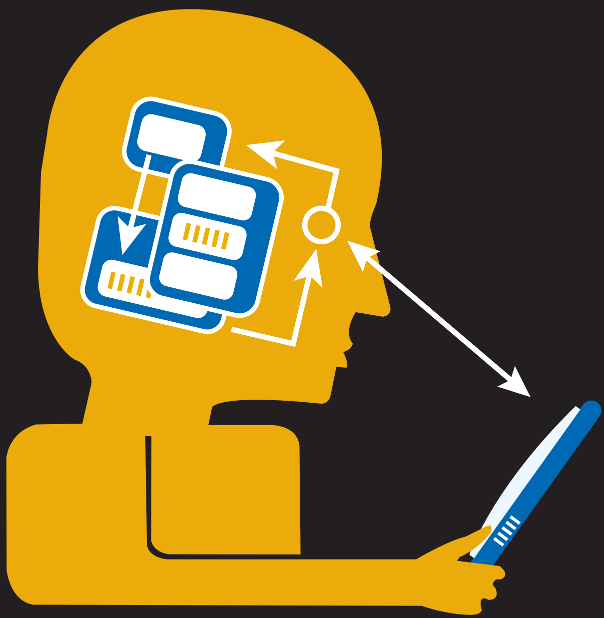

Minimal Mobile Human Computer Interaction

Abdallah El Ali

Minimal Mobile
Human Computer Interaction

Abdallah A. El Ali

Minimal Mobile Human Computer Interaction

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de
Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof.dr. D.C. van den Boom
ten overstaan van een door het college voor promoties ingestelde
commissie, in het openbaar te verdedigen in
de Agnietenkapel
op vrijdag 15 november 2013, te 12:00 uur

door

Abdallah El Ali

geboren te Riyadh, Saudi Arabia

Promotiecommissie

Promotor:
Prof. dr. L. Hardman
Co-promotor:
Dr. F. Nack

Overige leden:
Prof. dr. B. Wielinga
Prof. dr. ir. B. Kröse
Prof. dr. M. de Rijke
Prof. dr. S. Boll

Faculteit der Natuurwetenschappen, Wiskunde en Informatica

SIKS Dissertation Series No. 2013-35
The research reported in this thesis has been carried out under the auspices of SIKS, the Dutch Research School for Information and Knowledge Systems.

This research was supported by the Amsterdam Living Lab project under project number PID07071, and by the Eurostar SmartINSIDE project under project number E5262.

Cover by Hosni Auji
Printed by Offpage, Amsterdam

ISBN: 978-94-6182-364-9

Looking back...

4.5 years. 243 pages. I may have written this work, but an uncountable amount of thanks go to the ones who accompanied me in this (at times ridiculous) journey.

Family,
for being there every step of the way.

Frank, Lynda,
for being my mentors.

Nokia Research,
for fire-starting my research.

T-Labs,
for keeping my interests glowing.

Co-authors,
for the inspiration and enjoyable collaboration.

Ex-students,
for the dialogues and good work.

HCS, ILPS, & other SP labs,
for the times at Polder, Oerknal, and beyond.

Commelinstraat crew & the Usual Suspects,
for all the crazy times, nothing short of.

Existence,
for surviving you until now.

Contents

1 Introduction

In the last 20 years, the rise of new hardware devices, ubiquitous wireless networks, and widespread adoption of personal computing devices in everyday life, have allowed entry into a new technological era in Human Computer Interaction (HCI). During the early days of HCI, the context in which systems were used was strongly influenced by the place in which computers were set up. Personal computers were primarily used in office environments or in factories to automate production. In such environments, the context of use was relatively static, where computers were used for a limited set of functions (e.g., as arithmetic calculators). However, with the rise of mobile devices in the late 1980's and 1990's, this changed as devices became smaller, more portable with GPS and constant internet connectivity. By the early 90's, the first cellular phone to incorporate personal assistance and internet connectivity features (e.g., e-mail) was the Simon Personal Computer developed by BellSouth.[1] These portable computers are now called smartphones.

The constant change of physical and social context in a user's situation made possible by the portability of mobile devices also means that the user's attention becomes limited. If a user wishes to perform a task on her smartphone (e.g., calling, texting, or reading a map), this consumes her information processing resources. This makes it difficult to focus on the surrounding environment or a given social situation. This can result in user frustration, accidents, and an inefficient and ineffective means of interacting with smartphones. In other words, it can result in situational impairments and negatively affect the user's experience. In order to deal with sensory and information overload, researchers and designers in HCI have proposed different solutions. One solution from the field of context-aware computing is to make use of context-awareness, so that mobile devices can sense, learn from, and adapt to the user, thus freeing the user's attentional resources. Another solution from the field of multimodal interaction is to design and develop non-visual input techniques (such as gestural or speech input), so that the user need not rely on his or her visual sense while interacting with a smartphone in an urban setting.

This thesis draws from both of these solutions, where we introduce the concept of *minimal mobile HCI*. The goal here is to design mobile interactions that require *minimal* reliance on visual touchscreen interactions, so that they can be suitable for use in urban settings. To explain the motivation behind introducing this concept and how it can help design for good user experiences (Preece et al., 2002), some background is necessary. We

[1] http://www.businessweek.com/articles/2012-06-29/before-iphone-and-android-came-simon-the-first-smartphone; last retrieved: 01-08-2013

give an overview of the broad field of mobile HCI, and two other closely related fields: context-aware computing and multimodal interaction.

A visual illustration of each field that contributes to minimal mobile HCI is shown in Fig. 3.1. In mobile HCI, the concern is studying how users interact with mobile technology (e.g., optimizing touchscreen keyboard layout to increase user typing accuracy and efficiency). In Context-aware computing, the concern is to design and develop technologies that can sense user behavior and the environment to provide intelligent applications and services (e.g., urban lampposts that turn on when their proximity sensors are activated). In multimodal interaction, the concern is to design and develop technology that allows users to interact with technology in a natural way (e.g., through voice- or gesture-based interfaces). Minimal mobile HCI, by contrast, complements and makes use of mobile HCI, context-aware computing, and multimodal interaction, where the goal is to design and develop technology that makes minimal use of users' visual modality. Each of these fields will be explained in detail below.

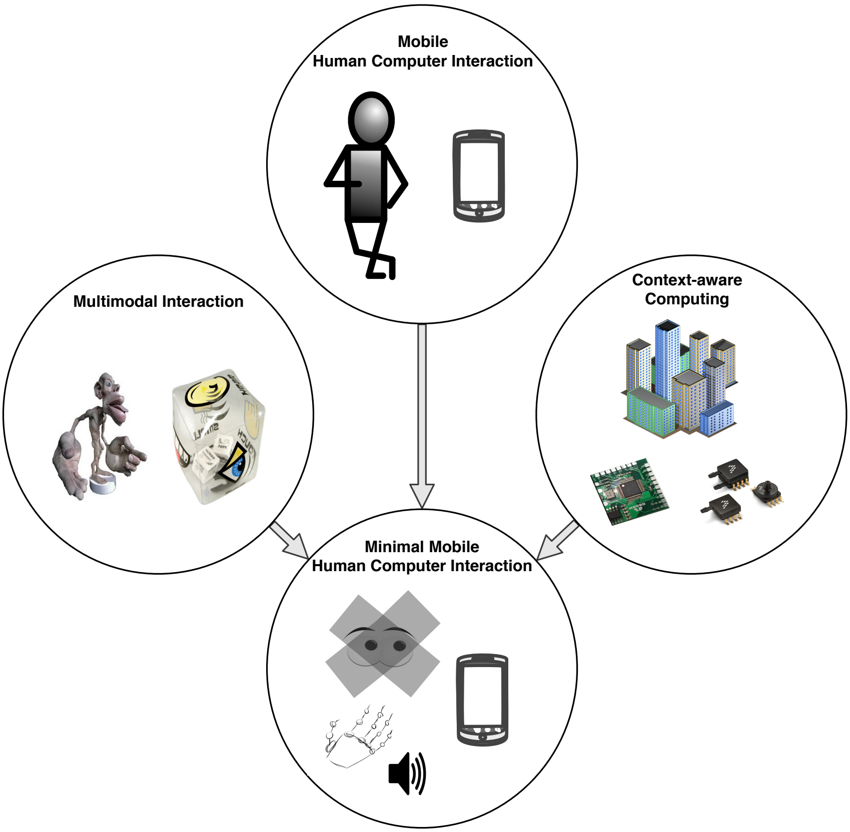

Figure 1.1: Overview of the different fields related to this thesis.

1.1 Background

1.1.1 Mobile Human Computer Interaction

Mobile HCI is defined as the *"study of the relationship (interaction) between people and mobile computer systems and applications that they use on a daily basis"* (Love, 2005, p. 2). Whether this involves the design of a new technique to visualize off-screen objects on a small mobile display, an application that allows uploading a photo to a social network, or doing collaborative work by sharing a document on one's mobile device, these are all examples of mobile HCI. For our purposes, we are concerned with understanding users of mobile devices, their capabilities and expectations, and how these influence the design of mobile systems or applications. For mobile interaction, there are distinctive aspects that pose interaction design challenges in providing the optimal user experience for users of mobile devices. As reviewed by Chittaro (2009), these include:

- **Hardware limitations:** Small screen (cf., Zwick et al. (2005); Kärkkäinen and Laarni (2002)), limited input and output channels

- **Perceptual limitations:** Noisy street, sunlight reflection, device is sometimes outside of the user's line of sight (Obrenovic et al., 2007; Oulasvirta et al., 2005)

- **Motor limitations:** Voluntary movements when inside a moving vehicle, and the fat-finger problem (higher portability by reducing the screen size comes at the expense of target accuracy and precision (Kane et al., 2008; Parhi et al., 2006))

- **Social issues:** Awkwardness (e.g., a phone ring while at a conference), or in some cultures, performing 3D motion gestures in front of strangers (Rico and Brewster, 2009)

- **Cognitive limitations:** Humans have a limited attention span (3-4 second bursts (Oulasvirta et al., 2005), high stress and perceived workload (multitasking and interruption issues (Mcfarlane and Mcfarlane, 1997)), limited memory (working memory 4-7 items (Baddeley et al., 1974; Baddeley, 2003))

Each of these aspects influences the adoption of mobile technologies, and are relevant for the future of mobile HCI. Essentially, a mobile context means users' cognitive (e.g., attention and memory), perceptual and motor resources are limited (Tamminen et al., 2004). For example, consider the situation where a person is talking with a friend, and simultaneously scrolling through a map to navigate to the right destination and crossing the street. This may pose safety issues. In short, more information presented to the user, while already in a dynamic urban setting, can incur further sensory and information overload. In short, visual attention in a mobile setting becomes a scarce cognitive resource. So what can be done about this? New hardware and sensing capabilities have resulted in major innovations in how we interact with computers, where two important trends in mobile computing and HCI can address the problem of information and sensory overload. These trends that have emerged are context-aware computing and multimodal interaction.

1.1.2 Context-aware Computing

Ubiquitous Computing

Over the last decade, the small form factor of mobile devices now allows users to carry computers with them and use them in a variety of situations (Bentley and Barrett, 2012). This gave rise to the challenge of how to make mobility transparent for the user, so that users can access information wherever they are. This provision of contextualized information anytime, anywhere, to the right persons as they go about their daily lives is part of this emerging paradigm dubbed as ubiquitous computing (Weiser, 1991), context-aware computing (Dey et al., 2001), pervasive computing (Ark and Selker, 1999), or more recently, everyware (Greenfield, 2006).

Irrespective of the name given, a central tenet of this paradigm is the promise of populating our everyday lives with context-aware services that make interaction with the world easier, more manageable, more enjoyable, and more efficient. This endeavor is made possible through embedding (at times personal and imperceptible) low-cost and low-power sensors and devices into our everyday environment. Mark Weiser stated in the early 90's (Weiser, 1991) that when computers become a part of everyday life, it is essential that they are easy to use, and ultimately disappear: *"The most profound technologies are those that disappear. They weave themselves into the fabric of everyday life until they are indistinguishable from it"* (p. 66). One significant example has been the widespread adoption of location-aware technologies such as GPS-enabled mobile devices (smartphones) and automotive GPS, which have now become inextricably embedded in our daily lives.

Urban Computing

Progress in ubiquitous computing gradually gave rise to a closely related area of research known as urban computing or urban HCI (Fischer and Hornecker, 2012; Paulos and Jenkins, 2005). Urban computing is concerned with the interaction between humans and public environments, such as cities, parks, or suburbs. It essentially refers to "the situation that is composed of the built environment, the interface and any associated computer system, and the social context" (Fischer and Hornecker, 2012, p. 307). Examples of research in this area (Paulos and Jenkins, 2005) include pedestrian navigation, citizen journalism, urban planning, rendezvousing, public displays, and window shopping.

One of the goals of urban computing is to enable so-called 'smart-cities', which make use of embedded sensors that allow seamless and invisible interaction in the user's daily life. However, for realizing smart cities, sensors need not be embedded directly in the environment. Smartphones, which are already equipped with a large number of sensors, can enable so-called 'smarter cities'. Interaction possibilities with smartphones, from social networks through video capture, route planning, and interaction with public displays, have transformed how we interact in an urban environment. Social networks allow us to connect with friends and strangers while on the go (e.g., Facebook[2] on smartphones); video capture allows us to capture and store content about the environment, enabling so-called citizen journalism (e.g., YouTube[3]); interaction with small and large public displays allows perva-

[2]https://www.facebook.com/; last retrieved: 01-08-2013
[3]http://www.youtube.com/; last retrieved: 01-08-2013

sive access to information in public places (Müller et al., 2010), and digital route planning influences the routes we take in a city (e.g., Google Maps[4]).

Wearable Computing

Another related area is a trend in computing that promises to deliver what are called Wearable Computers (Mann, 1997). Recently, efforts towards providing a good user experience for wearable computers has resulted in what is informally dubbed as Glanceable User Interfaces (UIs). The goal of such interfaces is to minimize attentional demands from the user. Examples include Head Mounted Displays (HMDs) (Lucero et al., 2013; Cakmakci and Rolland, 2006) such as Google Glass,[5] smart wrist watches (Raghunath and Narayanaswami, 2002) such as Apple's anticipated iWatch,[6] or other small sized touch devices (Baudisch and Chu, 2009).

For all these devices, they are meant to be worn by the user on a daily basis and for an indefinite amount of time. This means that the information presented to the user should not be always in the user's attentional spotlight, so as not to distract from other daily living tasks – they should be 'glanceable'. Some wearables, including Google Glass, additionally support voice-based interaction, so that the user need not devote full visual attention to performing interface tasks such as dialing a contact or requesting navigation instructions. HMDs like Google Glass or NotifEye (Lucero et al., 2013) provide good examples of minimal interaction, as they combine both non-visual interaction techniques with touch-based interaction. This combination ensures a good user experience, especially when users interact with them in outdoor, urban settings.

1.1.3 Multimodal Interaction

Multimodal Input

Multimodal user interfaces, in emphasizing human communication and motor skills, seek to make human computer interaction more natural and more effective (Turk and Robertson, 2000). Multimodal interaction refers to the situation where the user is provided with multiple modalities for interacting with a system, typically through natural means such as speech, gaze or gestures. Specifically, multimodal interfaces *"process two or more combined user input modes (such as speech, pen, touch, manual gesture, gaze, and head and body movements) in a coordinated manner with multimedia system output. They are a new class of interfaces that aim to recognize naturally occurring forms of human language and behavior, and which incorporate one or more recognition-based technologies (e.g. speech, pen, vision)"* (Oviatt, 2003, p. 286). Research in multimodal interaction is typically split between multimodal input (e.g., use of speech, gestures, eye gaze) to control a system or interface, and multimodal output which focuses on the type of feedback provided to users when they interact with a device or interface. In this thesis, we focus mainly on multimodal input.

[4] https://www.google.com/maps/; last retrieved: 01-08-2013

[5] http://www.google.com/glass/ ; last retrieved: 01-08-2013

[6] http://techcrunch.com/2013/04/01/apples-iwatch-is-actually-just-a-wrist-band-that-attaches-to-your-iphone-ipad/ ; last retrieved: 01-08-2013

Multimodal interaction is closely related to a recent trend in Human Computer Interaction that aims to provide what are called Natural User Interfaces (NUIs) (Jain et al., 2011). As stated by Jain et al. (2011), this class of interfaces enables users to interact with computers in the way we interact with the world. An important element of such natural interaction is the use of 3D gestures. However gestures alone do not suffice to allow seamless natural interaction, as the user still needs to receive feedback from the system on a performed gesture (Norman, 2010). This is usually complemented by the use of touchscreen buttons, menus, auditory or speech feedback, or some kind of visual feedback from the system. This is an important aspect for provision of minimal mobile interactions that make use of non-visual interaction techniques such as 3D gestural input. For such minimal interaction, complementing user interaction with minimum visual interaction (e.g., feedback on actions) is necessary to provide a minimally attention demanding user experience.

Eyes-Free Mobile Interaction

Interaction with smartphones is typically achieved through touchscreen interaction, where information is usually presented visually. However, user attention is a scarce resource during interaction (Oulasvirta et al., 2005), especially under mobile settings (e.g., crossing the street). To address the issue of limited visual attention for mobile users, researchers and designers have attempted to make use of non-visual modalities when designing interfaces. These interfaces rely on auditory (Vazquez-Alvarez and Brewster, 2011; Li et al., 2008; Zhao et al., 2007) or gestural/haptic interaction (Ashbrook et al., 2011; Ketabdar et al., 2010c; Baudisch and Chu, 2009) with smartphones in order to minimize the need for visual attention. This wave of research is known as multimodal 'eyes-free' interaction (Yi et al., 2012; Brewster et al., 2003).

Given the cognitive burdens associated with urban environments, and the social and interactional limitations of using smartphones in such settings, eyes-free interaction provided at least initially a suitable goal for mobile interaction designers. When analyzing the motivations behind using eyes-free interaction from a user-centered perspective, Yi et al. (2012) found that motivations fell along two dimensions: context dependency (independent vs. contextual) and physicality (physical vs. human). Under these dimensions, motivations were clustered:

1. **Environmental** (contextual + physical), which includes enabling operations under extreme lighting conditions and improving safety in task-switching
2. **Social** (contextual + human), which includes fostering social respect, avoiding interruption of human activities, and protecting private information
3. **Device** (independent + physical), which includes enabling operations with small or no screens and enabling multitasking on the same device
4. **Personal** (independent + human), which includes entertainment, desire for self-expression, and lower perceived effort

Given the foregoing issues, eyes-free interaction can be used to address the attentional demands that our smartphones require of us, as well as ease the multitasking costs and frustration that arises out of smartphone use in urban settings (Tamminen et al., 2004). Moreover, it provides one solution to the 'fat-finger' problem (cf., Baudisch and Chu (2009)), which occurs as a a result of the small screen size of mobile displays Chittaro (2006); Brewster (2002); Forman and Zahorjan (1994).

Evaluating Usability and User Experience

While there has been much work designing and developing such eyes-free interaction techniques, the *usability* and *user experience (UX)* issues associated with these techniques is an ongoing research effort within the HCI, ubiquitous computing as well as more generally the User Experience communities. Usability here is based on the ISO 9241-11 (1998) definition:

> *"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."*

Likewise for UX, we make use of the ISO 9241-210 (1994) definition:

> *"A person's perceptions and responses that result from the use or anticipated use of a product, system or service."*

This means that all the research carried out in this thesis involved rigorous user testing, whether quantitative (e.g., Likert-scale questionnaires, duration analysis) or qualitative (video analysis, focus groups, interviews). Investigating these issues in mobile interactions that make minimal use of visual touchscreen interaction is crucial to the adoption of such 'eyes-free' interaction methods, and is the topic of this thesis.

