect was found of placement on reported certainty for Design 3.

### 7.3.5. Preference and satisfaction

The mean reported value for visual design was 3.73. This means that users did like the designs in general, but had additional comments and remarks how they could be improved. The means for the three designs were similar. The results from Kruskal-Wallis test indicate that there was no statistically significant difference among the three designs in terms of visual design ratings (X2=3.168, p=0.205). Overall, users found AR apps useful, with a mean of 4.46 for usefulness. The results from Kruskal-Wallis test indicated that the three means for the different groups were very similar and not significantly different (X2=0.081, p=0.96).

Similar results were obtained when it comes to the mean satisfaction level (4.22), suggesting that the users were satisfied with the experience of browsing information in this way on the smartphone. Results from Kurskal-Wallis test showed that the three means among the groups were very similar and there were no significant differences among them (X2=0.471, p=0.79).

## 7.4. Discussion: Visual Coupling and Association

The results from the experiment confirm the observations and findings regarding association of virtual AR annotations and physical targets, expressed visually through the consolidated task model (Figure 7.1). There are significant differences in task performance between direct visual coupling and indirect visual coupling of physical and virtual objects. These observations emphasize the importance of the visual layout of AR annotations, which is essential in supporting users to associate physical targets and virtual AR annotations effectively and efficiently. The results from the experiment confirm that task accuracy, time completion, certainty and ease of use improve significantly when association relies on a direct visual match between the graphical variables of the AR annotation and the perceived visual characteristics of the reference target object. In such cases, performance is significantly better than in cases where users have to rely on indirect match, or associate abstract words with the inferred non-visual characteristics of the target object.

Overall, the low number of errors, time, high certainty and low difficulty indicates that all three designs support well the user in associating physical targets and virtual content. The overall improvement in task performance can be explained with the controlled nature of the laboratory environment. Lack of hand tremors, annotation movement or environmental factors (bright sunlight) (Herbst et al., 2008; Livingston, 2013), could have affected the data. Considering the average reported values for time, errors, certainty and difficulty, it is important to keep in mind that the experiment was carried out in "ideal" settings, where the effect of jitter, lightning conditions or other external variables, such as movement of the annotations, was purposefully excluded. This was necessary in order to investigate the effect of different types of visual and spatial coupling on association. Therefore, the average values for all measures could be used as a benchmark in the future when setting up quantitative usability goals and testing the performance of future designs for AR browsers. The average reported values could also be used to test the effect of various external confounding factors in actual context of use with non-ideal settings.

The results coordinate well with previous research within the field of Information Rich Virtual Environments (Bowman et al., 2003; Polys, 2006). When working with a large number of annotations for both visible and non-visible targets, users tend to adopt strategies in order to (physically and mentally) reduce the visual clutter in virtual space (e.g. Polys et al., 2006). Limiting search between physical and virtual spaces to only visible annotations and physical targets is extremely beneficial for AR browsers.

The results from the experiment confirm most of the predictions with respect to the influence of perceived visual characteristics and legibility and their influence on the association process. To expand on this, we will first revisit the hypotheses, set out in the beginning of the chapter (Table 7.12).

**Table 7.12. Summary of the results from the experiment**

| N | Hypothesis | Graphical notation | Status |
|---|---|---|---|
| **H1** | Task performance will improve when there is a direct visual match between the graphical variables of the annotation and the representation of the target object, compared to the use of pictograms where users have to rely on mental rotation. | D3 (CC) > D1 (P) | Accepted |
| **H2** | When placement is imprecise, task performance will improve when there is a direct visual match between the graphical variables of the annotation and the representation of the target object, compared to the use of pictograms where users have to rely on mental rotation. | D3 (CC) > D1 (P) | Accepted |
| **H3** | Task performance will improve when there is a direct visual match between the graphical variables of the annotation and the representation of the target object, compared to the use of keywords, where users have to rely on the inferred non-visual attributes of the target object. | D3 (CC) > D2 (K) | Rejected |
| **H4** | When placement is imprecise, task performance will improve when there is a direct visual match between the graphical variables of the annotation and the representation of the target object, compared to the use of keywords, where users have to rely on the inferred non-visual attributes of the target object. | D3 (CC) > D2 (K) | Accepted |
| **H5** | Task performance will improve when there is at least one visual match between the graphical variables of the annotation and the representation of the target object, compared to the use of keywords. | D1 (P) > D2 (K) | Rejected |
| **H6** | When placement is imprecise, task performance will improve when there is at least one visual match between the graphical variables of the annotation and the representation of the target object, compared to the use of keywords. | D1 (P) > D2 (K) | Accepted |

***Hypothesis 1 (D3 > D1, precise placement)***: The results for all measures confirm the hypothesis that task performance will improve significantly when colour-coding is used compared to pictograms. When annotations are precisely placed on top of their target object, colour-coding helps users to carry out tasks more effectively (higher success

rate) and efficiently (lower time on task). The participants reported higher certainty and lower difficulty for all tasks. Such findings can be related to the processes associated with stimulus-driven processes of visual attention (Matlin, 2013). The process is similar to working with photorealistic 3D environments, where users extract visual cues from the physical environment and match them to their virtual counterparts (Partala et al., 2010).

***Hypothesis 2 (D3 > D1, imprecise placement)***: The main assumption was that colour-coded design will support association better than complex pictograms irrespective of inaccuracies in placement of annotations. While no significant effect on task accuracy was found, Design 3 (CC) improved overall performance and allowed faster task completion time, higher certainty and lower difficulty when compared to Design 1 (P).

***Hypothesis 3 (D3 > D2, precise placement):*** It was hypothesized that a colour-coded design (D3) will improve task performance significantly compared to the use of keywords and leader lines (D2). The data from the experiment confirmed that subjective ranking of certainty was higher with D3 than with D2. This shows that users felt more certain that they have provided the right answer when using Design 3.

When annotations were precisely placed on top of their target objects no significant differences were found in terms of errors, time or reported difficulty between Design 3 (CC) and Design 2 (K). Despite expectations, both designs supported users well with associating targets and reference objects and led to low number of errors, time and experienced difficulty. One possible explanation is that the effect of the directional pointer on the association process was bigger than expected. Leader lines are used extensively in textbooks and digital graphics to support users with establishing a referential relationship between visual and textual elements (Gotzelmann et al., 2006). Leader lines are especially effective when they connect text with abstract and easily delineable forms, such as simple squares or, for instance, the various body parts in the human atlas. Since objects are abstract and easily delineable, users are typically able to effectively association each textual label with only one visual element (Hartmann et al., 2005). Because usually it might be difficult to delineate and distinguish among objects in urban environments, it was expected that the use of leader lines can result in ambiguities. However, the findings show that users are able to complete tasks equally well (low amount of time, low number of errors) with both the colour-coded and the directional pointer designs.

***Hypothesis 4 (D3 > D2, imprecise placement)***: As expected, Design 3 (CC) improved task accuracy, task time completion, certainty and difficulty significantly, compared to Design 2 (K) when annotations were imprecisely placed on top of physical targets. These results confirm the assumption that task performance improves when users have to rely on a direct visual match, rather than subjective interpretation and non-visual attributes of the target object.

***Hypothesis 5 (D1 > D2, precise placement)***: The main assumption was that task performance will improve when users have to rely on a pictogram (D1), rather than a keyword only (D2), even if they have to mentally rotate or interpret the pictogram. This assumption was based on findings from previous studies which indicate that photorealism and the use of 3D landmarks improves visual recognition in the real world as they offer more visual cues compared to 2D representations or text (e.g. Daft and Lengel, 1986; Elias and Paelke, 2008; Partala et al., 2010; Partala and Salminen, 2012). The results suggest that when annotations were precisely placed, Design 2 (K) outperformed Design 1 (P) and significantly reduced the number of errors that participants made. Design 2 (K) outperformed Design 1 (P) also with respect to task time completion, as it took participants a shorter amount of time to associate targets and virtual annotations with the directional pointer. No significant difference was found in the reported certainty and difficulty for both designs. One possible explanation for the observed data is that a directional pointer might be more suitable than using complex pictogram symbols when annotations are precisely placed on top of their target object. Direction pointers (leader lines) have been found extremely useful within Information Rich Virtual Environments (Bowman et al., 2003; Maass and Döllner, 2006; Polys, 2006). Complex symbols take more time to interpret and match, rather than a simple leader line.

***Hypothesis 6 (D1 > D2, imprecise placement)***: As expected, Design 1 (P) improved task accuracy performance and reported certainty significantly, compared to Design 2 (K). The findings relate well to observations by Bessa et al. (2006) who found out that geometry and contours were the most often used visual cues that facilitate people in relating pictures and physical targets. No significant difference was observed, however, in terms of mean task completion time between the two designs. While building details and parts attract most attention in urban scenes (Partala et al., 2010), one possible explanation is that, when symbols are very complex and include a lot of detail, it takes as much time to interpret them as it would if users had to rely simply on indirect matching mechanism. This is confirmed by the similarity in reported means for difficulty. No significant difference was found between the two designs.

Overall, the key implication from the results is that if positioning data acquisition for AR browsers improves in the future, both leader lines and colour-coding could be used to support effective and efficient association of visible target objects and AR virtual annotations. Both of these designs allow presenting additional information to users within the frame of the annotation. When annotations are precisely placed on top of their target object, the use of categorical symbols or more complex pictograms (landmarks) deteriorates performance, as they require mental rotation or more extensive visual search.

The results also suggest, however, that when annotations are not precisely placed on top of their physical counterpart, the use of pictograms or colour-coding would be more suitable in complex and unfamiliar urban environments. While colour-coding requires precise positioning data, the use of landmark pictograms could be implemented in order to improve performance with current AR browsers that cannot obtain immaculate positioning data.

## 7.5. Chapter Summary

This chapter described a laboratory-based experiment where 90 participants were tested with three alternative designs of AR browser annotations. The six hypotheses that were tested during the experiment (Section 7.2) were designed to investigate and understand better whether there is difference in performance when users have to rely on direct versus indirect visual coupling when associating virtual annotations and their reference target object. The results (Section 7.3) show that there is a significant difference in task performance. Time on task, accuracy, certainty and ease improve when users have to rely on a direct visual match between the target object and the virtual annotation. Such observations confirm empirically the essential role of visual salience and legibility within the association process.

# CHAPTER 8

# QUALITATIVE EVALUATION OF AR ANNOTATIONS

**8.1. Study Introduction and Goals**

Introduces the specific research questions that guided the field study

**8.2. Laboratory Pluralistic Walkthrough Evaluation**

Describes the evaluation of alternative AR annotations with domain-expert users, carried out through a pluralistic walkthrough

**8.3. Field-based Pluralistic Walkthrough Evaluation**

Describes in detail the design, analysis and findings from a field-based evaluation of alternative AR annotations

**8.4. Discussion**

Revisits the findings in relation to theory

**8.5. Chapter Summary**

Synthesis of the key points in the chapter

## 8.1. Study Introduction and Goals

To a large extent, the design of any information system is concerned with the properties of the user interface and how functionality and information is presented graphically. In addition, the presentation and selection of (type of) content is also a critical issue that designers have to address. Providing the right content to tourists in complex tourism environments is not trivial and the empirical evaluation described earlier indicated that currently AR browsers do not deliver relevant information (Chapter 6) to tourists. Addressing the dissatisfaction of users with provided content within AR annotations, it was considered critical to obtain further feedback related to improving the utility of such applications (Stage D, Table 4.2). To this end, several design mock-ups of AR annotations were developed and used as artefacts in two qualitative evaluations with a total of 19 participants. Both studies were implemented as collaborative pluralistic walkthrough sessions. The first was carried out with domain experts (in eTourism and Marketing) in a controlled environment. Considering the huge role of physical context on the usability and utility of AR browsers, as well as user familiarity, the second evaluation was conducted in the field in an unfamiliar urban tourism context with Human-Computer Interaction and Geo-Information Science domain experts. The method, procedures, materials, analysis and findings from the laboratory evaluation (Section 8.2) and field-based study (Section 8.3) are further described below.

## 8.2. Laboratory Pluralistic Walkthrough Evaluation

Given that the main purpose of this study is to contribute to Information Systems design theory, it was considered critical to obtain further feedback from domain-expert users. In line with recommendations for applying UCD to design of AR (Gabbard and Swan II, 2008) and in order to balance resources and richness of obtained data, it was decided that feedback will be obtained through a pluralistic walkthrough (Bias, 1994). Similar to a focus group, this method allows for interactive discussions among experts in a group who go collaboratively through a user interface (Stage D.2, Table 4.2). One of the key advantages was that the approach provided a platform where participants could validate and discuss each other's perspectives and opinions.

### 8.2.1. Materials

A pluralistic cognitive walkthrough and evaluation of design alternatives has to be carried out in the context of specific users and specific tasks (Bias, 1994). In preparation for the evaluation, several set of materials were prepared to be presented to evaluators in the group:

- Definition of the users of the system – the first step in conducting a walkthrough is a systematic identification of the user population of a product. While end users for this particular research project could vary widely in demographics and background, a special characteristics of users is that they are first-time visitors to a specific location.
- Definition of the tasks for the walkthrough – this involves identification of the tasks around which the walkthrough will be conducted.

In order to provide both descriptions, a presentation was prepared which comprised of an explanation of AR interfaces, the aim of the study, ground rules for the evaluation, and its purpose. An archetypical user and a simple scenario were used to capture user characteristics, goals and tasks and encourage the participants to evaluate the mock-ups in a specific context of use (Figure 8.1).

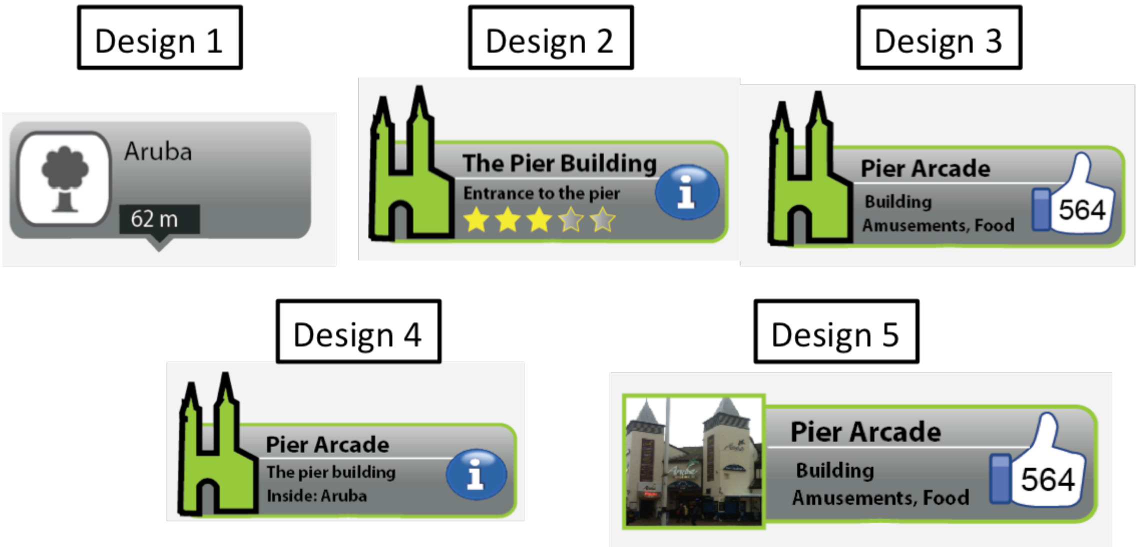
**Figure 8.1. The scenario used during the pluralistic walkthrough**

**Basic Scenario**

*"Jane is a tourist who just arrived in Bournemouth. It is her first time in the city and while she does not know anything about her surroundings, she has one day to explore it. She goes to the tourist information desk and they advise on a route. While following the route, Jane will use a smartphone AR application to obtain more information about her surroundings."*

During the introduction, two videos, capturing work with Junaio and LocalScope were also shown to the participants. Each video lasted for approx. 30-90 seconds. The main goal was to present a working AR browser, so that domain experts obtain a better idea of how such applications work.

**Figure 8.2. Annotation design alternatives for the Amusement Arcade, presented to domain experts for qualitative evaluation**




After the exploration of a number of alternative ideas, five annotation mock-ups were selected for presentation to the group (Figure 8.2). The digital mock-ups were created using Adobe Illustrator and PowerPoint. While it could also be useful to evaluate stand-alone AR annotations, important details, problems and concerns could be missed out when the representations are taken out of their context. This is why it was decided that the AR annotations would be presented as part of an overall interface (Figure 8.3), in the context for which they were designed.

**Figure 8.3. Annotation design alternatives for the Amusement Arcade, represented in context on the screen of the smartphone**

The mock-ups were available in printed format to each participant. They were also incorporated into a PowerPoint presentation and shown through a media projector on a wide interactive screen, positioned in front of the participants. Each participant had an individual document with the map of the selected location. Photos of the location, the selected POIs and additional description of the surroundings and other POIs were also available.

Following procedures for walkthrough evaluations (Bias, 1994; Nielsen and Mack, 1994), each evaluator went through the tasks and interfaces individually prior to starting the discussion. To this end, individual printed documents with the mock-ups and tasks were provided (Appendix 8). In essence, each participant was asked to answer questions in several categories: Association (e.g. *Can you determine what is the name of the building in front of you?*), Content (e.g. *What is your opinion about the symbols within the AR bubbles?),* Relevance (e.g. *Is the information provided within the AR bubbles relevant to the current situation?*), Preference (e.g. *Which interface / AR bubble do you like the most? Why?)*. A sample of the questions can be found in Appendix 8.

### 8.2.2. Participants

In order to obtain more focused feedback several key experts with background and current work in eTourism were invited. An invitation email was sent to key experts who were planning to attend the ENTER2013 eTourism conference. Nine experts attended the meeting. The profile for each participant can be found in Table 8.1.

**Table 8.1. Profile information of the expert participants in the evaluation**

| Test person | Gender | Country | Academic title | Background / Expertise |
|---|---|---|---|---|
| **P1** | M | Switzerland | Professor | eTourism |
| **P2** | M | Austria | Professor | eTourism, Sociology |
| **P3** | F | Switzerland | Post-doc | eTourism, eCommerce, eWord-of-Mouth |
| **P4** | F | Switzerland | Post-doc | eLearning, eTourism |
| **P5** | F | Spain | Professor | Augmented Reality, eTourism |
| **P6** | F | USA | Associate professor | Design and Communication |
| **P7** | F | China | Assistant professor | Tourism and Hospitality Management |
| **P8** | F | Ireland | PhD Researcher | eTourism, Marketing |
| **P9** | F | Finland | IT Project Leader | eTourism, eCommerce |

The table illustrates the broad background of the experts who had expert knowledge and practical experience with subjects such as Sociology, Augmented Reality, and eWord-of-Mouth.

### 8.2.3. Procedure

The evaluation took place in a specially equipped room at the Congress Innsbruck convention center (Figure 8.4) in Innsbruck, Austria. The meeting was moderated by the researcher and was recorded with a video camera by a second researcher. The meeting started with an introduction to the research and the aim of the discussion. The researcher also presented the scenario and the fictional location where the evaluation would take place. After making sure that the ground rules are clear, the researcher encouraged all participants to first evaluate each of the mock-ups individually.

**Figure 8.4. Screenshot from the focus group evaluation video recording**

After the individual evaluation, the discussion focused on the advantages and drawbacks for each of the AR interface mock-ups, where special focus was placed on potential usability problems with graphical design and content. The moderator made sure that all of the participants could comment and express their opinion about the individual AR mock-ups, and encouraged additional comments and feedback regarding issues that had not been discussed but seemed important for any of the domain experts. During the discussion the moderator made sure that each participant commented on the questions or remarks that were brought up by the other participants, by asking open-ended (neutral) questions, such as: "*What do you think about…*", "*How do you feel about…*", "*Could you comment on…*".

### 8.2.4. Findings

Three main themes emerged from the analysis: (1) content for visible points of interest, (2) supporting situational awareness through content for non-visible points of interest, (3) issues connected with the overall user experience with AR browsers.

#### *8.2.4.1. Content and visible points of interests*

A considerably large part of the discussion focused on association of annotations and target objects. The arguments and suggestions that experts provided emphasized the use of content, rather than placement of AR annotations, in order to facilitate association. In particular, discussion focused on the suitability of different elements of the provided content to achieve association between virtual and physical worlds. Categorical symbols, general keywords (e.g. *building*) and distance were heavily criticized for being redundant. All three were considered information assets that do not add value to the overall communication process, as they capture information which is already "*present in the world*" (P7). When the target object is within sight, such information can readily be extracted from the physical environment. Categorical symbols were also criticised for being ambiguous and all participants agreed that they might be misinterpreted, especially if AR browsers are used by international tourists. Pictograms and 3D models were preferred for supporting effective association.

The participants pointed out that the address for a POI instead of keywords or description is less useful, especially when the user is in an unfamiliar environment. Annotations containing postcodes or street names were deemed unhelpful, even in navigation scenarios, when tourists are trying to reach to a final destination. All experts agreed that each annotation should contain sufficient information to support decision-making. When tourists are trying to learn more about their surroundings, or make a decision whether to visit a destination, there was a common agreement that the name of the POI alone is insufficient to give a clear indication of the type of attraction/object. In such context, a short description or keywords were the preferred information assets that should be included in each annotation. A common suggestion was that the content of AR annotations should communicate how POIs are special and/or unique from a tourist point of view. Pictures, symbols and keywords should be used to indicate "*what do you get when you get there*" (P2).

#### *8.2.4.2. Awareness and non-visible points of interest*

All participants expressed the need for delivery of relevant information regarding both visible and non-visible targets that supports the current task/goal or decision-making process of the tourist. In particular, the participants agreed that it should be possible for users to distinguish quickly between annotations that relate to visible targets and annotations that relate to non-visible targets. Pictures and images of the POI were considered most suitable when it came to delivering content about non-visible targets. Participants agreed that an image could communicate information effectively about the type of attraction or target.

Navigation scenarios were also discussed. In particular, the participants focused on situations in unfamiliar environments where tourists might have difficulties locating where non-visible objects are. Experts discussed the suitability of directional arrows for wayfinding. There were, however, concerns that a simple arrow for non-visible objects does not represent well the direction in which the participant has to turn and might result in confusion. Eventually, experts agreed that arrows would be redundant, especially in view of the fact that the position of the AR annotation communicates well the overall direction in which the POI is in space.

While distance was considered irrelevant for visible POIs, experts considered this information asset useful for non-visible POIs. There was a general agreement that distance to POIs is an important factor that tourists take into account during decision-making and when optimising their route on a micro scale. In such situations, distance might be used as a proxy for walking time. There were, however, concerns that straight-line distance is not an accurate proxy for walking time (or time to reach a destination). In cases where the environment is more complex (e.g. a lot of turns), straight-line distance could lead to miscalculating the time it takes to reach a destination.

#### *8.2.4.3. Influence of content on tourists experiences*

Part of the discussion was dedicated to the influence of content on the perceived qualities and characteristics of POIs, and consequently on tourists' decision-making. Experts pointed out that care should be taken when selecting the picture of a POI, as the quality of the picture might influence the perception for that POI. A low quality picture might ultimately influence the tourist in thinking that the POI is less interesting and not worth attending.

Influencing the perception of the attractiveness of POIs was also brought up when the discussion switched to the social elements in the annotations. Marketing experts were concerned with the influence of including social media information and the effect it would have on the perception of users for specific attractions. This was an especially big concern when it came to comparing different types of attractions and when there is information about different attractions in one interface. As one expert pointed out:

> *"You should be careful with this. The Aruba building might have less likes than the beach, simply because one is free and the other one is not. This will ultimately influence how tourists see the attraction and their decision during the trip when they are standing there and browsing through the annotations" (P4)*

The experts pointed out that this design could influence the perception of a place, since it suggests that "there is nothing interesting" around. Inclusion of pictures and other elements would influence the feelings of tourists as they go about an unfamiliar place.

## 8.3. Field-based Pluralistic Walkthrough Evaluation

The overarching goal for any mobile information system is to deliver useful and usable information. In the context of tourism, a useful AR browser would deliver relevant information to tourists in (unfamiliar) urban environments (Chapter 2). Empirical field-based evaluation (Chapter 6) indicated that current AR browsers do not provide useful information to tourists. Additional expert evaluation (Section 8.2) suggested that information assets should differ for visible and non-visible object.

In order to investigate further what type of content AR browsers should deliver to tourists in unfamiliar urban environments a second evaluation of different design alternatives was carried out (Stage D.3, Table 4.2). The method followed the principles of pluralistic walkthrough (Bias, 1994), where domain experts evaluated several design alternatives and discussed the potential strengths and usability problems of each. In this sense, the procedure was similar to the pluralistic walkthrough described in the previous section (Section 8.2). However, apart from potential drawbacks, one of the goals was to obtain more focused feedback on the relevance of provided information within AR annotations. Evaluation of relevance and utility of information is best carried out when and where information needs arise (Wilson, 1992; Wilson, 2006). This is why the second pluralistic walkthrough was carried out on the field in actual context of use. In essence, 10 domain experts in Human-Computer Interaction and Geo-Information Science were asked to walk a pre-defined route and evaluate several design alternatives for AR browser annotations. The evaluation took place in Paris and was selected as a special field activity by the committee of the first workshop on Geo Human Computer Interaction (GeoHCI), organised as part of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI 2013). The procedure, route and locations, participants and results are further described.

### 8.3.1. Route and stops

The evaluation took place around the conference venue Université Paris-Dauphine in Paris, France. While a number of routes around the venue were considered, the final selection was based on the following criteria: (1) has to take up no more than 1 hour to walk, and (2) has to include different in nature urban points of interest (historical buildings, streets/avenues/boulevards, tourist attractions). Figure 8.5 shows the final route.

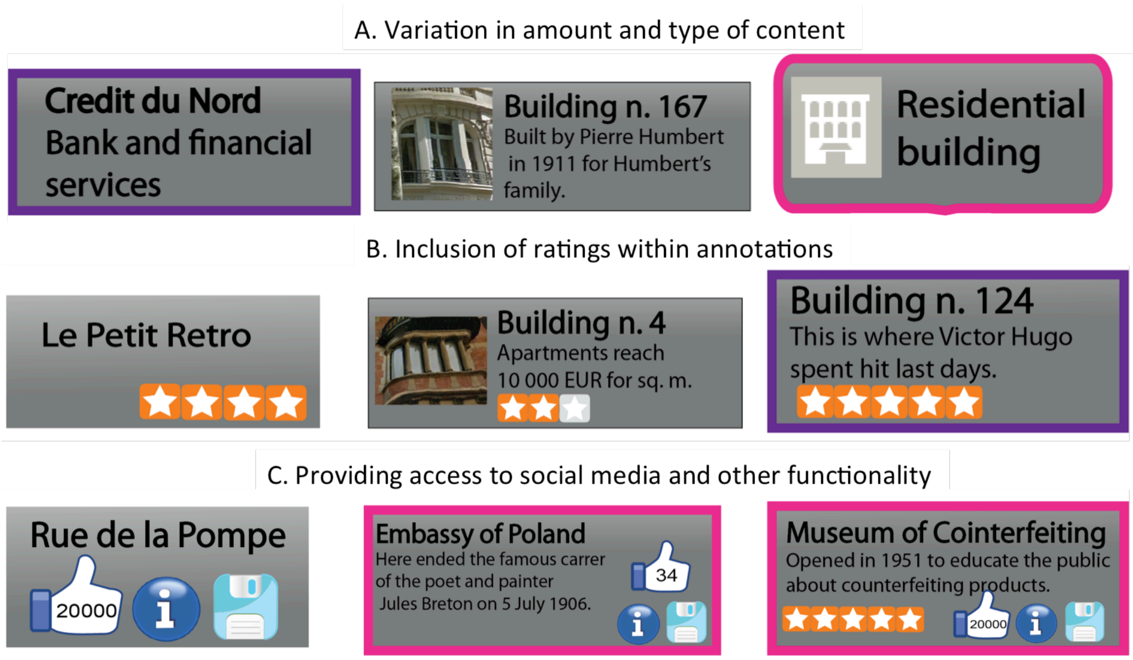
**Figure 8.5. The selected route for the field activity, including four main streets around the venue**


The route was selected so that it represents a typical urban environment, but at the same time provides visibility to interesting and important urban objects that could be augmented with information. More than 60 points of interest were considered for augmentation. The final route included 7 stops. Along the route, the participants were exposed to 55 different annotations.