From a practical perspective, while research efforts dedicated to providing complete eyes-free interaction solutions have shown such interaction techniques to be usable within research contexts (e.g., Vazquez-Alvarez and Brewster (2011); Williamson et al. (2007); Zhao et al. (2007); Brewster et al. (2003)), interaction designers still largely rely on the touchscreen model for smartphone interactions despite the lack of screen space on a mobile screen (Brewster, 2002; Forman and Zahorjan, 1994). This is used to support typical user tasks for mobile interaction, such as calling, texting, taking photos, or playing games. This can be seen from for example Android's Design Guidelines for gestural interaction[7] or Apple's iOS Human Interface Guidelines for User Experience.[8]

1.2 Research

1.2.1 Outline

This thesis examines whether we can reduce the attentional costs and ultimately improve the user experience associated with *smartphone* use. We introduce the concept of *minimal mobile human computer interaction*, a subset of eyes-free mobile interaction that allows minimal combination of the visual modality with other sensory modalities to minimize attentional demand, frustration, and situational impairments when users interact with

[7] http://developer.android.com/design/patterns/gestures.html; last retrieved: 01-08-2013
[8] http://developer.apple.com/; last retrieved: 01-08-2013

smartphones. Our goal is to design mobile interactions that require *minimal* reliance on touchscreen interactions, in order to enhance users' experience of interacting with their smartphones. Especially in an urban context where interaction costs (whether perceptual, cognitive, motor, or social) are higher.

> **Research statement:** This thesis investigates the usability and user experience issues associated with *minimal* mobile human computer interaction.

We carried our research in different domains: Urban Interaction (Chapter 1, 3), Playful Interaction (Chapter 2, 5), Task Independent Interaction (Chapter 4), User Authentication (Chapter 6). A summary of how each thesis part relates to each chapter and study, the system/technique under investigation, the domain, and the publication source, are shown in Table 1.1. The common denominator across all studies was ensuring rigorous user testing (whether in the laboratory or in the wild), with the goal of improving the usability and user experience (UX) of the designed minimal mobile interactions. To this end, we focus on two main themes of mobile interaction design:

1. Minimizing mobile interaction by making use of context-awareness (Part I)
2. Minimizing mobile interaction by making use of gestural input techniques (Part II).

	Chapter	Study / RQ	System / Technique	Domain	Publication
Part I	2	LMM RQ1	Location-aware Multimedia Messaging	Urban Exploration	El Ali et al. (2010)
	3	Playful LMM RQ2	Location-aware Multimedia Messaging	Playfulness, Urban	El Ali et al. (2011)
	4	Route Planner RQ3	Exploration-based Route Planner	Urban Exploration	El Ali et al. (2013b)
Part II	5	Gesture Errors RQ4	3D Gestural Interaction	Task-independent	El Ali et al. (2012)
	6	Playful Gestural Interaction RQ5	Magnet-based ADI	Playfulness, Music	El Ali and Ketabdar (2013)
	7	Gestural Authentication RQ 6	Magnet-based ADI	User Authentication	El Ali et al. (2013a)

Table 1.1: Summary of each thesis chapter.

1.2.2 Research Questions

In Part I of the thesis, we focus on context-aware computing. To design and evaluate minimal mobile interactions, we start by investigating the contextual factors associated with location-based media production and consumption by city residents (*LMM Study*). To do this, we evaluated a location-aware multimedia messaging (LMM) system that allows users to create multimedia content anchored to locations. The LMM system provides a suitable use case for investigating minimal mobile interaction, since the multimedia messages are automatically anchored to the location they were created at. This allows a more restricted but simpler form of interaction with user generated multimedia content. Additionally, the Augmented Reality (AR) presentation output of these messages makes interaction easier for users, particularly in an urban setting where user attentional resources are limited. We conducted a pilot ethnographic study to evaluate participant interaction with the LMM prototype. Given the early stage of our prototype and the short duration of the pilot tests, we followed up with another longitudinal study using instead a multi-modal diary method. We ask:

> **RQ 1:** How do users create and interact with urban location-based content, and how should this impact the design of future Location-aware Multimedia Messaging systems?

From our findings on how users interact with LMM systems (*LMM Study*), it became evident that the primary goal for users was to use the LMM system to facilitate playfulness. However, the simplicity of the context-awareness in the LMM app and AR presentation (i.e., the designed minimal mobile interactions) were limited in providing the desired playful experience that users wanted. To investigate this limitation, we took our findings from Study 1 and used them as a case study to analyze in more analytical detail the qualities of playful experiences (*Playful LMM Study*). Specifically, how can LMM systems provide fun and playful user experiences. Our analysis gave rise to three primary problems (listed below) that emerge when designing simple, playful interactions. From analyzing these problems, we drew design considerations for playful human computer interaction in urban settings, addressing explicitly:

1. How playful experiences can be inferred
2. How playful experiences be maintained over time
3. How playful experiences be measured

In this chapter, we ask:

> **RQ 2:** How can location-aware multimedia messaging (LMM) systems be used to support playful urban interactions?

While the *LMM Study* and *Playful LMM Study* focused on the overall user experience and elicited playfulness of multimedia messaging behavior at urban locations, these studies also showed that urban interactions take place across locations, rather than isolated locations. To account for the connectedness of urban interactions across locations, we followed up with the *Route Planner Study* to design a system that allows pedestrians to explore a city. To design such a system, we reasoned that the movement of city photographers can tell us something about interesting routes in a city. To maintain the requirement of designing minimal mobile interactions, we wanted to avoid burdening users in supplying lengthy user preferences. Therefore, we made use of a smartphone's context-aware capabilities (in this case, location sensing using GPS). To do this, we made use of the geotagged data provided by the photo sharing website Flickr.[9] We ask:

> **RQ 3:** How can we automatically generate routes to support pedestrians in exploring a city?

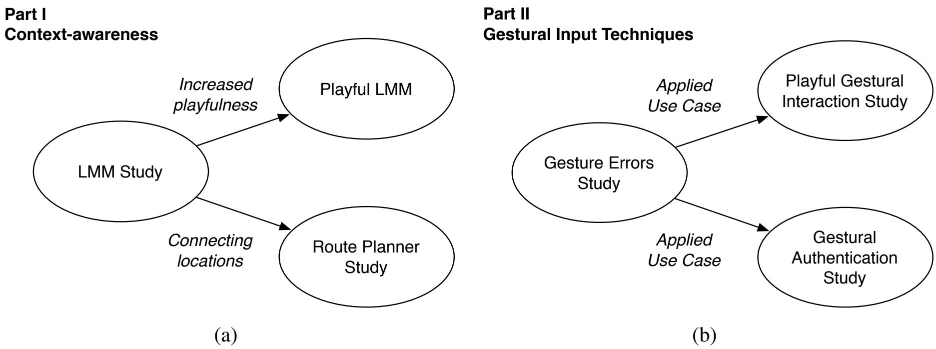

Figure 1.2: Relationship between the studies in this thesis.

The first three studies (*LMM Study, Playful LMM Study, Route Planner Study*) in Part I of the thesis showed that context-awareness can contribute to the design of minimal mobile interactions, which makes user interactions in urban settings simpler and more playful. A visual summary of the relationship between each study in Part I is shown in Fig. 1.2(a).

To investigate gestural input techniques, in the second theme of our thesis (Part II), we investigated 3D gesture-based interaction. This input interaction method allows for eyes-free mobile interaction, which frees the user's visual attention. However, such techniques are error-prone, which can incur additional processing costs from users when in a crowded setting or when encumbered. Therefore, we needed to understand how users deal with errors in 3D gesture-based interaction, and investigate which set of gestures provides the best user experience (*Gesture Errors Study*). In this work, we ask:

> **RQ 4:** What are the effects of unrecognized 3D gestures on user experience, and how do these affect the design of error-tolerant 3D gesture sets?

While the *Gesture Errors Study* focused on the usability of 3D gesture-based interaction, we mainly looked at task-independent interaction. Looking at task-independent interaction was necessary to investigate the usability issues associated with performing 3D gestures, as pairing with tasks may have influenced gesture performance and preference. Given the promise of gesture-based interaction, we revisited how this form of interaction can be applied in an actual domain. Moreover, given the problematic nature of supporting playful experiences uncovered in the *Playful LMM Study*, we revisited the domain of playful mobile interactions. For this study (*Playful Gestural Interaction Study*), we investigated how 3D gesture-based interaction can be synergistically coupled with minimal touchscreen interaction to facilitate playfulness. To do this, we looked at gestural control made possible by magnet-based Around Device Interaction (ADI), which allows users to interact with their smartphones by gesturing around the smartphone device using a properly-shaped magnet. Here, we investigated how gestures can be used to support playful music composition and gaming. Using three musical applications (Air Disc-Jockey, Air Guitar, Air

[9] http://www.flickr.com/; last retrieved: 01-08-2013

GuitaRhythm), we investigated whether this paradigm can be effectively used to support playful music composition and gaming on mobile devices. Here, we ask:

> **RQ 5:** How can 3D gestural interaction support playful music composition and gaming on smartphones?

As an additional example of 3D gesture-based interaction in daily mobile interactions, we look at a common task performed by mobile users: authentication. User authentication has a long history in computing, and represents an essential component of data security and privacy (Cranor and Garfinkel, 2005). Understanding how users interact using security methods is of paramount importance, given the common task of unlocking one's mobile device (e.g., PINs) or entering passwords to access private data. However, a fundamental challenge when designing and implementing any security method is to ensure that the method is both usable by users, and at the same time providing sufficiently strong security against any kind of adversarial attack. To preserve the requirement of designing minimal mobile interactions, we look at 3D gestural interaction (*Gestural Authentication Study*). This gestural interaction, as in the *Playful Gestural Interaction Study*, also makes use of magnet-based ADI, wherein users can gesture around a smartphone with a magnet to execute some smartphone function. In this case, 3D gestural authentication allows users to gesture for example their signature around the device, and if the performed signature matches the recorded signals of their initially recorded signature, then they would be granted access to their device. To investigate the security and usability tradeoff of this method, we designed and executed two separate usability and security studies. Here, we ask:

> **RQ 6:** How does 3D gestural interaction affect the usability and security of mobile gestural user authentication?

A visual summary of the relationship between each study in Part II is shown in Fig. 1.2(b).

1.3 Main Contributions

The main contributions of this thesis, split according to contribution type, are:

1.3.1 Empirical Findings

- **Contextual factors in LMM production and consumption:** In the *LMM Study*, we collected longitudinal data on people's multimedia messaging practices, and analyzed these reported activities using an episodic memory framework borrowed from Cognitive Science. Our analysis provides insight into the contextual factors governing what, where, how and why people create and consume location-based multimedia messages. This provided the groundwork for the need for minimal mobile interactions, given the high information load incurred on users in urban settings.

- **Digital information aids to support city exploration:** In addition to showing how exploration-based routes can be generated using sequence alignment methods , we also empirically investigated through the *Route Planner Study* the role that different digital information aids play in supporting users to explore a city. The results of this investigation provide user-driven findings into which digital information aids people would like to make use of when they want to explore a city.

- **Security verification of magnet-based gestural authentication:** Using a 3D gestural interaction framework for user authentication on smartphones, we replicate a study that investigates the security of this authentication method (*Gestural Authentication Study*). Our security analysis results provides further verification for the security of the magnetic gestural user authentication method.

1.3.2 Methods

- **Exploration-based city route planner method:** To support pedestrians in exploring a city (*Route Planner Study*), we adapted a sequence alignment method from bioinformatics for aligning DNA and protein, in order to align sequences of geotagged photos on a grid partitioned map. Thereafter, Dijkstra's shortest path algorithm was modified to connect these aligned photo sequences for generating exploration-based routes in small-sized cities that attract many tourists (e.g., Amsterdam).

- **Automated Wizard-of-Oz gesture recognizer method:** To investigate the usability and user experience of 3D gesture-based interaction using accelerometer- and gyroscope-equipped smartphones (*Gesture Errors Study*), we developed an automated Wizard-of-Oz method for 3D gesture recognition. This method, adapted from speech recognition research, allows interaction designers to quickly and easily test user frustration and tolerance to gesture recognition errors, as well as identify immediately which gesture sets provide the best user experience.

1.3.3 Mobile Interaction Design

- **Design recommendations for future LMM systems:** Based on our LMM prototype evaluation, and on the longitudinal data collected on multimedia messaging practices (*LMM Study*), we distill design recommendations (Section 2.8) that interaction designers can use for designing future context-aware systems that support recording and sharing urban experiences.

- **Design considerations for playful mobile HCI:** Our empirical findings on how people create and share experiences using LMM systems (*LMM Study*) led to a larger analytical exposition of design considerations for designing and evaluating playful human computer interactions (*Playful LMM Study*). From this case study, we distill three design considerations (Section 3.7) that interaction designers should take into account when designing and evaluating playful mobile interactions using LMM systems.

- **Design recommendations for error-tolerant 3D gestures:** Using our automated Wizard-of-Oz method for gesture recognition, we empirically tested (*Gesture Errors Study*) two iconic gesture sets for use in smartphone gesture-based interaction. Our investigation provides interaction designers with recommendations (Section 5.7) for which gesture sets are more robust to recognition failures, and which overall provide the best user experience.

- **Design recommendations for playful magnet-based ADI:** By empirically investigating how 3D gestural interaction can be used to support playful music composition and gaming (*Playful Gestural Interaction Study*), we arrived at a set of design recommendations for designing playful interactions. These recommendations (Section 6.8) can be used by interaction designers who wish to use 3D gesture-based interaction for creating playful user experiences in the music production domain, as well as designing engaging 3D gesture-based games.

- **Design recommendations for usable magnetic gestural authentication:** As part of our investigation into the security of 3D gestural user authentication (*Gestural Authentication Study*), we investigated the usability and user experience issues associated with using this method. We distill design recommendations (Section 7.9) that aid designers and developers in improving users' experience when using this 3D gesture-based method for privacy and security access.

1.4 Thesis Overview

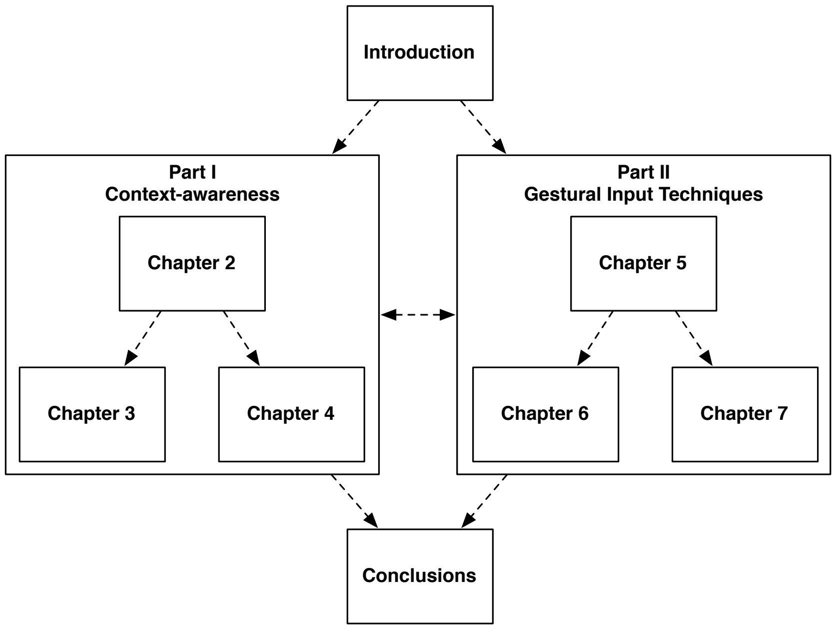

Figure 1.3: Alternative reading paths for the chapters in this thesis.

This thesis consists of six research chapters containing our core contributions plus a concluding chapter. Each of the research Chapters 2 to 7 can be read individually, as the contents of these chapters is largely independent of the other research chapters (except in the Related Work section where a dependency is explicitly specified). However, we recommend reading each part chronologically for a more complete coverage of each research theme. The alternative reading paths are shown in Fig. 1.3. Additionally, reading only this introduction chapter and the conclusions in Chapter 8 gives a dense summary of the whole thesis, and provides answers to the research questions. A summary of the research chapters is given below:

Part I: Context-awareness

Chapter 2 - Contextual Factors in Location-Aware Multimedia Messaging: We address the contextual factors surrounding location-based media production and consumption (El Ali et al., 2010). We present a pilot study evaluating a prototype Location-aware Multimedia Messaging (LMM) system, and a longitudinal multimodal diary study (*LMM Study*) to gather insight into the relevant contextual factors that influence users to use such systems in urban settings. The findings led to a set of design recommendations for designing and developing future LMM systems. Additionally, the observations made in this chapter serve as the groundwork for the analysis presented in Chapter 3.

Chapter 3 - Case Study: Designing for Playful HCI: Based on the findings of the *LMM Study*, this chapter presents an in-depth case study (El Ali et al., 2011) to address three design problems for inferring, maintaining, and measuring playful urban experiences. The chapter discusses in detail each of these design problems, and arrives at a set of design considerations for how each problem can be addressed to optimize playful interactions in urban location-based media production and consumption.

Chapter 4 - Automatic Exploration-based Route Planning: We address the problem of how to support city residents and tourists wishing to explore a city (El Ali et al., 2013b). We build an exploration-based route planner that leverages 5 years of geotagged photos taken from the photo sharing website Flickr (*Route Planner Study*). We evaluate our generated exploration-based route plans through a controlled laboratory study, as well as through a web survey. Drawing on experience questionnaire data, web survey responses, and user interviews, the findings led to a set of design recommendations for going towards automatic data-driven approaches to support city exploration, and the role different digital information aids play in supporting such exploration behavior.

Part II: Gestural Input Techniques

Chapter 5 - Effects of Error on Device-based Gesture Interaction: We focus on the usability and user experience of 3D gesture-based interaction (El Ali et al., 2012); specifically, what happens when 3D gestures performed by users are not recognized (*Gesture Errors Study*). We present a controlled laboratory study to arrive at which gesture set is more robust to recognition errors under varying error rates. Drawing on experiment logs, video observations, participants feedback, and a subjective workload assessment questionnaire, the findings led to a set of design recommendations on which gesture sets provide the best overall user experience, and thus most suitable for inclusion into today's smartphones.

Chapter 6 - Playful 3D Gestural Interaction: Revisiting the issue of supporting playful urban interactions presented in Chapter 2, this chapter makes use of a novel gesture-based paradigm in mobile HCI to provide such support (*Playful Gestural Interaction Study*). This chapter introduces 3D gestural interaction and its applied use in a playful, music-related context (El Ali and Ketabdar, 2013). Using three musical applications (Air Disc-Jockey, Air Guitar, Air GuitaRhythm), this chapter investigates whether 3D gestures can be effectively used to support playful music composition and gaming on mobile devices. Based on results from a controlled user study (measuring usability and user experience questionnaire responses, users direct feedback, and video observations), we arrive at a set of design recommendations to inform the design of future music-related smartphone applications that make use 3D gesture-based interaction.

Chapter 7 - Usability and Security Tradeoff in 3D Gestural Authentication: Another common task that pervades urban interactions is user authentication. Using the 3D gestural authentication technique introduced in Chapter 6, this chapter (*Gestural Authentication Study*) investigates the usability and security trade-off in 3D gestural authentication. To investigate this tradeoff (El Ali et al., 2013a), we replicate a controlled security study to assess the vulnerability of this authentication method against video-based shoulder surfing attacks. Additionally, in a separate controlled study, we measure the usability and user experience issues associated with performing air gestural signatures for smartphone security access. For this, we measure user experience using experience, subjective workload, and trust questionnaires as well as analyze performed signature durations. Our security analysis provides further validation of the security of this authentication method, and with our user experience research, we arrive at a set of design recommendations for optimizing the user experience of using this magnetic gestural authentication method.

1.5 Thesis Origins

The work presented in this thesis is based on a number of research papers, all of which are listed in the bibliography. These are:

Research Chapters

- **Chapter 2:** El Ali, A., Nack, F. & Hardman, L. (2010). Understanding contextual factors in location-aware multimedia messaging. In Proceedings of the 12th international conference on Multimodal Interfaces (ICMI-MLMI '10), 2010, Beijing, China.

- **Chapter 3:** El Ali, A., Nack, F. & Hardman, L. (2011). Good Times?! 3 Problems and Design Considerations for Playful HCI. In International Journal of Mobile Human Computer Interaction (IJMHCI), 3, 3, p.50-65.

- **Chapter 4:** El Ali, A., van Sas, S. & Nack, F. (2013). Photographer Paths: Sequence Alignment of Geotagged Photos for Exploration-based Route Planning. In proceedings of the 16th ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '13), 2013, San Antonio, Texas.
- **Chapter 5:** El Ali, A., Kildal, J. & Lantz, V. (2012). Fishing or a Z?: Investigating the Effects of Error on Mimetic and Alphabet Device-based Gesture Interaction. In Proceedings of the 14th international conference on Multimodal Interaction (ICMI '12), 2012, Santa Monica, California. [Best student paper award]
- **Chapter 6:** El Ali, A. & Ketabdar, H. (2013). Magnet-based Around Device Interaction for Playful Music Composition and Gaming. To be published in International Journal of Mobile Human Computer Interaction (IJMHCI).
- **Chapter 7:** El Ali, A., Ketabdar, H. & Nack, F. (2013). Investigating the Usability and Security Trade-off in Magnetic Gestural Authentication. Under peer-review in International Journal of Human Computer Studies (IJHCS).

Other Publications:

- Wolbert, M. & El Ali, A. (2013). Evaluating NFC and Touchscreen Interactions in Collaborative Mobile Pervasive Games. Poster accepted at MobileHCI '13. Munich, Germany.
- Bouwer, A., Nack, F. & El Ali, A.. (2012). Lost in Navigation: Evaluating a Mobile Map App for a Fair. In Proceedings of the 14th international conference on Multimodal Interaction (ICMI '12), 2012, Santa Monica, California.
- Holopainen, J., Lucero, A., Saarenp, H., Nummenmaa, El Ali, A., & Jokela, T. (2011). Social and Privacy Aspects of a System for Collaborative Public Expression. In Proceedings of Advances in Computer Entertainment Technology (ACE'11), Lisbon, Portugal.
- El Ali, A., Lucero, A. & Aaltonen, V. (2011). Multimodal Interaction Design in Collocated Mobile Phone Use. In MobileHCI '11. MobileGestures workshop, 2011, Stockholm, Sweden.
- El Ali, A. (2011). Studying and Designing for Mobile Social Awareness Cues in Urban Interactions. In MobileHCI '11 Extended Abstracts (Doctoral Consortium), Stockholm, Sweden.
- El Ali, A., Nack, F. & Hardman, L. (2010). Good Times?! Playful Aspects of Location-based Experience Capture. MobileHCI '10. Please Enjoy workshop, 2010, Lisbon, Portugal. [Nominated for best workshop paper]
- Nack, F., El Ali, A., van Kemenade, P., Overgoor, J. & van der Weij., B. (2010). A story to go, please. In Proceedings of the 3rd International Conference on Interactive Digital Storytelling (ICIDS '10), 2010, Edinburgh, Scotland.
- El Ali, A. & Nack, F. (2009). Touring in a Living Lab: some methodological considerations. MobileHCI '09. Mobile Living Labs 09 workshop, Bonn, Germany.

Part I

Context-Awareness

2 Identifying Contextual Factors in Location-Aware Multimedia Messaging

We investigate the contextual factors surrounding location-based media production and consumption. We present a pilot study evaluating a prototype Location-aware Multimedia Messaging (LMM) system, and a longitudinal multimodal diary study (*LMM Study*) to gather insight into the relevant contextual factors that influence users to use such systems in urban settings. The findings lead to a set of design recommendations for designing and developing future LMM systems. The work presented in this chapter was published as "Understanding contextual factors in location-aware multimedia messaging" in the proceedings of the 12th International Conference on Multimodal Interfaces (El Ali et al., 2010).

2.1 Introduction

In this chapter, we identify through an exploratory approach the contextual factors surrounding the production and consumption of location-aware multimedia messages (LMMs), with the aim of eliciting implications for the study and design of future LMM systems. Examples of these multimedia messages (MMs) include geo-tagged photos, text, video, audio. These LMMs are anchored to a location by some person, which can be perceived and interpreted by recipients by being at (approximately) the same place where the message was made. Given that locations within cities are rich sources of "historically and culturally situated practices and flows" (Williams and Dourish, 2006, p. 43), it is reasonable to assume that LMMs can reflect culturally entrenched aspects of people's experiences and make them visible at locations. To this end, we argue that an experience-centered framework is necessary to talk about and identify the contextual factors surround LMM.

How do users understand and make use of these LMM systems? What usability issues do emerging LMM technologies give rise to? Is our experience-centered framework suitable for addressing these usability issues? What are the relevant real-world contextual factors involved in creating multimedia messages at locations, and how can these inform the study and design of future context-aware LMM systems? In this chapter, our goal is to address the contextual factors surrounding LMM production and consumption in urban settings.

2.2 Research Questions

Our main research question in this chapter is:

> **RQ 1:** How do users create and interact with urban location-based content, and how should this impact the design of future Location-aware Multimedia Messaging (LMM) systems?

Specifically, which contextual factors are involved in LMM production under an experience-centered framework? And what are the implications for studying and designing future LMM systems? To answer these questions, we adopt an exploratory approach, one that is amenable to the subjective nuances of everyday human cognition and affect.

To build our experience-centered framework, we distinguished between two aspects of an experience: process and memory. An experience process (cf., Nack (2003)) is a sensory and perceptual process that some person undergoes (through direct participation or observation of events and situations) that results in a change in that person. Given the high variability in computationally modeling and predicting the process of an experience, here we look mainly at the memory of an experience. Based on the definition of episodic memory given by Tulving (1993), we define an experience memory as the result of an experiential process, which can be manipulated and actively recalled. It consists of one or more actors, and spatiotemporal, social, cognitive, and affective aspects. We use these aspects of an experience memory as a framework for studying LMM. To this end, we make a contribution towards identifying which contextual factors are important in studying LMM systems, and what kind of experience-enhancing mechanisms need to be supported.

The rest of this chapter is structured as follows: first, we provide a review of related work. Next, we describe our LMM prototype and the lessons learnt in using it in a pilot study. Then, we describe our multi-modal diary study, the category attribution task that was necessary for analyzing the diary results, and discuss the assimilated results. Finally, we draw design recommendations for the study and design of LMM systems and conclude.

2.3 Related Work

2.3.1 Sharing Experiences

Bentley and Metcalf (2009) investigated how to increase mobile social presence by mediating people's experiences through awareness of daily activities. They developed multiple mobile probes that relay either motion presence (whether a friend is moving or stationary), music presence (sending metadata of a song a person was listening to via SMS), and photo presence (immediate upload of photos to a person's feed during real-time phone conversing). They found activity sharing to be an effective means of sharing experiences.

Appan et al. (2004) investigated how to communicate everyday experiences using mobile devices by structuring everyday happenings using narrative structures. They found that imposing narrative structures was too rigid due to the 'unstructurable' nature of everyday experience content and associated annotations. They used an interactive event-based framework instead to elicit structured interaction for consumption of everyday experiences. As we show later, using an event-based framework may also be insufficient for capturing everyday experiences. A different approach was taken by Ståhl et al. (2009), where they placed an armband sensor around users' arms for days to collect movement and arousal levels which can then be transferred and visualized in their Affective Diary system. This system functioned like a real diary, where the highlighted emotion-color visualizations corresponded to different affective states of users' daylong activities, which subsequently made users more aware of their daily experiences.

Jacucci et al. (2007) used an ethnographic approach to understanding the social aspects of experience capture and sharing at a rally championship event. Their approach revealed a number of useful experience sharing implications including the importance of distinguishing between socially mediated multimedia expressions and expressions used as personal records, shared versus individual memory, and the importance of taking into account current situation-dependent factors communicating records of experiences. In short, the foregoing approaches provide support for the goal of handling everyday 'unstructured' experiences at locations.

2.3.2 Location-aware Messaging Systems

Previous work has focused primarily on location-aware systems that allow users to leave textual messages such as reminders or post-it notes at locations (Burrell and Gay, 2002; Griswold et al., 2003; Persson and Fagerberg, 2002; Sohn et al., 2005). While these systems support only text, GeoMedia (Papliatseyeu and Mayora Ibarra, 2008) permits attaching multimedia messages (as images, audio or video) to locations. The GeoMedia system however lacked a thorough user evaluation, leaving a gap to be addressed in the study of LMMs, and how they relate to experiences in mobile and ubiquitous environments.

The Place-its system (Sohn et al., 2005) was designed to study how location-aware reminders are used throughout a person's day, the relative importance of locations for reminders, and the effects of reminder message positional accuracy on the reminding process. While reminders may serve as triggers for experiences, the scope is rather narrow. The ActiveCampus application (Griswold et al., 2003) provided insights into how people living on a campus would use such location-aware messages, however, the restriction to a textual medium and an academic surrounding is insufficient for understanding the range of human experiences in everyday settings.

Both GeoNotes (Persson and Fagerberg, 2002) and E-graffiti (Burrell and Gay, 2002) were extensively studied in real-world usage contexts. Studying each provided insight into how people conceived of location-aware systems, the perceived usability of their location-aware functionality, and the relationship between an information and physical space. As in E-graffiti, we are also less interested in tackling the technical problems of context detection, but rather to focus more on evaluating user reception of a location-aware messaging system. Specifically, we want to focus on interesting and novel uses of such a system, and how that can enrich the human experience of being at a media-rich location. However, whereas GeoNotes and E-graffiti were existing application prototypes which were committed to certain design decisions (e.g., in GeoNotes commenting within a note or content-searching using a word-based search engine), we are more interested in the human perceptual conditions involved in LMM with sufficient flexibility to avoid commitment to any one design.

Put differently, our work differs in that we are interested in users' perception of how such systems should be or look like, and not in their reaction to committed design rationales. For example, in Geonotes, the connection between a space and a note was defined explicitly using place-labels, while for us we wanted our users to inform us about the causal relationship between media messages and the entities in a space. Also, while GeoNotes committed to certain types of metadata, we are interested in seeing what kind of metadata people would firstly fill in and then later desire to consume. Finally, we wished to study multimodal capture behavior that made use of various types of media (including but not limited to videos, songs, images), and not only location-aware text messages.

2.4 Pilot Study

To understand the experiential factors surrounding LMM, we took a developed prototype application that allows the annotation of locations using three different media types (text, drawing, and photos). The prototype was pilot-tested with 4 participants where an *in situ* interview method (Consolvo et al., 2007) was used to observe experience capture behavior. By annotating locations, the prototype allows users to capture their experiences, i.e., create a digital memory of an experience (Fig. 2.1(a)). The generated message remains anchored to the location it was created at for later viewing by anyone who has the application installed on their multimedia-enabled mobile device and is at the same place where the message was created.

2.4.1 LMM Prototype

Generation: The prototype application was installed on the Android Dev Phone 1. The initial screen consists of three functions: Create, Snap, and Explore. In Create, a user can create a free drawing (Fig. 2.1(b)) using touch-based input or type text using the device's keyboard. Here, the location and orientation of the device is retrieved and the user is presented with a camera-view where she can choose to draw or write something. In choosing either option, a snapshot of the camera view is subsequently used as a background canvas for the user to draw or write on. Once a user is finished, the annotated image can be saved. In Snap, a user is taken directly to a camera-view where she can snap a photograph.

Presentation: To view a message, a user has to be at the right position and orientation. In switching to Explore mode, a user is presented with a camera-view, where she is guided to a message by leading her to the creator's original position and orientation. An arrow is drawn on the screen to guide the user towards a message. To indicate the distance between the user's current position and that of the message, the color of the arrow changes within 200m of the message location. Once at the right position, the user can adjust her orientation by looking at a small green indicator arrow shown on the right or left edge of the screen. In doing so, the selected media message is overlaid on top of the camera-view (Fig. 2.1(c)).

2.4.2 Lessons Learned

While the approach of using a developed prototype provided direct user-feedback on experience capture, all the tested participants expressed that they had insufficient time to

(a) Planning at t_0 (b) Creation at t_1 (c) Viewing at t_n

Figure 2.1: Interaction with the prototype.

satisfactorily express themselves. Moreover, since the prototype was at its early design stages, users, in capturing their experiences using the provided media forms (drawings, text, photos), were limited by the presented technology. This created an 'experimental straw man[1]', where it was now unclear what kind of experience-eliciting behavior was being measured: did the users feel that their created LMMs were intrinsically tied to the existing functionality and interaction methods offered by the prototype application, or did they understand that the application was merely a probe into informed user-centric development of future context-aware LMM technology? These concerns are not new: previous work has addressed possible confounds in using location-aware messaging technology in its earlier stages (such as short battery life of the PDAs used by Griswold et al. (2003) or the sluggishness and effort required for carrying laptops to make messages in the work by Persson and Fagerberg (2002)). The limitations encountered in previous work and the problems that surfaced in the pilot study led us to revise the chosen method in favor of one that allows understanding LMM behavior for a longer duration and without predisposing users to the functionality and interaction modes of existing technology.

2.5 Multi-modal Diary Study

The lessons learnt from the pilot study resulted in a redesign of the investigation method. To alleviate the pilot study limitations, we set up a longitudinal multi-modal diary study (Amin et al., 2009; Ståhl et al., 2009) in order to investigate the contextual factors surrounding LMM production and consumption.

[1] A straw man is a reasoning fallacy that occurs when an opponent's position is misrepresented. To attack a straw man is in fact to create an illusion of having refuted a given proposition by attacking a superficially similar proposition (the straw man). For us, we adapt the straw man notion to describe misplaced measurement of something superficially similar to what actually should be measured.

2.5.1 Participants

Eight participants (6 male, 2 female) aged between 13-27 (M= 23; SD= 4.4) were recruited for the diary study. All participants were in their 20's, except for S6 who was 13 years old. The reason behind recruiting a young participant was to accommodate a different attitude to technology. Five of the participants had completed their bachelor's studies, one her master's studies, one pre-master's studies, and finally S6 had completed the first year of high-school. Three of the participants owned a smart mobile device. All however were familiar with viewing multimedia on such devices and GPS usage. All but S5 declared themselves as social, outgoing people.

Q1	Where are you right now?
Q2	Please explain why you made the media message at this place.
Q3	Please describe how you are feeling right now. (e.g., happy, sad, anxious, excitetd, lazy)
Q4	Please describe the environment around you.
Q5	Who are you with right now?
Q6	What were you doing before you made the media message?
Q7	Is there an event going on where you are (e.g., sunset, festival, live band, market, dinner)? If yes, please describe the event.
Q8	If yes to question 7, are you participating in this event, or did you only observe it?
Q9	If yes to question 7, is this the first time you participate/observe such an event?
Q10	Were you able to express what you wanted? If not, please state why you couldn't.
Q11	Was there something specific in the environment that you directed this message at? If yes, please state what it is.

Table 2.1: The second set of questions asked in the diary that pertain to the participant and her context.

2.5.2 Materials

Materials consisted of an information brochure, 8 custom-designed paper diaries, and a set of post-study interview questions. The diaries were custom-designed so that the diary each participant had to carry looked professional and hence would make participants take the study more seriously, in addition to ensuring that study questions were available for easy look-up. The diary included 2 pages of instructions and 2 pages that contained the 'question template': a set of questions that each participant had to answer after making a message. The question template was split into two parts: questions about the message made and a set of questions about the participant and her context. The first set of questions were: date, time, message media type (drawing, text, photo, video, audio recording, other), title of message, and whether the message is public or private. The message questions (see Table 2.1) were about: spatiotemporal aspects (Q1, Q4), social aspects (Q5), affective (Q3) and cognitive aspects (Q2, Q7, Q8, Q9, Q10, Q11). The interview consisted of the following questions: difficulty faced in filling in the diary, inspiring days and locations, media preference, environment awareness and overall experience of the past week, willingness to use a future context-aware LMM application, desire to view and write message metadata, and further participant additions.

2.5.3 Procedure

After reading the information brochure, participants were asked to fill in a personal information form along with a permission statement that permits the analysis and usage of their data. Afterwards, each participant was given a short demo of the LMM prototype, and asked to make two messages with it. This was done as a cautionary measure (as highlighted by Burrell and Gay (2002)) to ensure that participants understood what was meant by location-aware functionality. Each participant was given a personal diary and an oral explanation about the requirements of the study. Participants were required to carry the diary with them for approximately one week. They were asked to make a MM (photo, video, text, drawing, song or audio recording) twice per day, so that by the end of the week they had a total of 14 messages. Given the stringent nature of filling in the diary twice per day, participants were told that they are allowed to make 3 messages per day if they so desired, at the cost of a message on another day. This was done in order to make the testing conditions as natural as possible, under the assumption that there are days where one is more inspired to make messages than others.

The messages made by participants were restricted to public places, loosely defined as any place outside of their own homes. Upon making a message, participants were asked, if possible, to immediately answer the questions provided in the 'question template' in the diary. Since participants may not possess the necessary media capturing device at the time of making a message (e.g., a video camera), they were asked to instead provide an image-based or textual description as a surrogate for the actual message (e.g., a textual description or series of images depicting what a participant's video shot would have captured). At the end of the study, participants were asked to provide the actual MM either by e-mail or directly through a USB flash drive, return the diary, and sit through a ∼10 min. interview. Each interview was captured by means of a tripod-anchored digital camera. After the interview, as motivational measure, each participant was awarded a €20 note and thanked for their participation.

2.6 Categorization Task

The diary study resulted in 110 user-generated messages, where the interpretation of these was subjective. To understand the motivations offered behind the made multimedia messages, we categorized participants' motivations into domain (to what domain does a given location-aware message belong; e.g., entertainment, architecture) and task (for what purpose or task was the message created for; e.g., appreciation, criticism) categories. To ensure the domain and task categories we chose reliably group participants' message motivations, we needed to account for inter-coder reliability. Therefore, we set up a secondary categorization task that required participants (distinct from the participants tested in the pilot and diary study) to categorize the motivation responses provided by the diary-study participants. In order to decide on the best approximate categorization, a voting “winner-takes-all” procedure was applied where a message-classifying category with the most votes wins.

2.6.1 Participants

Six participants (3 male, 3 female) aged between 24-29 (*M*= 26; *SD*= 2) were recruited for the category attribution task. All participants had completed their bachelor’s studies.

2.6.2 Materials

The materials for the category attribution task were the 110 message motivations (i.e., why participants chose to make the message at a given place) and their corresponding media type, made by the 8 diary-study participants .

2.6.3 Procedure

Participants were contacted through e-mail, where they were provided with 110 message motivations and their corresponding media type. They were asked to categorize each message under both domain and task categories (Fig. 2.2 & Fig. 2.3), where multiple categories can classify a message. The first set of domain and task categories identified were used as exemplars for subsequent classification. However, if an exemplar category did not suitably classify a given message, participants were allowed to create new categories as needed.

2.7 Results & Discussion

The diary-study proved to be a powerful low-fidelity mechanism for studying LMM in real-world contexts without the intrinsic bias evoked from using an existing yet incomplete technology. From the 8 participants, 2 of them completed only 13 messages, which resulted in a total of 110 MMs. The results of the categorization task provided the basis for further analyzing the diary study data, where the categorization task results were directly assimilated into the diary study results. An equal number of responses to two distinct categories resulted in classifying the message as belonging to both. Below, we present and discuss participants’ media preferences, the identified domain and task categories, the difference between captured experiences and the experience of capture, the different aspects of captured experiences (using out experience-centered framework), and the relevant post-study interview responses given by participants.