### 8.3.2. Materials

To create the material for the field expert evaluation, several different design alternatives were sketched on paper. Different alternatives for the delivered type of information were considered. It was decided that the final annotation designs would vary in terms of (Figure 8.6): (1) type of delivered information assets and combinations thereof, (2) amount of delivered information, (3) visual design. The following types of information assets were included: name of POI, keywords, short description (1 sentence), long description (2-3 sentences), and social media assets (recommendations, reviews, ratings). The following types of visual designs were included: picture/landmark, leader line, directional pointer, colour-coded.

**Figure 8.6. A set of designs with different elements, type, and amount of content**




The design alternatives were implemented as mock-ups in the form of augmented photos. For the purpose, preliminary set of stops and targets were selected using Google Maps, Google Street View, and Google Earth. Additional photos were taken for each target two days prior to the field evaluation. These photos were then prepared (with Adobe Photoshop) and augmented with content through the use of Adobe Illustrator. Hence, the field evaluation did not require full-time connectivity to Internet.

### 8.3.3. Participants

It was expected that in each group there were a minimum of 2 and maximum of 5 experts. However, the event triggered more interest than expected. While more than 20 participants showed up, only 10 (5 in each group) participated actively in the evaluation. Table 8.2 lists the participants that took active part in the field evaluation.

**Table 8.2. Profile information of the expert participants in the evaluation**

| Test person | Gender | Academic title | Background / Expertise |
|---|---|---|---|
| **P1** | M | Associate professor | Human-Computer Interaction |
| **P2** | M | Post-doc | Geo-Information Science |
| **P3** | F | Post-doc | Geo-Information Science |
| **P4** | F | Post-doc | Geo-Information Science |
| **P5** | F | Post-doc | Human-Computer Interaction |
| **P6** | M | Professor | Human-Computer Interaction |
| **P7** | M | Associate professor | Human-Computer Interaction |
| **P8** | M | Post-doc | Geo-Information Science |
| **P9** | M | PhD Researcher | Human-Computer Interaction |
| **P10** | F | Post-doc | Geo-Information Science |

### 8.3.4. Procedure

After meeting at the first stop, the participants were given a short introduction and allowed to practice with two smartphones. Afterwards, half of the participants were assigned to one of two groups (Group 1 and Group 2). There was one moderator for each group. Group 1 was accompanied by the author of the thesis, while Group 2 was led by another researcher. Group 1 then moved to Location 2, while Group 1 carried out the first task for Location 1. Two mobile camcorders were used to obtain audio/video recordings from both groups.

At each location, the participants were asked to identify the points of interest they would like to learn more about prior to using the smartphone and formulate questions regarding their surroundings. Then, they were asked to look at the screen of the smartphone individually and carry out one matching task (Figure 8.7). Afterwards, they were encouraged to make comments about the provided information and additional questions were asked in order to stimulate feedback and discussion regarding the content and provided information.

**Figure 8.7. Participants interacting with the smartphone during the field evaluation**

When carrying out the matching tasks, the users were encouraged to reason about the interface and provide their feedback on the delivered information. The protocol included general and open-ended questions, aiming to stimulate feedback and the collaborative discussion among the experts: *"How do you feel about browsing information in this way", "How do you think that the content / interface could be improved"?*

After the final location, all participants returned to the University for a post-evaluation debriefing session. The main objective was for both groups to share their experiences and observations with respect to information delivery through AR annotations.

### 8.3.5. Findings

The overall attitude towards AR was positive; the participants enjoyed discussing the interfaces and took longer time to provide feedback. This resulted in more rich and detailed information. Overall, experts agreed that this type of application and visualization paradigm would be very useful in a new and unfamiliar environment, because it does not require the use of a guidebook. This was considered especially relevant for on-site visits, rather than for trip planning:

> *"I don't see this as a tool that you use to plan, it's something that you use while you are walking around" (P3).*

This section describes the key findings from the evaluation.

#### *8.3.5.1. Influence of the physical environment on information needs*

In order to assess the impact of perceived visual characteristics on information needs, the participants were asked to identify points of interest and formulate questions prior to using the smartphone. Many of the stops did not trigger specific information needs and users could not verbalise questions regarding their surroundings. In both groups, the most common comment was that "*nothing looks interesting"*. This was mainly because of the similarities of the physical entities along the route in terms of overall architecture, textures, contours and shapes.

After the first few stops and working with the augmented photos several times, the participants started noticing fine details within their surroundings. Along the route, it became evident that information retrieval is not only directed towards individual targets, but specific *elements* of the surroundings. For instance, questions referred to the flags, memorial displays on buildings, and individual architectural elements, such as ornaments on doors and windows.

Similar to observed behaviour during the field-based evaluation of existing AR browsers (Chapter 6), the participants used different visual cues to select targets in their environment that they would like to know more about. Colours (e.g. "*the white colour*") or shapes (e.g "*the strange shape*") of ornaments on windows, doors and walls, made things stand out, attracted attention and triggered information needs. Such visual cues were used to infer the non-visible attributes of targets. For instance, TP4 indicated that: "*…this one [building], funnily enough doesn't have these gate windows that they have on the windows…so it stands off…I suppose it is an embassy…because of the flag"*.

The lack of visually salient targets influenced the perception of users with respect to the cultural and touristic significance of the environment. In turn, this influenced their expectations with respect to availability of AR content. Participants in both groups shared that it is highly unlikely that they would point to the targets selected for the field evaluation. This was mainly because the overall context is uniform and targets do not stand out:

> *"In some cases you were pointing us to see things which I would have never spotted, and I never expected that I will find information about them" (P8).*
>
> *"How likely is it that I would point to that building? I wouldn't have picked it up!" (P3).*

The need to bring attention to specific targets through push-based information was suggested.

##### *8.3.5.2. Influence of perceived visible characteristics on spatial permanence of annotations*

The observed embodied interaction during the field activity mimicked the one observed during the mobile field study. When asked to comment about AR annotations, each participant first raised the device towards the visual centre of the target object. An interesting observation, however, was made by several participants with respect to annotations for streets:

> *"One thing was interesting …the name of the street, but I don't understand why show it from here and not earlier on…so what characterizes from this section onwards" (P6).*

Due to their continuous nature, participants expected that annotations for streets will appear along the path and not only on isolated locations.

##### *8.3.5.3. Preferred content in AR annotations*

The different types of content provided within the AR annotations triggered long and interesting discussions regarding the suitability of various information assets in the current context of use. There was no uniform preference for type of content and it became evident that users preferred access to more and different types of information assets.

Overall, users were dissatisfied with the amount and level of detail of the annotations that provided only symbols, names and keywords about the target object. The participants used such annotations to confirm their assumptions about the function of target objects but agreed that the provided information is not enough to satisfy contextual information needs. The common agreement was that a consequential and hierarchical provision of information could support better the decision-making process of tourists. Providing access to more detailed information (once annotations are tapped) was considered critical, especially in view of the small smartphone screen, coupled with the information needs of tourists in unfamiliar locations.

Specific keywords, such as "interesting" and "unique" triggered information needs and encouraged additional questions for the physical targets they annotated. For instance, during stop 3, participants were initially not interested to find out more about the building in front of them until they saw an annotation that contained the phrase "*interesting architecture"*. Apart from attracting attention to that physical target, the phrase triggered additional questions:

> *Cose I know it has an interesting architecture, I can see it, but when was it made? Why was it made this way? It doesn't tell me anything useful (P1).*

A similar situation occurred at stop 7. The participants in both groups expressed desire to learn more about a shop that did not seem interesting prior to them interacting with the AR annotations. The annotation contained information that the target is a "*popular*" dessert and candy store in Paris. Additional questions, regarding the history of the shop, the products and services that they offer and whether locals like to visit, arose.

Users were also attracted to superlatives, as they naturally searched for information that would make them understand why specific physical targets are important and interesting from a tourist point of view. For instance, the reference to the "*longest*" street in Paris attracted a lot of attention:

> *"I was wondering which one is the longest…and now I get the answer, I like that" (P5).*

This annotation came up also during the debriefing sessions and participant noted that they have remembered the information it provided. They were also positive that an AR interface should emphasize delivery of information that helps tourists understand which physical targets are unique in their surroundings and why.

Names of architects and other famous people spiked interest only when they were familiar, or when they were delivered together with additional information. Unfamiliar names of famous architects, such as "Pierre Humbert" and " Edouard Georg", did not seem to attract attention and were not considered relevant or interesting. In contrast, the annotations referring to familiar names, such as Victor Hugo or Mitt Romney, attracted attention and triggered additional questions. At stop 4, the participants in both groups spent considerable amount of time to discuss the annotation referring to FBI, but did not seem to remember or know the name of the actress the annotation referred to. However, participants agreed that this is enough information to spike interest and to trigger a desire to access the content of the annotation, and read more about the target object.

At each individual stop, users were asked to carry out one matching task, relating to a specific target object. The presence of additional AR annotations on the smartphone screen, however, attracted attention towards other physical objects in the surroundings and triggered additional questions. This effect was most obvious when users were exposed to the colour-coded design, as it highlighted visually different parts of the surroundings.

#### *8.3.5.4. Influence of legibility on perceived relevance of content*

Context-awareness and adaptation were considered fundamental for AR browsers used in urban tourism context, especially in view of the fact that tourists would require different types of information depending on the stage of their trip and the already visited locations. During the evaluation, it became evident that the relevance of content depends on the perceived non-visible attributes of targets. Opening hours, type of food and prices were considered relevant for food venues, while Wikipedia-type information was suggested for other types of points of interest. Reviews and recommendations were considered relevant only when they referred to services and food venues. In all other cases, the star icons included in the annotations were confusing. This was the case, especially when the annotations referred to different in function objects, such as restaurants, streets, and historical buildings.

The relevance of short (keywords) and longer (description) textual content also depended on the type of physical target. For instance, the fact that the patisserie was "preferred by the locals" was considered very relevant, important and interesting by the participants. This type of information also brought up many additional questions, such as "Why do they prefer it?", "What do they order there?".

The visual characteristics of the target object also influenced how participants interpreted content. For instance, at stop 4 the participants discussed an annotation which contained only an unfamiliar name. Using the architecture of the building as a visual cue, the participants concluded correctly that the name has to refer to the architect who designed the building.

It was considered critical that the user is provided with an option to change the delivered type of information or the type of augmented objects depending on the situation. Experts outlined several different scenarios, but focused on two main use cases. The first was driven by cognitive needs (learning), while the second was driven by physiological needs (e.g. hunger). Experts indicated that these situations would be supported by different types of information and the necessity to annotate and augment different elements and POIs in the environment.

#### *8.3.5.5. Preference for information about non-visible points of interest*

Participants in the field activity had a clear preference for accessing information about non-visible targets. All of the participants were interested to find out more about the surrounding area and the affordances it provides from a tourist point of view. The participants wanted to learn where nearby points of interest are located and how to "*get somewhere interesting*":

> *"So maybe there is a building and behind the building there is something very important …so you can't see this…maybe an alert that says..look, if you go around you can find something interesting…" (P1).*

They also wanted to use the AR browser to confirm the locations of non-visible targets that they were already familiar with. For instance, several participants expressed the need to locate famous attractions, such as the Eifel tower, or Champs Elise and learn how much time it takes to get there from their current location.

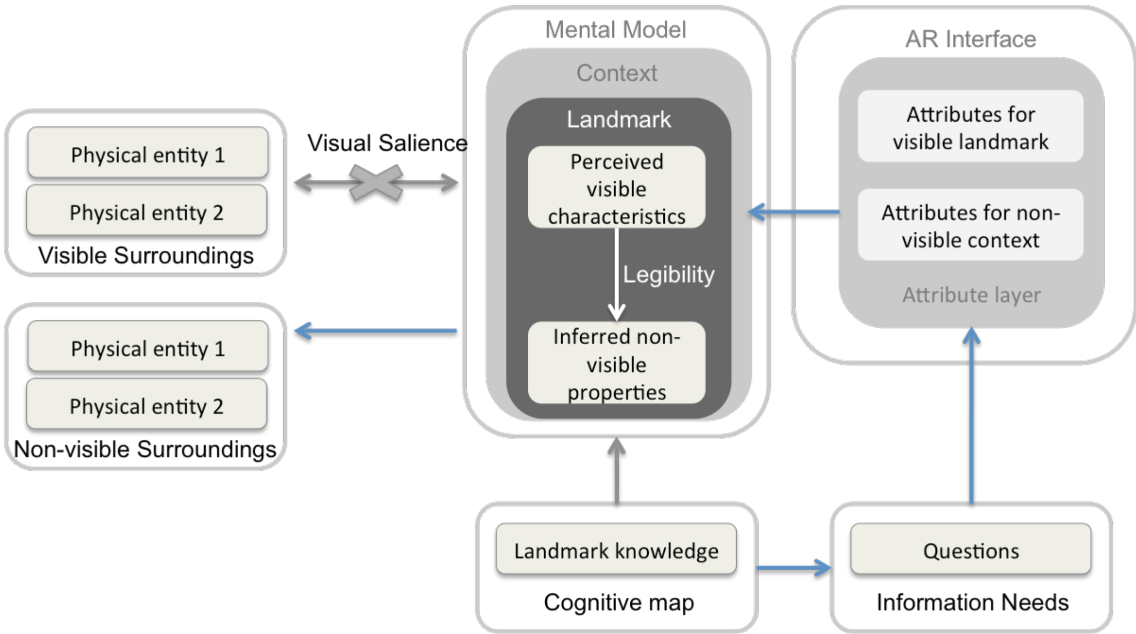
This observation was confirmed during the debriefing session, where it became evident that experts consider the provision of information about non-visible POIs critical in urban tourism context. Access to additional types of mLBSIs, such as maps, was also brought up.

## 8.4. Discussion

Both evaluations described in this chapter were used to obtain feedback and sample the opinions and attitudes of domain experts with respect to knowledge acquisition through AR browsers in urban tourism context. The rich findings served to identify key aspects, apart from association, that would influence the user experience and utility of AR browsers in unfamiliar urban environments. During both evaluations, it was observed that there is no ultimate preference for specific content type or way of presentation of information. Opinions and attitudes towards the different design alternatives and combinations of content varied. Two key themes emerge from the evaluation: (1) delivering content for visible POIs, and (2) supporting situation awareness through content for non-visible POIs.

A central requirement in both evaluations was to provide content for both visible and non-visible POIs. Domain experts pointed out that information for non-visible targets could facilitate and alleviate the decision-making process for tourists. The notion of providing such relevant information that is critical to the task has been referred to as supporting *situation awareness* (Endsley, 2000). Participants in the field evaluation expressed the need to be able to access information about non-visible targets that they are already familiar with. This suggests the influence of familiarity and accumulated knowledge on requirements of tourists with respect to delivery of information in AR browsers (Figure 8.8). Evaluators also identified that users should be able to distinguish quickly among annotations, depending on their visibility status. Visibility awareness has been discussed in literature for location-based mobile maps (Gardiner et al., 2009; Yin and Carswell, 2013). Empirical research has confirmed that enabling users to distinguish between visible and non-visible POIs on mobile maps enhances orientation and navigation (Fröhlich et al., 2006). The findings in this study confirm the need for visibility awareness in AR browsers.

**Figure 8.8. Influence of context annotations on attention and information needs towards non-visible physical objects**




Both evaluations suggest that information assets within annotations should be considered carefully when it comes to delivering useful content to tourists in unfamiliar urban context. This concurs with findings from Speiginer and MacIntyre (2014) who pose that the *level of detail* of delivered content should be considered carefully. Type of content and the level of details with which it is represented have the potential to influence tourists' perception towards a destination, hinder or enhance the formulation of information needs and interest in unfamiliar urban surroundings. In particular, the suitability and relevance of type of content (e.g. symbol, reviews, recommendations) depends on the inferred function and importance of the target object. Provided level of detail (and wording), on the other hand, influences interest and could hinder or enhance interest. Manipulating such parameters for AR browsers has not been discussed so far in literature and should be studied in more detail in the future.

Meaningful experience with space requires integration of both physical interaction (e.g. touching, pointing) with social and cultural contexts: "What am I touching?", "Why am I standing here"? (Barba, 2014, p. 44). This notion is confirmed and expanded in this study. Users of mobile AR need to acquire information that explains and imbues physical space with meaning. Additionally, in heavily built-up environments this process has to be selective in order to prevent information overload and lessen demand on attention. The study findings indicate that urban tourists require information that makes physical targets stand out from their (visually complex) context.

At each individual stop, users were asked to carry out one matching task, relating to a specific target object. Throughout the field evaluation, attention was often directed at other annotations that were not originally the focus of discussion. This effect was most obvious when users were exposed to the colour-coded design, as it highlighted visually different parts of the surroundings. This observation emphasises the role of *context annotations* for maximising information acquisition (Figure 8.8).

Delivering target and context annotations that spike interest is especially important in non-salient urban environments. During the field evaluation, the participants were exposed to an unfamiliar environment with low visual salience. The lack of visually salient landmarks and the inference about the cultural significance of the environment hindered the formulation of specific information needs and questions (Figure 8.8). Such findings could be examined through the lens of the Information Foraging theory. Within Information Science, the Information Foraging theory (Spink and Cole, 2006) has been used to examine human interaction with information retrieval (Kukka et al., 2011). A key concept in information foraging is that of *information scent*, which reflects the profitability of an information source in relation to other potential sources (Pirolli, 1999). Given a strong scent, the information forager can quickly reach their information goal. In the absence of one, the forager will search for new direction by sniffing for scent activities (Spink and Cole, 2006). The findings in this study confirm the key role of physical context as a background against which "information foraging" is carried out. The lack of visual cues in the environment prevented optimal selection of ''prey information'' or potential alternatives of important and interesting points of interest. In this context, lack of information assets (e.g. specific keywords, symbols) that trigger interest within the AR annotations hindered information needs formation.

## 8.5. Chapter Summary

This chapter described two qualitative empirical evaluations with domain expert users that aimed at obtaining further feedback regarding delivered content through AR annotations. The observations show that tourists form specific expectations towards delivered content based on inferred non-visual attributes (mainly function) of target objects. While AR browsers have a huge potential in augmenting the immediate visible surroundings of tourists, the obtained feedback shows clearly that there is a need to consider delivering information for non-visible POIs. The visual layout and annotation elements for both visible and non-visible target objects need to be considered carefully as they can influence the overall tourist experience and perception towards a destination. The results suggest that designers should consider not only the graphical presentation of information, but several other factors that influence perceived utility. These include level of detail, wording, movement and spatial permanence of AR annotations.

# CHAPTER 9

# USER-CENTERED DESIGN FRAMEWORK FOR SMARTPHONE AUGMENTED REALITY

**9.1. Introduction**

Introduces the specific research questions that guided the field study

↓

**9.2. The Revised Design Framework**

Describes the final user-centred design framework

↓

**9.3. User Requirements for AR Browsers**

Describes the key user requirements that have to be met when it comes to design for AR browsers in urban tourism context

↓

**9.4. Design Parameters and Taxonomy for AR Browsers**

Identifies the key design parameters that need to be considered when it comes to AR browsers

↓

**9.5. Design Guidelines for Smartphone AR Browsers**

Provides the guidelines extracted from the framework

↓

**9.6. Chapter Summary**

Synthesis of the key points in the chapter

## 9.1. Introduction

The main motivation behind this research project was the lack of design theory or frameworks and guidelines that prescribe how to design more usable and useful smartphone AR browsers for tourism. Placing the user in the centre of design, this study further adopted a User-Centred Design approach to investigate empirically how tourists use current (Chapter 6) and future (Chapter 7 and Chapter 8) AR browser annotations. Two critical components of any design theory are the specification of 1) scope and goals and 2) constructs (Gregor and Jones,2007). In line with the standard design theory generation process, this study started with identifying existing design knowledge (Chapter 5) relevant to AR browsers and knowledge acquisition by tourists. Treating AR browsers as tools that facilitate location-based knowledge acquisition, existing knowledge was captured in a theoretical framework, which identified important constructs and the relationships among them.

Evaluation of current AR browsers revealed that in complex unfamiliar urban environments tourists use visual cues and legibility of urban objects in order to associate AR annotations with their targets (Chapter 6). Further empirical testing confirmed that task performance improves when users rely on annotation designs that support direct, rather than indirect visual matching of annotations and physical targets (Chapter 7). Finally, empirical evaluation by domain experts suggested that the utility of AR browsers depends on supporting knowledge acquisition not only about the immediate visible surroundings, but also enhancing situation awareness by delivering content for non-visible POIs. Important aspects regarding expectations towards augmented objects and delivered content were further identified.

This chapter proposes a user-centred design framework (Stage E, Table 4.2) for analysis, design and evaluation of smartphone AR browsers by extending the initially developed theoretical framework. The framework is then used to derive the two critical components of design theories (Walls et al., 1992): 1) meta-requirements and 2) meta-designs (design principles).

It should be noted that the user requirements identified within this chapter are aspirational in nature. This is a direct result of the adopted pragmatist interpretivist approach in this study. The main benefit of the selected approach is that it allows capturing the elements and potential relationships among them for a very large and complex phenomenon. Considering the nature of the study and the selected target user group, the approach allowed gauging insights for a relatively short amount of time. The framework is aspirational in nature because it captures a multitude of elements and relationships which have to be further validated. Because of its extensive nature, it is recommended that requirements are considered for specific contexts and selected based on the goal for any particular information system.

The main drawback of the aspirational framework is connected with the selected qualitative approach. In order to gauge in-depth insights for specific contexts and participants, the data collected in all studies described throughout the thesis emphasise depth, versus breadth. Because of this, it is questionable to what extent the identified relationships generalise to other contexts and types of tourists.

Quantitative data and findings, gathered during the mobile field study and the laboratory experiment, were important in order to direct and focus research. The quantitative data collected during the mobile field study were used to identify behavioural patterns and usability problems. Those focused and directed the use of qualitative data, which were used to seek for clarification and explanation. The laboratory experiments were then organised in order to collect quantitative data and validate a small part of the conceptual framework. Further quantitative research is needed in order to validate the various parts of the framework. In this sense, the various elements are flexible and the relationships between them can be changed. The framework can be used to derive hypotheses and this was considered more valuable, rather than a closed and quantitative model which can be used in very narrow set of contexts and which will not be possible to use outside of the parameters of this study.

Insights based on qualitative data are also not suitable for generating quantitative predictions. Further quantitative research is needed in order to confirm and quantify the identified relationships and determine their strength. Quantitative data will allow transforming the framework from an aspirational tool to more rigorous model that can be used for prediction. Considering the nature of the studied phenomena, and the significant variance associated with different contexts and situations, it is recommended that future studies are focused in nature and look at each or several elements in isolation first.

First, a description of the general process of acquiring information through AR browser annotations is outlined. Then, the basic components of the framework are presented, followed by their interaction and the processes that unfold when users make use of AR annotations in unfamiliar urban environments. The developed framework has both practical and theoretical implications for the design and evaluation of AR interfaces. From a theoretical point of view, it raises a number of research questions with respect to design and interaction issues, user modelling and evaluation methods and techniques not only for AR browsers, but also for mLBSIs in general. From a practical point of view, it provides a way to extract the most important user requirements, as well as design guidelines for more useful and usable AR browsers. This chapter aims to explore in detail the implications relating to the optimization of design of AR browsers, based on the developed framework.

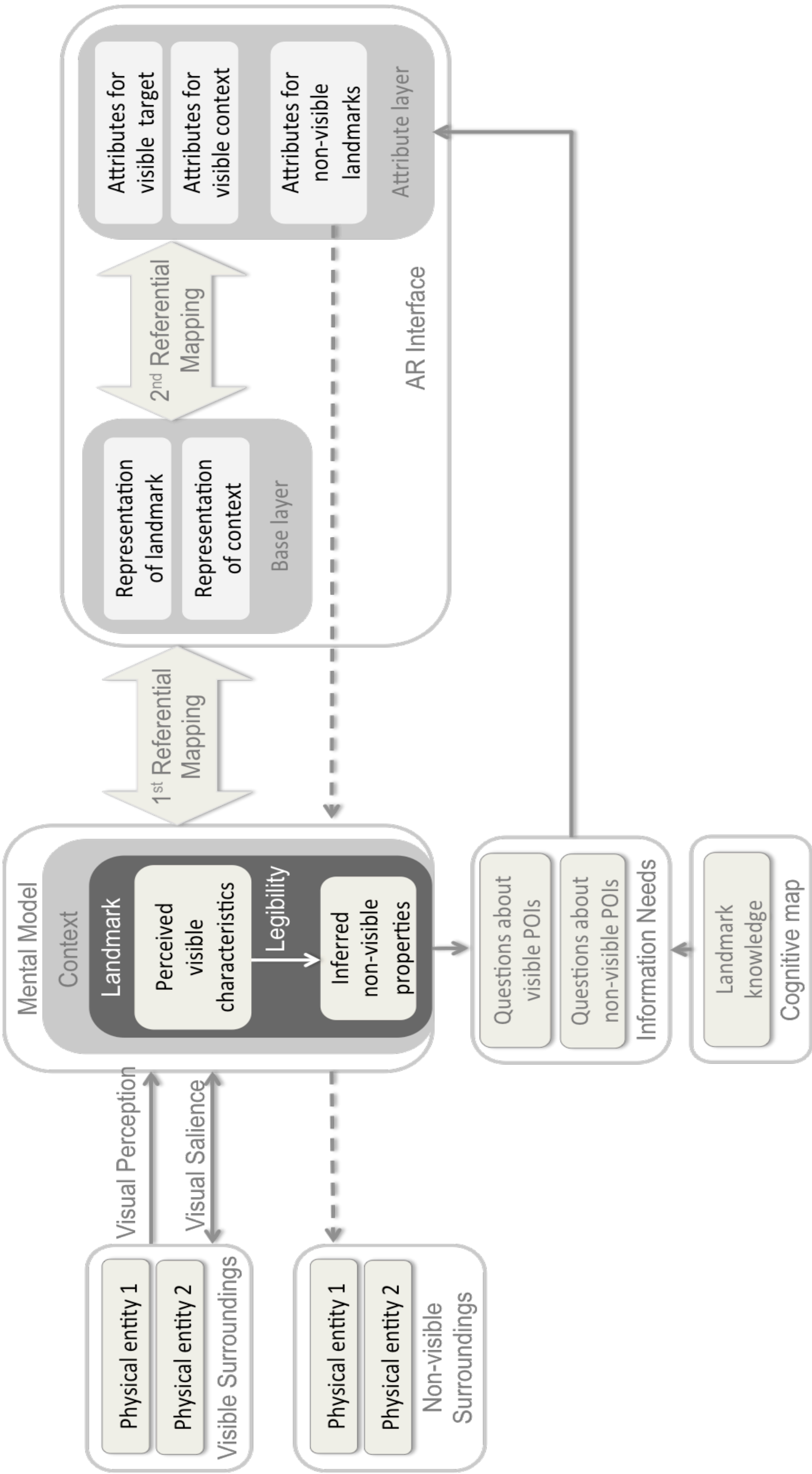
## 9.2. The Revised Design Framework

Chapter 3 reviewed the various frameworks that have been developed for design and analysis of mobile AR (Hansen, 2006; Alzahrani et al., 2011; Vincent et al., 2012). None of the frameworks, however, considers the user, their characteristics, knowledge and abilities. Likewise, there has been only limited discussion of the role of physical context and its influence on the design process of AR browsers. The findings in the empirical studies described in this thesis emphasize the importance of placing the user in the centre of design in order to ensure that smartphone AR browser annotations are usable and useful. This is also the primary aim of the developed user-centred design framework, illustrated in Figure 9.1.