2.7.1 Media Preferences for MMs

To identify what media types should be supported in LMM tools, participants were asked about their media preferences. From the 110 messages, the most prevalent media types were: photos (45%), text (24%), and songs (13%). The other media types (namely, video and audio recording), were each less than 10% of the total messages made. The lack of video recordings could have been due to the non-availability of the media capture device (e.g., handheld video camera). Only one participant made use of multiple media in a given

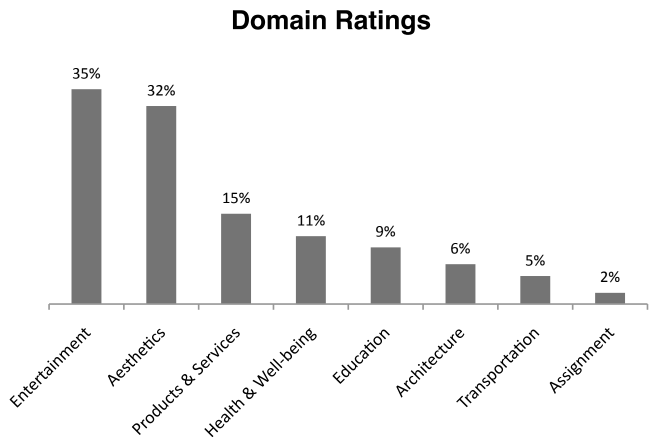

Figure 2.2: Distribution of domain categories (total = 114%) rated by participants (N=6) for 110 messages.

message, namely 'photo + text' pairs. Not surprisingly, the most chosen media type was photos, which require little cognitive effort to make. For photos made at locations, they can only give a unique perspective on the location, given the high iconic correspondence between a photograph of something at a location and the location itself. As one participant stated when asked about his media preferences *"In the beginning, it was photos, and during the week, because it wasn't that interesting, I used more text."* Indeed, if the location is not interesting or does not offer any unique perspectives to share with others, then a symbolic medium such as text can be used to express something beyond the qualities of the location itself. Also of interest is P2's remark on using songs because places can remind one of songs, but also because songs themselves can become surrogates for the memory of a place.

2.7.2 Identified and Rated LMM Domains and Tasks

From the initial set of identified domain categories, only 4 out of the 110 messages were problematic to classify. Upon closer inspection, the reason was due to messages where participants saw it as a duty to make a message (e.g., "Because I had to"). This led us to create an extra 'noise' category: Assignment. Indeed, such problems with participant motivation are sometimes unavoidable during requested study participation (Burrell and Gay, 2002). The highest density of messages fell into the Entertainment (35%) and Aesthetics (32%) domains (Fig. 2.2). Here, aesthetics was defined as something that offers sensori-emotional value (e.g., a beautiful scene), whereas entertainment something that offers amusement (e.g., a film). Only 17% of aesthetic messages were also classified as belonging to the entertainment domain, indicating that there is indeed a distinction to be made. Products & Services (15%) and Health & Well-being (11%), comprising around a third of total messages, are also domains typical of everyday experiences. Overall, the majority of the messages were about the entertainment and aesthetic domains.

Coincidentally, only 4 out of the 110 messages were difficult to classify into task

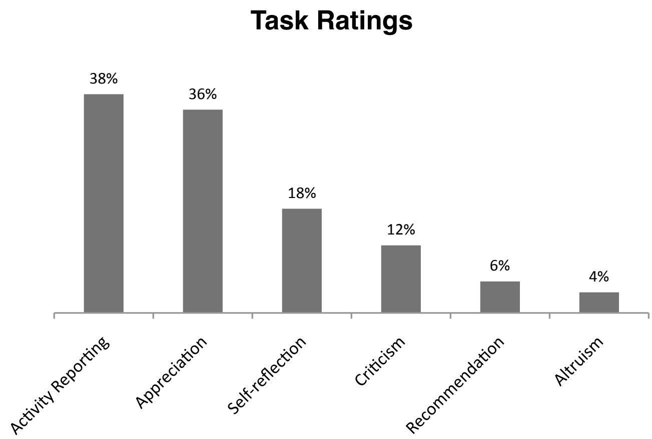

Figure 2.3: Distribution of task categories (total = 113%) rated by participants (N=6) for 110 messages.

categories. Here, the divergence was mainly between classifying messages as belonging to Appreciation or Criticism. For example, the motivation behind some messages (*P8: "It looks sad with the snow."*) can indeed be understood as both an appreciation statement and subsequent criticism of the state of affairs. Most messages were classified into the Activity-reporting (38%) and Appreciation (36%) task categories (Fig. 2.3). Activity reporting[2] (i.e., reporting to people what you did) and Appreciation (i.e., enjoying the qualities of something) paralleled mostly the classification of messages into Aesthetic and Entertainment categories, where Activity reporting messages fell into Entertainment, and Appreciation messages into Aesthetics. Self-reflection (i.e., reflecting on one's own actions or feelings) (18%) indicated that reflecting on one's self is also typical of everyday experiences that warrant capturing. Overall, the findings show that activity-reporting and appreciation were the most prevalent task categories.

2.7.3 Captured Experiences vs. Experience of Capture

When comparing the analyzed data with our participants' experience with the diary during the post-study interview, the importance in distinguishing between captured experiences (i.e., experience memory) and the experience of capture itself (i.e., experience process) became clear. Whereas captured experiences are information 'about' an experience (*cf.,* the answers to the diary questions), the experience itself is a process emergent from an undertaken activity (*cf.,* the actual experience of using media capture devices and filling in diaries to capture experiences)). For capturing experiences, the aim is to provide an adequate *representation* of a real-world experience that took place (e.g., a community-rated image-based experience sample of a person who parked her bike to photo capture something special in the surrounding scenery). For the experience process however, the

[2]Activity-reporting is broader than citizen journalism, which describes activities that pertain to nation- or worldwide events, and not necessarily personal events.

aim is to subject users to conditions in everyday settings that would strongly correlate to (if not cause) a desired type of experience while *interacting* with a system (e.g., equipping an LMM system with an adaptive notification system that learns never to interrupt users about newly created content while driving vehicles).

For the latter, the concern is less about what context is needed to sufficiently re-contextualize the experience of others, but instead about the scoped interaction between the user(s) and the system, where the user experience takes place during the interaction process itself. For the experience process then, we feel the emphasis should be on modeling the user and anticipated interaction with the system. This requires accounting for not only (context-dependent) multimodal input and output support (Chittaro, 2009), but also the extent the system can make sound predictions about a user's current state to sustain and enhance the flow of interaction (Kapoor and Horvitz, 2008). For example, captioning a LMM such as a photo through textual input might interrupt the user's current experience, whereas a voice command label that achieves the same function may occasion a more seamless interaction experience. Additionally, for notification, the system would need to temporally adapt to when users would be most receptive to receiving LMMs, so that the notified LMM can intersect itself gracefully between the user's cognitive and digital life.

2.7.4 Captured Experiences

To understand the different facets of LMM, the results of the diary questions were clustered according to the different aspects of an experience. These are discussed in detail below.

Spatiotemporal Aspects:
For the spatial aspects, participants were asked about where they were when they made a message (Q1, Q4), giving an indication about their experience at a place. This resulted in the following grouping: Urban (39%), an outdoor setting in the city, such as being on the street; Public Place (21%), an indoor public place such as a café or bar; University/School (17%); Nature (7%), being at a park or nature reserve; Friends/Family Home (6%), at the home of a friend or family member; Home (6%); Transport Vehicle (3%), inside a transportation vehicle such as a tram or metro. Most messages were made in an urban setting, public place, or at the university,[3] providing an indication as to the kinds of places future experience-capture technology would be used in. Also, despite that participants were asked to make messages outside of their homes, a few did not comply, which were classified under the Home category.

For domain and task dependencies in an urban setting, most urban messages fell into the aesthetics domain category (62.9%) and appreciation task category (49%), which highlights the tight correspondence between being outdoors and aesthetic appreciation. Not surprisingly, when controlling for a university/school setting, many of the messages fell into the Entertainment domain category (42%) and Activity Reporting task category (53%), which shows that using such a technology in an academic setting does not necessarily pertain to education. Finally, many of the messages were about Activity Reporting (39%) when controlling for Public Place, which is reminiscent of micro-blogging behavior (e.g.,

[3]While arguably a university/school is a public place, the distinction was made here to highlight possible differences between making a message in a non-academic setting and an academic one.

Twitter[4] feeds).

With respect to the spatiotemporal aspects, we were only interested in whether certain days affected participants' LMM behavior, and not in specific dates and times. For participants' behavior, P3 and P5 almost exclusively made messages in an urban outdoor environment (78.6% and 71.4%, respectively). Curiously enough, when these participants were later interviewed about whether there were more inspiring days (temporal dimension) or locations (spatial dimension) in making a message, they reported the following: *P3: "Yes, not a particular day, but of an inspiring moment [asked about location] I wouldn't say it was because of the location, it was a matter of coincidence"; P5: "Yes, definitely the weekend [asked about location]; yes, I found that I like changes in my everyday routine places, and when I encounter something that I like a lot that's changed, that's something that inspires me but doesn't happen everyday."* From the 8 participants, 3 of them stated that the location did not provide a source of inspiration, but rather it was coincidental inspiration. However, for inspiring days, all participants agreed that events provided a source of inspiration, where events included their weekend activities, such as going out for a drink. Overall, these findings are consistent with the findings of Sohn et al. (2005) and Burrell and Gay (2002), who found that the location, in and of itself, is perhaps not an essential part of context, though certainly useful as a trigger for an experience.

Social Aspects:

Participants were asked about whether they wanted their MMs to be public (visible to anyone at approximately the same location it was made) or private (viewable to only specified networks) (Burrell and Gay, 2002), as well as who they were with at the time of making the message (Q5). Most messages were made public (71%) and the rest private (29%). In analyzing who a participant was with, we defined a person as a single friend or family member and a group as a collection of friends or family members.[5] Nearly half of the messages were made while a participant was alone (46%), compared to being with a group (30%) and with a single other person (25%). However, this might reflect a participant's personality or age; for example P3 made all but one message when alone, and S4 and S8 made more than half of their messages alone (57%). By contrast P6, the 13 year old participant, made most messages while in a group (64%) – this may be because at a younger age, a teenager is usually surrounded by people at home and at school. In considering the domain and task categories for messages made alone, the highest percentage was for the Aesthetics domain (36%), and the highest percentages for the Appreciation (34%) and Activity-reporting (34%) tasks.

The foregoing results illustrate the difference between public and private messages, and messages made alone or with others. While alone-messages dominated our findings, most of these messages (76%) were nevertheless made public. This is in contrast to the findings of Burrell and Gay (2002), who found that notes posted voluntarily were mostly made private. This may be due to their misleading conceptual model that resulted in users treating the E-graffiti system as a limited e-mail system, where E-mail messages are generally addressed to a few private individuals. In assessing the dependencies between the social and spatial aspects, it was interesting to see that half of the messages made alone

[4] http://www.twitter.com; last retrieved: 01-08-2013

[5] Strangers also counted, but there had to be at least one friend or family member for group classification.

were made at an urban outdoor setting (50%), which also comprised half of the total number of song messages made (50%). While this may have been a coincidence, it is also not unlikely that when walking outdoors, participants still feel the need to record their experience, even if alone (e.g., *P3: "To have a memory of this special location"*).

Affective and Cognitive Aspects:
The mood responses (Q3) of participants were classified according to valence (positive, negative, neutral, ambivalent) and arousal (high, moderate, low), in accordance with the circumplex model of emotion (Russell, 1980). We used this model as an instrument for easy and relevant classification of participant's responses according to the valence and arousal dimensions. Most messages were made when participants were in a positive mood (46%) or highly aroused (46%), where only around half overlapped between these two factors (54%). Negatively valenced (32%) and low-arousal (33%) affective contexts were also prevalent in participants' responses, compared to neutrally valenced moods (16%) and ambivalently valenced moods (8%) on the one hand, and moderate arousal levels (22%) on the other. It was interesting to see a tendency between being alone and being in a negatively valenced mood (60%), whereas from all messages made in a group, most tended to be positive (55%). For the overlap between negatively valenced moods and being alone, the diary may have functioned as a cathartic outlet for them to express their negative mood, which is also typical of web 2.0 social behavior (Chapman and Lahav, 2008). This is further supported by the observation that most negatively valenced moods (74%) resulted in messages that were made public. Together, these findings highlight the variability in mood states in everyday contexts, which do not exhibit strong overlap between the location of the experience and the MM.

With respect to the cognitive aspects (Q2, Q7, Q8, Q9, Q10), we were mainly concerned with the causal trigger of a participant's experience (i.e., what in the perceptual environment captured a participant's attention and thereafter served as a trigger to make a message). Moreover, we wanted to investigate the causal dependencies, if any, that exist between prior participant activity and the created messages. Most messages did not surface a direct causal relation between prior activity and message creation (65%). Yet when there was such a direct dependency (36%), messages tended to fall into either Activity Reporting (39%) or Appreciation (28%). Related to the causal relation between prior activity and message creation were participants' responses during the post-study interview about whether or not the diary made them more aware of their daily environment. All participants reported that indeed it did make them more aware insofar as they had to plan where to make a message. As one participant (P2) reported, the diary, if it were a pervasive mobile tool, would not make a difference in raising awareness if it were embedded in daily life. In contrast, P8 stated that the diary was effective in raising awareness by making him contemplate over the beautiful parts of the city he normally takes for granted. This raises the question of whether continuous cognitive access should be designed in an experience-capture tool, so that deliberate planning behavior becomes the norm. Such a mechanism can serve as a persuasion tool to not only create meaningful MMs (cf., the work by Singh et al. (2009), where they use a game-theoretic approach to study selfish user media contribution behavior for designing user incentive mechanisms), but also to raise perceptual awareness of the daily environment.

For the trigger of a message (Q11), most participants reported that there was something in the environment they directed their message at (60%), however it was surprising that many said there was nothing they directed their messages at (40%). Closer analysis led to distinguishing between three types of triggers and subsequent message classification: Situation (57%), Object(s) (33%), and Person(s) (10%). Here, a situation was defined as a collection of objects that are a pretext for an event(s) or caused by an event(s). Given this typology, it was assumed that if a participant did not direct her message at something specific in the environment, then the trigger of the message was a situation. The high frequency of situation-triggers is consistent with the findings of Persson and Fagerberg (2002), who found that situation-related chat outnumbered object-related chat.

Post-study Interview Responses

Viewing and Adding Experience-based Metadata:
During the post-study interview, participants were asked about what kind of metadata (information similar to that asked in the diary) they would like to see if they were using an application that supported LMM. Afterwards, they were asked about their willingness to add this metadata themselves. 5 participants reported they would like to view such metadata, specifically to see the following: a person's mood, who that person was with, and the event, if any, that relates to the message. When asked about viewing metadata, P7 stated: *"Some information might be fun to have, like who a person was with, and what event is happening. I would like a context between the message and an event, because the event might no longer be there, and then you would not know it happened at a location, so then it might not make sense."* One participant expressed that he would like such information, but only upon request (*"Not at first sight, that would ruin my personal view of their message. But it should be available if wanted...why the message was made, what did the person want to express."*). The last two participants found it unimportant to view metadata other than standard attributes such as names, date, and time; *P6 [in response to what metadata s/he would like to see]: "Date and time would be nice to see so you know it's a winter photo, and for the private messages to see the name of the person so I know who it is."*

Alongside viewing experiential metadata, we also inquired about participants' preferred methods of being notified about messages at locations. After exposure to the diary for around 1 week, it seemed reasonable to assume they can tell us about their notification preferences, despite that the study's focus was not on MM notification. Notification in this context means adaptive filtering of messages to participants' current situation and interests. All but one participant mentioned they would like the future LMM tool to automatically adapt the presentation of messages to their current situation. Only 2 participants, P1 and P2, specified explicitly the kind of adaptation they would like: filtering by current mood and by date, respectively. The other 5 did not explicitly specify the type of filter, but stated that adaptivity would be the preferred method of handling the hypothetically large number of messages at locations. Despite that most participants did not have any clear idea how this would be possible, they mentioned that the application adaptivity should depend on the situation they are in, so that it does not become obtrusive (*P6: "If I'm walking, then I'd like to search myself, but if I'm biking, I'd like notification of what there is. For example, great nature photos."* This indicated that application adaptivity may be best considered as itself context-dependent. The one participant who did not endorse application adaptivity stated that s/he would like to make queries herself through a search function.

With respect to writing metadata, one participant (P7) mentioned s/he would fill this kind of information, four participants said it would be too much effort (P1, P2, P3, P5), and three participants (P4, P6, P8) said it is contingent on the situation. The latter case is typified by P8's response: *"If it would be of any use to me as a user, let's say I filled in 10 of these experiences, and it would say something about what I would like in particular, that would be a nice application to me, so it all depends on the use."* However, most participants (even the ones who thought it would be too much effort to fill in such information) stated that after some time, to make viewing messages more interesting, would start filling in the metadata. This indicates that the problem of filling in metadata can be partially alleviated if potential users are aware of the consumption benefits provided by the metadata (such as more fully grasping the original experience of the LMM creator).

To take a closer look at participants' metadata writing behavior, participants' responses were analyzed syntactically according to word count for two factors: the motivation description length for a created message and the environment description length (see Table 2.2). These two factors were chosen because they generally require elaborate responses to be contextually meaningful, and therefore are indicative of efforts from participants to fill in media metadata in general. It is interesting to notice the discrepancy in P3's motivation description length (98%), which is at odds with his later response of finding it takes too much effort to fill in the metadata, especially given his relatively high mean word count scores. Also interesting is P7's high discrepancy across environment description lengths (78%); when asked about filling in metadata, s/he said "It's difficult, but yes probably I would fill it in, actually these are a reasonable number of questions; like tagging who you're with, we do that already." While this kind of analysis gives an indication over participants' efforts and attitudes towards filling in metadata, it may be difficult to generalize these findings to real application usage.

	Motivation Length			Environment Desc. Length		
Participant	Mean	SD	SD %	Mean	SD	SD %
1	13.5	7.6	57	12.3	7.6	62
2	6.2	2.8	44	9.8	5.1	52
3	9.2	9.0	98	15.1	9.6	64
4	6.4	2.6	40	6.3	2.0	32
5	9.1	5.5	61	8.4	6.2	74
6	6.6	3.7	57	15.0	7.1	47
7	7.9	4.9	61	3.9	3.1	78
8	7.0	1.7	24	2.9	2.1	71
Mean	8.2			9.2		

Table 2.2: Syntactic mean description lengths, standard deviations, and percentage of standard deviations across participants for a) reasons provided for created messages b) description of the environment in which the message was created.

Potential Application Usage

Finally, during the post-study interview, participants were asked about what type of functionality and interaction they expect from future LMM tools. One participant drew the analogy between such a future application and the microblogging platform Twitter (*P1: "I* *would compare such a device to Twitter, so if there was a device that can instantly post to Twitter a multimedia message, that would be nice, also might be nice to have it just like a diary, to keep a record of what you've done or what you've seen."*). This latter part of his statement indicates the potential for LMM applications to behave like life-logging applications such as the Affective Diary by Ståhl et al. (2009). Participant responses tended to cover standard online social network interaction: All participants stated their preference for 'click and share'-type features, indicating that the easier the application used for sharing, the better. Also, nearly all participants mentioned they would like to comment on other messages (as in Facebook[6]). Related to this, participants expect to be able to edit their own messages as well as delete them. One participant (P3) stressed the importance of having an optional expiration date for messages, in reference to messages that concern a temporary problem that will likely be resolved in the near future (*P3: "To warn/alert others for glass on the street, which can cut into tires"*).

When asked whether or not they would actually use an LMM tool if one is available, all but one participant said they would. However, all reported that they would not make as many messages; only when the occasion arises for them to express something worth sharing. The one participant who reported that s/he wouldn't use it explained that s/he spent her whole life in Amsterdam, so if s/he would spend time abroad, then s/he would.

2.7.5 Study Limitations

There were two main problems with the diary study: first, making two messages per day for one week may impose an unnatural demand on participants. In other words, participants had to sometimes invest cognitive effort in making messages, resulting in noise (*cf.*, Assignment task category). Related to this, participants were not always able to immediately answer the diary questions (e.g., snapping a photo while walking outdoors), waiting instead until the next opportune moment to do so (e.g., reaching home). The second problem was the availability of media capture devices. Despite that participants were told to capture anything they wished, so long as they provided a description of what they wanted to capture, a few participants mentioned they could not express themselves because they lacked the right media-capture tools (e.g., handheld photo camera).

2.8 Design Recommendations

From the above findings, we derive a set of design recommendations for designing and studying LMM systems.

1. **Embedding playful elements can improve user engagement with LMM systems.** From the pilot study, our users raised the issue that the LMM prototype was useful, but not fun and engaging. Interaction designers should consider how to increase user engagement and the overall UX of LMM systems by embedding playful elements (e.g., through gamification (Deterding et al., 2011)).

[6]http://www.facebook.com; last retrieved: 01-08-2013

2. **Predominant domain and task categories in LMM.** Aesthetics and entertainment domain categories and appreciation and activity-reporting task categories predominate experience capture behavior (Sec. 2.7: Identified and Rated LMM Domains and Tasks). This provides a starting point for tailoring future LMM tools to the right target groups (e.g., park visitors, exhibition goers).

3. **Application personalization is itself context-dependent**: When participants were asked about filtering messages, many expressed they would like messages to be shown in accordance with their current situation (Sec. 2.7: Viewing & Writing Metadata). More importantly, the tension between self-initiated queries and application adaptivity was itself largely a matter of context. This highlights that future LMM applications should not only account for personalized content, but the personalization itself should learn from and therefore adapt to the user's context.

4. **Perceptual alteration of places can be mediated by explicit experience-capture planning behavior:** To make certain perceptual qualities of locations more salient, experience-altering mechanisms that persuade users to consciously plan their capture behavior should be explicitly designed into the system (Section 3.2.2 & 4.3.3: Affective & Cognitive Aspects). Whether this mechanism operates through the provision of implicit ambient feedback or functions as an explicit interaction method (e.g., through a competitive game), remains an open question.

5. **Location-based experience capture also means open access to all:** Given the prevalence of expressions made public irrespective of mood or social context (Section 4.3.3: Social & Affective & Cognitive Aspects), it seems open, public access is the default in future experience capture behavior. This means future experience-aware systems should take into account not only privacy concerns, but also support open-access functionality (e.g., public ratings or comments) when designing the mediation between creators and viewers.

6. **Location-based experience capture methods expected to follow basic social network behavior standards:** Users have ingrained expectations about current experience-centric technology (Section 3.2.2 & 4.3.4: Potential Application Usage). Users mentioned they would like to be able to comment on expressions (e.g., locations or buildings) and evaluate them. Moreover, they would like expressions to have an optional expiration date and be able to edit and delete them. For accessing expressions, they desired a non-obtrusive notification system. Together, these show that location-based experience capture should comply with basic social network standards.

2.9 Conclusions

We have taken preliminary steps towards understanding the contextual factors surrounding LMM behavior and how that relates to capturing and consuming experiences. Using an exploratory approach, we were able to derive implications for the study and design of future LMM systems. The collected data in the diary study hinted at the inherent complexity and multidimensional nature of everyday human experiences, where subjective reports did not always offer patterned clues into how to build technology that can support capturing and communicating experiences. Nevertheless, this complexity provides further support for the importance of studying users and their behavior under real-world contexts.

From our findings on how users interact with LMM systems, it became evident that the primary goal for users was to use the LMM system to facilitate playfulness (see Section 2.8). However, the simplicity of the context-awareness in the LMM prototype and AR presentation (i.e., the designed minimal mobile interactions) were limited in providing the desired playful experience that users wanted. To investigate this limitation, we built on the findings from this study and used them as a case study to analyze in the next chapter (Chapter 3) in more analytical detail how to design effective playful interactions.

3

Case Study: Designing for Playful HCI

Based on the findings of the *LMM Study*, this chapter presents an in-depth case study to address three design problems for inferring, maintaining, and measuring playful urban experiences. We analyze each of these design problems, and arrive at a set of design considerations for how each problem can be addressed to optimize playful interactions in urban location-based media production and consumption. The work presented in this chapter was published as "Good Times?! 3 Problems and Design Considerations for Playful HCI" in the International Journal of Mobile Human Computer Interaction (El Ali et al., 2011).

3.1 Introduction

Figure 3.1: A mockup illustrating the photo Nicole took of the Corniche seaside and the corresponding annotations she added.

On a sunny afternoon in mid-July, Nicole and Nick are tourists shopping around Nejmeh Square in downtown Beirut, Lebanon. While Nick insists on seeing the cultural offerings of Saifi Village, a village completely rebuilt as a New Urbanist-style neighborhood after it's *destruction during the civil war, Nicole has a different notion of what is fun and enjoyable. Familiar with her interests in warm, foreign cities, Nicole's mobile device sets her to experience 'fun' places nearby, suggesting several lively cafés along the Corniche, a seaside walkway with a glittering view of the Mediterranean. Skeptical about the suggestion, she makes a predefined gesture instructing her device to show her different multimedia (photos, songs, videos, text) that reflect people's experiences there. The device presents her with a dizzying nexus of visual and musical perspectives captured by people enjoying themselves, complementing each multimedia message with related past and future events. Leaving Nick, she makes her way toward the Corniche until she reaches a café, where she sits outdoors, happily absorbing the scorching sun rays. Wondering where Nick went, she decides to capture her current experience. She takes a photo of the clear blue sky and sea (Fig. 1), which she annotates with the song by The Cure 'Play for Today' and writes: "That's New Urbanist-style culture too!!" While she awaits her hookah and drink, she scans through other people's experiences at the café she is at, only to realize the place attracts mainly an older crowd, which is no fun at all.*

The preceding scenario illustrates the need for defining computational methods that facilitate tourists with contextualized and media-based access to information while they freely explore a city. As discussed in Chapter 1 under Section 1.1.2, the provision of contextualized information anytime, anywhere, to the right persons as they go about their daily lives is part of this emerging paradigm known broadly as context-aware computing (Dey et al., 2001), where a major step towards achieving that vision has been the widespread use of location-aware technologies such as GPS-enabled smartphones. Yet with our cities becoming interfaces for computational experimentation that are intermixed with human activities, we need systems that go beyond location-awareness and towards context-awareness. In other words, we need to know more about context (Dey, 2001), its inference from human activity, and how that feeds into our everyday experiences. As Bellotti and Edwards (2001) state, inference and adaptation to human intent in context-aware systems is at best an approximation of the real human and social intentions of people. This raises the need to further explore the kinds of services and usability issues brought forth under real-world usage contexts.

One important shift from computing for the desktop to computing for the world is that systems need no longer be about work-related activities, but also about fun and playful[1] endeavors (Cramer et al., 2010). To realize the system that Nicole in the introductory scenario uses, context-aware systems need to know not only about locations, but about people's lived experiences and their relationship(s) to the location(s) they took place at. To this end, we make use of Location-aware Multimedia Messaging (LMM) systems. Such systems allow people to create multimedia messages (photos, text, video, audio, etc.) that are anchored to a place, which can be received/perceived and interpreted by other people at (approximately) the same place where the message was made (*cf.,* Nicole's photo portrait in the above scenario made at the café on the Corniche). Given that locations within cities are rich sources of "historically and culturally situated practices and flows" (Williams and Dourish, 2006, p. 43), it is reasonable to assume that LMMs can reflect culturally entrenched aspects of people's experiences at locations.

[1] Throughout this chapter, we will use the concepts of fun and playfulness interchangeably.

3.2 Research Questions

Given the above scenario, how can a system 'know' what fun or playful experiences are, in general and idiosyncratically as in Nicole's case of not enjoying older crowds? What kind of contextual elements can be automatically acquired (e.g., date, time, place) to infer playful experiences, and are these contextual elements rich enough to disambiguate the meaning of a user's activity, with and beyond interaction with the system (Dourish, 2004)? Should playful experiences be coded as representations to be used as information that the system makes use of (as in Nicole's device), or should fun be understood as an enjoyable open-ended interaction dialogue between a human and machine (Cramer et al., 2010)? If the latter, what kind of mechanisms need to be in place to ensure that not only the information presented is 'about' fun and playful experiences, but the human-machine interaction is itself an enjoyable experience? If fun and enjoyable experiences can indeed be predicted and catered for, how can this be measured? In this chapter, our main research question is:

> **RQ 2:** How can location-aware multimedia messaging (LMM) systems be used to support playful urban interactions?

Below we will try to address the above question, where the rest of this chapter is structured as follows: first, we provide definitions for an experience in general and a playful experience in particular. Next, we discuss in detail that inferring playful experiences largely depends on whether context is viewed under a positivist or phenomenological lens. Then, we briefly describe past research efforts with using a LMM prototype that allows capturing experiences into different media forms and discuss how the playful experience of capture can be maintained. Afterwards, we briefly highlight common methodological problems that arise when measuring experiences of people under mobile and ubiquitous environments. In response to each of the mentioned problems, we draw three design considerations for the study and design for playful experiences under mobile and ubiquitous environments. Finally, we present our conclusions.

3.3 What is a Playful Experience?

We agree with Law et al. (2009) when they state that the high degree of mutual consensus in the current Human-Computer Interaction (HCI) community over the importance of studying and designing for the user experience (UX) is truly intriguing. The trend towards thoroughly investigating the UX is in part a reaction to the traditional HCI usability frameworks that take user cognition and performance as key aspects in the interaction between humans and machines. With the advent of mobile and ubiquitous computing, human-technology interactions, even if they *involve* work settings, need not be *about* work (Cramer et al., 2010; Greenfield, 2006). This computing for everyday 'non-serious' life has shifted the attention of HCI towards user affect and sensations, where the user experience has become a desirable thing to have during the interaction with a system. Yet what exactly is this user experience? As Law et al. (2009, p. 719) write: "...UX is seen as *something* desirable, though what exactly *something* means remains open and debatable." They move on to argue that embracing a unified definition of the user experience can reap valuable scientific and (industrial) design benefits by: a) facilitating scientific discourse within and across disciplines to avoid communicational breakdown b) aiding the operationalization and evaluation of experience-centric applications c) helping understand the term, its importance and scope.

The term 'user experience' is already pregnantly associated with a wide range of fuzzy and dynamic concepts, with attached attributes such as pleasure, joy, pride, etc. (Law et al., 2009; Hassenzahl and Tractinsky, 2006). In a survey conducted to arrive at a unified definition of UX, Law et al. (2009) found that the elements of the UX provided by their participants largely conformed to the ISO definition of UX (ISO DIS 9241-210:1994, 1994), which states: "A person's perceptions and responses that result from the use or anticipated use of a product, system or service." We find that the ISO definition to be accurate, in part because it provides an abstraction without appeal to specific affective attributes (such as fun or pleasure). However, we find that its accuracy comes at the cost of being overly general in aiding the study and design of context-aware systems. Also, while the definition provides the future aspect of anticipating an experience, it is missing the retrospective aspect of looking at a finished experience.

In attempt to understand experiences, we took the present and past relational temporal properties of experience into account, allowing us to distinguish between prospective experiences (i.e., experiences as they are currently happening) and the retrospective understanding of experiences (i.e., the mental time travel to an experience episode in the past (Tulving, 2002)). This is in line with how Hassenzahl and Sandweg (2004) understand an experience, where they make a distinction between instant utility (a moment in product use within a larger experience episode) and remembered utility (a retrospective summary assessment of the product use experience). Not surprisingly, when they asked their participants how they felt towards a product after they used it, they found that remembered utility is not necessarily the sum of all measured instant utilities. As will be explained later, the decision to view an experience from within (during its occurrence) or from without (after its elapse) also relates to which epistemological stance (positivist or phenomenological) one adopts in conceptualizing and reasoning about the world.[2]

We distinguish between the process of an experience and the memory of an experience, where playful experiences are a subset of both. We define the process of an experience (based on Nack (2003)) as a sensory and perceptual process that some person undergoes (either through direct participation or observation of events and situations) that results in a change in that person. The high variability and subjective interpretation involved in predicting an experiential process indicates that it is useful to retrospectively capture a given experience; in other words, to consider the memory of an experience. Based on the definition of episodic memory given by Tulving (1993), we define an experience memory as the result of an experiential process, which can be manipulated and actively recalled. The memory of an experience consists of one or more actors, a spatiotemporal aspect, a social aspect, a cognitive aspect, and an affective aspect. We use these aspects of an experience memory as a basis for studying experience capture using LMMs. This approach is similar to the one employed by Wigelius and Väätäjä (2009), where they made use of five dimensions of the user experience to study and design for mobile work (e.g., mobile news journalism): social, spatial, temporal, infrastructural, and task context. However, while

[2]While the field of epistemology involves more than just positivism and phenomenology, we are here concerned with only these two.

Wigelius and Väätäjä (2009) separate the characteristics of the user and system from the contextual factors involved, in our understanding of experiences we treat contextual factors as part and parcel of the user's memory of a past episode.

A playful experience, when understood as a process, is characterized by amusement, risk, challenge, flow, tension, and/or negative affect (Csikszentmihalyi, 1990; Nacke et al., 2009). However, we believe that only amusement, which is an affective reaction to a 'playful' activity, is a sufficient condition for playful experiences. While the other attributes (such as tension, risk, flow) can frequently occur in playful experiences, each by themselves, unlike amusement, do not uniquely give rise to a playful experience. According to the definition of playfulness provided by Cramer et al. (2010, p. 1), playfulness refers to "non-utilitarian (but not necessarily non-useful) aspects of interactions that provide pleasure or amusement." While we do not fully agree with Cramer et al. (2010) that a playful experience is non-utilitarian (as playful experiences serve a practical goal of making one feel better as well as aid child learning and development), we do agree that playfulness is largely based on how an activity is approached, rather than an essential property of the activity itself (Csikszentmihalyi, 1990). While this indicates that playfulness is a mental state brought forth by users to an activity, it does not entail that playful interactions cannot be anticipated for particular user groups and explicitly designed for. In other words, if coupled human-system activities frequently draw users toward playfulness during the interaction process, the designer can reason backwards and identify what it is about the interaction that prompted the playfulness in the first place. As will be shown later, the problem of cleanly delineating the cause of a phenomenon (in this case playfulness) for intelligent inference is subject to what notion of context is adopted.

For fun and playfulness, we believe that the most common elicitors of playful experiences are games (e.g., board games, video games), where most games tend to be challenging, create tension, a sense of flow, induce positive and negative affect, and evoke amusement (Nacke et al., 2009; Poels et al., 2007). However, something like The World's Deepest Bin,[3] a bin that makes an elongated sound to indicate depth when someone throws something in it, only elicits brief amusement. Nevertheless, interacting with the bin qualifies as a playful experience because it elicits amusement. What characteristics of playful experiences (e.g., tension, amusement) are to be elicited in users depends largely on the purpose of the system: is the system designed to carry out tasks that are useful or serious (e.g., a context-aware tourist guide (Cheverst et al., 2000) or context-aware firefighter system (Jiang et al., 2004)), or is it meant to entertain (e.g., a location-based game (Benford et al., 2005) or a virtual storyteller (Lim and Aylett, 2009))? While the purpose of the system can aid in helping designers conceptualize the kind of playful experiences desired in interacting with the system, the real problem is how, if at all, can a system infer a playful experience when it happens.

3.4 The Inference Problem

How can a system automatically detect and recognize an experience as playful? What kind of contextual clues are necessary for a system to draw this kind of inference? The

[3]One of several initiatives taken by Volkswagen to improve people's behavior: http://www.thefuntheory.com/worlds-deepest-bin, last retrieved on 01-08-2013

answer to these questions we believe lies in revisiting the concept of ‘context’. Dourish (2004) argues that the notion of context in ubiquitous computing varies with respect to two distinct schools of thought: positivism and phenomenology.

3.4.1 Positivist vs. Phenomenological Theories

Positivist theories, tracing back to sociologist Auguste Comte (1880), derive from a rich, rational and empirical history that takes the scientific method as the sole arbiter of objectively attainable knowledge. This epistemological stance seeks to reduce complex social phenomena into objective, clearly identifiable descriptions and patterns that are idealized abstractions of the observed social instances and situations that make up such phenomena. Phenomenological theories on the other hand, tracing back to Edmund Husserl (1917), are essentially subjective and qualitative. Objective reality according to the phenomenologists is always channeled through the interpretive lens of human perception and action; as Dourish (2004, p. 21) writes, “social facts are emergent properties of interactions, not pre-given or absolute, but negotiated, contested, and subject to continual processes of interpretation and reinterpretation.”

According to Dourish (2004), the positivist account of context renders context as a representational problem whereas the phenomenological account makes context an interactional one. The representational problem is essentially concerned with how context (such as location, time or date) can be encoded and represented in a system so that the system can intelligently tune its behavior according to what values these precoded contextual factors take in a given situation. The main assumption here is that human activity and context can be cleanly separated. For example, the lighting of a room (a contextual factor) is seen as independent of the series of actions required to make coffee (activity) in the room.

By contrast, the interactional problem is primarily concerned with how and why people, through interacting with one another, can establish and maintain a shared understanding of their actions and the context they occur in? To revisit the coffee example, the phenomenological take on it would be that the lighting of the room and the coffee making within it are inseparable; they are tightly woven into an activity-context coupling that give a unified experience, without which that particular experience could not be said to have happened. For Dourish (2004), this underscores the distinction between viewing context as a set of stable properties that are independent of human actions and viewing context as an emergent set of features that are dynamically generated through common-sense reasoning and culturally entrenched beliefs about the world throughout the course of interaction. In other words, while positivism strives for universals (attained through the method of induction), phenomenology contests that the richness of particulars is irreducible to abstraction.

3.4.2 Playful Representation or Interaction?

How do the two accounts of context fare into our understanding of playful experiences? In the context of LMM, we make the distinction between playfulness as an information-rich post-hoc representation (*cf.*, experience memory and the positivist claim) and playfulness as interaction (*cf.*, experience process and the phenomenological claim). To illustrate, the kind of playfulness that Nicole’s mobile system in the opening scenario affords is retrospective, where the system representation of experiences is composed of a clearly identifiable collection of past, personal and publicized multimedia messages that have been annotated as 'fun'. The very act of conceding that labeling these multimedia messages with an identifiable label such as 'fun' is possible arises from a positivist understanding of the world. Following the sequence of Nicole's activities, the representational vehicle (the media presentation of other people's experiences) which subserves the subsequent experiential process that she undergoes when sitting down at the seaside café (namely, absorbing the sun rays and making a multimedia message of her own) is seemingly no longer within the scope of her interaction with the system (Dourish, 2001). This happens despite that causally, the system representation is what brought her to have the experience at the seaside café in the first place.

Following Nicole's interaction with the system to its interactional finish point, we see that the situation changes when Nicole consults her device while she awaits her hookah: the system's presentation of an older crowd, mistaken about Nicole's notion of fun, has now interfered with and altered her current joyful experience. This unanticipated system response can be seen as a flaw when explicitly designing playful human-mobile interactions, where 'playfulness' is scoped only between the interactional possibilities that rest between the user and the system. We believe this reflects the deeper issue of whether to treat playfulness as a representational problem independent of the actual activity process involved in playfully perceiving and acting upon it, or on the other extreme, letting the playful process bleed into interaction windows where the interaction is no longer playful. It is this problem of scoping that makes inferring playful experience a hard problem. Since the context-sensitive variables precoded into the system representation do not account for and update dynamically with the unfolding of the human-system interaction process, inferring playfulness becomes entangled between the system representation and the human interaction with this representation, leaving the system with poor inferential precision.

3.5 The Maintenance Problem

In related work (see 2), we investigated the contextual factors involved in LMM (El Ali et al., 2010). Part of this effort involved field testing a LMM prototype application that allows leaving multimedia messages at locations using three different media types: text, drawing, and photos. The prototype was pilot-tested with 4 participants where an *in situ* interview method (Consolvo et al., 2007) was used to measure participants' user experience. By annotating locations, the prototype lets users capture their experiences by allowing them to create a digital *memory* snapshot of it. The generated message remains anchored to the location it was created at for later viewing by anyone who has the application installed on their smartphone and is at the same place where the message was created.