The framework captures the process of information acquisition by tourists in unfamiliar urban environments. It proposes two new factors that have to be considered when designing AR browsers: visual perception (salience) and legibility of physical objects in urban environments. In addition, by incorporating empirical knowledge about work with AR browsers, the framework deconstructs the AR interface to the most important elements that designers and researchers need to consider. In particular, the base layer is further deconstructed to (1) representation of physical target, and (2) representation of context. Likewise, the framework proposes that there are three types of AR annotations, depending on the task of the user. At any given moment, designers need to consider how to communicate the attributes of: (1) the visible target object that users want to obtain information about, (2) the visible context, and (3) non-visible landmarks.

.

**Figure 9.1. The revised and expanded user-centred design framework for design of AR browsers in urban tourism context.**



Apart from previously discussed legibility requirements (Jankowski et al., 2010), the framework facilitates designers to focus on the two key requirements for AR annotations: (1) match the perceived physical characteristics of target objects to ensure usable annotations, and (2) predict the information needs of users to enhance utility of delivered content. In essence, it emphasizes the important role and influence of physical context on the usability and utility of AR browsers. While physical context has been discussed in literature, its role has been constricted to a simple background against which augmented content is overlaid.

The results from this study indicate that physical context plays an active role in determining the usability and utility of AR browsers. Several important parameters, including visibility, visual salience, and legibility of urban environments have been identified as important aspects of context of use that have to be considered when it comes to information acquisition in unfamiliar settings. These factors influence: (1) when and whether users will interact with the AR browser, (2) their expectations with regard to available content, (3) the association process of AR annotations and physical targets. All of these are further discussed.

### 9.2.1. Interaction triggers with the AR browser and expectations for content

The way urban environments are (visually) perceived, and the inferences users make about physical targets influence interaction with AR browsers in several important ways.  Visually salient physical targets attract the attention of the user. Once focused on a specific physical target, users will use different visual cues to determine the non-visual attributes of that target, such as its function, cultural significance and importance, a process referred to as legibility. This process, in turn, influences when and whether users will interact with the AR browser to search for content. Without any other stimuli (e.g. social context, push-based information) users will not interact with the AR browser if the legibility of the physical target implies that the object is not culturally important, or significant from a tourist point of view.

The visual characteristics and legibility of physical targets also influences the expectations of tourists with respect to available AR content. When the visual characteristics of the target signify a culturally important and/or interesting from a tourist point of view target, users expect that they will be able to find content about this physical target. When tourists are located in a visually non-salient environment, they require guidance and content for culturally important non-visible physical targets. The lack of content for visible and non-visible targets that are perceived important from a tourist point of view will ultimately lead to lower perceived utility of the AR browser.

Familiarity and already acquired landmark knowledge is one parameter that influences this process. If users are able to recognise physical targets that they have already identified as important, then they will expect to find content about them, irrespective of their visual characteristics or legibility. In addition, landmark knowledge will influence questions and needs for information about non-visible targets. Users that have knowledge about culturally important landmarks will expect to find this content in the AR browser, even if such targets are not visible from their current location.

### 9.2.2. Association of virtual annotations and physical targets

The empirical studies indicate that there has to be at least one (direct or indirect) visual match between the perceived characteristics of the physical target and the AR annotation in order for users to associate them effectively (Chapter 6 and Chapter 7). This is why the perceived visual characteristics of the target object, as well as its legibility have a direct influence on the association process. When the smartphone is raised towards a physical target, tourists will first try to match specific visual cues, observed in physical space, with visual cues in virtual space (direct visual match). If this process fails, tourists will try to match the content of the AR annotation with inferred non-visible properties of the target object (indirect visual match).

From a design point of view, there are several key elements that will determine the effectiveness and efficiency of the association process. If the base layer (representation of the physical world) is unaltered, the association process will be mainly influenced by the characteristics of the attribute layer (AR annotations). In particular, users will search the virtual space until there is a positive match between the target and the attributes of at least one AR annotation. Here, the attributes of other AR annotations (attributes for visible context) play a critical role. When all annotations have the same visual attributes (layout, the same symbols, same names), users will have to make a cognitive effort in order to eliminate and select only one annotation that refers to the target object.

It is also important to consider the characteristics of the base layer and how it influences the association process. With unaltered video feed, users will be able to make the first referential mapping successfully. The success of the second referential mapping, then, will depend only on the characteristics of the attribute layer. However, this might not be the case when the base layer has been digitally altered. For instance, if the base layer has been generalised and abstracted (similar to a 3D map), the association process will depend more heavily on an effective first referential mapping. Users will be able to match virtual annotations with their targets only if they can first match the physical target with its virtual representation on the smartphone screen. More abstract representations of the base layer would require more time for the first referential mapping. This would, however, make the second referential mapping easier, especially in more visually complex urban environments.

### 9.2.3. Information needs and queries

The inferred non-visual properties of the target object will influence the information needs of users. For instance, participants that infer two different types of function for the same physical target (church versus disco club) will have different types of questions and will look for different information within AR annotations. Inferences for legibility require a conscious or subconscious mental effort and will depend on the attention of the user. As discussed earlier, attention in large-scale physical environments is limited and directed towards visually salient features of the environment. This is why it is logical to assume that the information needs of the user will depend on the visibility of physical targets. Partial visibility might lead to a different set of assumptions about the target physical object, and therefore impact information needs.

Apart from visibility and visual salience, the information needs of users will also be influenced by already acquired landmark knowledge. On one hand, already acquired knowledge about the environment might trigger information search directed towards non-visible physical targets. On the other hand, knowledge about landmarks will also influence information needs with respect to visible physical targets. For instance, if the tourist has learned about the historical and cultural significance of a landmark beforehand, their questions will be different (e.g. Why is this important?), rather than the questions of tourists who do not have this information (e.g. What is this?).

When an information need is formed, users will try to satisfy it by referring to the virtual information contained in the attribute layer. Lack of specific information that answers the object-based questions of tourists will influence the perceived utility of AR browsers. Legibility also influences the relevance and usefulness of provided content. In particular, the perceived function of target objects (e.g. restaurant, historical building) influences the requirements of tourists and their expectations with respect to different types of content. For instance, reviews and ratings are considered necessary, useful and relevant only for specific types of physical objects (e.g. restaurants, cafes, food venues).

### 9.2.4. Embodied interaction and spatial permanence of annotations

During all empirical studies, it was observed that when tourists start interaction with the AR browser, they will initially point the device towards the visual centre of the physical target. However, it was also observed that the visual characteristics (spatial layout) of the target influenced embodied interaction with the device and expectations for spatial permanence. Users expect that annotations for individual discrete objects are placed over or nearby the target object. Expectations are different when it comes to spatial or linear physical objects, such as streets or squares. In such cases, users expect to find annotations in the visual centre (from their current location) of the physical target. They also expect that annotations appear along (movement) the feature and not only at one specific location.

## 9.3. User Requirements for AR Annotations

The key aim of this study is to provide recommendations with respect to improving the overall usability and utility of AR browsers. Identifying key user requirements for effective delivery of content through AR annotations is crucial towards achieving this aim. The results suggest that, in order for AR browsers to be useful and usable for tourists, a number of specific user requirements have to be met. This section discusses user requirements captured through the revised conceptual framework.

***R.1. The AR interface should prioritise on providing information for visible physical targets.***

Just like more traditional paper-based media, current smartphone AR browsers rely on manual search inquiries from users. In order for users to acquire information, they need to first point the phone towards an object and interact with the application to select the right "channel" or "layer" with content. The empirical studies described in this thesis reveal the large influence of physical context on this process. In particular, tourists would most likely interact with the smartphone screen and raise their hand if there are *visible* physical entities that they would like to learn about. Most frequently this interaction will be triggered by visually salient physical targets that have attracted the attention of the user. This is why an AR browser should prioritise delivery of content about visible targets.

***R.2. The AR interface should provide information for both visible and non-visible physical targets.***

While prioritising on AR annotations for visible targets, AR browsers should deliver information about non-visible physical targets as well. The results from the evaluation of alternative designs of AR annotations with domain experts (Chapter 8) confirm the necessity for enhancing situation awareness through such annotations, especially in non-salient urban environments. Already acquired landmark knowledge about important points of interest could trigger interaction in search for information about non-visible targets. Due to learnt behaviours and habits with using mobile map-based services, it is likely that tourists would prefer to use a map in such situations. However, the delivery of AR annotations for non-visible targets could enhance incidental learning and knowledge acquisition. Lack of content for important POIs that are not visible from the current location of the user could lower the perceived utility of AR browsers.

***R.3. It should be possible for users to distinguish visually between annotations for visible and non-visible targets***

In order to understand the information delivered through an AR browser, tourists have to unambiguously match each physical target with only one physical object. This process becomes very long and difficult if tourists are not able to distinguish between the AR annotations that refer to visible targets and AR annotations that refer to non-visible targets. In such a situation, the tourist will try to match all of the annotations on the screen with their physical targets. This process will take more time and require a large cognitive effort. To alleviate the situation, the user has to be able to distinguish immediately (once the information has loaded on the screen) between the annotations that refer to visible targets and those that communicate the attributes of non-visible objects.

***R.4. The AR interface should support effective and efficient association between AR annotations and visible physical targets.***

Unambiguous association of virtual annotations and physical targets is necessary in order for users to be able to make sense of the provided information on the smartphone screen. If tourists are not able to match virtual AR annotations with their reference target, then the AR browser becomes useless and very difficult to use. Until now, literature has discussed association in general and there is a surprising lack of research that investigates user requirements in detail or debates the role of context and its influence on the process. The developed user-centred design framework highlights the key link between perceived visible and non-visible properties of the environment and the design of AR annotations. A key requirement for designers is to ensure that users are able to match AR annotations with *visible* physical targets. This process can be facilitated if there is at least *one visual (direct or indirect) match* between the AR annotation and the target object. The fact that users will try to match AR annotations with their target objects emphasizes the requirement to ensure that the design of AR annotations for visible and non-visible targets is visually different. The laboratory experiments confirm that users are most successful and required less time and effort when the visual elements of the virtual annotations matched directly at least one of the perceived visual characteristics of the virtual annotation and the representation of the target object on the screen of the smartphone.

***R.5. The AR browser should acquire contextual information about the visibility of physical targets and adapt the representation of AR annotations***

Association between virtual and physical spaces is mainly influenced by the perceived visible characteristics of physical targets. The extent to which a target object is (fully or partially) visible from the current position of the user is one of the key factors that might influence knowledge acquisition through AR browsers. Apart from acquiring and tracking the current position of the user (location-awareness), AR browsers need to acquire contextual information, infer and track the current visibility of physical targets with respect to the user (visibility-awareness). Changes in visibility to target objects should be reflected in the representation of the virtual AR annotation in order to ensure efficient and effective association between virtual and physical spaces.

***R.6. The provided digital content should match the perceived non-visible properties of urban spaces and objects***

Legibility of urban objects (or their inferred non-visible properties) is one property that plays a central role in influencing expectations of tourists with regard to the type of available content. As discussed earlier, users use visual cues to infer non-visible properties of urban spaces and objects, such as their importance and whether they are significant and interesting to learn about from a tourist point of view. The AR browser should match expectations by providing content for physical targets that tourists consider important and interesting. Familiarity, expressed as already acquired landmark knowledge, is one factor that influences this process. It is especially important to provide content for points of interest that tourists might have learned about from other information sources and consider important. The lack of such content and AR annotations for such objects might lead to mistrust and confusion.

***R.7. Placement and spatial permanence of AR annotations should match the spatial layout of physical targets***

Unlike other mobile Location-Based Services, AR browsers have the potential to eliminate the gap between virtual and physical worlds. The use of such interfaces, however, enforces the impression that information should be "tied" to physical objects. This is the reason why users expect that AR annotations will match the spatial layout of the physical target. They also expect that AR annotations will "move" along the target, while its representation is still on the screen of the smartphone. This is why different rules need to be set for discrete (e.g. buildings), continuous linear (e.g. streets, rivers) and spatial (e.g. squares) entities.

***R.8. AR annotations should facilitate decision-making by providing useful information***

While association is critical, the design of AR annotations should maximise knowledge acquisition by providing enough information for tourists to support informed and fast decision-making. The elements contained within AR annotations should be considered carefully to avoid redundancy, or communicating information that can already be visually perceived or extracted from the physical environment (e.g. the name of a coffee shop). If such information is not used (or necessary) for association, then it only takes up valuable screen space. More importantly, it would lead to lower perceived utility and annoyance. The AR interface should provide useful information that can be utilised immediately for decision-making and micro management of a route or the experience in unfamiliar urban environments. The data obtained during the first (Chapter 6) and second field studies (Chapter 8) suggest that tourists considered most useful those AR annotations that communicated and explained how physical targets, locations or objects are special (unique), interesting and important from a tourist point of view.

***R.9. AR annotations should enhance knowledge acquisition by providing relevant information***

Apart from supporting decision-making, AR browsers should deliver relevant information that enhances learning of unfamiliar environments. Within unfamiliar environments, relevance depends on how closely the provided content matches the questions that tourists have formulated, expressed as specific object-based queries. The information within AR annotations should provide quick answers to those contextual questions. When space is limited, users should be able to infer that they will be able to find those answers by interacting with the smartphone display and sequentially accessing further information about the physical target. Already acquired knowledge and inferred non-visible characteristics of physical targets will ultimately influence the questions that tourists will look answers for. This is why familiarity and legibility are two factors that influence perceived relevance of information.

***R.10. Users should be able to transition effectively and efficiently among different types of mobile location-based service interfaces***

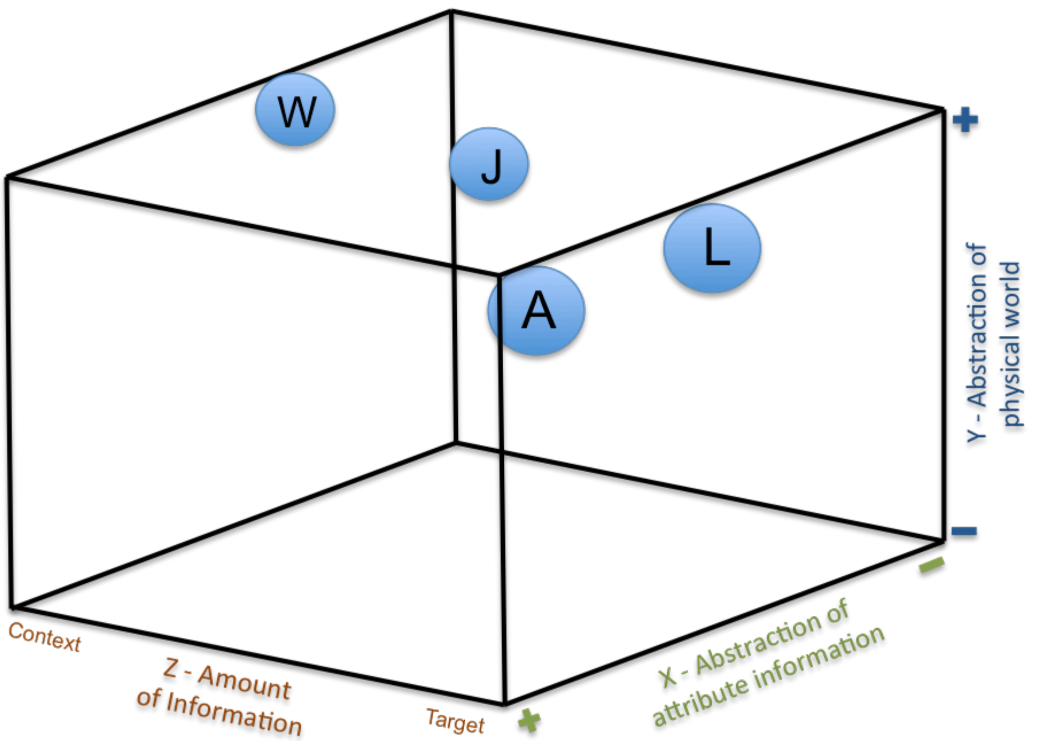
AR browsers can support situation awareness by providing AR annotations for non-visible physical targets. The position of such AR annotations can communicate the direction in which those targets are located. Apart from landmark knowledge, however, tourists will ultimately need to acquire information about paths (route knowledge) and the relation among POIs (survey knowledge). In order to enhance fast knowledge acquisition and help tourists build an accurate mental representation (cognitive map) of a large-scale urban environment, information should be presented through different location-based interfaces, such as 2D and 3D maps, lists or more traditional tour guide interfaces, which provide more information about individual points of interest or a larger territory. For this to happen, users should be able to transition quickly and effectively among different types of interfaces, without losing their sense of "place" in the overall application.

## 9.4. Design Parameters and Taxonomy for AR Browsers

Chapter 5 discussed the vast design space for AR annotations. A big design space becomes problematic due to the lack of empirical knowledge how design elements influence the usability of AR browsers. Designers are often forced to make decisions blindly, without knowing how specific elements will impact the end user experience. The developed user-centred design framework provides a new way to investigate the design space for AR and optimize the design of AR browsers as a medium to deliver geospatial information.

On a more general level, the developed user-centred design framework emphasizes three main high-level design parameters that will ultimately impact the usability and perceived utility of AR browsers: (1) abstraction level of base layer (y), (2) abstraction level of attribute layer (x), and (3) amount of information (z). It is important to note that these three design parameters are inter-connected. Figure 9.2 illustrates this interdependency.

**Figure 9.2. A three-dimensional design space for AR browsers**




Presenting the design space in this form provides a useful way to examine the characteristics of current AR browsers and their position within the cube. The figure illustrates the positions of the four AR browsers used during the first empirical evaluation of existing commercial AR browsers: AcrossAir (A), Junaio (J), Wikitude (W), and LocalScope (L). The position of each AR browser within the cube reflects its characteristics and potential to ensure the 1st referential mapping (y-axis), the 2nd referential mapping (z-axis) or to potentially satisfy the information needs of the user through a balanced amount of information for targets and context entities (x-axis). The three high-level design parameters are connected with the processes that they need to support:

***1) The abstraction level of the representation of physical world / base layer (y-axis) and first referential mapping***

The first referential mapping depends on the visual coupling of the perceived real-world entities and the base layer. Therefore, the abstraction level (or visual characteristics) of the base layer will influence most significantly whether users are able to successfully carry out the first referential mapping. For instance, current AR browsers use an unaltered video feed of the environment. When no alterations are made to this video feed, this base layer is said to be realistic or not abstracted. This is the case for all of the evaluated commercial AR browser solutions and this is why they are situated in the far end of the y-axis.

***2) The abstraction level of attribute information / attribute layer (z-axis) and second referential mapping***

The second referential mapping (matching annotations with the representation of the target object on the base layer) depends on the abstraction level of the attribute layer. In a more abstract attribute layer (high level of abstraction) the AR annotations contain only keywords, or categorical symbols. This is the case with the evaluated existing commercial AR browsers: AcrossAir (A), Wikitude (W), Junaio (J) and LocalScope (L). In comparison, images and pictograms are associated with lower level of abstraction, because they capture the visual appearance of the actual physical target. Lower level of abstraction means that represented attributes match directly the perceived visual characteristics of the target object. Ultimately, then, a lower level of abstraction means faster and more effective second referential mapping (and overall association of AR annotations and physical targets).

***3) The amount of information about individual objects (targets) and context (x-axis) and information needs***

In addition, an AR browser interface can also be characterized by the amount of information that it communicates with respect to an individual target (a landmark the user would like to acquire information about) and its context (all other entities, which fall within the viewport of the device). This parameter could also be manipulated and adjusted.

The presented taxonomy of AR annotations is a useful tool that allows analysis and optimisation of alternative designs. The 3D cube allows us to consider the compromises that need to be made during the design of AR browsers. For instance, providing more information about context objects (that are not the focus of attention) leaves less space on the screen of the smartphone for communicating information about the target object. Movement within the 3D design space illustrates that there is no optimal position and balance have to be found for different types of tasks and situations. One possible solution is to provide options for automatically or manually changing the positions of the AR annotations within the 3D design space (or their overall properties).

## 9.5. Design Guidelines for Smartphone AR Browsers

In order to be usable and useful, designers need to consider each element of AR annotations. This section suggests guidelines for design of AR browser annotations, based on the identified user requirements.

### 9.5.1. Satisfying the information needs of tourists

***DG.1.1. Augment the right entities***

The mobile field study and the field activity resulted in users selecting different in type and nature POIs that they were interested to learn about. Visual salience is one property of urban environments that explains why this is the case. Once attention is focused on a specific target, legibility influences the way users perceive that target and their expectations regarding the availability of content. Lack of content about specific targets that are considered important and/or interesting will lower the perceived usefulness of the interface. Therefore, it is important that the AR browser provides information about the right physical entities that users would expect to find information about.

Apart from whole physical structures, various elements of the environment, such as signs, windows, and different architectural elements could attract the attention of the tourist and trigger information needs. This is why there is a clear need to provide more detailed information about such entities. This is especially important not only for large and famous touristic places, but also when users are roaming around in non-salient environments, where an overall uniform context (buildings with similar shape, contours, architecture) could make small details stand out.

***DG.1.2. Satisfy the contextual information needs of users***

The primary purpose of the attribute layer is to capture information that is not present in the physical environment and could not be obtained without the smartphone device. This is why it is extremely important that AR annotations provide information that answers the context-based questions of the user. The empirical studies suggest that legibility, or the assumptions that users make for visible targets, will influence the perceived utility of delivered content. It is important that the information captured within AR annotations answers such questions.

***DG.1.3. Maximize information flow***

Hand-held mLBSIs are used spontaneously and each use session is very short due to limited attention resources, normally extending for only several seconds (Oulasvirta et al. 2005). This time limit is shorter for AR browsers because information is acquired in awkward positions as tourists need to spend time with an extended arm. Therefore, it is important to maximize the information flow within individual use sessions.

Maximizing information flow does not necessarily mean that the amount of information within individual annotations should be increased. During the qualitative evaluation (Chapter 8), users rarely read longer descriptions for individual annotations when they had to consult the AR display with extended arm. This is why each element within the AR view should be considered carefully, so that only the elements that communicate the maximum information per unit of space should be included. Users should also have access to additional information, if they decide to refer to it.

One possible strategy for increasing the amount of delivered information is to maximize the number of annotations that appear on the screen. Since the primary attention of the user will be directed towards one annotation at a time, context annotations should be self-explanatory and the information contained within them should be easy to understand. At the same time, they should be visually salient, attract the attention of the user and increase the desire to learn about the environment. All content should be balanced and merge well with the physical representation of the surroundings (base layer) and the target annotation.

*DG.1.4. Deliver information for visible and non-visible physical targets*

Visual salience has also implications regarding the type of targets that should be augmented, especially on hand-held devices. Due to the characteristics of hand-held AR, users need to point the device in a specific direction in order to obtain augmented content. Visual salience suggests that this process depends on the visible characteristics of the surroundings. Therefore, the accent in developing smartphone AR browsers should be on communicating information about the visible surroundings of the user. Information about non-visible distant targets could be communicated more effectively through a combination of other mLBSIs. However, the empirical studies also emphasized the need for information acquisition for non-visible targets. In order to maximize information flow, the AR browser should provide information for both visible and non-visible targets.

*DG.1.5. Support visibility-based, rather than distance-based filtering of information*

The empirical findings suggested that distance-based filtering in AR browsers is not only an under-utilized function, but leads to difficulties and confusion when users want to reduce the amount of annotations on display (Chapter 6). Difficulties with estimating distances in less familiar environments have long been documented in literature (e.g. Kirasic et al., 1984). Providing a function that filters out information based on the visbility status of physical entities could prevent such difficulties, save time and be less cognitively demanding for tourists.

### 9.5.2. Ensuring effective association

*DG.2.1. Determine the target for augmentation*

In order to provide useful information for tourists, the AR browser has to deliver annotations that answer their specific questions. As discussed earlier, those information needs can be directed towards visible and non-visible points of interest. The AR browser must somehow detect what is the target that should be augmented with information. In the case of visible physical objects, users will point the device towards the visual centre of the target. In this case, all objects that fall outside of the centre of the smartphone screen should be considered as context. In certain situations, however, tourists might be more interested to find out information about non-visible physical targets, rather than what is in their immediately visible surroundings. This is why users might be allowed to specify that. A tappable button that switches on and off the virtual attribute layer for non-visible targets might be one way to achieve this.

***DG.2.2. When the base layer is not digitally altered, ensure effective second referential mapping for visible targets***

AR browsers communicate information primarily through the content of the AR annotations they deliver on the screen of the smartphone. It is important that the user can make (at least in their mind) a connection between the delivered information and the physical target this information refers to. Otherwise, an AR browser ends up delivering a lot of "floating around virtual bubbles" that provide useless content.

Within the field of AR, only a few studies have discussed the importance of the second referential mapping or how it can be achieved. Normally, this is discussed in the context of placement of the AR annotations, or the ***spatial link*** between the representation of the physical target (base layer) and the AR annotations (attribute layer) (e.g. Vincent et al., 2012). The empirical studies confirm that the position of the annotation is important when it comes to association, as users expect that an annotation for a discrete object should be placed over or near that object. In complex urban environments, however, placement alone is not enough for effective association. The identified relationships in the framework suggest that, in order to communicate a *strong, one-to-one link* between virtual and physical space, *at least one* of the elements of the virtual attribute layer has to match visually the perceived (visible or non-visible) characteristics of the target object. This process is also referred to as ***visual coupling***. To date, commercial AR browsers have tried to achieve this visual coupling through delivery of abstract symbols or keywords within AR annotations. Such elements increase cognitive load and time, as the user first needs to interpret the symbol / keyword and then use it to match the result with inferred non-visual properties of physical targets (e.g. function).

A key implication is that designers and developers need to consider how objects in urban environments are perceived and interpreted in context. Visible graphic variables (e.g. colour, contour) are more suitable to be used as a matching parameter. A number of different approaches can be considered when it comes to visual coupling. Names and keywords can be used if they are physically present and visible from the current location of the user. Pictograms (landmarks) can be used when the target object is a building with a distinctive shape and contour.

When ensuring effective second referential mapping, the surrounding physical context has to be considered as well. The selected matching parameter has to be unique (e.g. unique shape in the surroundings). Annotations that contain a graphical variable common to more than one (or all) of the physical targets in the surroundings might lead to ambiguity. For instance, if there are more buildings within the same scene that have similar form, including a pictogram in the annotations that matches the contour of the target could lead to confusion. In this context, unique names on buildings are potentially the most successful matching parameter. However, they have to be relevant to the whole structure, and be visible from the current position of the user. With buildings that are similar to their context, pictures of specific elements within the building can be used for an effective visual match. The visual coupling can be relaxed for complex environments where the annotated feature has a spatial (e.g. square) or linear (e.g. street) characteristics.