3.5.1 Fun, But Not Useful

After briefly explaining how the prototype works and how to use it, we let participants at a café create multimedia messages in all three supported media types: drawings, text, photos. For the drawn expressions, two of the participants drew a cup of coffee to show that you can get coffee at the cafeteria. The other two made graffiti expressions, where their drawings augmented parts of the environment. For the drawings, we found that drawings were meant only as fun digital augmentations on the physical environment. When asked about his/her drawing message, S1 explained: *"Well that ['Dancer in the Dark' poster] is a poster that I enjoy looking at a lot when I'm drinking, and I always wondered about the frame, so I wanted to draw lines around it, but to do it freely. Doesn't have a purpose but it looks nice."*

For the textual messages, participants used text for: recommending items (e.g., S4: *"You should try the green tea"*), a means for self-expression (e.g., S1: *"Beer Perspective"* and S2: *"Things are looking up"*), or as a warning to others (e.g., S3: *"Don't confuse gravy with soup"*). For the photo expressions, two of the participants took a photo of the experimenter, and the other two a photo of the street. All photo messages made were used as a means to contrast the present with the future that others will witness (e.g., friends viewing photos of them with the experimenter at a later time).

When participants were later asked about their overall experience with the LMM prototype, they all reported that it was fun to doodle over the environment and leave photos to share with public and private networks, but did not find either of them to be useful. On the other hand, they all found that it is useful to share text messages (such as recommendations) with others at a place. Using text for practical purposes is in line with what Persson and Fagerberg (2002) found in evaluating the GeoNotes messaging system and what Burrell and Gay (2002) found for the E-graffiti system. The lack of usefulness in drawing or capturing photos in the LMM prototype hinted that perhaps an incentive mechanism that motivates users to use the application is needed to ensure that the experience of capture using the LMM application is perceived as not only fun, but also useful (cf., discussion by Greenberg and Buxton (2008) on why designed systems must first be deemed useful, and only then usable). Equipping a system with persuasive techniques to increase personal and social gain has been explored in social media networks (e.g., Cherubini et al. (2010); Singh et al. (2009)), where users are provided with a strong incentive to make contributions of a certain type (e.g., high quality media contributions). Likewise, if game-theoretic elements are designed into the interaction process, the playful aspects of using LMM can be maintained beyond amusement reactions, insofar as the LMM contribution behavior of users is reinforced with personal and social rewards.

3.6 The Measurement Problem

Finding an appropriate testing methodology to understand playful experiences that can unlock suitable interaction methods in mobile and ubiquitous settings poses a real challenge. This challenge is amplified by the difficulty in probing into the inner subjectivity of the cognitive and emotional lives of people under changing contexts and while on the move. There has been several successful attempts at measuring user's experiences, especially during interaction while immobile. Much work in this respect has focused on interaction with digital (video-)games (e.g., (Nacke et al., 2009; Bernhaupt et al., 2008; Mandryk et al., 2006)).

3.6.1 Subjective and Objective Experience Measures

Broadly, experience measurements can be broken down into subjective and objective measures (Bardzell et al., 2008; Greenberg and Buxton, 2008). Subjective measures typically involve self-reports of a given experience, where methods for obtaining them typically include interviews, surveys, and ethnomethodological techniques in general (Kuniavsky, 2003). A popular measure to measure the gaming experience has been the use of the Game Experience Questionnaire (GEQ) (Nacke et al., 2009). Objective measures, by contrast, evaluate *observable* aspects of a person's experience independent of that person's perception. These can range from observations of human posture and gait, button press count and task completion time, to physiological measurements such as Electroencephalogram (EEG) recordings, Galvanic Skin Response (GSR) recordings, Electromyography (EMG) recordings, or eye movement capture using Eye-tracking hardware (Nacke et al., 2009; Bardzell et al., 2008). Such objective metrics however are difficult to generalize to mobile and ubiquitous environments (Kellar et al., 2005), where not only is the user's location subject to change, but also the context at a given location.[4]

One methodology that promises to deal with the fuzzy nature of user testing in the wild is the Living Lab methodology (de Leon et al., 2006; Eriksson et al., 2005). El Ali and Nack (2009, p. 23) defined the Living Lab methodology as research grounds for the testing and evaluation of humans interacting with technology in natural settings, to "better inform design decisions sprouted from what real-life users want, so that technology development becomes an intimate three-way dance between designers, developers, and users." Two challenges to this ambitious research agenda raised by El Ali and Nack (2009), the risk of over-measurement and under-measurement, warrant recapitulation here. While these considerations are fairly general, they are stated here to underscore the importance of choosing the right testing methods for measuring experiences in mobile and ubiquitous environments.

3.6.2 Over-measurement and Under-measurement

Over-measurement can occur when a user is left to freely use a mobile and/or ubiquitous experience-centric application while on the move. Without informed understanding of what *kind* of data is being collected, extraction of meaning from the continuous flux of data (e.g., interaction history logs) proceeds in an ad hoc manner, and thus risks a loss in interpretation and quality of drawn implications. Consider Nicole's complex behavior in the introductory scenario, where she initially accepted the seaside walkway recommendation from her device, but retracted the recommendation later in light of new information about the café she is at. Without being explicitly informed about what kinds of media she, or people like her, find enjoyable and fun, it would not be possible for a system to adequately adapt to her needs. This indicates that interaction behavior should be constrained to a small number of measurable units that provide (partial) immunity from the unpredictable nature of un-

[4]There are exceptions to this: mobile Electrocardiograph (ECG) can measure heart rate while a person is moving, the wearable EOG goggles (Bulling et al., 2009) can measure (saccadic) eye movements in everyday interactions, and Brain-Computer Interfaces (BCIs) such as wearable EEG can measure brain electrical activity during daily interactions (Casson et al., 2008). While indeed these kinds of tools permit objective measurement, they are not without problems: a) the collected signals are difficult to interpret (especially in noisy environments) and b) these devices are not always feasible for use in user tests.

supervised human-technology interaction. Without minimal supervision exerted on testing conditions during system evaluation and early development, caution should be exercised concerning whether or not the elicited knowledge is trustworthy enough to solicit informed understanding and design of mobile and ubiquitous systems.

At the other end of the spectrum, rigorously controlled laboratory testing can result in *under-measurement*, where the main problems are: a) testing is confined to the walls of the laboratory. This means that ‘natural’, mobile behavior is by necessity beyond the scope of the method b) only a handful of experiential variables can be measured. This is due to the complexity and error-proneness of developing multidimensional designs that can properly incorporate several independent variables and tease out the possible effects on the dependent variables of interest. Together, these problems make controlled laboratory testing, by itself, insufficient for measuring playful experiences in mobile and ubiquitous environments.

Given the two highlighted problems, how can a middle-ground be reached for evaluating experiences in unconstrained environments? One immediate response provided by El Ali and Nack (2009) is to split the evaluation process into two phases: subjective observation and objective measurement. In the observation phase, the researcher employs outdoor, *subjective* observational methods during the early design stages of application development as a means of reducing the phenomenon dimensionality down to a few *objectively* measurable variables. During the second phase, depending on their nature, these variables can be experimentally teased out under rigorously controlled indoor environments. There are two promising augmentations to the early observation phase, well-suited for dealing with the difficulties in evaluating context-aware applications under mobile and ubiquitous environments: using Urban Pervasive Infrastructure (UPI) methods (Kostakos et al., 2010, 2009) and context-aware Experience Sampling Methods (ESMs) (Consolvo and Walker, 2003; Froehlich et al., 2007).

3.6.3 UPI Methods and ESMs

Without going into excessive detail, the UPI methods defined by Kostakos et al. (2009) are built on the premise that the city can be viewed as a system, where the variables of interest are the combination of people, space, and technology that together aid in studying and deploying urban pervasive applications.[5] These methods deal with five characteristics of the UPI: mobility (e.g., human distance travelled or visit duration), social structure (e.g., social network analysis metrics such as degree of separation), spatial structure (e.g., space syntax metrics such as integration), temporal rhythms (e.g., time-based distributions of people’s activities), and facts and figures (e.g., statistical characteristics such as number of devices detected at a defined area).

Focusing on the above characteristics, Kostakos et al. (2009) have developed methods of observation and analysis that reveal real-world values under these metrics. For example, in their ‘augmented gatecount’ observation method, gatecounts (using Bluetooth scanners) are used to define the flows of people at several sampled locations within a city. The main point here is that these concepts, metrics and methods can considerably aid in gaining an understanding of a city objectively, which in turn aids in the early design stages of appli-

[5] In this context, ‘urban pervasive applications’ is synonymous with ubiquitous applications deployed in a city.

cation development. To ground it in context of playful experiences, the understanding of a city afforded by the UPI methods can identify spatial and social clusters in a city where people meet for entertainment purposes (e.g., the movies or the park), which provides support for narrowing down the objective of playful applications to the right target group or spatial structure.

Other methods that are useful in evaluating and narrowing down the early design space of mobile and ubiquitous application development are Experience Sampling Methods (ESMs) (Consolvo and Walker, 2003). ESMs work by alerting participants each day to fill out brief questionnaires about their current activities and feelings. Sampling experiences throughout the course of a day make ESMs a great tool to evaluate a given application *in situ*. Moreover, unlike classical self-report techniques, ESMs do not require participants to recall anything and hence reduce cognitive load. Typical studies with ESMs involve a minimum of 30 participants, and are longitudinal. The longitudinal aspect also means the analysis of collected structured data from participant responses is amenable to statistical analyses. Together, these characteristics of ESMs make them not only invaluable tools in uncovering current usage of mobile and ubiquitous applications, but practical methods of investigating human 'technology' needs under different, real-world contexts. An exemplary translation into the opening scenario would be interval-dependent or event-dependent sampling of Nicole's experience of playfulness with her environment and/or with the device. By sampling Nicole's experiences, her device is able to build a predictive user model that probabilistically *knows* what things she finds fun, and can tailor the media presentation accordingly.

To sum up, while measuring experiences is a difficult endeavor, deliberation over and choosing the right testing methodology can be extremely useful in aiding the design and development of mobile and ubiquitous applications meant to elicit playful experiences. Moreover, the process of designing for playfulness can strongly benefit from objective social, spatial, and temporal analysis using UPI methods as well as user models built from contingent experience sampling from users across a given timespan.

3.7 Design Considerations

For each of the problems highlighted above (the inference problem, the maintenance problem, and the measurement problem), we provide design considerations that we believe are relevant in the study and design for playful experiences under mobile and ubiquitous environments:

1. **Experience Representation vs. Interaction Experience**: As stated in Section 3.4, a distinction can be made between an experience representation, which is information 'about' an experience, and the experience itself, which is a process emergent from an undertaken activity. This reflects the difference in how one understands context. Under a positivist view, the focus is on capturing experiences while under a phenomenological view the focus is on eliciting experiences through coupled activity-context pairs. For capturing experiences, the aim is to provide an adequate representation of any experience that took place, of which playful experiences are an instance. This requires a computational method for annotating the media-based

experience representations with the right kind of information (e.g., affective information about the degree of fun had) for later intelligent retrieval (*cf.,* Nicole's device suggesting fun places nearby given her request of fun things to do).

For eliciting experiences, the aim is to subject users to activities and contexts that would strongly correlate to (if not cause) a desired *type* of experience (e.g., experiencing trust when interacting with a system). The concern here is not about which contextual elements are supported so as to sufficiently re-contextualize the experience of others, but rather about the scoped playful interaction between the user(s) and the system, where the user experience takes place during the interaction process itself. For example, the act of shaking a mobile device to indicate a change in preference for presented location recommendations can itself be a playful experience. In the domain of LMM, one way of enhancing the playful experience would be to provide the right kind of multimodal input and output support (Chittaro, 2009). For example, labeling a media expression (e.g., a photo) by means of textual input (*cf.,* Section 3.5) might be more intrusive and interruptive of a playful experience, whereas a voice command of 'fun' that achieves the same function can occasion a more seamless interaction experience. In short, researchers and designers alike should be aware of which epistemological stance (positivist or phenomenological) they commit to when studying and designing for experiences in general and playful experiences in particular.

2. **Incentive Mechanisms as Mediators of Continuous Playfulness:** We mentioned in Section 3.5 that our pilot study participants had reported that their interaction with the LMM prototype for doodling and photo-capture was fun but not useful. This led us to consider that, at least for LMMs, users require an incentive to interact with the system that transcends merely playful interaction. In other words, the fun things such as tension and challenge, risk and unpredictability, positive and negative affect, have to be deliberately embedded in the interaction process. However, the fun aspects should be secondary to the user task of documenting and sharing their experiences as multimedia messages. Simply put, the perceived usefulness of a system should be treated as a first-class citizen.

   Notwithstanding the importance of usability issues, this raises an important issue of whether the user should be made aware of the real goal of the performed task (i.e., task transparency), and in what domains does it actually matter to apply such persuasive techniques. For example, implicit ambient light feedback is a useful mechanism to unobtrusively indicate excess electricity consumption during the day. A promising approach for applying incentive mechanisms in the context of LMM is to utilize game-theoretic approaches (Singh et al., 2009) to create competitive game-like environments that persuade users to perform a given task, such as tagging or rating people's generated messages (cf., Facebook's[6] 'Like' button). This would not only motivate users to collaboratively rank the generated content, but given the competitive element, would make the experience of doing so fun and engaging.

3. **Balancing Testing Methodologies when Measuring Playfulness:** Measuring fun and playfulness is by now a well-known slippery endeavor (Cramer et al., 2010). As

[6] `http://www.facebook.com`; last retrieved: 01-08-2013

mentioned in Section 3.6, the difficulty arises in deciding to test users in a natural setting, where objective experiential data is hard to acquire. At the other extreme, controlled testing permits objective measurement at the cost of narrowing explanatory scope. While there is no clear prescription for the most effective approach to evaluating experiences, it is likely that a gradual progression from unconstrained to controlled testing in the course of mobile and ubiquitous application design and development is an effective means to measure experience. More concretely, during early design stages, outdoor testing of mobile users can help yield design implications that help narrow down the set of observable phenomena to a few variables, which can then be experimentally teased out in a more controlled environment.

As we have suggested, there are two promising methods to augment understanding, analysis, and narrowing down of the early design space: UPI methods and ESMs. While UPI methods permit objective measurement and analysis of structures (social, spatial, temporal) within the city, ESMs can help shed light into individual human-technology needs under certain places and times. Due to the importance of objective measurement and analysis on the one hand, and the need to systematically understand human subjective responses on the other, we believe that a combination of both methods can strongly aid in both understanding the playground of existing playful interactions, and the subsequent development of future-generation mobile and ubiquitous tools to enhance these interactions. For example, the duration of a visit at a particular site in a park with a particular social setting (characterized for example by a minimum person co-occurence frequency count) can be used as a trigger for unobtrusively sampling a person's experience. That person's response includes both the receptivity to the sampling interruption as well as the content of interruption (e.g., what activities he was engaged in at that moment and with how many people). This response in turn can on the one hand provide a useful feedback loop (Kostakos et al., 2009) into the quality and capacity of objectively measuring and inferring people's activities from such measurements, and on the other hand shed light into what kinds of experiences these people undergo at certain locations within a city (such as the park).

3.8 Conclusions

In looking at what playful experiences are, how they can be inferred, how the experience of capturing them can be motivated and maintained, and how to measure them, we have underscored what we believe to be fundamental problems underlying the scientific study of playful experiences in mobile and ubiquitous environments. Drawing on past research efforts and an envisioned LMM usage scenario, we hope to have drawn attention to the importance of thoroughly examining the different aspects of playful experiences (inference, capture-maintenance, measurement) when designing LMM systems to be used under ubiquitous environments.

As highlighted in the introductory scenario, there are a myriad of cognitive and affective factors intermixed with the system interaction that are difficult to experimentally and computationally disentangle. This in part stems from which epistemological stance (positivist or phenomenological) one chooses to adopt in practicing HCI (Section 3.4). Intermixing the two views, at least in LMM, makes it difficult for a system to automatically acquire the right kind of experiential information (e.g., media tagged or rated as fun that corresponds to how fun an experience was) and to intelligently retrieve this information in the right situation (*cf.,* Nicole's desire to experience something fun), while at the same time ensuring that interaction with and cognitive processing of this information is itself enjoyable. The latter point, as we mentioned (Section 3.5), can be mediated by explicitly incorporating fun and enjoyable game-like elements in the experience capture process. Lastly, we considered the problems that arise in measuring experiences in general and playful ones in particular (Section 3.6), and argued that a gradual progression from controlled to out-in-the-wild testing provides a systematic methodology which can aid in understanding the playground for future experience-centered mobile and ubiquitous systems.

In response to the highlighted problems, we have furnished playful HCI with three design considerations (experience representation is not the same as interaction experience, incentive mechanisms can be mediators of playfulness, and measuring playfulness requires a balance in testing methodology choice) that together serve as useful guidelines for scientifically studying and designing playful experiences in mobile and ubiquitous environments. The need for clear guidelines has been well-articulated by Greenfield (2006, pg. 232) when he wrote back in 2006: "Much of the discourse around ubiquitous computing has to date been of the descriptive variety...but however useful such descriptive methodologies are, they're not particularly well suited to discussions of what ought to be (or ought not to be) built." Yet to what extent it is possible to truly design and build mobile and ubiquitous systems that carry out the task of capturing experiences while making the experience of capture itself fun and enjoyable remains an open question.

While Chapter 2 and 3 focused on the overall user experience and elicited playfulness of multimedia messaging behavior at urban locations, these studies also showed that urban interactions take place across locations, rather than isolated locations. To account for the connectedness of urban interactions across locations, in the next chapter (Chapter 4) we follow up with a study that investigates the design and evaluation of a system that allows pedestrians to explore a city.

4

Automatic Exploration-based Route Planning

We investigate the problem of how to support city resident and tourists wishing to explore a city. We introduce an approach borrowed and adapted from bioinformatics to build an exploration-based route planner that leverages 5 years of geotagged photos taken from the photo sharing website Flickr (*Route Planner Study*). We evaluate our generated exploration-based route plans through a controlled laboratory study, as well as through a web survey. Drawing on experience questionnaire data, web survey responses, and user interviews, the findings led to a set of design recommendations for going towards automatic data-driven approaches to support city exploration, and the role different digital information aids play in supporting such exploration behavior. The work presented in this chapter was published as "Photographer Paths: Sequence Alignment of Geotagged Photos for Exploration-based Route Planning" in the Proceedings of the 16th ACM Conference on Computer Supported Cooperative Work and Social Computing (El Ali et al., 2013b).

4.1 Introduction

We are not always in a hurry to get from point A to point B. Sometimes we take a longer route because it is more scenic, more interesting, or simply to avoid the mundane (Hochmair, 2004). While expert tour guides (e.g., Lonely Planet[1]) tell us what to see and do, they are geared towards recommending destinations and tour guide offers, not generating route plans or journeys. In fact, research has focused extensively on tourism, and replete with how to develop mobile technology (or electronic mobile guides) to support travelers and tourists in 'what to do' (Kenteris et al., 2011; Brown and Chalmers, 2003). For current route planning services (e.g., Google Maps[2]), the generated routes are tailored towards providing shortest paths between any two locations. However, city pedestrians, whether tourists or locals, may not always want the fastest route – this is strengthened when for example

[1]http://www.lonelyplanet.com/ ; last retrieved: 01-08-2013
[2]http://www.maps.google.com/ ; last retrieved: 01-08-2013

considering the buzz surrounding Foursquare's[3] launch of the Explore functionality that recommends places based on friends' check-ins. In other words, given the right context and time, people do wish to wander into hitherto unfamiliar or unconventional paths. However, there is surprisingly little work in CSCW and geographic HCI (Hecht et al., 2011) that addresses this gap: how can we build city route planners that automatically compute route plans based not on efficiency, but on people's trailing city experiences? Importantly, how do these experiences influence our route preferences and perception of urban spaces?

With the unbridled adoption of location-aware mobile devices that permit geotagging multimedia content, places and routes can now be ubiquitously micro-profiled with geotagged user-generated content. This geotagged data comes from mobile social media services (e.g., Flickr[4], Twitter[5]), and relates to the actions and experiences of thousands of people at different locations. In line with a recent SIG meeting discussing the research opportunities of geographic HCI and the rise and use of User-Generated Content (UGC) (Hecht et al., 2011), we believe this geotagged data can be used not only for revealing the social dynamics and urban flow of cities (Kostakos et al., 2009), but also unlock fragments of user intentions and experiences at places and transitions between them. This data can provide a latent source for generating exploration-based routes traversed in a city that are not based on travel efficiency.