***DG.2.3. Consider different representations for target and context annotations***

Users can only focus attention on a limited number of annotations at a time and too much information would lead to cognitive overload. When the annotation for the target object is visually different from the annotations that relate to the context, then users will be faster in focusing immediately on the most relevant content. The AR browser could guide the users' attention by providing slightly different annotations for targets and context. For instance, one approach would be to deliver information about the target object through an annotation that has a directional pointer. Apart from making the information about the target stand out, this is also beneficial, as the annotation does not occlude the target object and other physical targets that might fall within that region.

***DG.2.4. Consider manipulating the base layer to ensure effective second referential mapping***

The empirical studies indicate that the second referential mapping is more effective when it relies on a direct visual match, rather than on inferred non-visual properties of the target object. However, the perceived visible characteristics of the target might vary in time, for example, due to changing environmental conditions, lightning conditions, or the position and orientation of the user. While they are more measurable and objective than the inferred non-visible characteristics of the target, it still might be difficult to ensure second referential mapping based on perceived visual properties of the target object. This might also be the case when the physical target is not visually different from the surrounding context. Buildings or objects in the surroundings might have similar shapes, contours, colours, and/or textures. In such situations designers should consider manipulating and altering the representation of the target object on the smartphone screen (i.e. manipulating the base layer). Visual salience modulation (Vaes et al., 2013), adjusting slight details in specific image regions (e.g. Dong et al., 2011), rendering the real world in a non-photorealistic way (Fischer etal. 2005; Takeuchi ,and Perlin, 2012), photorealistic virtual models (Lee et al., 2012) or the colour-coding technique (e.g. matching the colour of a semi-transparent overlay with the colour of the annotation) could be used. The laboratory experiment (Chapter 7) showed that the latter is a very effective approach for augmentation and improves association significantly, even if annotations are not precisely placed on top of their physical target object.

***DG.2.5. If the base layer is digitally altered (manipulated), ensure successful first referential mapping for visible targets***

This design guideline is especially important if the base layer is somehow manipulated, for example, to ensure a more effective second referential mapping (see DG.2.4). If the first referential mapping is lost, users would be able to associated virtual information with its representation on the screen, but it would be difficult for them to relate this information to real-world features. This is especially the case with non-photorealistic rendering of the base layer (Fischer et al., 2005; Takeuchi and Perlin, 2012). Design for successful first referential mapping has been discussed widely in cartographic and GIS literature (Chapter 5). Empirical studies with augmented panoramas where the highlighting technique was used suggest that "people have such powerful capabilities for visual search and recognition that any highlighting should be designed carefully so that it does not compromise the convenience of free inspection of the details by the users themselves" (Vaittinen et al. 2013, p.201). Handheld devices allow users to inspect the physical environment freely, without interference. However, manipulation of the base layer could prevent users from relating actual physical targets with their representation on the smartphone screen.

***DG.2.6. Consider a different visual layout for AR annotations for visible and non-visible targets***

Throughout the empirical evaluations documented in this study it became clear that visitors to unfamiliar environments require provision of information about non-visible targets. There are a number of situations in which it is beneficial to include annotations for non-visible targets. For instance, the target object might be inside or behind a building. Likewise, in certain situations, e.g. during navigation, it is important to communicate that a certain POI is in the right direction, but not visible from the current position of the user.

The framework explains the different cognitive and perceptual processes involved in inference about visible and non-visible targets. The attributes of AR annotations for visible targets have to be considered carefully in order to allow users to match the AR annotation with its reference object. On the other hand, AR attributes for non-visible targets are not required for association. If all annotations are the same the user will be required to scan through both (physical and virtual) spaces in order to determine whether each of the annotations communicate information about visible or non-visible targets. To save time and cognitive resources, the graphical design and/or the content of the annotations should be manipulated so that it is immediately clear which annotations should be "used" for visual match and which should be used as a reference to non-visible targets. In this context, already developed guidelines for communicating depth in AR interfaces can be considered (Livingston et al. 2003).

While it seems that non-visible targets are easier to design for, care should be taken when determining both their layout and content. When the target object is not-visible from the current position of the user, and previously unknown, the content and layout of the annotation would determine the first impression of the user about that POI. This is why the attributes of the annotation, and the quality of the provided information, become critical for non-visible objects and can influence the overall experience with an unfamiliar environment. Lack of suitable information might prevent users from forming an adequate mental representation about the target object.

***DG.2.7. Match the spatial layout of the surroundings***

The type of physical target influences users' expectations with respect to where and when annotations should appear on the screen. Expectations are mainly influenced by the layout (spatial, linear, discrete) of the annotated physical target. For discrete physical object, users expect that the annotation should be static and appear in the visual centre of the target. For spatial entities, such as squares, users expect to find annotations near the ground or within the visual centre of the feature. For linear entities, such as streets, users expect that annotations will move together with the user along those features. This is why rules have to be defined with respect to spatial permanence for annotations.

### 9.5.3. Influence and control over perception of urban environments

***DG.3.1. Guide the attention of the user towards specific physical entities***

One key implication from the developed framework is that not all objects that might be important for the task of the user will be visually salient, trigger information needs and interaction with the hand-held display. In such situations, push-based notifications could be used (Beer et al. 2007) to encourage the start of a user session. Once a user session has started, models for determining automatically the visual salience of urban environments (Itti and Koch, 2001) might be applied. Such models are often based on neurobiological concepts of visual attention (Itti and Koch, 2001; Winter et al., 2005 – in Meng et al., 2005) and could determine perceptually salient (e.g. colour hue, colour value, orientation) characteristics of objects. Salient features could then be augmented with content. In non-salient environments, designers should consider manipulating the representation of physical objects in order to guide the attention of the user. Studies show that altering the details of augmented panoramas has successfully led to increased attention towards specific target objects (Vaittinen et al., 2013). Different visualization techniques (e.g. Veas et al. 2011) should be considered and evaluated to determine which is most successful with smartphone AR browsers.

***DG.3.2. Guide the attention of the user towards virtual entities***

The benefits of providing different representations for the target and context annotations were discussed earlier in this chapter (DG.2.3). Apart from attracting attention to specific target annotations, visualization techniques could aid in highlighting content within context annotations. This could be beneficial when there is a chance that users will miss out on important information about physical targets that are not the object of interaction. AR annotations that refer to different types of POIs but use similar information assets and graphics (e.g. the same symbols, same colours) hinder decision-making (Chapter 6). Once the user has started a use session, designers have the opportunity to deliver information about important points of interest, even if they are not the focus of attention of the tourist. Visual layout and content should be re-considered in order to attract attention, and communicate difference among POIs where and when necessary.

***DG.3.3. Avoid redundancy of information***

The primary purpose of the attribute layer is to communicate "more than the visible", or information that is not available within the physical environment. This is why it is important that content, which is not used for association, does not mimic already available information in the physical environment, or communicates information that the user has already inferred from the visible characteristics of the target. Redundancy leads to confusion and lowers the perceived utility of AR browsers. For example, locational or macro geospatial information (e.g. the name of the destination where the user is currently in) should be avoided.

***DG.3.4. Emphasize the uniqueness of physical objects***

The empirical results (Chapter 6 and Chapter 8) confirm the need to provide information about uniqueness of POIs. This is especially important when users are within a non-salient physical environment, or when they perceive the current settings as residential and far away from the locus of the tourist region. Keywords such as "interesting", or "popular", trigger interest and influence the perception towards specific urban entities. Such information could change the perception for perceived non-visible attributes of POIs and make them more memorable (and the destination as a whole) for tourists.

## 9.6. Chapter Summary

Following up on the empirical work described in this thesis, this chapter presented the revised user-centred design framework for design and evaluation of AR browsers (Section 9.2). It also described 10 key user requirements (Section 9.3) that have to be met in order to make AR browsers more useful and usable. Finally, the chapter provided design guidelines for design parameters (Section 9.4) that could be used, as well as 16 guidelines that need to be followed (Section 9.5) for developing more usable and useful AR browsers.

# CHAPTER 10

# CONCLUSION AND FUTURE WORK

**10.1. Introduction**

Introduction to the chapter

**10.2. Research Objectives Revisited**

Revises the main goals and questions for the study

**10.3. Thesis Contributions**

Describes the major contributions of the thesis

**10.4. Limitations and Directions for Future Research**

Reflective analysis of the thesis and directions for future work

**10.5. Conclusion**

Synthesis of the key points in the chapter

## 10.1. Introduction

Just like many urban residents, tourists require timely and fast access to relevant spatial and attribute information that supports effectively their decision-making process, but also enhances the experience with a destination. Smartphone Augmented Reality interfaces hold a great promise to provide relevant content in an easy and engaging manner. Early research has identified a number of challenges connected with tracking and registration, content delivery and representation of information. Thus far, however, efforts to elicit user requirements and provide guidelines in order to facilitate the design of AR browsers have been scarce. It is important to address this gap not only because of the potential of AR as an information delivery channel, but in order to provide further direction for current design and development efforts within academia and industry. Therefore, located within Information Systems Design and Human-Computer Interaction, the main aim of this study was to make a theoretical contribution by generating user-centred design knowledge expressed as the qualities and characteristics that Augmented Reality browsers should possess in order to meet user requirements in urban tourism context. This final chapter provides a synthesis of the findings and how they address each research objective. The chapter then proceeds with a discussion of the theoretical and practical contribution of this thesis. The chapter concludes with an evaluation of the research and directions for future work.

## 10.2. Research Objectives Revisited

Two critical components of design theories (Walls et al., 1992) are: 1) meta-requirements and 2) meta-solutions (design principles). Design principles aim to propose meaningful recommendations and “rules of thumb” that suggest how to satisfy user requirements and ensure a uniform experience and usable and useful interfaces, irrespective of platform (Nielsen, 1994; Fling, 2009). In order to make a contribution to design theory for AR browsers, five main objectives were identified, revisited below before presenting the contributions of the thesis.

***Research Objective 1: Explore the role of AR browsers in supporting (geospatial) information acquisition in large-scale urban tourism destinations.***

Augmented Reality browsers present a novel way to deliver information to tourists about large-scale physical environments. Until now, research and development has been based on the assumption that AR will be the killer app for tourism. However, the process of acquiring, storing and using spatial information is complex. Researchers have failed to discuss what is the role of AR in the overall (geospatial) information acquisition process in large-scale unfamiliar environments. This understanding ultimately helps researchers and designers to appreciate the actual advantages (and limitations) of using AR browsers to deliver information to tourists in urban tourism destinations.

In order to investigate the advantages of AR browsers for tourists, Chapter 5 discussed the key elements and cognitive processes that underpin the progression of geospatial knowledge acquisition. The importance of the three different types of geospatial knowledge (landmark, route and survey) for tourists were discussed. After review and comparison of different ways to provide information in large-scale urban environments through mobile Location-based services, the chapter identified the key role and advantage of smartphone AR browsers to support effective and efficient landmark (declarative) information acquisition.

***Research Objective 2: Examine the main problems that influence the usability and utility of AR browsers used in urban tourism destinations.***

Much of current research that investigates the usability and utility of AR browsers is often focused on navigation and wayfinding goals and tasks, or carried out with users who are already familiar with their environment. The empirical findings presented in Chapter 6, suggest that there are a number of problems that tourists experience when they want to acquire landmark (declarative) knowledge through AR browsers in large-scale unfamiliar urban tourism destinations.

**Problems with association of AR annotations and physical targets**

AR browsers deliver information through geo-tagged virtual bubbles, called AR annotations. Findings from the field-based empirical evaluation of four commercial applications showed that AR browsers do not support effective and efficient information acquisition through current AR annotations. Half of the time, tourists made errors when they tried to match virtual annotations with their physical target. As a result, the participants in the study were unable to effectively find information about points of interests in their immediate physical surroundings. The observations and analysis indicated that, in order to use AR browsers, tourists need to invest significant amount of time, cognitive and physical effort. Tourists were slow in associating AR annotations with their physical target and had to invest a lot of attention and cognitive resources in order to understand the mappings between virtual and physical space. As a result, the use of AR browsers was associated with higher physical effort, as tourists needed to spend extra time with extended arm.

**Problems with type and amount of augmented physical targets**

The results from the first mobile evaluation indicated that tourists are dissatisfied with the type of objects that current AR browsers augment. On a number of occasions tourists were unable to find AR content for points of interest they considered important and interesting from a tourist point of view. Partially, this problem occurred because tourists were unable to associate annotations with their physical targets. Most of the time, however, this problem arose because the annotation was missing. In such cases, participants took additional time and made a conscious physical and cognitive effort to look for the annotation that will deliver the desired content. In almost all of the test sessions, the inability to find the annotation for a specific object resulted in confusion and mistrust towards the mobile application.

The already available digital content that AR browsers deliver was also criticised. All of the fourteen test persons expected to find more information about specific tourist attractions in the city. The wide availability of content about local restaurants, cafes and shops was considered less valuable during sightseeing and discovery. The provided content influenced the perceived utility of AR browsers and participants expressed concerns about the overall quality, completeness and accuracy of information.

**Problems with finding relevant and useful content**

The empirical results presented in Chapter 6 and Chapter 8 indicated that tourists expect to find answers to specific questions triggered by observing a point of interest. The type of content delivered within AR annotations (names, symbols, distance, keywords) was criticised when participants were unable to find such information. Problems with content influenced the perceived utility of AR browsers. Provided content, such as addresses and distance to points of interests, was considered irrelevant, as it did not facilitate the on-site decision-making process. The feedback of all participants emphasized the need to access information that will provide answers to their location- and object-based questions. Participants also criticised the content delivered through AR browsers for being redundant, superficial and non-informative.

In addition, the observations and analysis revealed that current content does not support well decision-making in unfamiliar urban destinations. During the first mobile field study, the participants were asked to use the AR browser to optimise their route and select a point of interest they would like to visit. The provided content, however, could not support users in making an informed decision and almost all (8/10) of the participants resolved to use the physical, rather than the virtual space, to make a decision.

Apart from utility, available content also influenced inferences that users made with respect to physical targets. Points of interest were discarded when the TPs could not understand the content within AR annotations, e.g. the name, what the symbol stands for or the description of the object.

**Problems with awareness for available content**

Information Systems implemented on handheld devices require explicit interaction from users. This is especially evident when we compare smartphone and head-mounted AR displays. Head-mounted displays provide continuous augmentation of the environment, while smartphone AR requires users to take out the device and raise it vertically towards their surroundings. These specific interaction requirements might lead to situations where users are not aware about available AR content. Results from the qualitative field evaluation (Chapter 8) showed that this is a significant problem, especially in non-salient urban environments that provide little affordance or information scent. In such cases, users miss out on available information about important points of interest. While this issue is relevant to all ISs implemented on handheld devices, AR interfaces exacerbate this problem, as users need to know in advance in what direction they should point the device in order to obtain useful and relevant content.

**Application-specific usability problems**

The first empirical evaluation, presented in Chapter 6, revealed additional usability problems mainly due to technical issues. All of the AR browsers were slow to load content and crashed occasionally. A number of application-specific problems were also documented. In general, movement of annotations was a minor problem that did not affect performance with AR browsers. However, excessive movement of annotations in LocalScope required that users stand very still in order to work with the application.

Since users did not have experience with other AR browsers, this feature did not lead to negative feedback.

Size and overlap of annotations were mainly problematic for Wikitude users. Overlap prevented participants to access the content for certain AR annotations. The same was the effect of size, as annotations were too small and users were forced to tap several times on the screen before being able to access additional content.

***Research Objective 3: Investigate how key context of use factors influence the usability and utility of AR browsers.***

Usability is only meaningful when it is investigated in relation to representative users, with representative tasks, and in actual context of use. Usability and utility of AR browsers have mainly been investigated in relation to different types of goals (navigation, wayfinding), users (urban residents familiar with the environment) or environments (natural surroundings). As a result, understanding of actual context of use and the factors that determine usability and utility of AR browsers when used to acquire (geospatial) information in urban tourism has been limited. This study set out to investigate and propose meaningful ways to improve the usability and utility of AR browsers to support geospatial knowledge acquisition in urban tourism destinations. To this end, it was important to identify the key context of use factors that influence usability and utility.

Physical context has been discussed in a number of studies and researchers have emphasized the need to understand the role of "reality" in AR interfaces (Chapter 5). The findings emphasize the key role of physical context. As opposed to previous research, this is the first study to identify and discuss the importance of perceived physical space in unfamiliar urban environments. In addition, and even more important when it comes to user-centred design, the empirical evaluations revealed not only which context parameters influence usability, but also *how* and to *what extent* they determine whether users are able to acquire information effectively and efficiently through AR interfaces.

**The influence of context on association**

Summarising the empirical results described in Chapter 6 and Chapter 7, it was observed that association of AR annotations and physical targets was mainly influenced by two properties of physical objects: ***visual appearance*** and ***legibility***. When trying to match virtual annotations and physical targets, participants first referred to the available visual cues in the physical and virtual spaces. During this process, participants tried to use physical visual cues, such as colours and textures, physical name on display, form, contour and shapes and match them to the visual characteristics of virtual AR annotations. Association was most effective (least errors, higher certainty) and efficient (lowest time, lowest difficulty) in the cases where there was at least one direct match between the perceived visible characteristics of the physical object (e.g. colour, physical name on display) and the visual elements of the virtual annotation (name, colours, pictures).

In the absence of a direct visual match, users relied on ***indirect visual match*** strategies. In other words, they tried to match the elements of the annotation with the ***inferred non-visible attributes*** of the target object. This worked in situations where users could infer correctly at least one of the non-visible attributes of an object (e.g. the function of a building) and match it with an element within the AR annotation (e.g. the symbol for the Information Centre). Further laboratory testing (Chapter 7) confirmed that performance deteriorates when the visual characteristics of the building, or its inferred non-visual attributes, did not correspond to information within the AR annotations.

The identified relationship between perceived visible and non-visible properties of physical targets and association requires that we re-examine the role of environmental and physical context within the AR interface. The findings suggest that visual salience is an important property that has to be considered when it comes to design of AR browsers. Whether an object is visually salient depends on the characteristics of the user, the physical target and the surroundings. It also depends on the visibility of the target object. So far, ***visibility*** has been used as a contextual factor that is used for information filtering and determines whether content should be displayed (or not) on the screen of the smartphone device (Julier et al., 2002; Kruiff et al., 2010). As opposed to previous studies, this study reveals the importance of visibility as a contextual factor that influences association of physical targets and virtual AR content. This means that, rather than considering visibility as a binary property (visible/non-visible), it is important to consider the extent to which a target is visible (fully visible, partially visible) from the current location of the user. In addition, other environmental factors, such as lightning level, could influence perceived visible characteristics of the physical target and, in turn, hinder or facilitate association of targets and virtual annotations.

Finally, the use of inferences and non-visual properties of physical objects suggests that acquired ***landmark knowledge*** will influence significantly the association process of physical targets and virtual AR annotations. In particular, instead of relying on spontaneous inferences about physical objects, users could use acquired knowledge (e.g. names and functions of buildings and points of interest) to match annotations with physical targets. This observation is extremely interesting to investigate further in the context of other types of visual displays, as the relationship between familiarity and the use of geospatial technologies remains unexplained.

**The influence of context on perceived utility of AR browsers**

It has long been recognised that mobile information needs of tourists are mainly influenced by their location. While an important parameter, this study confirms observations from previous research that the use of location, expressed as absolute geographical coordinates, is an insufficient sole determinant of information needs and, therefore, cannot be used as single predictor of utility for mobile ISs. The findings from this study emphasize the role of physical context and its influence on information needs, and therefore, perceived utility of AR browsers. Both the visible and non-visible properties of physical entities influence the ***information needs*** of tourists, expressed as questions that they sought answers to through the smartphone device. Lack of content that answers such questions influences the overall utility of AR browsers.

Tourists' ***expectations*** were also shaped by visible cues and inferred non-visible properties of their physical environment. In particular, especially within unfamiliar environments, visual cues are used consciously and subconsciously to infer the non-visual properties of physical objects, and therefore as signifiers of importance and uniqueness of points of interest. Thus, the appearance of the physical environment shapes to a large extent the expectations of tourists with respect to available content that they should be able to access through smartphone AR browsers. Lack of such content influences the perceived utility of mobile ISs.

**The influence of context on interaction**

Despite differences in terms of use of visual cues, it is clear that the visible characteristics of the environment play a significant role and determine which landmarks will attract the attention of the user. However, not all visually salient objects and elements will trigger interaction with the smartphone AR browser. Apart from the visible characteristics of physical entities, their non-visible properties also play an important role as triggers that will result in users interacting with the AR browser. Inferred functional and cultural significance of physical targets will determine whether users are willing to dedicate cognitive and physical effort in order to learn about their surroundings.

***Research Objective 4: Identify the key user requirements that need to be satisfied in order to improve the usability and utility of AR browsers.***

Identifying key user requirements (meta-requirements) is crucial towards making a theoretical contribution to Information Systems Design theory. Chapter 9 presented the key user requirements that have to be satisfied in order to prevent problems and ensure usable and useful AR browsers for tourists in urban tourism context.

Elicited user requirements indicate that in order to provide a seamless and immersive user experience, designers need to consider *how* virtual content is presented on the screen of the smartphone. In order to be usable, a key user requirement that has to be satisfied is to ensure that users are able to match AR annotations with *visible* physical targets. This process can be facilitated if there is at least *one direct (visual or inferred) match* between the AR annotation and the target object. The content should also match the perceived non-visible characteristics of targets, as users expect to find information for points of interest that they consider important and/or interesting from a tourist point of view.

While content for both visible and non-visible points of interest should be provided, a key requirement is to prioritise design and information for *visible* physical targets. Delivered AR annotations should maximise information acquisition by providing enough content that supports informed decision-making within the current immediate visible settings. Redundancy of information (between virtual and physical spaces) should be minimised. Delivered content should match or influence the perception of tourists by providing information that explains how and why physical targets are unique, interesting and important from a tourist point of view.

Finally, AR browsers should support situation awareness by providing visually different AR annotations for non-visible physical targets. Their relative position can communicate the direction in which those targets are located.

***Research Objective 5: Propose key design parameters that could be used to improve the usability and utility of AR browsers.***

**Placement of annotations**

Much of current AR research focuses on developing different algorithms directed at precise placement of AR annotations. The results described in the thesis (Chapter 6 and Chapter 7) indicate that even if immaculate placement is achieved, tourists might be unable to associate AR annotations with their reference target object. In many situations, association was successful despite the fact that annotations were misaligned. Further laboratory-based evaluation (Chapter 7) showed that placement is critical only when users have to rely on indirect match between virtual annotations and physical targets. When users had to infer the function of the target object and match it with the keywords used within AR annotations, they were only able to do so if annotations were precisely placed on top of the physical target. Placement, however, did not influence success, time, certainty and difficulty significantly when users relied on direct visual match between annotations and physical targets. Therefore, the results from both the field and laboratory-based evaluations suggest that designers need to consider alternative design parameters in order to ensure effective association of content and physical environment.

**Visual layout and abstraction level of annotations**

Graphical (visual) layout of annotations for smartphone AR browsers has scarcely been discussed in literature. The few existing studies that describe and evaluate different graphical variables (opacity of background, colour of background, colour of font) for AR annotations were mainly directed at ensuring legibility on wearable (HMD) displays. Until now, layout has not been considered when it comes to effective association of virtual annotations and physical targets. The results from this study show that the process of matching physical and virtual spaces in AR browsers is heavily influenced and dependent on the visual layout of AR annotations. As discussed earlier, association is most successful when users are able to directly match the perceived visible characteristics of the physical object (e.g. colour, physical name on display) and the elements of the virtual annotation (colours, shape, symbols). The key implication from further empirical testing (Chapter 7) suggested that, if positioning data is accurate, colour could be used to support effective and efficient association of target objects and AR virtual annotations. Pictograms (landmark symbols) could be used to support association when positioning data is not precise and error prone. Care should be taken, however, that pictograms (landmarks) are simple and do not require mental rotation or visual search. Since the process relies on direct visual matching, performance is influenced mainly by the abstraction level of the pictogram. More realistic pictograms will ensure more effective and efficient association.

**Visual layout and abstraction level of base layer**

The low number of errors, time, high certainty and low difficulty observed during the laboratory experiment, described in Chapter 7, indicated that different graphical variables could be used to support effective and efficient work with AR annotations. Overall, however, association was subtly influenced by the relationship between the abstraction level of the base layer (representation of physical world) and the abstraction level of the attribute layer (AR annotations). The results suggested that task performance was most effective and efficient when graphical variables of both base and attribute layers were manipulated to have similar visual appearance (colours).

**Content of annotations**

Perceived relevance and usefulness of content delivered through AR browsers was investigated and reported in Chapter 6 and Chapter 8. The results from the empirical studies described in this thesis suggested that content should satisfy the information needs of users by providing information that answers their specific object-based questions. In addition, content needs to be considered carefully in view of findings that reveal the impact of redundancy on perceived utility of AR browsers. Delivered information within AR annotations has to be selected and represented in a way that enriches and adds to the perception and knowledge of users about their immediate visible and non-visible surroundings. Type and level of detail of information impact the perception of tourists for specific points of interest and the overall affordances that the physical environment offers. Therefore, content needs to explain why specific tangible (points of interest) and non-tangible (e.g. events) entities are unique and special from a tourist point of view.

**Leader lines and directional pointers**

Leader lines and directional pointers are extremely valuable in abstract 3D graphics (e.g. graphs and charts, anatomy drawings). The results from the laboratory-based experiment (Chapter 7) confirm the advantages in using directional pointers, as they facilitate users to effectively and efficiently associate virtual AR annotations with physical targets. However, when placement of annotations is imprecise, the use of directional pointers confused users and led to increased task time and mental effort. The use of directional pointers for non-visible targets might lead to misunderstanding and errors.

***Research Objective 6: Capture the key constructs and relationships that determine usability and utility of AR browsers in a conceptual user-centred design framework that facilitates the design and evaluation of AR browsers.***

The literature review (Chapter 2 and Chapter 3) observed a lack of coherence in the definitions, approach and methods used in AR and HCI to design AR annotations. Therefore, one of the key objectives of this study was to develop a conceptual framework that captures the process of using AR annotations in unfamiliar urban environments and provides a mechanism to evaluate and optimize their design. The main objective that the framework tried to achieve was related to identifying the key constructs (and relationships between them) that are important to consider when designing, developing and evaluating AR browsers.

Several key activities were undertaken to achieve this objective, all in line with the general process of generating design knowledge through Information Systems design theory development (Section 3.4.1) and user-cented design (Section 3.4.2). First, relevant design knowledge, captured in theories, guidelines, models and frameworks was identified from available literature in several domains, including geo-information science, environmental psychology, information science and tourism literature. This process helped in preliminary identification of the most important constructs and the relationships between them, captured in a conceptual theoretical framework, described in Chapter 5. The identified relationships and different constructs of the theoretical framework were re-examined after obtaining empirical data and analysing it. This process is described in the discussion sections of Chapter 6, 7 and 8. Finally, the obtained new empirical findings were incorporated in a re-visited version of the framework, described in Chapter 9. The final framework examines the relationship between three key constructs, namely the user, the AR interface and the context of use. Apart from the high-level interactions among those three elements, the framework identifies the role of various sub-components that determine and influence the usability and utility of AR browsers.