In this chapter, we focus on sequences of geotagged photos, which we show can allow computing city paths that represent the history of where the photographers of these photos have been. By using this latent information on photographer paths, we believe this unlocks novel application and research avenues for data-driven exploration-based route planners.

4.2 Research Questions

Our main research question is:

> **RQ 3:** How can we automatically generate routes to support pedestrians in exploring a city?

Specifically, what existing data sources and which methods can be used to generate such routes, and how are these routes perceived in comparison with both the popular and fast routes in a city? While our target user group is city residents (defined as having lived in the city for at least one year), our contributions as will be shown later also apply to tourists who wish to discover off-beat paths when visiting a city.

To define what constitutes an interesting walkable route, we reasoned that the mobility behavior of city photographers tells us something about worthy alternative routes in a city. The underlying assumption here is that locations of photographs are potentially interesting as the photographer(s) found it worthwhile to take a picture there. For this purpose, the image photo-sharing site Flickr provides a suitable data source given that many images are geotagged.[6] Here, we focus on users that do not have any specific interest or do not want to

[3] https://foursquare.com/ ; last retrieved: 01-08-2013
[4] http://www.flickr.com/ ; last retrieved: 01-08-2013
[5] http://www.twitter.com/ ; last retrieved: 01-08-2013
[6] Around 520,00 geotagged photos tagged with 'Interesting' in Amsterdam alone (retrieved on 30-05-2012).

supply this interest and they just want to be given an interesting route from A to B, which we suspect city photographers (be they locals or tourists) can help unravel.

To avoid making the user supply preferences, we wanted to automatically generate routes based on where people traveled within a city. But not every route may be interesting, so we focused on routes made up of locations were people took pictures, given our assumption that taking a picture somewhere depicts an interesting location. One such route made out of photographs from a single photographer is insufficient, so ideally we want multiple photographers that took pictures at the same locations in the same sequence, i.e. took the same route and found similar things photo-worthy along the same locations. Thus, we needed a method that handles not only where photographers have been, but importantly, in what order they have been there and to what extent their movements resemble the movements of other city photographers. To achieve this, we use sequence alignment (SA) methods. These methods are borrowed from bioinformatics and later adapted to time geography to systematically analyze and explore the sequential dimension of human spatial and temporal activity (Shoval and Isaacson, 2007). We hypothesize that the aligned routes traversed by multiple city photographers (or 'photographer paths') provide desirable paths for pedestrians wishing to explore a (familiar) city. Furthermore, while we are concerned with route planning using both mobile devices and desktops, here we focus on pre-trip route plans, which usually involves viewing routes on a desktop.

Our work yields two main research contributions: a) a novel data-driven methodology for generating walkable route plans based on photographers' paths in a city and b) an empirical understanding (based on quantitative and qualitative assessments) of how users perceive these photographer paths in comparison with today's efficiency driven route planners and popular routes. Additionally, we provide a preliminary investigation on the role that digital information aids on a map (e.g., Points-of-Interest (POIs), photos, comments, etc.) play in influencing people's decisions about which route to take for exploring a city.

The rest of the chapter is structured as follows: we give a review of related work, followed by our Photographer Paths approach and alignment experiments. We then present a user study (consisting of a lab and web-based study) to evaluate the different route plans and importance of digital information aids in influencing users' perception of city routes. We then present and discuss our results, and finally conclude.

4.3 Related Work

Given our interest in both generating and consuming UGC-generated routes, this chapter draws from various related work, including time geography, urban modeling techniques, and importantly route planners.

4.3.1 Time Geography

Time geography dates back to Hägerstrand (1970), who stressed the importance of taking into account temporal factors in spatial human activities. This gave rise to a space-time path visualization which shows the movement of an individual graphically in the spatial-temporal environment when one collapses the 3D space and uses perpendicular direction on a 2D map to represent time.

Essentially, time geography seeks to analyze patterns of human activity using space-time paths in an objective, structural manner (e.g., aligning sequences of activities by visitors to the Old City of Akko (Shoval and Isaacson, 2007)). The idea behind this is to visualize human movement and interactions between individuals on a 2-D plane where the x- and y- axis represent geographic coordinates (longitude and latitude, respectively) and the z-axis represents time. This so-called space-time "aquarium" is used for analysis and evaluation of social dynamics and activity distribution across space and time. This is useful for analysis of aligned sequences of human activity, where the activity of concern here is the photo-taking behavior by photographers of the geotagged images retrieved from the Flickr photo-sharing site. In short, we use these representational methods to analyze sequences of photo-taking activities, where we later use alignments for generating walkable city routes based on these photographer paths.

4.3.2 Photo-based City Modeling

Given the iconic correspondence between photographs and reality, we believe photo sharing services like Flickr provide a window into the unique perspectives of city photographers. If we consider Flickr photo features, thousands of photos shared by photographers come contextualized with textual user-defined tags and descriptions, geotags (latitudes and longitude coordinates), and time-stamps (date and time of day).

Snavely et al. (2008) used the varied photos taken by multiple photographers of the same scene along a path as controls for image-based rendering, allowing automatic computation of orbits, panoramas, canonical views, and optimal paths between 3D scene views. Relatedly, Tuite et al. (2011) used a game-based crowdsourcing approach to constructing 3D building models, based on contributions from a community of photographers around the world. In this work however, we are not concerned with 3D scene views (e.g., Google Street View[7]), only with the generation and perception of route plans plotted on a 2D map.

Using Flickr data alone, computational approaches have been developed to understand tourist site attractiveness based on geotagged photos (Girardin et al., 2008), constructing travel itineraries (De Choudhury et al., 2010) and landmark-based travel route recommendations (Kurashima et al., 2010), and generating personalized Point-of-Interest (POI) recommendations based on the user's travel history in other cities (Clements et al., 2010b). All these approaches however focus primarily on describing locations and/or landmarks at these locations, and not on within-city routes that connect them irrespective of popular landmarks. Closer to the present approach, Okuyama and Yanai (2011) mine sequences of locations from Flickr geotags – however, their focus is on recommending the most popular tourist places in a city.

4.3.3 Non-efficiency Driven Route Planners

Relevant here is whether there is work on route planners that go beyond finding routes that optimize commute efficiency. Lu et al. (2010) developed a system to automatically generate travel plans based on millions of geotagged photos and travelogues, which was tailored towards providing city tourists with popular attractions/landmarks and popular routes between them. Relatedly, Cheng et al. (2011) mined people's attributes from photos to

[7] www.google.com/streetview/ ; last retrieved: 01-08-2013

provide personalized travel route recommendations; however, their method was aimed at finding personalized hotspots, not for exploring off-beat paths in a city.

Arase et al. (2010) categorized travel trips from people based on geotagged images taken and the accompanying tags and photo titles, allowing development of an application for searching frequent trip patterns. While the goal here was catering for users that wish to learn more about the most frequently visited places in a city, we are interested in automatically computing route plans for exploring a city based on sequences of photographers' movements. Relatedly, Hochmair (2004) present a method comprising a user survey and subsequent clustering analysis to classify route selection criteria for bicyclists. Here, they found that bicyclists most favored fast and safe routes, followed by simple and attractive ones in an urban environment. Finally, using a crowdsourcing approach, Zhang et al. (2012) developed the Mobi system which allows people to collaboratively create and edit itinerary plans in cities, thus showing the merits of human computation tasks to provide rich plans. In our work however, we try to automate the process of providing exploration-based route plans in a city.

4.4 Photographer Paths

4.4.1 Approach: MSA of Arbitrarily Long Sequences

To align the geotagged photos, we used the ClustalTXY (Wilson, 2008) alignment software. ClustalTXY is suitable for social science research, as it makes full use of multiple pairwise alignments, where alignments are computed for similarity in parallel – in other words, it makes use of a progressive heuristic to apply multiple sequence alignment (MSA) (Wilson, 2008). Furthermore, ClustalTXY allows representing up to 12-character words, which allows us to uniquely represent small map regions containing the geotagged photos.

MSA is done in 3 stages: first, pairwise alignments are computed for all sequences. Then these aligned sequences are grouped together in a dendogram based on similarity. Finally, the dendogram is used as a guide for multiple alignment. To deal with differences in sequence length, ClustalTXY adds gap openings and extensions to sequences. Opening is the process of adding a gap between two previously gapless words and extension is the process of adding another gap in between two words which already had a gap.

Throughout the chapter, 'words' are synonymous with 'locations' and 'nodes', where a given term is used depending on the context of discussion. The more aligned sequences that contain the same words, the more popular is a particular word. Thus, the most interesting sequences are distilled by finding matching sequences of popular words in the alignment results. In our approach, we map each location in a sequence to a cell in a partitioned grid map where each cell corresponds to an indexed location unit (e.g., 125 x 125 m cell). For example, a route containing 5 locations would thus be `BcEfSgQlQn`, where `Bc` constitutes the first word (i.e., a location). Furthermore, all repeated words were trimmed down to one (e.g., `FyEjEjEjYfWyFs` would become `FyEjYfWyFs`). We use a simple grid-based approach instead of a mean-shift clustering approach (cf., (Clements et al., 2010a)) in order to allow for locations photographers visited that may not otherwise contain many data points. We then apply MSA to the photographer routes (consisting of sequences of their photos' locations) to find the aligned location sequences. These are used for selecting matching segments of sequences across photographers – we call these exact matches *photographer route segments* (PRSs).

4.4.2 Dataset

We used the Flickr API to retrieve geotagged photos within Amsterdam, The Netherlands (17.3 km N-S and 24.7 km E-W)[8] over a 5-year period (Jan. 2006 - Dec. 2010), with the following attributes: owner ID, photo ID, date and time-stamp, tags, latitude, longitude and the accuracy of the coordinate. This resulted in a database of 426,372 photos.

4.4.3 Preprocessing

We included in our database only photos with geo-coordinates with near-street accuracy or better (accuracy 14-16 in Flickr attributes). We inferred the sequences taken by photographers from the time and geotags of their photos. Each photo in the sequence had to be taken within 4 hours from the previous photo. Sequences were constrained to having at least two or more different locations (or nodes), where each location corresponds to a cell on the grid. Given early experiments, we used a grid cell size of 125 x 125 m. We also now focused our grid on the city center of Amsterdam as most routes were in this area and this would speed up alignment computation. The city center could be described using a grid of 26 by 26 cells, so 2-letter words were sufficient. These steps resulted in a dataset of 1691 routes, which had an average length of 9.92 words ($\min = 2$, $\max = 124$). There were 1130 unique photographers, where on average each photographer contributed 1.50 routes to the dataset.

4.4.4 Sequence Alignment

Main parameters in MSA are gap opening and extension values. In bioinformatics these values correspond to a penalty for splitting a DNA or protein sequence, which needs to be restricted in order to retain informative groups of sequences of nucleotides or amino acids. In our case this analogy does not hold and we want to match as many words as possible, therefore we set both values to 0. Alignment for this 125m dataset took approximately 7 hours on a single core server.

To find photographers paths from PRSs, we set constraints for selecting PRSs having *at least 4* photographers having *at least* 2 aligned nodes (or locations/words). This choice was motivated by the resulting PRSs from our 5-year dataset (see Fig. 4.1), where having more photographers per route segment took precedence over number of locations (or nodes). These 2 or more aligned nodes form the PRSs. Photographers could have made different photos in between nodes, but they must have visited the locations in the same order and within 4 hours between each visited location. After applying these constraints, we had 231 PRSs (visualized in Fig. 4.2) with an average length of 2.61 nodes and a maximum PRS length of 4 nodes.

[8] The area is based on the Amsterdam region as currently defined in the Flickr API (bounding box: 4.7572, 52.3178, 5.0320, 52.4281; centroid: 4.8932, 52.3731).

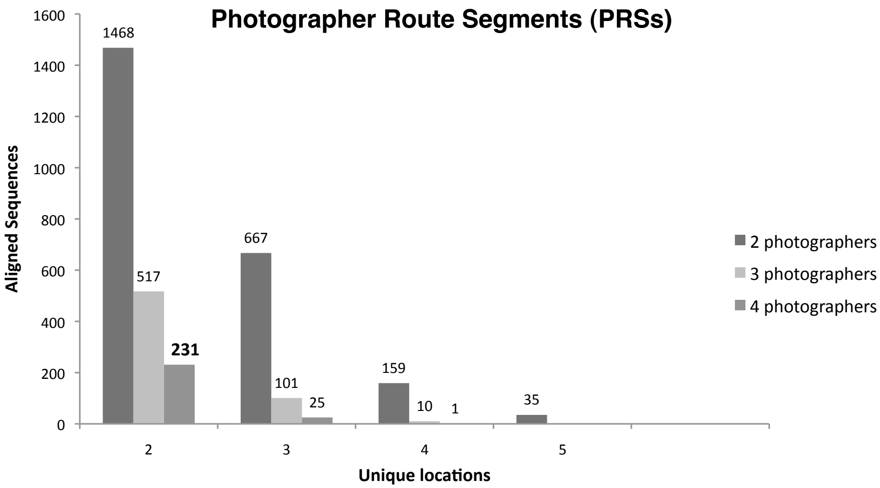

Figure 4.1: Aligned sequences (PRSs) in Amsterdam over a 5-year period for different numbers of unique photographers and locations. Our PRS set choice value ('231 sequences') is shown in bold.

4.4.5 PRS Aggregation

Next step was to develop a method which uses these PRSs to generate routes from a given start location to a user specified destination. We used an implementation of Dijkstra's shortest path algorithm[9] to find the shortest route along the PRSs. Recall that a PRS is a transition between two or more locations/nodes based on sequences of at least 4 photographers, where a PRS is calculated from the aligned sequences of the ClustalTXY alignment.

Dijkstra's algorithm requires a network of edges between nodes, with a specified cost for traversing each edge. We thus had to specify how our PRSs would both connect within themselves and to each other. Every edge cost is set to the distance between the nodes. However, if all nodes were to be connected with each other, then Dijkstra's algorithm would simply output the direct connection between the start node and the end node as a route, so we set a maximum distance for edges between and within PRSs. Dijkstra's algorithm finds the shortest path between nodes, but we wanted to steer the algorithm to make use of as many transitions between nodes within each PRS as possible, even if this meant a detour, because these transitions are more representative of the actual paths of photographers. To solve this, we required that at least two nodes were used in each PRS, thus at least one edge within a PRS is always used. After this hard constraint, Dijkstra's algorithm connects a PRS to another PRS, because using a third node within the original PRS will usually result in a longer route. The final route would thus be made out of PRSs where only two nodes within each PRS are used. To maximize the number of nodes within each PRS, we gave discounts [range 0-1] to the distances of every edge (beyond the first edge) used within a

[9] http://code.activestate.com/recipes/119466-dijkstras-algorithm-for-shortest-paths/ ; last retrieved: 01-08-2013

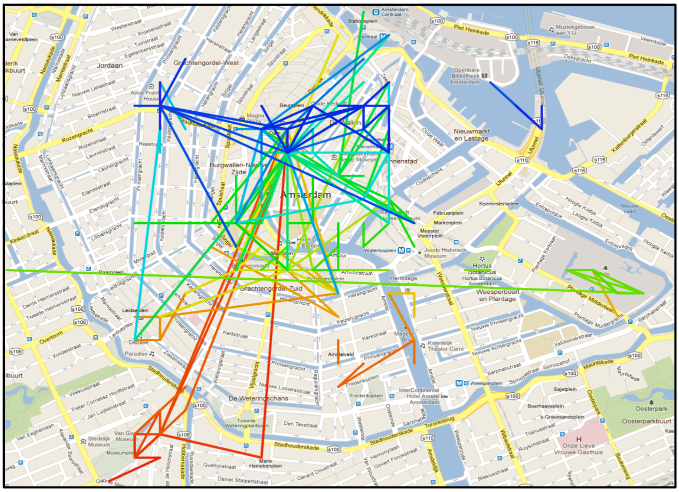

Figure 4.2: 231 PRSs of alignments of 4 photographers and 2 unique locations in Amsterdam city center. *Best seen in color.*

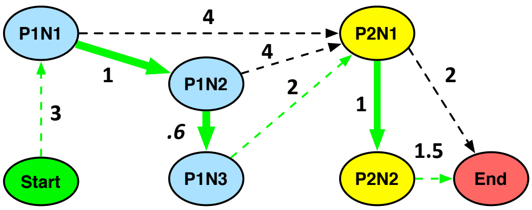

Figure 4.3: Example of PRS aggregation. 'P1N2' stands for node 2 of PRS 1. Numbers indicate inter-node distance. *Best seen in color.*

PRS, forcing Dijkstra's algorithm to incorporate extra edges within the PRS.

A simplified PRS aggregation task using these methods is shown in Fig. 4.3. The thick solid lines show the edges between the nodes within PRSs, while the thin dashed lines show the connections between the PRSs and the user specified start and end nodes. Dijkstras algorithm would normally find the following shortest path Start-P1N1-P2N1-End with a cost of 9, but due to the constraint of at least two nodes per PRS and the discounted edge cost (0.6 (cost) * 1 (original weight) = 0.6; shown in italics) between P1N2-P1N3, a different route is selected. The recommended route (green edges or Start-P1N1-P1N2-P1N3-P2N1-P2N2-End) will now make use of all the PRS edges.

We applied this algorithm on our chosen PRS set (4 photographers, 2 locations), to create two different photographer routes in the city center of Amsterdam: one made up of 9 PRSs with 11 total connections (where black route segments are gaps filled for completing the route), that runs from Central Station to Museumplein (CM). The other was made up of 4 PRSs with 6 total connections (again black route segments are route gaps filled), and runs from Waterlooplein to Westerkerk (WW). These routes are visualized in Fig. 4.4.

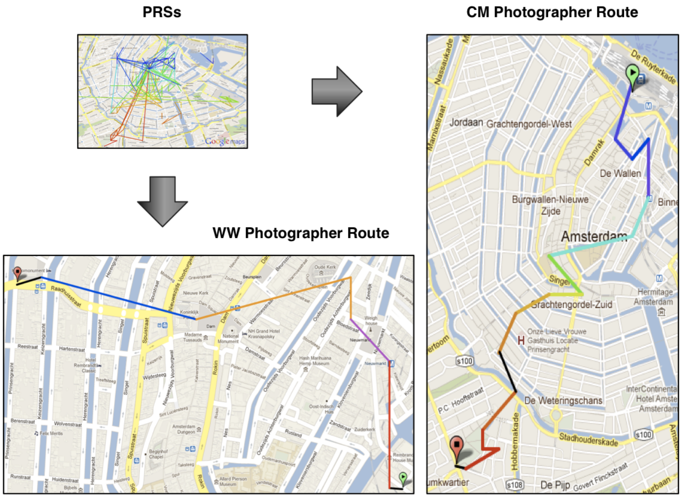

Figure 4.4: Our chosen PRS set after applying the modified Dijkstra's algorithm resulted in two 'crude' photographer routes (where individual PRSs are color coded): a) Central Station to Museumplein (CM) route b) Waterlooplein to Westerkerk (WW) route. *Best seen in color.*

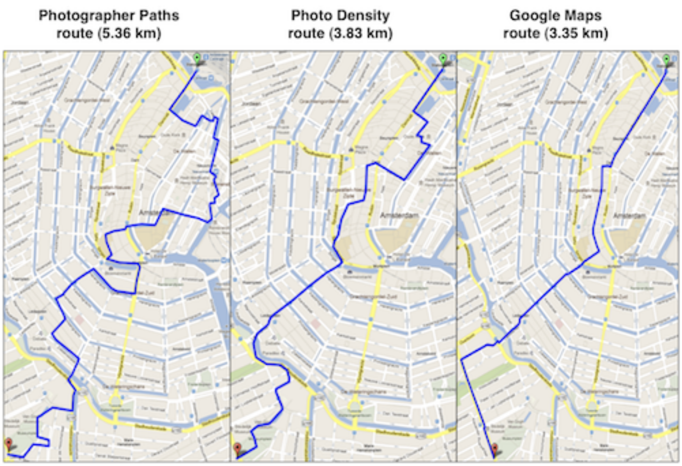

Figure 4.5: Visual comparison of the generated routes from Central Station to Museumplein. *Best seen in color.*

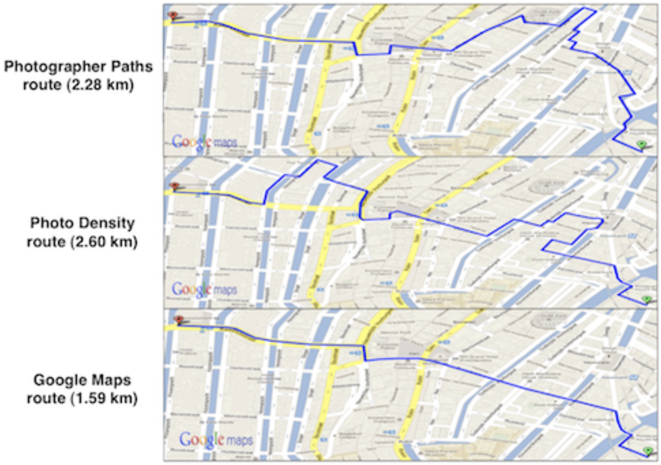

Figure 4.6: Visual comparison of the generated routes from Waterlooplein to Westerkerk. *Best seen in color.*

4.4.6 Results

To turn the 'crude' photographer routes given by our adapted Dijkstra's algorithm into walkable routes, these routes were mapped to a Google Maps map where we took the shortest walking distance between each route node (or location). This resulted in walkable photographer paths. The Photographer Paths (PP), Photograph Density (PD), and Google Maps (GM) route variations for our chosen two routes, Central Station to Museumplein (CM) and Waterlooplein to Westerkerk (WW) route are shown in Fig. 4.5 and Fig. 4.6, respectively. Explanation and motivation for the PD and GM route variations is given below.