One of the key advantages of the developed framework is that, due to its explanatory nature, it can be used to generate hypotheses regarding different design alternatives and how well they could support work with AR browsers.

***Research Objective 7: Propose design principles for developing AR browsers used in tourism context.***

Identifying meta-solutions or design principles for AR browsers is essential towards contributing to Information Systems Design theory. Visual displays should present information in a way that enhances cognitive and physical activities and facilitates users to make informed and correct decisions. More importantly in the context of this research, AR is a visualisation that can not only enhance, but also directly influence the perception of physical space for tourists in unfamiliar urban destinations. To this end, however, designers need to consider carefully both the perceptual properties of selected visual variables, as well the type of information that is presented in AR browsers. The provided guidelines, described in Chapter 9, emphasize the importance of considering both the design of AR annotations, as well as the characteristics of the representation of the physical world as part of the AR interface.

When it comes to design of AR annotations, the AR browser needs to be adaptive with respect to the intent of the user and the target that should be augmented with content. In order to support effective information acquisition, developers should consider different designs for AR annotations that refer to visible targets, visible context and non-visible points of interest. The design of *target AR annotations*, which refer to the target that triggered an information need, has to support effective second referential mapping (association). Design for *context AR annotations* (visible and non-visible) should emphasize attracting the attention of the user to the referent object.

Apart from careful selection of design parameter for AR annotations, designers should also consider manipulating the incoming video feed (base layer). This is important in order to ensure effective second referential mapping in visually uniform urban environments. Altering the characteristics of the base layer could also be used to attract attention to specific physical objects and points of interest that otherwise might be missed. As discussed in Chapter 9, designers should also take care that the visual representation of the base layer supports users in effectively carrying out the first referential mapping (association between the physical world and the representation of the physical world on the smartphone screen).

## 10.3. Thesis Contributions

### 10.3.1. Contributions to Information Systems Design Theory

Located within Information Systems Design and Human-Computer Interaction, the main aim of this study was to make a theoretical contribution to Information Systems Design theory through generating new user-centred design knowledge expressed as the qualities and characteristics that Augmented Reality browsers should possess in order to meet user requirements in urban tourism context. In line with the process of Information Systems Design Theory (ISDT) generation, this study reviewed and identified relevant existing design knowledge and theories in several disciplines. While existing research has provided descriptions of other theoretical design frameworks (Vincent et al., 2011; Alzahrani et al., 2012), the existing models and frameworks are mainly developed from a technical point of view. Despite the critical need for empirical, user-centred approach when it comes to design of Augmented Reality (Swan II and Gabbard, 2005; Gabbard and Swan II, 2008; Dunser et al., 2008) and context-aware smartphone information systems (Greenberg, 2001; Dourish, 2004; Oulasvirta, 2012) such frameworks and models have failed to consider design of AR browsers from a user-centred perspective.

Adopting a unique approach towards AR browsers as (visual) tools that can enhance and support (geo)spatial knowledge acquisition, this study first identified relevant knowledge in several disciplines that was captured in a conceptual design framework. Building on previous research and theories in AR (Vincent et al., 2011) and geo-information science (Imholf, 1975; Kraak and Ormeling, 2010), the framework was used to deconstruct the AR interface and identify potentially important design parameters and context of use factors that could influence the usability and utility of AR browsers in urban tourism context. In line with ISDT generation (Nunamaker and Chen, 1991; Walls et al., 1992; Gregor and Jones, 2007; Gregor, 2009; Gregor and Hevner, 2013; Gregory and Mautermann, 2014), the framework was then used to drive further empirical data acquisition through observations and experiments.

The empirical studies described in this thesis enhance our understanding of the relationship among users, context of use and design of AR interfaces. Consistent with studies outside the tourism and AR domains (Nasar et al., 2005; Caduff and Timpf, 2008), this research revealed the influence of perceived visual and non-visual appearance of physical environments on the association process between virtual and physical spaces. It was discovered that visual perception of large-scale physical environments plays a fundamental role and influences the information needs and expectations of tourists regarding delivered content through smartphone AR browsers.

The study also makes several important empirical contributions to research within the Information Systems domain. The first field-based mobile user study was conducted with commercial AR browsers (Chapter 6). It resulted with an extensive and rich dataset that captured work and embodied interaction with smartphone AR browsers in unfamiliar urban environments. The analysis extended substantially findings from previous empirical studies with AR browsers in everyday settings (Ganapathy et al., 2011) and tourism context (Toh et al., 2011; Linaza et al., 2012; Kourouthanassis et al., 2014). In particular, the study documented the problems that tourists experience to carry out association of virtual AR annotations and physical targets in different types of urban environments. As opposed to previous studies where evaluation has been mainly subjective, the study used a set of objective HCI measures (time and errors) to investigate the severity and impact of such problems. In addition, a number of other limitations were documented, including the lack of useful AR content.

Analysis of the obtained data suggested that tourists use two main strategies in order to relate virtual and physical space: direct and indirect visual matching. Such findings are consistent with previous findings that document the use of mobile 2D and 3D location-based interfaces and information rich virtual environments (Daft and Lengel, 1986; Elias and Paelke, 2008; Oulasvirta et al., 2009; Partala et al., 2010; Partala and Salminen, 2012). Further empirical testing (Chapter 7) confirmed and enhanced our understanding of the process. In particular, visual design of annotatios and the use of type of content influenced significantly task performance (time, errors, certainty and difficulty) with AR browsers. Laboratory testing with 90 participants confirmed that users need at least one direct visual match in order to carry out the association process between virtual and physical spaces.

Beyond association, the research found support for the role of physical context in the formulation of information needs in urban environments. In particular, the role of physical context and tourists' familiarity was discussed. The findings support and expand previous research and theories about mobile information needs (Church and Smith, 2009) and meaning making in urban tourism context (Barba, 2014). Previous research has discussed the role of physical context (location) as a trigger for information needs and information search behaviour. The observations confirm the importance of physical context, discussing the specific influence of how visible and non-visible surroundings are perceived on information needs. Moreover, the notion of meaning-making in physical space (Barba, 2014) was confirmed as users of mobile AR expressed the need to acquire information that explains and imbues physical space with meaning.

The obtained empirical data were used to revise the developed original theoretical model and propose a new user-centred design framework for design of smartphone Augmented Reality browsers used in urban tourism context. The framework examines interaction with AR browsers and accommodates existing theories to explain the process of information acquisition in unfamiliar environments. The framework is of high value and relevance to researchers as it can be used to support the planning of experimental and user-based studies.

Considering the unique and multi-disciplinary approach undertaken in this study and the scope of the obtained results, the thesis has smaller contributions relevant to new knowledge within the fields of Augmented Reality, Mobile Human-Computer Interaction and Geo-Information Science.

### 10.3.2. Contributions to Augmented Reality

The main contribution to the multi-disciplinary field of Augmented Reality lies in generating new empirical knowledge relevant to the influence of context of use on usability and utility of AR browsers. The need for understanding how reality influences the overall usability and utility of AR interfaces has been recognised (e.g. Livingston, 2013). Existing research within the field has proposed models and frameworks that consider context simply as background that should be augmented with virtual information (e.g. Kalkofen et al., 2009). The findings in this study reveal the active role of physical context in determining the usability and utility of AR browsers and therefore make a strong theoretical contribution to design of Augmented Reality interfaces. In addition, the developed conceptual design framework provides a novel perspective towards AR interfaces as tools to acquire (geospatial) knowledge, and highlights the three key design components for AR browsers (abstraction of base layer, abstraction of attribute layer, amount of information). This original perspective allows a thorough understanding of the design space of AR browsers (Chapter 9), and a more meaningful classification of existing AR annotations.

### 10.3.3. Contributions to Mobile Human-Computer Interaction

The contributions of this study to Mobile HCI are mainly empirical in nature. Within the field of Human-Computer Interaction, empirical contributions can be quantitative, qualitative or mixed in nature and consist of "new findings based on systematically observed data" (Wobbrock, 2012, p. 1). This study documented empirical data obtained from a total of 112 participants through four empirical studies. The main purpose was to provide new data and reveal formerly undocumented insights about human behaviour and interaction with smartphone Augmented Reality browsers in urban tourism context. The obtained knowledge was captured in a new user-centred design framework, which constitutes a descriptive and predictive tool for mobile interaction in urban tourism destinations. As a descriptive tool, the framework can be used by Mobile HCI researchers to provide a new perspective and way of thinking about potential design problems and processes that occur when geospatial knowledge acquisition is mediated through smartphone technology. In addition, the framework can be used as predictive tool, to generate hypotheses which could be explored further in future research.

### 10.3.4. Geo-Information Systems Design

This study relied heavily on existing theories, frameworks and models within environmental psychology and geo-information science. Literature revealed that, while familiarity is a key construct, there is still little understanding of whether and how it actually influences geospatial knowledge acquisition and design of smartphone geo-information tools. Findings suggested that familiarity influences both the usability and utility of AR browsers and should be considered during evaluation of smartphone visual displays. More importantly, the developed framework and empirical observations documented in this study discuss *how* familiarity influences knowledge acquisition in urban environments and, therefore, advances theory in mobile geo-information systems design. The thesis also has methodological contributions, as it presents a new way to evaluate familiarity in urban environments.

### 10.3.5. Practical contributions

From a practical point of view, the research is geared towards providing tangible help to developers of smartphone ISs not familiar with mobile ISs for tourism, or AR and their implications for effective delivery of information in outdoor environments.

## 10.4. Limitations and Directions for Future Research

This thesis has contributed to the emerging area of Mobile Information Systems Design concerned with communicating information in a useful and usable manner to users who require fast access to spatial and temporal data about their immediate physical surroundings. With all research being the product of a number of compromises, researchers benefit from evaluating their approach and findings using operational and empirical adequacy criteria. In this regard, the findings in this thesis are associated with a number of limitations that are noted in this section.

The appropriateness of the epistemological and methodological decisions that were made throughout the thesis was thoroughly justified based on their appropriateness with respect to the main aim of the study. The overall UCD research methodology provided valuable framework for investigating behaviour and interaction with smartphone AR browsers in urban tourism context. Key decisions, methods, measures, data acquisition and analysis techniques were also considered in light of the approaches in previous research carried out in Human-Computer Interaction, Augmented Reality, Geo-Information Science and eTourism. Previous research, as well as the objectives set out in the thesis, guided the design of the empirical studies described in Chapter 6, 7 and 8. As already discussed in the relevant chapters, a number of decisions had to be made, mainly driven by time and resource availability, which ultimately influenced the richness of the obtained data.

The sampling criteria for all studies were based on the selection of the method, the purpose of each empirical study and the adopted (qualitative and quantitative) analysis techniques. In terms of sampling, the first mobile field evaluation set out strict criteria with respect to the recruited participants and only representative test subjects were allowed to participate. While only 14 participants took part in the study, it yielded sufficient data for quantitative and qualitative analysis. While no correlations were found in this study, additional research could utilize a larger sample in order to address how variations in different user characteristics affect information needs and AR design.

For instance, studies could examine the role of tourist preferences and interests, as well as other factors such as cultural background.

The thesis addresses the needs of tourists, or people who roam around in unfamiliar environments. Since potentially anyone can be a tourist, the main aim is to generate a universal design theory, or design principles, applicable to all people, to the greatest extent possible (Schneiderman, 2000). This means that the final AR browsers should be usable by virtually anyone. This concept, which addresses the access of information and communication technologies by anyone, is called *universal usability* and has been heavily advocated by Ben Schneiderman (2000).

In order to gauge insights from a wider group of users, recruitment was directed at a group of users with different characteristics. However, due to the limited resources and time, it was not possible, nor needed, to unearth requirements from all possible populations. Therefore, the user requirements described at the end of the study (Chapter 9) relate to tourists who:

- engage with an environment they have little knowledge about
- visit a destination for leisure purposes
- are aged between 18 and 61 years
- have a high-school and university degree
- speak English fluently
- travel alone
- tech savvy and are fluent in using a mobile device
- have lived in the UK or have background knowledge of its culture
- have no special knowledge of or experience with using AR
- have no special domain knowledge or interests
- have no physical or cognitive disabilities

In this context, further research is needed to uncover the requirements of users where changes in demographics (elderly or children), education background (e.g. no degree), physical abilities (people with disabilities), cultural background and different language proficiency (e.g. do not speak English) or social context (e.g. family with children, couple) could have influence on the final user requirements.

The main drawbacks of mobile field studies include the influence of confounding factors and noise in the data. In order to ensure the internal validity of the data, evaluation in the first mobile field study had to be limited to specific temporal, task, and physical context (a pre-defined route). Therefore, the influence of social or temporal context of use factors could not be explored empirically. Further empirical research could be carried out to reveal the influence of travel companions, time of day or year, or various types of tourism destinations (e.g. coastal, natural) on the perceived utility and usability of AR browsers.

The second empirical evaluation, described in Chapter 7, used a more traditional laboratory-based approach in order to ensure the external validity of findings related to observed mobile interaction in urban environments. The use of augmented photos during the laboratory-based evaluation was suitable and appropriate as tools that simulate AR, and they served as useful artefacts for evaluation. Due to the set objectives and research questions, it was found critical that the evaluation of AR during the laboratory-based experiment considers annotations for visible physical targets only. The obtained findings were thus more focused and useful as they confirmed key relationships in the conceptual design framework. Further research could address the role of additional factors, such as movement of annotations, lightning conditions, and visibility of target objects and their actual influence on usability and utility of AR browsers.

Qualitative evaluation with domain expert users carried out in laboratory settings and on the field was then described in Chapter 8. The design of all questions and materials that were used during the evaluation was directed towards obtaining further feedback regarding the design of AR annotations used in urban tourism context. Therefore, the questions and evaluation tasks were rather open and exploratory in nature. Further research could address evaluating the utility of different types of content presented to tourists in a more structured and quantitative manner.

A pluralistic walkthrough is typically carried out in a room, outside of the context of where the mobile IS will be used. In order to obtain results grounded in actual context of use, the pluralistic walkthrough was carried out outside, in an unfamiliar large-scale environment. Naturally, the results have to be understood and considered with regard to those contexts and pertaining to a specific type of user population. Gaining a truly holistic picture of all of the potential issues and evaluating all content types through a pluralistic walkthrough is unrealistic. Nevertheless, the aim was to gain understanding as extensively as possible in a very early stage of development and using an approach that has not been used before in AR research. In this context, and considering the growing popularity of AR browsers, any information about potential issues and problems, as well as user requirements is important.

Cognitive walkthrough has been applied successfully to evaluation of tourist guides (Almeida et al., 2007). This approach is especially suitable when the needs and potential problems of intermittent or novel users are assessed. This makes cognitive walkthrough suitable for exploring the potential problems that tourists experience with AR browsers, as tourists fall within this category of users. However, cognitive walkthroughs have also many drawbacks, inherent to other expert usability inspection methods.

During a pluralistic walkthrough, the focus is always on first time users and how they will react to the system. This was considered suitable considering that the target user group is tourists. Such users will be intermittent in their use of the system and might forget, and therefore need to re-learn, how to use an AR interface. Therefore, the results do not address issues or requirements that might occur with more experienced, or the so-called *power users*, who have had significant exposure to the technology and might use AR every day on a number of occasions.

Like other methods that involve multiple people discussing a topic (e.g. focus groups), pluralistic walkthroughs benefit from a collaborative discussion (Nielsen, 1993), which is very valuable for identifying usability problems and solutions. However, this also means that participants within the group might be influenced by others' opinions and refrain from being honest. To prevent this, an introduction is given in the beginning where participants are asked to be honest and provide as much detailed feedback as possible. Outside of this, there was no way to limit the influence of this factor. Pluralistic walkthroughs involve multiple groups (Dix et al., 2004). Usually those include:

- representative users (from the target user group)
- usability or HCI experts
- programmers or designers

Following such recommendations, the pluralistic walkthrough was carried out with a group which comprised of domain experts relevant to the design of AR interfaces for tourism, including HCI and geo-information experts, as well as tourism and marketing domain specialists. At the same time, the participants were placed in an unfamiliar environment, trying to simulate the experience of being a tourist.

One of the main disadvantages is that opinions and feedback are influenced by the expert domain knowledge available to users. However, considering that design theory for AR information systems requires expert knowledge within HCI and Geo-Information Science, it was needed to gauge insights and feedback from experts within those disciplines.

One of the main disadvantages of a pluralistic walkthrough is that feedback and the evaluation itself is focused on a range of specific pre-defined tasks and situations. This means that issues outside of the selected tasks might not be detected and effectively ignored. Indeed, the pluralistic walkthrough focused on matching tasks that required participants to associate virtual and physical worlds. Considering the results from the first mobile field study, described in Chapter 6, it was essential that evaluation is further carried out with those range of tasks. Therefore, other tasks were purposefully excluded. Further research could evaluate issues or perceived benefits that arise from additional tasks, such as search and browse.

Finally, smartphone AR is still an extremely young field, and there remain many avenues for future investigation that can contribute to the overall usefulness and acceptance of this special type of mobile ISs in tourism. An important characteristic of the developed design framework is that on a high level it captures constructs and processes relevant to information acquisition through location-based interfaces in general. Therefore, in the future, it could be used to analyse and improve the design of other types of mobile context-aware interfaces that are used on-site. More importantly, the framework could be used to analyse, evaluate and design for effective and seamless transition among different types of interfaces. This process concerns a newly emerging area within Augmented Reality concerned with the design of *transitional interfaces* (Trevisan et al., 2011). Further research could investigate empirically the use of AR browsers in combination with different types of mLBS interfaces, such as 2D and 3D maps. Another strand of research could explore the suitability of the framework and proposed design for different current and future form factors, such as eyeware, fixed digital binoculars or projective displays. This is especially important in view of the fact that tourists, as well as everyday users of mobile ISs, often need different types of visual displays to effectively acquire geospatial knowledge about their environment.

## 10.5. Epilogue

This thesis provides practical and theoretical directions for design of smartphone AR browsers. To this end, it investigates empirically the use of existing and future (prototypical) AR solutions in urban tourism context. The study adopts a unique approach towards AR design, as it treats such interfaces as external visual tools that enhance geospatial knowledge acquisition in unfamiliar urban environments. As such, it represents an original and innovative piece of research that has a number of contributions to both theory and practice.

Considering the wide popularity of AR browsers, and the significant potential of such tools to enhance the experience of users, the topic is likely to gain more traction and prominence in both academia and industry. In particular, it is expected that design knowledge generation will increase as more and more researchers and practitioners incorporate and apply human-centred approaches to development. User-Centred Design has proved to be an effective methodology in this study. More importantly, it directed research and emphasized the need for a multi-disciplinary approach to design of AR interfaces as tools that mediate the experience of tourists.

By adopting such novel and original approach, this study contributed to Information Systems Design theory through identification user requirements, as well as the development of a predictive and analytical user-centred design framework for AR browsers used in urban tourism context. On a more general level, the framework could be used to drive research in several closely related disciplines, including Augmented Reality, Geo-Information Science and Mobile Human-Computer Interaction. In summary, this thesis provided directions for design of more usable and useful Augmented Reality interfaces. In this sense, it brings us one step closer to the actual vision for ubiquitous and pervasive computing, and in particular augmented smart cities, where physical and virtual fabrics co-exist and intertwine seamlessly to provide meaningful, memorable and unique tourist destination experiences.

# REFERENCES


Abras, C., Maloney-Krichmar, D., and Preece, J., 2004. User-Centered Design. *In:* W. Bainbridge (ed.) *Encyclopedia of Human-Computer Interaction*. London: SAGE Publications.

AcrossAir, 2014. Available http://www.acrossair.com/ . [Accessed 10 January 2014].

Ahmed, T. and Cox, J., 2014. *Developing Information Systems: Practical Guide for IT Professionals*. London, UK: BCS Learning & Development Ltd.

Ajanki, A., Billinghurst, M., Gamper, H., Jarvenpaa, T., Kandemir, M., Kaski, S., and Koskela, M., Kurimo, M., Laaksonen, J., Puolamaki, K., Ruokolainen, T., and Tossavainen, T., 2010. An augmented reality interface to contextual information. *Virtual Reality*, 15(2-3), 1-13.

Allen, J. and Chudley, J., 2013. *Smashing UX Design*. Chichester: John Wiley and Sons.

Almeida, L., Amdal, I., Beires, N. , Boualem, M., Boves, L., den Os, E., Filoche, P., Gomes, R., Knudsen, J., Kvale, K., Rugelbak, J., Tallec, C., and Warakagoda, N., 2002. The MUST guide to Paris: Implementation and expert evaluation of a multimodal tourist guide to Paris. In: *ISCA Tutorial and Research Workshop (IDS'2002)*, Kloster Irsee, Germany.

Alzahrani, A., Loke, S. and Lu, H., 2011. A formal model for advanced physical annotations. *In: Proceedings of the IEEE 9th International Conference on Dependable, Autonomic and Secure Computing*. Sydney, Australia, IEEE, 170–177.

Antunes, P. and André, P., 2006. A conceptual framework for the design of geo-collaborative systems. *Group Decision and Negotiation*, 15 (3), 273-295.

Appleyard, D., 1969. Why buildings are known: a predictive tool for architects and planners. *Environment and Behavior*, 1(2), 131–156.

Arazy, O., Kumar, N. and Shapira, B. 2010. A theory-driven design framework for social recommender systems. *Journal of the Association for Information Systems*, 11(9), 455-490.

Avison, D. and Fitzgerald, G., 2006. *Information Systems Development: Methodologies, Techniques and Tools*. London: McGraw Hill.

Avison, D., and Fitzgerald, G., 2002. *Information Systems Development*. 4th ed. London: McGraw Hill Companies.

Axure, 2014. Available http://www.axure.com/ [Accessed 13 September 2014].

Azuma, R. and Furmanski, C., 2003. Evaluating label placement for Augmented Reality view management. *In: Proceedings of the 2nd IEEE and ACM Symposium on Mixed and Augmented Reality (ISMAR '03)*, Tokyo, Japan, ACM Press, 66- 75.

Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S. and MacIntyre, 2001. Recent advances in Augmented Reality. *IEEE Computer Graphics & Applications*, 21 (6), 34–47.

Baldauf, M., Fröhlich, P., Masuch, K., and Grechenig, T., 2011. Comparing viewing and filtering techniques for mobile urban exploration. *Journal of Location Based Services*, 5 (1), 38-57.

Banga, C. and Weinhold, J., 2014. *Essential Mobile Interaction Design: Perfecting Interface Design in Mobile Apps*. Boston, USA: Addison-Wesley.

Barnard, L., Yi, J. S., Jacko, J., and Sears, A., 2007. Capturing the effects of context on human performance in mobile computing systems. *Personal and Ubiquitous Computing*, 11(2), 81-96.

Baus, J., Cheverst, K. and Kray, C., 2005. Survey of map-based mobile guides. In: Meng, D., Reichenbacher, T. and Zipf, A., eds., *Map-Based Mobile Services*, Springer, 193-209.

Becker, J. and Bjorn, N., 2007. Epistemological perspectives on IS research: a framework for analyzing and systematizing epistemological assumptions. *Information Systems Journal*, 17(2), 197–214.

Bell, B., Feiner, S. and Höllerer, T., 2001. View management for Augmented Reality. *In: Proceedings of the annual ACM symposium on user interface software and technology (UIST '01)*, ACM Press, 101-110.

Bell, B., Feiner, S. and Höllerer, T., 2005. Maintaining visibility constraints for view management in 3D user interfaces. *In:* Stock, O. and Zancanaro, M., eds., *Multimodal Intelligent Information Presentation*, Springer, 255-277.

Bellotti, V., and Edwards, K., 2001. Intelligibility and accountability: human considerations in context-aware systems. *Human Computer Interaction*, 16 (1), 193-212.

Bevan, N. and MacLead, M., 1994. Usability measurement in context. *Behaviour and Information Technology*, 13, 132-145.

Bevan, N., 1995. Measuring usability as quality of use. *Software Quality Journal*, 4, 115-150.

Bevan, N., 2009. Extending quality in use to provide a framework for usability measurement. *In:* Kurosu, M., ed., *Human-Centred Design*, 5619, Springer, 13-22.

Beyer, H. and Holtzblatt, K., 1998. *Contextual Design*. London: Academic Press.

Bias, R., 1994. The pluralistic usability walkthrough: coordinated empathies. In: Nielsen, J. and Mack, R., eds., *Usability Inspection Methods*. Los Angelis: John Wiley and Sons.

Biesta, G., 2010. Pragmatism and the philosophical foundations of mixed methods research. *In:* A. Tashakkori & C. Teddlie, eds, *Sage Handbook of Mixed Methods in Social and Behavioral Research*. 2nd ed. Thousand Oaks: Sage Publications, 95-118.

Billinghurst, M., Grasset, R., and Looser, J., 2005. Designing augmented reality interfaces. In: *Proceedings of the International ACM SIGCHI Conference on Computer Graphics – Learning through computer generated visualisation, 39 (1), 17 – 22.*

Bimber, O. and Raskar, R., 2005. *Spatial Augmented Reality: Merging Real and Virtual Worlds*. Massachusetts: A.K.Peters.

Blajenkova, O., Motes, M. and Kozhevnikov, M., 2005. Individual differences in the representations of novel environments. *Journal of Environmental Psychology*, 25(1), 97–109.

Bluestein, N. and Acredolo, L., 1979. Developmental changes in map-reading skills. *Child Development*, 50, 691–697.

Boulding, K., Downs, R. and Stea, D., 2005. *Image and Environment: Cognitive Mapping and Spatial Behavior*. New York: Transaction Publishers.

Bowman, D., North, C., Chen, J., Polys, N., Pyla, P., Yilmaz, U., 2003. Information-rich virtual environments: theory, tools, and research agenda. In: *Proceedings of the ACM symposium on Virtual reality software and technology (VRST '03)*. ACM, New York, NY, USA, 81-90.

Bowser, A., Hansen, D., Raphael, J., Reid, M., Gamett, R., He, Y., Rotman, D. and Preece, J., 2013. Prototyping in PLACE: a new scalable approach to developing location-based apps and games. *In: Proceedings of the SIGCHI Conference on Human Factors in Computing*, Paris, France, ACM, 1519-1528.

Bradley, N. and Dunlop, M., 2005. Towards a multidisciplinary model of context to support context-aware computing. *International Journal of Human-Computer Interaction*, 20 (4), 403-466.

Brimicombe, A. and Li, C., 2010. *Location-Based Services and Geo-Information Engineering*. Chichester, West Sussex: John Wiley and Sons Ltd.

Brown, B. and Chalmers, M., 2003. Tourism and mobile technology. In: Kuutti, K. & Karsten, H., eds. *Proceedings of the 8th European Conference on Computer Supported Cooperative Work*, Helsinki, Finland, Kluwer Academic Press.

Bruyas, M., Le Breton, B. and Pauzie, A., 1998. Ergonomic guidelines for the design of pictorial information. *International Journal of Industrial Ergonomics*, 21(5), 407–413.

Buhalis, D. and Law, R., 2008. Progress in information technology and tourism management: 20 years on and 10 years after the Internet - the state of eTourism research. *Tourism Management*, 29 (4), 609-623.

Burrell, G. and Morgan, G, 1979. *Sociological Paradigms and Organisational Analysis*. London: Heinemann.

Bush, V., 1945. *As We May Think*. The Atlantic Monthly, 176, 101–108.

Byrd, T. A., Cossick, K. L. and Zmud, R. W., 1992. A synthesis of research on requirements analysis and knowledge acquisition techniques. *MIS Quarterly*, 16(1), 117-138.

Cabral, D., Orso, Y., El-Khouri, Y., Belio, M., Gamberini, L. and Jacucci, G., 2014. The role of location-based event browsers in collaborative behaviours: an explorative study. *In: Proceedings from the Nordic Conference on Human-Computer Interaction (NordiCHI'04)*, Gothenburg, Sweden, ACM.

Caduff, D. and Timpf, S., 2008. On the assessment of landmark salience for human navigation. *Cognitive Processing*, 9(4), 249–67.

Card, S., Thomas, P., and Newell, A., 1983. *The Psychology of Human Computer Interaction*. New York: Lawrence Erlbaum Associates.

Carroll, J., 2000. *Making use: Scenario-Based Design of Human-Computer Interactions*. Massachusetts: MIT Press.

Case, D., 2012. *Looking for Information: A Survey of Research on Information Seeking, Needs, and Behaviour*. Bingley: Emerald Group Publishing Limited.

Chalmers, D. and Knight, R., 1985. The reliability of ratings of the familiarity of environmental stimuli: A generalizability analysis. *Environment and Behavior*, 17(2), 223–238.

Chang, K., Chang, C., Hou, H., Sung, Y., Chao, H. and Lee, C., 2014. Development and behavioural pattern analysis of a mobile guide system with augmented reality for painting appreciation instruction in an art museum. *Computers and Education*, 71, 185-197.

Cheng, K. and Tsai, C., 2013. Affordances of Augmented Reality in science learning: suggestions for future research. *Journal of Scientific Education & Technology*, 22 (2), 449 – 462.

Cheverst, K., Davis, N., Mitchell, K. and Smith, P., 2000. Providing tailored (context-aware) information to city visitors. *Adaptive Hypermedia and Adaptive Systems*, 1892, 73–85.

Cheverst, K., Davis, N., Mitchell, K., Friday, A. and Efstratiou, C., 1998. Developing a context-aware electronic tourist guide: some issues and experiences. *In: Proceedings of the SIGCHI conference on Human Factors in Computing Systems (CHI'00)*, The Hague, Netherlands, ACM, 17-24.

Cheverst, K., Mitchell, K. and Davies, N., 2002. Exploring context-aware information push. *Personal and Ubiquitous Computing*, 6 (4), 276-281.

Chipman, S., Schraagen, J. and Shalin, V., 2000. Introduction to cognitive task analysis. *In:* Schraagen, J., Chipman, S. and Shalin, V., eds., *Cognitive Task Analysis*, New York: Lawrence Erlbaum Associates, 3–24.

Choi, J., Jang and B., Kim, 2010. Organizing and presenting geospatial tags in location-based augmented reality. *Personal and Ubiquitous Computing*, 15 (6), 641-647.

Choudary, O., Charvillat, R., Grigoras, R. and Gurdjos, P., 2009. MARCH: mobile augmented reality for cultural heritage. *In: Proceedings of the 17th ACM International Conference on Multimedia*, Beijing, China, ACM, 1023-1024.

Christensen, J., Marks, J. and Shieber, S., 1992. Labelling point features on maps and diagrams. Center for Research in Computing Technology, Harvard University, TR-25-92

Christensen, J., Sussman, S., Levvy, S., Bennett, W., Wolf, T. and Kellogg, W., 2006. Too much information. *Queue*, 4 (6), 50-57.

Chua, A., Balkunje, R., and Goh, H.-L., 2011. The influence of user context on mobile information needs. *In: Proceedings from the International Conference on Advanced Information Networking and Applications*, IEEE, 721-726.

Church, K. and Smyth, B., 2009. Understanding the intent behind mobile information needs. *In: Proceedings of the 14th international conference on Intelligent user interfaces*, ACM, 247–256.

Cockton, G., Woolrych, A. and Lavery, D., 2009. Inspection-based evaluations. *In:* Sears, A. and Jacko, J., eds., *Human-Computer Interaction Handbook: Development Process*, Boca Raton, FL: Taylor and Francis Group.

Cooper, A. and Reimann, R., 2014. *About Face: The Essential of Interaction Design*. 4th ed., Indianapolis, IN, USA: John Wiley and Sons.

Craik, K. and Appleyard, D., 1980. Streets of San Francisco: Brunswik's lens model applied to urban inference and assessment. *Journal of Social Issues*, 36, 72–85.

Crandall, B., Klein, G. and Hoffman, R., 2006. *Working Minds: A Practitioner's Guide to Cognitive Task Analysis*, Cambridge, MA: MIT Press.

Creswell, J. and Plano-Clark, V., 2011. *Designing and Conducting Mixed Methods Research*. 2nd ed., Thousand Oaks, CA, USA: SAGE Publications.

Creswell, J., 2012. *Qualitative Inquiry and Research Design: Choosing Among Five Approaches*, Los Angelis: SAGE Publications.

Creswell, J., 2013. *Research Design: Qualitative, Quantitative and Mixed Methods Approaches*. Thousand Oaks, CA, USA: SAGE Publications.

Cynapse, 2014. Available http://cynapse.com/localscope/. [Accessed 10 January 2014]

Daft, R. and Lengel, R., 1986. Organizational information requirements, media richness and structural design. *Management Science*, 32(5), 554–571.

Davies, C., Li, C. and Albrecht, J., 2010. Human understanding of space. *In:* Hacklay, M. and Nivala, A., eds., *Interacting with Geospatial Technologies*, Chichester: Wiley-Blackwell, 19-35.

Delikostidis, I. and van Elzakker, C. P. J., 2009. Geo-identification and pedestrian navigation with geo-mobile applications. *In:* Gartner, G. and Rehl, K., eds. *Location-based Services and Telecartography II*. Berlin: Springer, 185-206.

Dervin, B., 1984. The human side of information: perspectives for communication, Norwood, N.J: Ablex.

Dewey, J., 1929. *Experience and Nature*. New York: W. W. Norton & Company, Inc.

Dey, A., 2001. Understanding and using context. *Personal and Ubiquitous Computing*, 5 (1), 12-34.

Dickinson, J., Ghali, K., Cherrett, T., Speed, C., Davies, N. and Norgate, S., 2013. Tourism and the smartphone app: capabilities, emerging practice and scope in the travel domain. *Current Issues in Tourism*, 17 (1), 84-101.

Dijk, S., 2002. Towards an evaluation of quality for label placement methods. *International Journal of Geographical Information Science,* 16(7), 641-661.

Dix A., Finlay J. E., Abowd G. D. and Beale R., 1998. *Human-Computer Interaction*. 2nd ed. London: Prentice Hall Europe

Dourish, P., 2004. What we talk about when we talk about context. *Personal and Ubiquitous Computing*, 8 (1), 19-30.

Downs, R. and Stea, D., 1973. *Image and Environment*, Chicago: Aldine.

Duncan, J.S., 1973. Landscape Taste as a Symbol of Group Identity: A Westchester County Village. *Geographical Review*, 63, 334–355.

Dunser, A., Grasset, R. and Billinghurst, M., 2008. A survey of evaluation techniques used in augmented reality studies. *In: Proceedings from SIGGRAPH Asia*, ACM Press.

Elias, B. and Paelke, V., 2008. User-centered design of landmark visualizations. *In:* Meng, L., Zipf, A. and Reichenbacher, T., eds. *Map-based Mobile Services: Theories, Practices and Implementations*. Berlin: Springer-Verlag, 33–56.

Emmanoulidis, C., Koutsiamanis, R. and Tasidou, A., 2013. Mobile guides: taxonomy of architectures, context awareness, technologies and applications. *Journal of Network and Computer Applications*, 36 (1), 103-125.

Endsley, M., 2000. Theoretical underpinnings of situation awareness: a critical review. *In:* Endsley, M. and Garland, D., eds., *Situation Awareness Analysis and Measurement*. Mahwah: Lawrence Erlbaum Associates. 3 -28.

eTips, 2014. Available https://www.etips.com/. [Accessed 10 January 2014]

Evans, G., 1980. Environmental cognition. *Psychological Bulletin*, 88, 259–287.

Fedosov, A. and Misslinger, S., 2014. Location based experience design for mobile augmented reality. *In: Proceedings of the 2014 ACM SIGCHI symposium on Engineering interactive computing systems*, Rome, Italy, ACM, 185-188.

Feiner, S., MacIntyre, S., Höllerer, T., Webster, A., 1997. A touring machine: prototyping 3D mobile augmented reality systems for exploring the urban environment. *Personal and Ubiquitous Computing*, 1 (4), 208-217.

Fesenmaier, D. and Jeng, J., 2000. Assessing structure in the pleasure trip planning process. *Tourism Analysis*, 5 (1), 13-27.

Fesenmaier, D. and Vogt, C., 2008. Evaluating the utility of touristic information sources for planning Midwest vacation travel. *Journal of Travel & Tourism Marketing*, 1 (2), 1-18.

Field, A., 2013. *Discovering Statistics Using IBM SPSS Statistics*. 4th ed., Los Angelis: SAGE Publications.

Fitzmaurice, G., 1993. Situated information spaces: spatially aware palmtop computers. *Communication of the ACM*, 36 (7), 38-49.

Fletcher, J., Fyall, A., Gilbert, D., 2013. *Tourism Principles and Practice*. Harlow: Pearson Education Limited.

Fling, B., 2009. *Mobile Design and Development*. Boston: O'Reilly.

Fodness, D. and Murray, B., 1997. Tourist information search. *Annals of Tourism Research*, 24 (3), 503-523.

Folstad, A., Anda, B. and Sjoberg, D., 2012. The usability inspection performance of work-domain experts: an empirical study. *Interacting with Computers*, 22 (2), 75-87.

Foursquare, 2014. Available https://foursquare.com/. [Accessed 10 January 2014].

Fremlin, G. and Robinson, A., 1998. Nine short essays about maps. *Cartographica*, 35 (1-2), 1-12.

Fröhlich, P., Simon, R. and Baillie, L., 2008. Mobile Spatial Interaction. *Personal and Ubiquitous Computing*, 13(1), 251-253.

Fröhlich, P., Simon, R., Baillie, L. and Anegg, H., 2006. Comparing conceptual designs for mobile access to geo-spatial information. *In: Proceedings from the 8th*

*International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI'06)*, Espoo, Finland, ACM, 109-112.

Fuchs, H., Livingston, M., Raskar, R., Colucci, D., Keller, K., State, A., Crawford, J., Rademacher, H., Samuel, H., Anthony, D. and Meyer., A., 1998. Augmented reality visualization for laparoscopic surgery. *Medical Image Computing and Computer Aided Intervention*, 1496, 934-943.

Gabbard, J. and Swan II, J., 2008. Usability engineering for Augmented Reality: employing user-based studies to inform design. *IEEE Transactions on Visualization and Computer Graphics*, 14 (3), 513-525.

Gabbard, J., Swan II, J. E., Hix, D., Kim, S.-J., Fitch, G., 2007. Active text drawing styles for outdoor augmented reality: a user-based study and design implications. *In: IEEE Proceedings of Virtual Reality Conference*, Blacksburg, USA, 35-42.

Ganapathy, S., Anderson, G. and Kozintzev, I. 2011. Empirical evaluation of augmented information presentation on small form factors – navigation assistant scenario. *In: Proceedings of the International Symposium on Virtual Reality Innovation*. Singapore, 1-12.

Gardiner, K., Yin., and Carswell, J., 2009. EgoViz – a mobile based spatial interaction system. *In:* Carswell, J., Fotheringham, S. and McArdle, G., eds., *Web and Wireless Geographical Information Systems*, 5886, Springer, 135-152.

Garrett, J., 2011. *Elements of User Experience: User-Centred Design for Web*. Chicago, USA: New Riders.

Gavalas, D., Konstantopoulos, C., Mastakas, K. and Pantziou, G., 2014. Mobile recommender systems in tourism. *Journal of Network and Computer Applications*, 39, 319-333.

Gaver, W., 1991. Technology affordances. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, New Orleans, LU, 79-84.

Geiger, P., Schichler, M., Pryss, R., Schobel, J. and Reichert, M., 2014. Location-based mobile augmented reality applications: challenges, examples and lessons learned. *In: Proceedings of the 10th International Conference on Web Information Systems and Technologies (WEBIST'14)*, Barcelona, Spain.

Gibson, J., 1979. *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.

Goldkuhl, G., 2011. Pragmatism vs interpretivism in qualitative information systems research. *European Journal of Information Systems*, 21(2), 135-146.

Gorlenko, L. and Merrick, R., 2003. No wires attached: usability challenges in the connected mobile world. *IBM Systems Journal*, 42 (4), 639–651.

Götzelmann, T., Gotze, M., Ali, K., Hartmann, K. and Strothotte, T., 2007. Annotating images through adaptation: an integrated text authoring and illustration framework. *Journal of WSCG*, 15 (1-3), 115 – 122.

Gould, J., and Lewis, C., 1985. Designing for usability: key principles and what designers think. *Communications of the ACM*, 28 (3), 300-311.

Grasset, R., Langlotz, T., Kalkofen, D., Tatzgern, M. and Schmalstieg, D. 2012. Image-driven view management for augmented reality browsers. *In: Proceedings of the*

*International Symposium on Mixed and Augmented Reality (ISMAR'12)*. Atlanta, USA. IEEE Computer Society, 177–186.

Greenberg, S., 2001. Context as a dynamic construct. *Human Computer Interaction*, 16 (2), 257-268.

Gregor, S, 2009. Building theory in the sciences of the artificial. *In: Proceedings of the 4th International Conference on Design Science Research in Information Systems and Technology*, Philadelphia, USA.

Gregor, S. and Hevner, A., 2013. Positioning and presenting design science research for maximum impact. *MIS Quarterly*, 37 (2), 337-335.

Gregor, S. and Jones, D., 2007. The anatomy of a design theory. *Information Systems*, 8 (5), 312-335.

Gregor, S., 2006. The nature of design theory in Information Systems. *MIS Quarterly*, 30(3), 611-642.

Gregory, R. and Mautermann, J., 2014. Heuristic theorizing: proactively generating design theories. *Information Systems Research*, 25 (3), 639-653.

Grudin, J., 1992. Utility and usability: research issues and development contexts. *Interacting with Computers*, 4 (2), 209-217.

Grudin, J., 2012. A moving target – the evaluation of Human-Computer Interaction. *In*: Jacko, J., ed., *Human-Computer Interaction Handbook*, 3rd ed., New York: Taylor & Francis.

Gursoy, D. and McCleary, K., 2004. An integrative model of tourists' information search behavior. *Annals of Tourism Research*, 31 (2), 353–373.

Hac, A., 2014. *Mobile Telecommunications Protocols for Data Networks*. London: CreateSpace Independent Publishing.

Hackos, J. and Redish, J., 1998. *User and Task Analysis for Interface Design*. New York: John Wiley and Sons.

Haklay, M. and Nivala, A.-M., 2010. User-Centred Design. In: Haklay, M., ed., *Interacting with Geospatial Technologies*. Chichester, West Sussex: Wiley-Blackwell, 91-106.

Halvey, M., Keane, M. and Smyth, B., 2006. Time based patterns in mobile-internet surfing. *In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'06)*, Montreal, Canada, ACM, 31-34.

Hamai, K., 2001. Location information is critical to wireless. *Wireless Design & Development*, 9 (11), 34-35.

Hansen, F., 2006. Ubiquitous annotation systems: technologies and challenges. *In: Proceedings of the 17th Conference on Hypertext and Hypermedia (HYPERTEXT'06)*, ACM, 121-132.

Harrison, J. and Howard, W., 1972. The role of meaning in the urban image. *Environment And Behavior*, 4, 389–412.

Harrison, R., Flood, R. and Duce, D., 2013. Usability of mobile applications: literature review and rationale for a new usability model. *Journal of Interaction Science*, 1 (1), 1-16.

Harrison, S., Sengers, P., and Tatar, D., 2007. The three paradigms of HCI. In: In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'07)*, San Jose, California, USA, ACM Press, 1-18.

Hart, R. and Moore, G., 1973. The development of spatial cognition: A review. In: Downs, R. and Stea, D., eds. *Image and Environment: Cognitive mapping and spatial behaviour*. Chicago, USA: Aldine Transaction, pp. 246–288.

Hart, S. and Staveland, LE. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In: Hancock, P. and Meshkati, N., eds. *Human Mental Workload*, eds., Amsterdam: North Holland Press.

Hartmann, K., Gotzelmann, T., Ali, K. and Strothotte, T., 2005. Metrics for functional and aesthetic label layouts. *Smart Graphics*, 3638, 115–126.

Henrysson, A. and Ollila, M., 2004. UMAR - Ubiquitous mobile augmented reality. In: *Proceedings of the 3rd international conference on Mobile and Ubiquitous Multimedia (MUM'04)*, Maryland, USA, ACM, 41-45.

Heo, J., Ham, D., Park, S. Song, C. and Yoon, W., 2009. A framework for evaluating the usability of mobile phones based on multi-level, hierarchical model of usability factors. *Interacting with Computers*, 21 (4), 263-275.

Herbst, I., Braun, A.-K., McCall, R. and Broll, W., 2008. TimeWarp: interactive time travel with mobile mixed reality game. *In: Proceedings of the 10th International Conference on Human-Computer Interaction with Mobile devices and services*. Amsterdam, the Netherlands: ACM Press, 235-244.

Hevner, A., March, S., Park, J. and Ram, S., 2004. Design science in Information Systems research. *MIS Quarterly*, 28(1): 75-105.

Hinze, A. and Voisard, A., 2003. Location- and time-based information delivery in tourism. *In: Proceedings of the 8th International Symposium in Spatial and Temporal Databases*. Springer, 489-507.

Hinze, A., and Buchanan, G., 2005. Context-awareness in mobile tourism information systems: challenges for user interaction. *In: Proceedings of the Workshop on Context in Mobile HCI*, Salzburg, Austria, ACM Press, 257-266.

Hinze, A., Chang, C. and Nichols, D., 2010. Contextual queries express mobile information needs. *In: Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI'10)*. San Francisco, USA, ACM, 327–336.

Hirschheim, R. and Klein, H., 1989. Four paradigms of information systems development. *Communications of the ACM*, 32(10), 1199-1216.

Hirtle, S. and Jonides, H., 1985. Evidence of hierarchies in cognitive maps. *Memory & Cognition*, 13(3), 208–217.

Höllerer, T. and Feiner, S., 2004. Mobile Augmented Reality. In: H. Karimi and A. Hammad, eds., *Telegeoinformatics: Location-based Computing and Services*. New York: Taylor & Francis Books Ltd.

Holtzblatt, K., Wendell, J. and Wood, S., 2005. *Rapid Contextual Design: A How-to Guide to Key Techniques for User-Centered Design*. San Francisco: Morgan Kaufmann.

Höpken, W., Fuchs, M., Zanker, M. and Beer, T., 2010. Context-based adaptation of mobile applications in tourism. *Information Technology & Tourism*, 12 (2), 175-195.

Hua, H., 2014. Past and future of wearable augmented reality displays and their applications. In: Barret, H., Greivenkamp, J. and Dereniak, E., eds., *Proceedings of SPIE 9186, Fifty Years of Optical Sciences at the University of Arizona*, San Diego, CA, USA, SPIE Digital Library.

IBM, 2014. Available http://www-01.ibm.com/software/uk/analytics/spss/ [Accessed 13 September 2014].

Imhof, E., 1975. Positioning names on maps. *The American Cartographer, 2(2), 128–144.*

Ingwersen, P. and Jarvelin, K., 2005. *The Turn*. Dordrecht, The Netherlands: Springer.

Intille, S., Bao, L., Tapia, E. and Rondoni, J., 2004. Acquiring in-situ training data for context-aware ubiquitous computing applications. *In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'04)*, The Hague, Netherlands, ACM, 1-8.

Ishiguro, Y. and Rekimoto, J., 2011. Peripheral vision annotation: noninterference information presentation method for mobile Augmented Reality. *In: Proceedings of the 2nd Augmented Human International Conference.* Tokyo, Japan, ACM.

ISO 9241-210, 2010. Ergonomics of human-system interaction – Part 210: Human-centred design for interactive systems [Online]. Available http://www.iso.org/iso/catalogue_detail.htm?csnumber=52075. [Accessed 12 May 2012].

ISO 28564-1, 2010. Public Information Guidance Systems – Part 1: Design principles and element requirements for location plans, maps and diagrams [Online]. Available http://www.iso.org/iso/catalogue_detail.htm?csnumber=44762 [Accessed 12 May 2012].

Jakobsson, A., 2003. User requirements for mobile topographic maps [Online]. GiMoDig Project Deliverable D2.1.1. Available http://lib.tkk.fi/Diss/2006/isbn9512282062/article5.pdf

Jankowski, J., Samp, K., Irzynska, I., Jozwowicz, M. and Decker, S., 2010. Integrating text with video and 3D graphics: the effects of text drawing styles on text readability. *In: Proceedings of the 13th Conference on Human Facotrs in Computer Systems, (CHI'10)*, Atlanta, USA, 1321-1330.

John, B. and Packer, H., 1995. Learning and using the cognitive walk- through method: a case study approach. In: Katz, I., Mack, R., and Marks, L., eds., *Proceedings of ACM CHI'95 Conference on Human Factors in Computing Systems (CHI'95)*, ACM Press, 429–436.

Johnson-Laird, P. and Byrne, R., 1991. *Deduction*. Hove, UK: Lawrence Erbaum Associates.

Johnson, B., Anthony J. Onwuegbuzie, T. and Lisa A., 2007. Toward a definition of Mixed Methods research. *Journal of Mixed Methods Research*, 1 (2), 112-133

Jones, C., 2013. *Geographic Information Systems and Computer Cartography*. New York: Taylor and Franics.

Julier, S. and Rosenblum, D., 2000. BARS: Battlefield augmented reality system. In: *Proceedings of the NATO Symposium on Information Processing Techniques for Military Systems*, Istanbul, Turkey.

Jumisko- Pyykkö, S. and Utriainen, T., 2011. A hybrid method for quality evaluation in the context of use for mobile (3D) television. *Multimedia Tools and Applications*, 55, 185–225.

Jumisko-Pyykkö, S. and Vainio, T., 2010. Framing the Context of Use for Mobile HCI. *International Journal of Mobile Human Computer Interaction*, 2 (4), 1-28.

Junaio, 2014. Available https://www.junaio.com/. [Accessed 10 January 2014].

Juniper Research, 2014. Press Release: Mobile Augmented Reality revenues to exceed $1bn annually by 2015 [Online]. Available http://www.juniperresearch.com/viewpressrelease.php?pr=427 . [Accessed 5 September 2014].

Kaasinen, E., 2005. User acceptance of location-aware mobile guides based on seven field studies. *Behaviour and Information Technology*, 24 (1), 37–49.

Kah, J., Vogt, C. and MacKay, K., 2011. Place-based information technology on vacations. *Tourism Geographies*, 13 (2), 209-233.

Kalkofen, D., Mendez, E. and Schmalstieg, D., 2009. Comprehensible visualization for augmented reality. *IEEE Transactions on Visualization and Computer Graphics*, 15 (2), 193-204.

Kamar, A., 2003. A mobile tour guide (m-ToGuide). Deliverable 1.4, Project Final Report, IST-2001-36004.

Karat, J., 1997. Evolving the scope of user-centered design. *Communications of the ACM*, 40(7), 33-38.

Kässi, J., Krause, C. M., Kovanen, J. & Sarjakoski, L. T. (2013). Effects of positioning aids on understanding the relationship between a mobile map and the environment. *Human Technology*, 9 (1), 92-108.

Keil, J., Michael, Z. and Wuest, H. 2011. The house of Olbrich – and augmented reality tour through architectural history. In: *Proceedings of the 10th IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, Basel, Switzerland. IEEE Computer Society, 15-18.

Kenteris, M., Gavalas, D. and Economou, D., 2006. A novel method for the development of personalized mobile tourist applications. In: P. Salvador, ed., *Proceedings of the 5th IASTED International Conference on Communication Systems and Networks (CSN'2006)*, Palma de Mallorca, Spain, ACTA Press, 208-212.

Kim, J.-B. and Park, C., 2011. Development of mobile AR tour application for the National Palace Museum of Korea. In: R. Schumaker, ed., *Virtual and Mixed Reality*, Part I, HCII 2011. Springer-Verlag, 55-60.

Kim, S., Moon, S., Park, J., Han, S., 2009. Efficient annotation visualization using distinctive features. In: *Proceedings from the Symposium on Human Interface*, San Diego, CA, USA, Springer, 295-303.

Kitchin, R., 2014. *The Data Revolution: Big Data Open Data, Data Infrastructures and Their Consequences*. Los Angelis: SAGE Publications.

Kjeldskov, J. and Graham, C., 2003. A review of mobile HCI research methods. In: *Proceedings of the 5th International Symposium on Human-Computer Interaction with Mobile Devices and Services (MobileHCI'03)*, ACM, 317-335.

Kjeldskov, J. and Paay, J., 2010. Indexicallity: understanding mobile human-computer interaction in context. *ACM Transactions on Human-Computer Interaction*, 17 (4), 1-28.

Kjeldskov, J. and Paay, J., 2012. A longitudinal review of Mobile HCI research methods. In: *Proceedings of the 14th international conference on Human-computer interaction with mobile devices and services*, San Francisco, USA, ACM, 69-78.

Kjeldskov, J. and Skov, M., 2014. Was it worth the hassle?: ten years of mobile HCI research discussions on lab and field evaluations. In: *Proceedings from the 16th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI'14)*, Toronto, Canada, ACM, 43-52.

Kjeldskov, J., 2003. Lessons from being there: interface design for mobile augmented reality. In: Anderson, P. and L. Qvortrup, eds., *Virtual Applications: Applications With Virtual Inhabited 3D Worlds*, Berlin: Springer-Verlag, 159-188.

Kling, B., 1977. The Organizational Context of Software Designs. *MIS Quarterly*, 1(4), 41–52.

Koeppel, I., 2000. What are location-based services from a GIS perspective [Online]. Available: itu.dk/stud/speciale/ckt/.../What%20are%20Loaction%20Services.doc [Accessed 14 May 2013].

Költringer, C. and Wöber, K., 2010. Information needs of city travellers. In: Mazanec, J. and Wöber, K., eds. *Analysing International City Tourism.* New York: Springer, 109-126.

Kooper, R. and MacIntyre, B., 2003. Browsing the real world-wide web: maintaining awareness of virtual information in an AR information space. *International Journal of Human Computer Interaction*, 16 (3), 425-446.

Kourouthanassis, P., Boletsis, K., Bardaki, C. and Chasanidou, D., 2014. Tourist responses to mobile augmented reality travel guides: the role of emotions on adoption behaviour. *Pervasive and Mobile Computing*, [In Press], DOI: http://dx.doi.org/10.1016/j.pmcj.2014.08.009

Kraak, M.J. and Ormeling, F., 2010. *Cartography: Visualization of Geospatial Data*. Harlow: Pearson Education Press.

Kramer, R., Modsching, M., Ten Hagen, K., Gretzel, U., 2006. A field trial to elicit individual preferences in the context of a mobile Dynamic Tour Guide. In: Hitz, M., Sigala, M. and Murphy, J., eds., *Information and Communication Technologies in Tourism*, 100-111.

Krikelas, J., 1983. Information-seeking behavior: patterns and concepts. *Drexel Library Quarterly*, 19, 5–20.

Krogstie, J., Lyytinen, K., Opdahl, A., Pernici, B., Siau, K. and Smolander, K, 2003. Mobile Information Systems - research challenges on the conceptual and logical level. In: Olive, A., Yoshikawa, M. & Yu, E., eds. *Advanced Conceptual Modelling Techniques*. New York: Springer, 124–135.

Kruijff, E., Swan, J. and Feiner, S., 2010. Perceptual issues in Augmented Reality revisited. In: *Proceedings of the 9th IEEE International Symposium on Mixed and Augmented Reality (ISMAR2010)*, Seoul, South Korea, 3-12.

Krumm, J., Davies, N. and Narayanaswami, C. 2008. User-generated content. Pervasive Computing, 7 (4), 10-11.

Kuechler, W. and Vaishnavi, V., 2008. On theory development in design science research: anatomy of a research project. *European Journal of Information Systems*, 17(5), 1-23.

Kuipers, B., 1982. The "map in the head" metaphor. *Environment and Behavior*, 14, 202–220.

Kukka, H., Kostakos, V., Ojala, T., Ylipulli, J., Suopajarvi, T., Jurmu, M. and Hosio, S., 2011. This is not classified: everyday information seeking and encountering in smart urban spaces. *Personal and Ubiquitous Computing*, 17 (1), 15 – 27.

Kumar, R., 2014. *Research Methodology: A Step-By-Step Guide for Beginners*. 3rd ed., Los Angelis, USA: SAGE Publications.

Langlotz, T., Grubert, J. and Grasset, R., 2013. Augmented reality browsers. *Communications of the ACM,* 56 (11), 4–36.

Langlotz, T., Nguyen, T., Schmalstieg, D. and Grasset, R., 2014. Next-generation Augmented Reality browsers: rich, seamless and adaptive. *Proceedings of the IEEE*, 102 (2), 155-169.

Layar, 2014. Available https://www.layar.com/. [Accessed 10 January 2014].

Lazar, J., Feng, J. and Hochheiser, H., 2009. *Research Methods in Human-Computer Interaction*. Chichester: John Wiley and Sons Ltd.

Lechner, M., 2013. ARML 2.0 in the context of existing AR data formats. In: *Proceedings of the 6th Workshop on Software Engineering and Architectures for Realtime Interactive Systems (SEARIS)*, Orlando, Florida, ACM, 41-47.

Levine, M., 1982. You-are-here maps: psychological considerations. *Environment and Behavior*, 14(2), 221–237.

Lewis, C. and Wharton, C., 1997. Cognitive walkthroughs. In: Helander, M., Landauer, T. and Prabhu, P., eds., *Handbook of Human- Computer Interaction*, 2nd ed, Elsevier, 717–732.

Leykin, A. and Tuceryan, 2004. Automatic determination of text readability over textured backgrounds for Augmented Reality systems. In: *Proceedings of the 3rd IEEE and ACM Symposium on Mixed and Augmented Reality (ISMAR'04)*, Arlington, USA, ACM Press, 224-230.

Linaza, M., Marimon, D., Carrasco, P., Alvarez, C., Montesa, J., Aguilar, J. and Diez, G. 2012. Evaluation of mobile augmented reality applications for tourist destinations. In: M. Fuchs, F. Ricci, & L. Cantoni, eds. *Information and Communication Technologies in Tourism, Proceedings of the International Conference on Information and Communication Technologies in Tourism ENTER2012*. Helsingborg, Sweden. Springer, 260–271.

Livingston, D., 2001. Adult's informal learning: definitions, findings, gaps and future research. In: *WALL Working Papers*, 21, 1-49.

Livingston, M. 2013. Issues in human factors evaluations of augmented reality systems. In: Huang, W., Alem, L. & Livingston, M., eds. *Human Factors in Augmented Reality Environments*. New York: Springer, 3-10.

Lobben, A., 2004. tasks, strategies, and cognitive orocesses associated with navigational map reading: a review perspective. *The Professional Geographer*, 56 (2), 270–281.

Long, S., Kooper, R., Abowd, G. and Atkeson, C., 1996. Rapid prototyping of mobile context-aware application: the cyberguide case study. In: *Proceedings of the 2nd ACM International Conference on Mobile Computing And Networking (MobiCom)*, Las Vegas, Nevada, USA.

Longley, P., Goodchild, P., Maguire, D. and Rhind, D., 2010. *Geographic Information Systems and Science*. Danvers, USA: John Wiley & Sons.

Looije, R., te Brake, G. and Neerincx, M., 2007. Usability engineering for mobile maps. In: *Proceedings of the 4th international conference on mobile technology, applications, and systems*. Singapore, 532-539.

Luley, P., Perko, R., Weinzerl, J., Paletta, L. and Almer, A., 2011. Mobile Augmented Reality for Tourists – MARFT. In: G. Gartner and F. Ortag, eds, *Advances in Location-Based Services*, Spinger, 24-37.

Lynch, D., 1960. The Image of the City, MA: MIT Press.

Maass, S. and Döllner, J., 2006. Efficient view nanagement for dynamic annotation placement in virtual landscapes. In: Butz, A., Fisher, A. and Olivier, P., eds, *Smart Graphics*, 4073, 1–12.

MacKenzie, S., 2003. Motor behaviour models for human-computer interaction. In: Carroll, M. (ed.) *HCI models, theories, and frameworks: Toward a multidisciplinary science*, pp. 27-54. San Francisco: Morgan Kaufmann.

Madden, L., 2011. *Professional Augmented Reality Browsers for Smartphones*. Chichester, UK: John Wiley and Sons Ltd.

Magllwo, J., Cohen, R., Allen, G. and Rodrig, J., 1996. The impact of wayfinders goal on learning a new environment: different types of spatial knowledge as goals. *Journal of Environmental Psychology*, 15 (1), 65-75.

Maguire, M., 2001. Methods to support human-centred design. *International Journal of Human-Computer Interaction*, 55 (4), 587-634.

Makita, K., Kanbara, M. and Yokoya, N., 2009. View management of annotations for wearable Augmented Reality. *In: Proceedings of the IEEE International Conference on Multimedia and Expo*, ICME'09, 982-985.

Marimon, D., Sarasua, C., Carrasco, P., Álvarez, R., Montesa, J., Adamek, T., and Romero, I., 2009. MobiAR: Tourist Experiences through Mobile Augmented Reality. In: *Proceedings from the International Summit on Network and Electronic Media (NEM),* Barcelona, Spain.

Markus, M., Majchrzak, A., Gasser, L., 2002. A design theory for systems that support Emergent Knowledge Processes. *MIS Quarterly*, 26 (3), 179-212.

Martin, D., Alzua-Sorzabal, A. and Lamsfus, C., 2011. A contextual geofencing mobile tourism service. In: Law, R., Fuchs, F. and Ricci, F., eds, *Proceedings of the International Conference on Information and Communication Technologies in Tourism, ENTER2011*. Innsbruck, Austria: Springer.

Matlin, M., 2013. *Cognitive Psychology*, 8th ed., New York: John Wiley and Sons.

Matsuda, K., 2010. Domesti/city [Online]. Available at: http://keiichimatsuda.com/kmatsuda_domesti-city.pdf.

Matthews, T., Pierce, J. and Tang, J., 2009. No smartphone is an island: the impact of places, situations and other devices on smart phone use. *IBM Research Report*, RJ10452, 1-10.

Maxwell, J. A. and Loomis, D., 2003. Mixed method design: an alternative approach. In: Tashakkori, A. and Teddlie, C., eds., *Mixed Methods in Social & Behavioural Research*. Thousand Oaks: SAGE Publications.

Mayhew, D., 1999. *The Usability Engineering Lifecycle*. San Francisco: Morgan Kaufmann Publishers.

McFedries, P., 2014. The city as computer. *IEEE Spectrum*, 51 (2), 28.

McGenere, J. and Ho, W., 2000. Affordances: clarifying and evolving a concept. In: *Proceedings of the Graphics Interface*, Montreal, Canada.

Melles, G., 2008. New pragmatism and the vocabulary and metaphors of scholarly design research. *Design Issues*, 24 (4), 88-101.

Mendez, E., 2010. On the usage of context for Augmented Realiy visualization [Online]. PhD Thesis, Institute for Computer Graphics, Gratz University of Technology, Available http://www.icg.tugraz.at/Members/mendez/Pueblications/on-the-usage-of-context-for-augmented-reality-visualization [Accessed 4 February 2012].

Meng, L., 2008. The state of the art of map-based mobile services. In: Meng, L., Zipf, A. and Reichenbacher, T., eds. *Map-based Mobile Services: Theories, Practices and Implementations*. Berlin: Springer-Verlag.

Michael, K. and Michael, M., 2011. The social and behavioural implications of location-based services. *Journal of Location-Based Services*, 5(3-4), 121-137.

Milgram, P., Takemura, H., Utsumi, A. and Kishino, F., 1994. Augmented Reality: a call for displays on the Reality-Virtuality continuum. SPIE, *Telemanipulator and Telepresence Technologies*, 2351, 282-292.

Mizzaro, S., 1997. Relevance: the whole history. *Journal of the American Society for Information Science*, 48 (9), 810-832.

Möhring, M., Lessig, C. and Bimber, O., 2004. Video see-through AR on consumer Cell-Phones. In: *Proceedings of the 3rd IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR '04)*, Arlington, USA, ACM, 252-253.

MySwitzerland, 2014. Available http://www.myswitzerland.com/. [Accessed 10 January 2014].

Nasar, J., Stamps, A. and Hanyu, K., 2005. Form and function in public buildings. *Journal of Environmental Psychology*, 25 (2),159–165.

Neil, T., 2014. *Mobile Design Pattern Gallery: UI Patterns for Smartphone Apps*. North Sebastopol, CA, USA: O'Reilly Media.

Neuhofer, B., Buhalis, D. and Ladkin, A., 2013. A typology of technology enhanced tourism experiences. *International Journal of Tourism Research*, 6 (4), 340-350.

Nielsen, J. and Budiu, R., 2013. *Mobile Usability*. Berkley: New Riders.

Nielsen, J. and Mack, R., 1994. *Usability Inspection Methods*. Los Angelis: John Wiley & Sons.

Nielsen, J., 1993. *Usability Engineering*. Academic Press, London, U.K.

Nieminen, M., Mannonen, P. and Turkki, L., 2004. User-centered concept development process for emerging technologies. In: *Proceedings from the 3rd Nordic Conference on Human-Computer Interaction (NordiCHI'04)*, Tampere, Finland, ACM, 225–228.

Norman, D. and Draper, S., 1986. *User Centered System Design*. Hillsdale, New Jersey: Lawrance Erlbaum Associates.

Norman, D., 1983. Some observations on mental models. In: Gentner, D. and Stevens, A., eds, *Mental Models*, LEA Publishing, 7–14.

Norman, D., 1988. *The Design of Everyday Things*. Cambridge, MA: MIT Press.

Norman, D., 1999. Affordances, conventions and design. *Interactions*, 6 (3), 38-43.

Nunamaker, J. and Chen, M., 1991. Systems development in Information Systems research. *Journal of Management Information Systems*, 7 (3), 89-106.

O'Grady, M., O'Hare, G. and Sas, C., 2005. Mobile agents for mobile tourism: a user evaluation of Gulliver's Genie. *Interacting with Computers*, 17 (4), 343-366.

O'Hara, K., Mitchell, A. and Vorbau, A., 2007. Consuming video on mobile devices. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'07)*, San Jose, California, USA, ACM Press, 857-866.

Oertel, B., Steinmuller, K. and Kuom, M., 2002. Mobile Multimedia Services for Tourism. In: Woeber, K., Frew, A. and Hitz, M., eds. *Proceedings of the International Conference on Information and Communication Technologies in Tourism, ENTER2002*. Innsbruck, Austria. Springer, 265-274.

Olsson, T. and Salo, M., 2011. Online user survey on current mobile augmented reality applications. In: *Proceedings of the 10th IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. Basel, Switzerland. IEEE Computer Society, 75–84.

Olsson, T., Ihamäki, P., Lagerstam, E., Ventä-Olkkonen, L. and Väänänen-Vainio-Mattila, K., 2009. User expectations for mobile mixed reality services: an initial user study. In: *Proceedings from the European Conference on Cognitive Ergonomics: Designing beyond the Product*, ACM Press, Helsinki, Finland.

Olsson, T., Lagerstam, E., Karkkainen, T. and Vaananen-Vainio-Mattila, 2013. Expected user experience of mobile augmented reality services: a user study in the context of shopping centres. *Pervasive and Ubiquitous Computing*, 17 (2), 287-304.

Orr, G., 2010. A review of literature in mobile learning: affordances and constraints. In: *Proceedings of the 6th IEEE International Conference on Wireless, Mobile, and Ubiquitous Technologies in Education*, IEEE Computer Society, 107 – 111.

Osviannikov, I., Arbib, M. and Mcneil, T., 1999. Annotation technology. International *Journal of Human Computer Studies*, 50, 329–362.

Ottoson, T., 1987. *Map-Reading and Wayfinging*. Gottebord: Acta Universitatis Gothoburgensis.

Oulasvirta, A., 2012. Rethinking experimental designs for field evaluations. *Pervasive Computing*, 11 (4), 60-67.

Oulasvirta, A., Estlander, S. and Nurminen, A., 2009. Embodied interaction with a 3D versus 2D mobile map. *Personal and Ubiquitous Computing*, 13 (3), 303-320.

Oulasvirta, A., Tamminen, S. and Hook, K., 2005. Comparing two approaches to context: realism and constructivism. In: *Proceedings of the 4th Decennial Conference on Critical Computing: Between Sense and Sensibility*, ACM, 195-198.

Paay, J., Kjeldskov, J., Howard, S. and Dave, B., 2009. Out on the town: a socio-physical approach to the design of a context-aware urban guide. *ACM Transactions on Human Computer Interaction*, 16 (2), 701-734.

Pakanen, M., Huhtala, J. and Häkkilä, J., 2011. Location visualization in social media applications. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Vancouver, Cadana, ACM, 2439–2442.

Papagiannakis, G., Schertenleib, S., O'Kennedy, B., Arevalo-Poizat, M., Magnenat-Thalmann, N., Stoddart, A. and Thalmann, D., 2005. Mixing Virtual and Real scenes in the site of ancient Pompeii. *Computer Animation and Virtual Worlds*, 16 (1), 11-24.

Parker, C. and Tomitsch, M., 2014. Data visualization trends in mobile augmented reality applications. In: *Proceedings of the 7th International Symposium on Visual Information Communication and Interaction (VINCI'14)*, Sydney, Australia, 228-231.

Partala, T. and Salminen, M., 2012. User experience of photorealistic urban pedestrian navigation. In: *Proceedings of the International Working Conference on Advanced Visual Interfaces (AVI'12)*, Capri Island, Italy.

Pascoe, J., 1998. Adding generic contextual capabilities to wearable computers. In: *Proceedings from the Second International Symposium on Wearable Computers*, Pittsburg, USA, IEEE Press, 92-99.

Pearce, P., 2011. *Tourist Behaviour and the Contemporary World*. Bristol: Channel View Publications.

Pereira, F., Silva, D., Abreu, P. and Pinho, A., 2014. Augmented reality mobile tourism application. In: Rocha, A., ed., *New Perspectives in Information Systems and Technologies*, 2, 175-185.

Pine, B. J. and Gilmore, J. H., 1999. *The Experience Economy: Work is Theatre and Every Business a Stage*. Boston, MA: Harvard Business School Press.

Pirolli, P., 2007. *Information Foraging Theory: Adaptive Interaction with Information*. Oxford University Press, New York.

Polys, N., 2006. Display techniques in Information Rich Virtual Environments. PhD Dissertation, Virginia Polytechnic Insitute and State University, Blacksburg, Virginia, USA.

Presson, C. and Montello, D., 1988. Points of reference in spatial cognition: stalking the elusive landmark. *British Journal of Developmental Psychology*, 6, 378–381.

QSR International, 2014. Available http://www.qsrinternational.com/products_nvivo.aspx [Accessed 14 September 2014].

Raper, J., Gartner, G., Karimi, H. and Rizos, C., 2007. A critical evaluation of location based services and their potential. *Journal of Location Based Services*, 1 (1), 5–45.

Raper, J., Gartner, G., Karimi, H. and Rizos, C., 2011. Applications of location-based services: a selected review. *Journal of Location Based Services*, 1 (2), 89–111.

Raptis, D., Tselios, N. and Avouris, N., 2005. Context-based design of mobile applications for museums: a survey of existing practices. In: *Proceedings of the 7th International Symposium on Mobile Human Computer Interaction*, 153-160.

Raubal, M. and Winter, S., 2002. Enriching wayfinding instructions with local landmarks. In: *Proceedings of the Second International Conference on Geographic Information Science.* Boulder, CA, USA, 243–259.

Rehrl, K., Hausler, E., Leitinger, S. and Bell, D., 2014. Pedestrian navigation with augmented reality, voice and digital map: final results from in-situ study assessing performance of user experience. *Journal of Location Based Services*, 8 (2), 75-96.

Reichenbacher, T., 2004. Mobile Cartography – Adaptive Visualisation of Geographic Information on Mobile Devices. *PhD Thesis*, Munchen: Technischen Universität München.

Roehl, W. and Fesenmaier, D., 1992. Risk perceptions and pleasure travel: an exploratory analysis. *Journal of Travel Research*, 30 (4), 17-26.

Roode, D., 2003. Information Systems Research: a matter of choice. *South African Computer Journal*, 30 (1-2), 23-45.

Roschelle, J., 2003. Unlocking the learning value of wireless mobile devices. *Journal of Computer Assisted Learning,* 16, 260 – 272.

Roschelle, J., Tatar, D., Chaudhury, S., Dimitriadis, Y., Patton, C. and DiGiano, C., 2007. Ink, improvisation, and interactive engagement: learning with tablets. *Computer*, 40 (9), 42 – 48.

Rosenbaum, S. and Kantner, L., 2007. Field usability testing: method, not compromise. In: *Proceedings from the International Professional Communication Conference IPCC 2007*, Seattle, USA, 1-7.

Royse, D., 1969. *Social inferences via environmental cues*. Massachusetts Institute of Technology, Cambridge, MA.

Sanderson, P. and Fisher, C., 1994. Exploratory sequential data-analysis: foundations. *Human-Computer Interaction*, 9 (3-4), 251–317.

Sandor, C. and Klinker, G., 2009. Lessons learned in designing ubiquitous augmented reality user interfaces. In: Zaharis, P. and Ang, C., eds., *Human-Computer Interaction: Concepts, Methodologies, Tools and Applications*. Hershey: Information Science Reference, 629–644.

Sarjakoski, L. and Nivala, A., 2005. Adaptation to context – a way to improve the usability of mobile maps. In: Meng, D., Reichenbacher, T. and Zipf, A., eds, *Map-Based Mobile Services*, Berlin: Springer-Verlag, 107-123.

Schilit, B. and Theimer, M., 1994. Disseminating active map information to mobile hosts. IEEE Network, 3 (2), 22-32.

Schleicher, R., Westermann, T. and Reichmuth,R., 2014. Mobile Human-Computer Interaction. In: Möller, S. and Raake, A., eds, *Quality of Experience*, Springer, 339-349.

Schmidt, A., Beigl, M. and Gellersen, H.-W., 1999. There is more to context than location. *Computers and Graphics*, 23 (6), 893-901.

Schneiderman, B., Plaisant, S., Cohen, M. and Jacobs, S., 2013. *Designing the user interface: strategies for effective human-computer interaction*. Harlow, Essex: Pearson Education Limited.

Schneidermann, B., 2000. Universal Usability. *Communications of the ACM,* 43 (5), 84 – 91.

Schul, P. and Crompton, J., 1983. Search behaviour of international vacationers: travel-specific lifestyle and sociodemographic variables. *Journal of Travel Research*, 22 (2), 25-30.

Seo, B., Kim, K. and Park, J., 2011. Augmented reality-based on-site tour guide: a study in Gyengbokgung. In: Koch, R. And Huang, F., eds, *Proceedings from the Computer Vision – ACCV 2010 Workshops*, Springer, 6469, 276-285.

Simon, H., 1960. The Sciences of the Artificial. Cambridge, MA: MIT Press.

Shackel, B., 1981. The concept of usability. In: *Proceedings of IBM Software and Information Usability Symposium*, Poughkepsie, New York, USA, IBM Corporation.

Shackel, B., 2009. Usability – context, framework, definition, design and evaluation. *Interacting with Computers*, 21 (5-6), 339-346.

Shadish, W., Cook, T. and Campbell, D., 2002. *Experimental and Quasi- experimental Design for Generalized Causal Inference*. Boston: Houghton-Mifflin.

Shaner, J., 2013. Smartphones, tablets and GPS accuracy [Online]. ESRI ArcGIS Blog. Available: http://blogs.esri.com/esri/arcgis/2013/07/15/smartphones-tablets-and-gps-accuracy/ [Accessed 20 September 2013].

Sholl, M., 1987. Cognitive maps as orienting schemata. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 13 (4), 615-628.

Siegel, A. and White, S., 1975. The development of spatial representations of large-scale environments. *Advances in Child Development and Behaviour*, 10, 9–55.

Smith, A., 1997. *Human Computer Factors: A study of Users and Information Systems*. McGraw-Hill.

Speiginer, G. and MacIntyre, B., 2014. Ethereal: a toolkit for spatially adaptive augmented reality content. In: *Proceedings of the Adjunct Publication of the 27th annual ACM Symposium on User Interface Software and Technology*, Honolulu, Hawaii, ACM, 113-114.

Spink, A. and Cole, C., 2006. Human information behavior: integrating diverse approaches and information use. *Journal of the American Society of Information Science and Technology*, 57 (1), 25–35.

Spohrer, J., 1999. Information in places. *IBM Systems Journal*, 38(4), 602-628.

Spool, J., Rubin, J. and Chisnell, D., 2008. *Usability Testing: How to Plan, Design and Conduct Effective Tests*. Indianapolis: Wiley Publishing.

SPRXmobile, 2009. We launched Layar: the first augmented reality browser for mobile [Online]. Available http://www.sprxmobile.com/we-launched-layar-worlds-first-augmented-reality-browser-for-mobile/. [Accessed 14 September 2013].

Stern, E. and Leiser, D., 1988. Levels of Spatial Knowledge and Urban Travel Modeling. *Geographical Analysis*, 20 (2), 140–155.

Steward, S. and Vogt, C., 1999. A case-based approach to understanding vacation planning. *Leisure Sciences*, 21, 79-95.

Straub, D., 2012. Does MIS have native theories? *MIS Quarterly*, 36 (2), iii-xii.

Sullivan, L., 1918. *Kindergarten chats and other writings*, New York: Dover.

Swan II, J. and Gabbard, J., 2005. A survey of user-based experimentation in Augmented Reality. In: *1st International Conference on Virtual Reality (VR '05), HCI International*, 1-9.

Tamminen, S., Oulasvirta, A., Toiskallio, K. and Kankainen, A., 2004. Understanding mobile contexts. *Personal and Ubiquitous Computing*, 8 (2), 135-143.

Tan, E., Foo, S., Goh, D. and Theng, Y.-L., 2009. TILES: classifying contextual information for mobile tourism applications. In: *Aslib Proceedings: New Information Perspectives*, 565-586.

Teddlie, C. and Tashakkori, A., 2010. Overview of contemporary issues in mixed methods research. In: Tashakkori, A. and Teddlie, C., eds., *Mixed Methods in Social & Behavioural Research*. Thousand Oaks: SAGE Publications.

Thanedar, V. and Höllerer, T., 2004. Semi-automated placement of annotations in videos. Technical Report 2004 – 11, UC, Santa Barbara, USA.

Tiefenbacher, P., Lehment, N. and Rigoll, G., 2014. Augmented reality evaluations: a concept utilizing virtual reality. In: Shumaker, R. & Lackey, S., eds., *Virtual Augmented and Mixed Reality*, 8525, Springer, 226-236.

Toh, Y., Jeung, J., and Pan, Y., 2010. A combined user research process for designing mobile AR guide in cultural heritage. In: *Proceedings from the 9th IEEE International Symposium on Mixed and Augmented Reality (ISMAR'10)*, Seoul, South Korea, IEEE Press, 71-72.

Tokusho, Y. and Feiner, S., 2009. Prototyping an outdoor mobile Augmented Reality street view application. In: *Proceedings of ISMAR 2009 - Let's Go Out: Workshop on Outdoor Mixed and Augmented Reality*, Orlando, Florida.

Tolman, E., 1948. Cognitive maps in rats and men. *Psychological Review*, 55, 189–208.

Tough, A., 1979. The Adult's Learning Projects: A Fresh Approach to Theory and Practice in Adult Learning, Toronto: OISE Press.

Trauth, M. 2001. The choice of qualitative research methods in IS. In: Trauth, M. (ed), *Qualitative research in IS: Issues and trends*. Hershey: Idea Group.

Treisman, A. and Gelade, G., 1980. A feature integration theory of attention. *Cognitive Psychology*, 12 (1), 97-136.

Trevisan, D., Carvalho, F., Raposo, A., Freitas, C. and Nedel, L., 2011. Exploring the design of transitional hybrid user interface. *SBC Journal on 3D Interactive Systems*, 2 (1), 2-13.

Tsakonas, G. and Papatheodorou, C., 2006. Analysing and evaluating usefulness and usability in electronic information services. *Journal of Information Science*, 31 (5), 400-419.

Turunen, T., Pyssysalo, T. and Röning, J., 2010. Mobile AR requirements for location based social networks. *The International Journal of Virtual Reality*, 9 (4), 67–78.

Tversky, B., 2005. Functional significance of visuospatial representations. In: Shah, P. & Miyake, A., eds. *The Cambridge Handbook on Visuospatial Thinking*. Cambridge University Press.

Umlauft, M., Pospischil, G., Niklfeld, G., and Michlayr, E., 2003. Lol@, a mobile tourist guide for UMTS. PInformation *Technology and Tourism*, 5(3), 151-164.

UNWTO, 2007. Understanding Tourism: basic glossary [Online]. Available: [Accessed 12 September 2012].

Van Krevelen, D. and Poelman, R., 2010. A survey of augmented reality technologies, applications and limitations. *International Journal of Virtual Reality*, 9 (2), 1–20.

Van Someren, M., Barnard, Y. and Sandberg, J., 1994. *The Think Aloud Method: A practical guide to modelling cognitive processes*, London: Academic Press.

Villiers, M., 2005. Thee approaches as pillars for interpretive Information Systems research: development research, action research and grounded theory. In: *Proceedings of the 2005 Annual Research Conference of the South African Institute of Computer Scientists and Information Technologies on IT research in Developing Countries (SAICSIT), p. 1-10.*

Vincent, T., Nigay, L. and Kurata, T., 2012. Classifying handheld Augmented Reality: three categories linked by spatial mappings. *Paper presented at the workshop Classifying the AR Presentation Space at the International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 5, Atlanta, USA.

Virranteus, K., Markkula, J., Garmash, A., Terziyan, V., Veijalainen, J., Katanosov, A. and Tirri, H., 2001. Developing GIS-supported location-based services. In: *Proceedings of the 2nd International Conference Web Information Systems Engineering*, Kyoto, Japan, 66-75.

VisitEstonia, 2014. Available http://www.visitestonia.com/en/. [Accessed 10 January 2014].

VisitTuscay, 2010. Tuscay+: the first augmented reality tourism application [Online]. Available http://www.turismo.intoscana.it/allthingstuscany/aroundtuscany/tuscany-the-first-augmented-reality-tourism-application/. [Accessed 10 January 2014]

Vlahakis, V., Karigiannis, J., Tsotros, M., Gounaris, M., Almeida, L., Stricker, D., Gleue, T., Christou, I., Carlucci, R. and Ioannidis, N., 2001. ARCHEOGUIDE: First results of an Augmented Reality, Mobile Computing System in Cultural Heritage Sites. In: *VAST, Proceedings of the conference on Virtual reality, archeology, and cultural heritage*, Glyfada, Greece, ACM, 131-140.

Walls, J., Widemeyer, G. and Sawy, O., 1992. Building an information system design theory for Vigilant EIS. *Information Systems Research*, 3 (1), 36-59.

Walls, J., Widmeyer, G. and Sawy, O., 2004. Assessing Information System Design Thery in perspective: how useful was our 1992 initial rendition? *Journal of Information Technology Theory and Application*, 6 (2), 43-58.

Walther-Franks, B. and Malaka, R., 2008. Evaluation of an augmented photograph-based pedestrian navigation system. In: A. Butz, Fisher, B, Kruger, A. and Olivier, P., eds., *Smart Graphics*, 5166, 94-105.

Watzman, S., 2012. Visual design principles for usable interfaces. In: : Jacko, J., ed., *Human-Computer Interaction Handbook*, 3rd ed., New York: Taylor & Francis, 263–285.

Weber, R., 2003. The problem of the problem. *MIS Quarterly*, 27 (1), iii-ix.

Weiser, M., 1991. The computer for the 21st century. *Scientific American*, 94-110.

Wiesenhofer, S., Feiertag, H., Ray, M., Paletta, L., Luley, P., Almer, A., Schardt, M., Ringert, J. and Beyer, P., 2007. Mobile city explorer: and innovative GPS and camera phone based travel assistant for city tourists. Location-Based Services and Telecartography, *Lecture Notes in Geoinformation and Cartography*, Section V, 557-573.

Wikitude, 2014. Available https://www.wikitude.com/. [Accessed 10 January 2014]

Wilson, T., 1992. Information needs and uses: fifty years of progress? In: Vickery, B., ed., *Fifty years of information progress: a Journal of Documentation review*, London: Aslib, 15-51.

Wilson, T., 2006. On user studies and information needs. *Journal of Documentation*, 62(6), 658–670.

Winters, N. and Price, S., 2004. Mobile HCI and the learning context: an exploration. In: *Proceedings from the Context in Mobile HCI Workshop at the International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI'04),* Glasgow, Scotland.

Wither, J., DiVerdi, S. and Höllerer, T., 2009. Annotation in outdoor augmented reality. *Computers and Graphics*, 33 (6), 679–689.

Wobbrock, J., 2012. Research contributions types in Human-Computer Interaction [Online]. Available http://www.cs.umd.edu/hcil/chi/ResearchContributionTypes.pdf [Accessed 23 July 2012]

Wood, M., 2003. Some personal reflections on change…the past and future of Cartography. *Cartographic Journal*, 40 (2), 111-115.

Yap, L., Bessho, M., Koshizuka, N. and Sakamura, K., 2012. User-generated content for location-based services: a review. In: Lazakidou, A, ed., *Virtual Communities, Social Networks and Collaboration*, New York: Springer, 163-179.

Yell, 2014. Available http://www.yell.com/. [Accessed 10 January 2014].

Yelp, 2014. Available http://www.yelp.com/. [Accessed 10 January 2014].

Yin, J. and Carswell, J., 2013. Spatial search techniques for mobile 3D queries in sensor web environments. *International Journal of Geo-Information*, 2 (1), 135-154.

Yovcheva, Z., Buhalis, D., Gatzidis, C. and van Elzakker, C., 2013a. Towards meaningful augmentation of the cityscape: new challenges for mobile GeoHCI. *In: Proceedings from the GeoHCI workshop @CHI2013*, Paris, France, 27–28 April 2013.

Yovcheva, Z., Buhalis, D., Gatzidis, C. and van Elzakker, C., 2013b. Proposing a design framework for smartphone AR browsers used in unfamiliar urban tourism destinations. *In: Proceedings of the Augmented Reality MobileHCI 2013 workshop*, Munich, 27–30 August, 2013, Germany.

Yovcheva, Z., Buhalis, D., Gatzidis, C. and van Elzakker, C., 2014. Empirical evaluation of smartphone Augmented Reality browsers in urban tourism destination context. *International Journal of Mobile Human Computer Interaction*, 6 (2), 10-31.

Zhu, W. and Owen, C., 2008. Design of the PromoPad: an automated augmented reality shopping assistant. *Journal of Organizational and End User Computing*, 20 (3), 32-48.

Zipf, A., 2002. User-adaptive maps for Location-Based Services (LBS) for tourism. In: Woeber, K., Frew, A. and Hitz, M., eds. Innsbruck, Austria. Springer, 1-10.

# APPENDICES

# Appendix 1 - Criteria for selection of smartphone AR applications

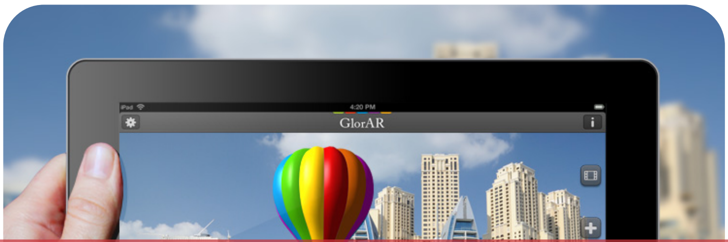
| Application | Type | Availability of content | Annotation Visualization | | Annotation Content | | | |
|---|---|---|---|---|---|---|---|---|
| | | | Layout | Placement | Name | Description | Distance | Other |
| **Acrossair** | AR | Y | Rectangle, mainly text, image | Floating | Y | N | Y | Address, icon showing source |
| **Augmented Reality UK** | AR | Y | Square, mainly image, text | Floating | N | N | Y | Distance in km and miles |
| **Cyclopedia** | AR | Y | Rectangle, mainly text, bottom | Bottom screen | Y | N | Y | None |
| **eTips** | AR-view | N | Rectangle, text | Floating | Y | N | Y | None |
| **GeoTravel Guide** | AR | Y | Circle, text | Floating | Y | N | Y | None |
| **Junaio** | AR Browser | Y | Rectangle, mainly text, image | Floating | Y | Y | N | Icon showing type of POI |
| **Layar** | AR Browser | Y | Rectangle, mainly text, image | Floating | Y | Y | Y | None |
| **Localscope** | AR | Y | Rectangle, text only | Bottom screen, connector line | Y | N | Y | Address, telephone number |
| **London AR Guide** | AR | N | Rectangle, mainly text, image | Floating | Y | N | Y | Icon indicates type of POI |
| **mobeedo** | AR | Y | Square, image | Floating | Y | N | N | Location accuracy and altitude |
| **mTrip** | AR-view | N | Square, mainly image, text | Floating | Y | N | N | Icon indicate type of POI |

| Application | Type | Availability of content | Annotation Visualization | | Annotation Content | | | |
|---|---|---|---|---|---|---|---|---|
| | | | Layout | Placement | Name | Description | Distance | Other |
| **Reality 2.0** | AR | Y | 3D icon with name | Floating in centre | Y | N | N | None |
| **Robotvision** | AR | Y | Rectangle, text only | Floating | Y | N | Y | None |
| **Tripwolf** | AR-view | Y | Icon, text | Floating | Y | N | Y | Icon indicates type of POI |
| **ubique** | AR | Y | Icon | Connected to a map view | N | N | N | None |
| **WhereMark** | AR | Y | Rectangle, mainly text, icon | Floating | Y | N | Y | Address, distance in km and miles |
| **WhereTo?** | AR-view | Y | Rectangle, mainly text, icon | Floating | Y | N | N | Icon indicates type of POI |
| **Wikitude** | AR Browser | Y | Icon and name | Floating | Y | N | N | None |
| **WorldViewer** | AR | Y | Dot, text | Floating | Y | N | Y | None |
| **Yell** | AR-view | Y | Rectangle, text | Floating | Y | N | Y | Telephone number |
| **Yelp** | AR-view | Y | Rectangle, text, image | Floating | Y | N | Y | Rating, reviews, type, money |

# Appendix 2 - Invitation to participate in the field study

ournemouth niversity

School of Tourism




# INVITATION

## TAKE PART IN A MOBILE STUDY

**Dear new colleagues,**

My name is Zory and I am a researcher at the School of Tourism at BU.

I would like to invite you to participate in a mobile experiment we are running. We are trying to find out how **the latest smartphone technologies** can help travellers explore a new place in a more easy and natural way.

During the study I will ask you to take a short walk with me in the city centre (similar to a **guided tour**), visiting some key attractions and sights. I will ask you to look at the screen of a smartphone. Afterwards, we will sit down for **coffee/tea** and we will discuss your experience.

This will take app. one and a half hours and you will be given **£10** as a recompense for your efforts.

You do **NOT need to prepare** anything in advance, have experience with smartphones or know how they work. The study is **anonymous** and the data will be confidential (I will not report your name or personal information in my final thesis).

I am completely flexible with the dates and time and we can run this whenever it is comfortable for you between **11th September - 27th September**.

If you would like to take part just send me a short message on
- **e-mail**: zyovcheva@bournemouth.ac.uk
- **facebook**: Zornitza Yovcheva
- **mobile**: +44 (0) 7411 716045
- **skype**: z.yovcheva

# Appendix 3 – Consent form for mobile field study

## User-Centred Design of Smartphone AR for Tourists

**Design of Smartphone Augmented Reality Information Systems in Urban Tourism Context**

**Project Background**
This research investigates the problems that tourists experience with smartphone Augmented Reality (AR) applications when at new (unfamiliar) place. The main aim is to use the obtained feedback in order to improve the design of smartphone AR in order to facilitate visitors to a new place to obtain information in an easy and more natural way.

**Consent**
I agree to take part in the above Bournemouth University research project. I have had the project explained to me. I understand that agreeing to take part means that I am willing to:

- Perform tasks with a smartphone in an outdoor environment
- Be interviewed by the researcher
- Allow the interview to be videotaped/audiotaped
- Complete a background questionnaire

**Data Protection**
I understand that any information I provide is confidential, and that my name or personal details will be disclosed in any reports, or at any other party.

I consent to the videotapes being showed to other researchers and interested professionals.

I consent to the use of the videotapes in publications.

I agree to Bournemouth University recording and processing this information about me.

**Withdraw from the study**
I understand that my participation is voluntary, that I can choose not to participate in part or all of the project.

I understand that I can withdraw at any stage of the project without being penalized or disadvantaged in any way.

Name:
…………………………………………………………………………………….

Signature: ………………………………………………………………

Date:……………………………..

## Appendix 4 - Background questionnaire used to obtain addition information about participants in the mobile field study

# User-Centred Design of Smartphone AR for Tourists

### Background Questionnaire

**General information**

What is you **gender**?_______________________

What is your **age**?_________________________

What is your **nationality**?___________________

What is your **mother language**?______________

What is you subject of **study** or area of **expertise**?
_________________________________________

Please tick the boxes that apply to you:

| | |
|---|---|
| I am left-handed | |
| I am right-handed | |

Is your vision **corrected** now?

| | |
|---|---|
| No | |
| Yes, I am wearing contact lenses/ glasses | |

**Residence**

For how long have you **lived** in the **UK**?_______________

In which **city** do you live currently?___________________

For how long have you lived at this **location**?____________

Have you ever visited **Bournemouth** city centre before? ☐ YES ☐ NO

If you answered yes to the previous question, when was the last time you **visited Bournemouth city centre**?

| | |
|---|---|
| Earlier today | |
| Yesterday | |
| Last week | |
| Last month | |
| Other | |

**Smartphone use and experience**

Do you own a **smartphone**?

| | |
|---|---|
| No | |
| Yes, I have a smartphone with a touch screen | |
| Yes, I have a smartphone with a keypad | |

Please write down the **brand** of your smartphone:_________________________________

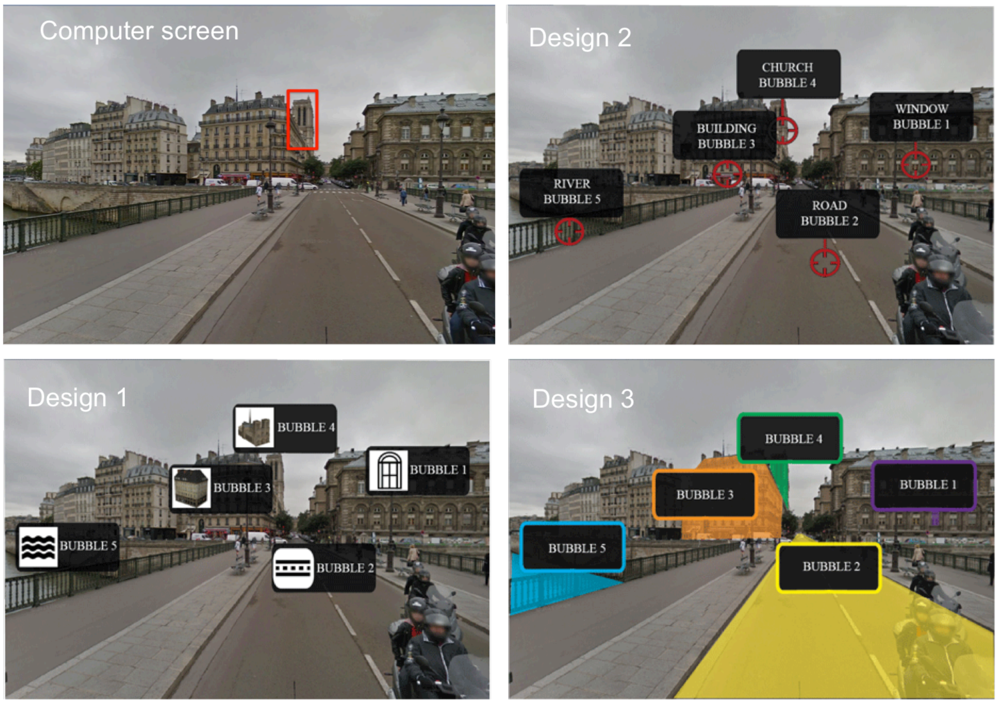
Tick the boxes below if you have **used** your smartphone to:

| | |
|---|---|
| Play games | |
| Work with maps | |
| Work with augmented reality apps | |
| Twitter / Facebook | |
| Browse photos and pictures | |
| Watch videos | |
| Read Wikipedia articles | |

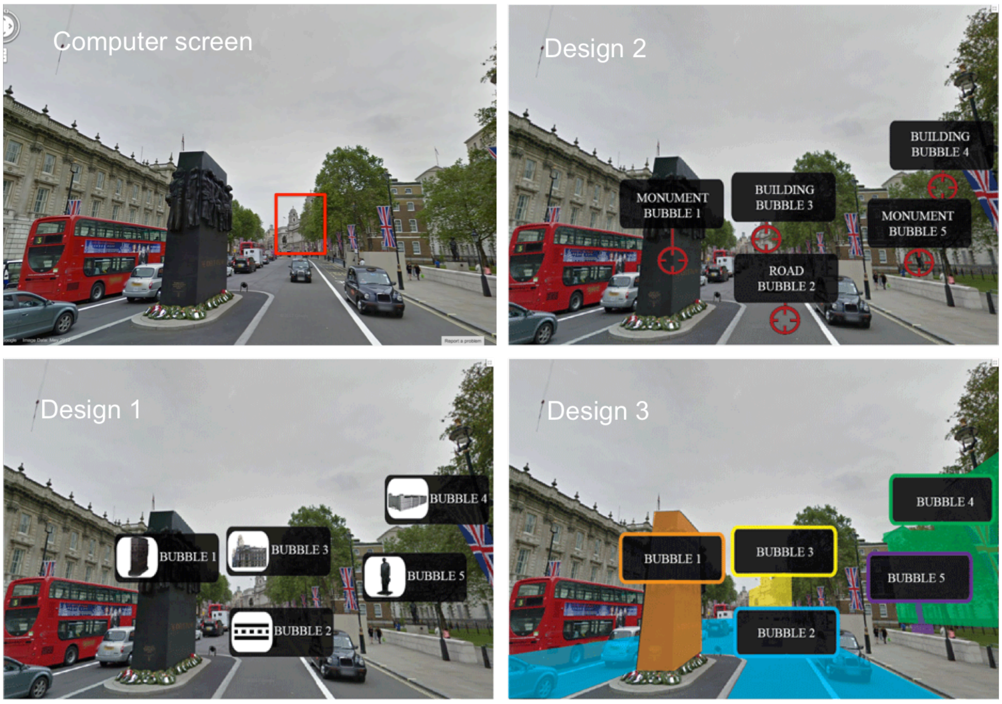
## Appendix 5 – Locations and mock-ups used for the laboratory experiment

The pictures show the materials presented to each participant on the computer screen (top left) and the smartphone screen in group 1 (bottom left), group 2 (top right) and group 3 (bottom left). The target for association has been indicated with a red square. The pictures show the materials used for the first 5 matching tasks.

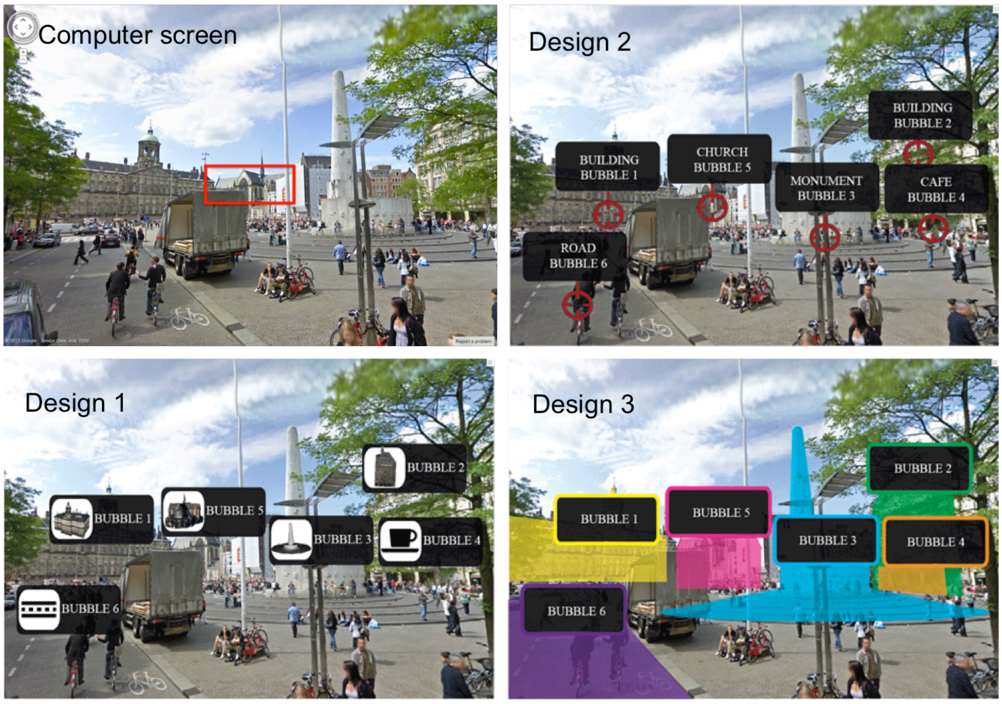
Task 1.



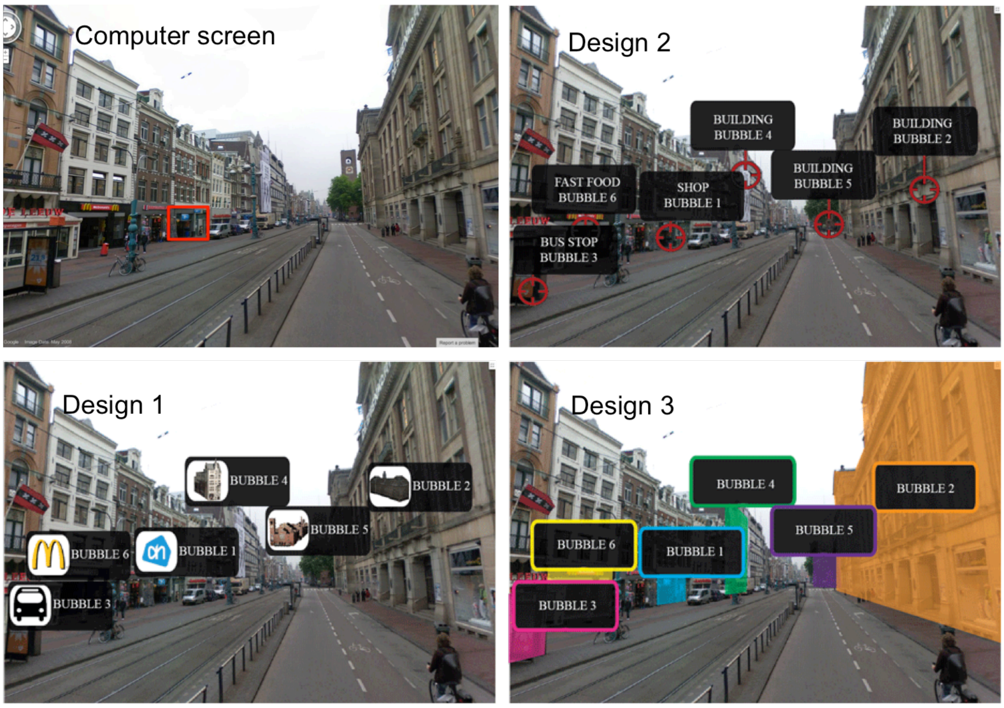
Task 2.



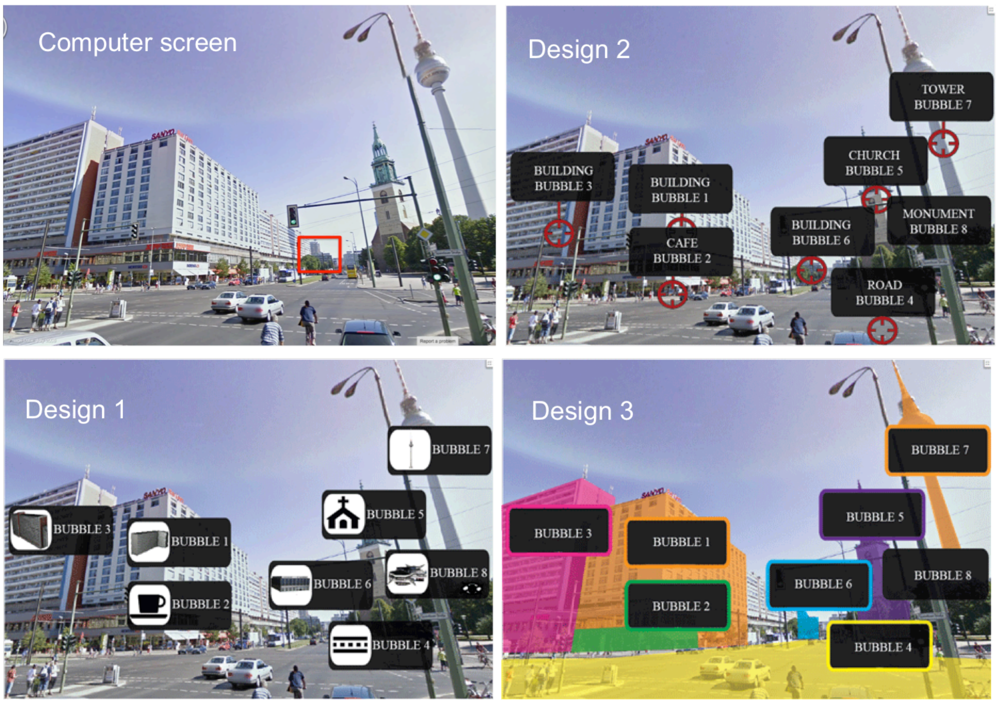

Task 3.




Task 4.




Task 5.




# Appendix 6 – Consent form used prior to the laboratory experiment

# User-Centred Design of Smartphone AR for Tourists

**Project Background**
This research investigates the design of smartphone Augmented Reality (AR) applications for tourists when at new (unfamiliar) place.

**Consent**
I agree to take part in the above Bournemouth University research project. I have had the project explained to me. I understand that agreeing to take part means that I am willing to:

- Perform tasks with a smartphone in an indoor environment
- Be interviewed by the researcher
- Complete a background questionnaire

**Data Protection**
I understand that any personal information I provide is **confidential**, and that my name or personal details will not be disclosed in any reports, or at any other party.

I agree to Bournemouth University recording and processing the information about me.

**Withdraw from the study**
I understand that my participation is **voluntary**, that I can choose not to participate in part or all of the project.
I understand that I can withdraw at any stage of the project without being penalized or disadvantaged in any way.

Name: ……………………………………………………………………………………….

Gender: Female / Male Age: ____________________

Is your vision corrected now:
O Yes, I am wearing glasses / contact lenses
O No
Do you use a smartphone?
O Yes, every day O Yes, several times a week O Yes, once a week O Yes, less than once a week
O No, I have just tried it out once O No, I have never used a smartphone

Do you use Augmented Reality applications (e.g. Layar, Wikitude, Junaio)?
O Yes, every day O Yes, several times a week O Yes, once a week O Yes, less than once a week
O No, I have just tried it out once O No, I have never used Augmented Reality

Signature: ………………………………………………………………

Date:……………………………..

# Appendix 7 – Protocol for collecting data during lab experiment

**Form for individual evaluation of AR designs**

Date: ______________

Participant ID: _________ Name: __________________________________________Test starts at: _________________

| Test Task | Actual task | Bubble N:/Success | Time | Certainty | Difficulty | Comments |
|---|---|---|---|---|---|---|
| 1 | | | | | | |
| 2 | | | | | | |
| 3 | | | | | | |
| 4 | | | | | | |
| 5 | | | | | | |
| 6 | | | | | | |
| 7 | | | | | | |
| 8 | | | | | | |
| 9 | | | | | | |
| 10 | | | | | | |
| 11 | | | | | | |
| 12 | | | | | | |
| 13 | | | | | | |
| 14 | | | | | | |
| 15 | | | | | | |
| 16 | | | | | | |
| 17 | | | | | | |
| 18 | | | | | | |
| 19 | | | | | | |
| 20 | | | | | | |

Test ends at: ___________________

Comments:_____________________________________________________________________________

# Appendix 8 – Example of the questionnaire used during the pluralistic walkthrough evaluation

## AR Annotations Evaluation

Thank you for agreeing to participate in this study!

The main aim of today's evaluation is to obtain feedback and improve the design of smartphone AR annotations for tourists in unfamiliar urban environments.

Please try to answer the questions provided in this document. It is important to have in mind that evaluation is not directed at your own experience and skills. There are no right answers. Please provide your honest feedback and opinion.

Task 1. Look at the information provided on the screen of the smartphone and answer the questions below.


Q1: Can you determine what are **the visible physical objects around you**? How?

________________________________________________________________________

________________________________________________________________________

Q2: Is it clear what other types of points of interest (**not visible from the current location**) are around you? How?

________________________________________________________________________

________________________________________________________________________

Q3: Can you determine what is the name of **the building in front of you**?

________________________________________________________________________

________________________________________________________________________

Q4: What is your opinion about the **symbols** in the AR bubbles? Why?

Q5: What do you think about **the names** of the points of interest? Why?

Q6: What do you think about the information about the **distance**? Why?

Q7: Is the information provided within the AR bubbles **relevant** to the current situation? Why?

Q8: Is the information in the bubbles **useful**? Why?

Q9: Which interface do you like the most? Why?

Q10: Additional comments…