4.5 User Evaluation

4.5.1 Laboratory-based study

Study Design

We wanted to evaluate whether our Photographer Paths (PP) route variations are indeed preferred by users for exploration-based route planning. While previous work (e.g., Kjeldskov et al. (2005)) addressed how to evaluate the usability of electronic mobile guides (which may include route planners), there are no established standards on how to best evaluate a service that provides alternative walkable city routes from a human-centered perspective, especially since POI selection accuracy and routing efficiency are not suitable measures for the desirability of the service. However there is work that addresses similar problems. Schöning et al. (2008) evaluated location-based stories generated automatically from Wikipedia[10] by making use of a Likert-type questionnaire. Suh et al. (2011) evaluated their mobile guide in a cultural heritage setting by means of participant observation, questionnaires, and semi-structured interviews. Kurashima et al. (2010) used a quantitative approach where they compared their photographer behavior model against three probabilistic models to account for the accuracy of their personalized route recommendations.

To evaluate whether our route variations are perceived by users as desirable alternatives to current efficiency-based route planners, we chose a user-centered mixed-methods approach that includes both quantitative and qualitative measures. While our target user group included both city residents and tourists, here we focus on expert evaluations from city residents. To test whether our approach provides not just a novel method for generating routes, but routes that city residents would rate as preferable for an exploration scenario, we chose to compare our generated route variation with two other route variations that have a similar start and end destination. Our baseline comparison was a route based on the density of photographs taken per grid cell, where we assumed that this would provide a route plan through the most touristic hotspots within Amsterdam. This was chosen instead of a route that connects a density of all POIs as a POI-density based route would require further differentiating between the kinds of POIs, which is not the aim of a route planner that generates routes automatically without requesting user preferences. For each scenario, participants (who were city residents) had to evaluate the routes in Amsterdam.

Two routes were tested, each with 3 variations: Photographer Paths (PP) route, a Photo Density (PD) variation as baseline, and a Google Maps (GM) efficiency-based route variation. For each route, participants were given scenarios. For the first scenario, participants had to imagine being in the company of local friends on a sunny Saturday between 14-15:00 o'clock, where they wished to walk from Central Station to Museumplein (CM route). For the second scenario, participants had to imagine being in the company of a friend (a local) on a cloudy, Sunday evening between 19-20:00, where this friend just returned from a vacation and they now wished to catch up at a café near Westerkerk (WW route). While both scenarios emphasized there was time to spare, we expected participants to favor efficiency in the WW route.

PD route was created by drawing a path between grid cells containing the highest density of geotagged photos taken in Amsterdam in 5 years, for the hour corresponding to each scenario given to participants (14-15:00 and 19-20:00, respectively). The restriction by hour was set so paths between cells remains meaningful, as plotting a 5-year dataset of geopoints makes it difficult to differentiate between choosing one cell over another. This route served as a popular and touristic route baseline by which to measure our PP route against.

We set up a within-subjects experimental design, where route variation is the independent variable, and measured dependent variables are: a) perceived quality of the presented route variations for each route (CM and WW) b) participants' route preferences and c) subjective reports on what they thought about the generated routes. To measure perceived quality of the route variations, we adapted the AttrakDiff2™ (Hassenzahl et al., 2003) questionnaire[11] so that participants can reflect on the presented routes and give us a quantita-

[10] http://www.wikipedia.org/ ; last retrieved: 01-08-2013

[11] AttrakDiff2™ is a questionnaire originally developed to measure the perceived attractiveness of interactive products based on hedonic and pragmatic qualities. However, the measured bipolar qualities that apply to interactive products can also apply to city routes, making for a suitable domain generalization.

tive measure of the hedonic and pragmatic aspects of each route variation. AttrakDiff2 measures pragmatic and hedonic qualities by allowing participants to provide ratings on a 7-point semantic differential scale for 28 attributes,[12] resulting in 4 quality dimensions: 1) Pragmatic Quality (PQ), which measures usability of a product (or in our case routes). Here, PQ gives insight into how well users can achieve their goal given each route 2) Hedonic Quality - Identification (HQ-I), which gives insight into the extent that users can identify with a given route 3) Hedonic Quality - Stimulation (HQ-S), which gives insight into the extent that a route stimulates users with novelty and enables personal growth 4) Attractiveness (ATT), which provides a global value and quality perception of a route, or in other words, perceived attractiveness.

To get further insight into participants' perception of the route variations, we had a two part semi-structured interview at the end of each testing session, where users could give their feedback directly on their route preferences and inform us what information types they find valuable in deciding whether or not a route affords exploration. For the first part, participants were asked to give their opinion on which route variation they preferred for each route, and what they thought about routes based on photographer paths. In the second part of the interview, they were provided with examples of different digital information aids and asked which (or a combination of) they found useful for exploring a city. These were: a plain Google Maps route, Foursquare POIs (that include short textual comments left by others) along a route, a route with Flickr photos, our PP route segments (made up of PRSs) that shows via color coding the different route segments that make up the photographer paths (see Fig. 4.4), and a route showing multiple photo geopoints (i.e., PD route). Finally, they were asked about the applied potential of this kind of route planning service.

The need to investigate information types (even if not the primary aim of our study) relates to the need for transparency and intelligibility in ubiquitous computing systems (Vermeulen, 2010). On one hand, to make a fair user perception comparison between routes generated by route planners means that further information cannot be provided on a route variation. This is because we risk comparing the effects of information type on route preference, and not the quality of the route itself. On the other hand, in an actual route planner system, users should be given the option to understand 'why' a given route is generated, which is why we had to simultaneously investigate digital information aids in different kinds of media.

Participants

15 participants (10 male, 5 female) aged between 21-35 (M_{age} = 29.2; SD_{age}= 3.3) were recruited. Only participants who had lived in Amsterdam for at least one year were recruited, to ensure that they were able to adequately judge the presented route variations. Our participant sample spanned 9 different nationalities. Most participants claimed to know Amsterdam fairly well (10/15), where the rest knew it either very well (2/15) or just average (3/15). Many (10/15) had a technical background (e.g., Computer Science), and all were familiar with route planning services, where most (10/15) reported using route planners at least once a week.

[12]Only one attribute-pair was changed to fit our study: Technical-Human was replaced with Slower-Faster for the PQ dimension.

Prototype, Setup & Procedure

To test the route variations, an interactive web-based prototype route planner interface[13] was shown to each participant. The interface was adaptable to mobile devices, but testing route preferences on a mobile device was not important as participants were selecting a route based on pre-trip preferences, which usually involves viewing routes on a desktop. The study was conducted at the User Experience lab at XYZ university. Each session took approximately 45 min. to complete. To facilitate discussion and eventual consensus regarding our interview questions amongst participants, participants were interviewed in groups of three. For the first part of the study, each participant was seated in front of a laptop, where they each interacted (zooming, panning) with the route planner interface. For the interview, participants were allowed and encouraged to discuss and answer the questions in a collaborative manner.

Before the study session, each participant filled a background information form, signed an informed consent form, and read through detailed instructions for performing the task. In each task, a participant had to inspect 3 route variations (PP, PD, GM) for each route (CM, WW). The order of presentation of both routes and route variations were counter-balanced and then randomized (after first presented variation) across participants. After inspecting each route variation, a participant had to fill in the 'same' adapted AttrakDiff2™ questionnaire, marking their responses with the corresponding route variation number so their responses where relative to one another. Participants were asked to give their first, spontaneous response. After all three participants finished inspecting all route variations, they were given the semi-structured interview. After the interview, each participant was offered a small monetary reward and thanked for their time.

4.5.2 Web Survey

To test if there is an immediate difference between the generated route variations (PP, PD, and GM), we constructed a short web-based survey to compare each of the route variations for the CM and WW routes. This survey was meant to be short and easy to fill, and to provide a rough idea of whether people consider route plans based on alternative city routes useful, and to collect data on what they find useful digital information aids for exploration. Basic demographic information (age, gender, years in Amsterdam, familiarity with route planners) was asked, and thereafter participants had to choose which route variations they would follow given our two respective scenarios (as in the lab-based study). Here however, there were two main differences to the lab study: a) maps showing each route were static images, and so participants could not zoom in on a location and b) all routes and route variations were shown on a single screen, so order effects were not accounted for.

For the survey, 82 participants (55 male, 27 female) aged between 17-62 (M_{age}= 27.6; SD_{age}= 6.1) responded. Most (44/82) lived in Amsterdam for more than 3 years, some (15/82) between 1-3 years, and the rest either less than a year (11/82) or only visited Amsterdam before (12/82). All were familiar with route planners and GPS-based systems. All participants were considered here, as we were interested in immediate reactions to the presented variations and questions about digital information aids.

[13]The prototype given to participants can be found here: http://staff.science.uva.nl/∼elali/routestudy/welcome.php

	Dimension	Route Variation	M	SD	CI	P-value	η_p^2
CM Route	**PQ**	**PP**	-1.5	.1	[-2,-1]	**p<.001**	
		PD	.4	1.0	[-.1,.9]	F(2,28) = 38.9	.7
		GM	1.7	1.1	[1.2,2.2]		
	HQ-I	PP	.5	.8	[.1,.9]	p = .262	
		PD	.4	.7	[.04,.7]	F(2,28) = 1.4	.1
		GM	-.01	.9	[-.5,.4]		
	HQ-S	**PP**	1.2	.8	[.8,1.6]	**p<.001**	
		PD	.2	.9	[.1,.3]	F(1.3,19.4) = 21.3	.6
		GM	-1.1	1.1	[-1.7,-.6]	(corr. G-G ϵ= .69)	
	ATT	***PP***	.9	1.2	[.4,1.6]	**p<.05**	
		PD	.8	1.0	[.3,1.3]	F(2,28) = 4.8	.2
		GM	-.4	1.4	[-1,.3]		
WW Route	**PQ**	***PP***	-.9	1.4	[-1.6,-.2]	**p<0.01**	
		PD	-1	1.2	[-1.6,-.4]	F(2,28)	.7
		GM	2.3	.9	[1.9,2.8]	= 38.18	
	HQ-I	PP	.1	.6	[-.2,.4]	p = .591	
		PD	.3	.5	[.05,.5]	F(2,28) = .5	.04
		GM	.3	.7	[-.01,.7]		
	HQ-S	***PP***	.5	1.0	[-.01,1]	**p<.001**	
		PD	.7	.9	[.2,1.3]	F(2,28) = 17.3	.5
		GM	-1.4	1.1	[-2,-.9]		
	ATT	PP	.3	1.1	[-.3,.8]	p = .877	
		PD	.3	1.0	[-.2,.9]	F(2,28) = 1.3	.01
		GM	.5	1.0	[-.04,1]		

Table 4.1: Descriptive statistics (N=15) for each route variation (PP, PD, GM) under each tested route (CM and WW) for each AttrakDiff2 dimension: PQ = Pragmatic Quality, HQ-I: Hedonic Quality-Identity, HQ-S: Hedonic Quality-Stimulation, ATT: Attractiveness.

4.5.3 Results

Perceived Route Quality

Responses on the modified AttrakDiff2 in the lab study were analyzed within groups, per generated route. For each category, one-way repeated measures ANOVA tests were conducted comparing results from all route variations. Means, standard deviations, confidence intervals, significance, and (partial) eta-squared values for each tested route variation for each route (CM and WW) are shown in Table 6.1. For the CM route, a repeated measures ANOVA showed significant differences in the responses across quality dimensions for only PQ, HQ-S, and ATT. For the WW route, a repeated measures ANOVA showed significant differences in the responses across quality dimensions for only PQ and HQ-S. Post-hoc pairwise comparisons (with Bonferroni correction[14]) between each variation (PP, PD, GM) were conducted in every case. Where significant, dimensions are represented in bold, and where a particular pairwise comparison is not significant, the dimensions are in (additional) italics.

For the CM route, participants perceived clear differences across all route variations

[14]Backward-corrected SPSS© Bonferroni adjusted p-values are reported.

for the PQ. It is not surprising that the GM variation scored the highest here, given that the route is based on efficiency. Nor is it surprising that our PP variation scored the lowest, given the length of the route. However, it is surprising to see that the PD variation significantly differed from both the PP and GM variation, as it is close in length to the GM variation, yet still considered not pragmatic. This can be partially explained by the fact that the PD route variation runs through all the heavily touristic areas, which may not be very practical to take. For the PP and PD variations, responses for HQ-I were around zero, however there were no significant differences between any of the variations. This suggests that all participants identified with the route variations similarly, but perhaps overall it may be that those variations were not ones that our sample of participants identified with.

For the HQ-S, there were significant differences between all tested route variations, where our PP variation scored the highest. This finding is in line with our hypothesis that photographer paths can provide a stimulating and novel route variation in a city. The mean responses for the PD variation are around zero, which shows that for our participants who are living in Amsterdam, the popular aspect of the PD variation was not stimulating (but nor was it found to be as completely dull and boring as the GM route). Finally, there were significant differences in ATT between PP and GM and between PD and GM route variations, however not between PP and PD. The fastest GM variation was rated the lowest. Here, we would have expected our PP variation to be rated higher than the PD variation, however it could be that participants found the PD variation attractive depending on the company they are with (given the scenario); as they later mention in the interviews, they could be with touristic friends in which case they would find the popular route variations more attractive for the sake of tourism.

For the WW route, participants perceived clear differences for the PQ dimension between the GM route variation and both our PP variation and the PD variation. The GM route variation was not surprisingly rated as the most pragmatic, however here the PD variation was on par with our PP variation. For the HQ-I, there were again no significant differences between any of the route variations, were all responses had a mean around zero. This again suggests that identifying with the generated routes was not important for our participants. For the HQ-S, the GM variation was again rated the lowest, however this time there were no significant differences between our PP variation and the PD variation. This was likely due to the overall short distance between Waterlooplein and Westerkerk, where little room was left for identifying stimulating qualities of the routes. Importantly, in line with the scenario, participants here did not value stimulating qualities of the route variations, as one going to a café with a friend back from vacation on a cloudy day is not a situation that affords a stimulating route variation. This is further confirmed by considering the ATT dimension, where the GM route variation was now rated higher than it was for the CM variation, which indicates that participants valued the short distance for the GM variation. However, given that there were again no significant differences, it seems to still be that all route variations are attractive enough to take, even under efficiency-driven scenarios.

Route Preference

All but one participant in the lab study found the scenarios quite realistic. The one participant objected that it would have been clearer if the gender of friends in the scenarios was given. In the lab-based study, after inspecting all route variations for each route, participants were asked which variation they would follow (if any) and why (convenient, interesting, scenic, or other response). For the CM route, most participants (9/15) chose to follow our PP route variation, where the rest picked the PD (4/15) and GM (2/15) variations. For all cases where participants picked the PP variation, they stated it was more scenic. This is in line with our hypothesis that the PP variation provides a favorable route to take in a city one has time to explore. During the interviews, participants mentioned that they found the PP variation attractive and suitable for the scenario (P3: *"I liked the second route variation [PP] given it was nice weather so perfect for exploration."*; P12: *"One of the routes [PP] was long and took many detours, and I thought that was a very attractive route!"*).

Also some reported that the PD route variation (as well as the GM variant) to be very touristic, which they did not like (P5: *"I would obviously not go through the Kalverstraat as it is very touristic.. the first [GM] and last [PD] route took me through there, and I think that is not really Amsterdam."*). Still, others found the PP route too long (P10: *"I thought the [PP] route was too long, my feet would fall off!"*) and preferred something in the middle between the PP and GM variations (P14: *"If I want to see the city, I would go for the [PD] one."*). In the web survey, most participants said they would follow the GM route (40/82), followed by the PD route (23/82) and the PP route (10/82), where the rest chose neither (9/82). The difference in findings here can be explained either because participants did not inspect the routes carefully or take the scenarios seriously, or perhaps more likely those that lived in Amsterdam already knew their own specific routes, which they take by commuting by bicycle and largely avoiding touristic areas (where both PP and PD route variations pass through). Unlike the lab study, the experimenter could not steer participants to stick with the scenario – this was evident by most comments left in the free-form box form (e.g., *"I would not easily walk these routes... who in amsterdam walks? ;)"*; *"I would use cycling routes."*). Nevertheless, they highlight that alternative route plans may take time to be accepted as city exploration aids (alongside routes suggested by expert travel guides) in both familiar and unfamiliar cities.

For the WW route, most participants (10/15) in the lab study favored the GM route variation, followed by the PD variation (4/15) and in one case neither one. For all participants that picked GM, they stated it was more convenient. This was not surprising under the WW scenario, where it was a cloudy evening and meeting a friend returning from a vacation. Indeed, later in the interviews this was mentioned explicitly (P3: *"You are going for coffee so you just want to get there, unlike in the first [CM] scenario where it is a nice day and you have time."*). Here, it was also mentioned that our PP route variation and PD both went off to the Red Light District, which may not be desirable for them (P11: *"Second [GM] route was very simple and straightforward... Red light part is more interesting for tourists."*). In the web survey, most participants said they would follow the GM route (67/82), followed by the PD route (6/82) and the PP route (3/82), where the rest chose neither (6/82). Favoring the GM variation for the WW scenario is in line with the findings of the lab-based study.

Digital Information Aids for City Exploration

During the post-test interview, participants were asked about digital information aids they would find useful when deciding to take a route in a city for exploration purposes. They were asked explicitly about what they thought about information that tells you how many persons (particularly city photographers) took a given route segment over a certain time period (e.g., 1 year). Around half (8/15) found such information useful for exploring a city one already knows, some where not quite sure (4/15), some thought it depended on who those photographers were (2/15), and one thought this is not something for him. In the second part of the interview (after participants saw examples of different digital information aids), many (10/15) found our PP information type attractive (e.g., P4: *"Nice to find corners that I don't really know."*). This was visualized by coloring the different segments of a route on a map and stating how many photographers took a segment over time (for us, 4 photographers through 2 locations over a 5 year period). However, many of those participants (6/10) also stated that they would combine this information with photos (3/6) and POIs (3/6). While some mentioned that the PP information is useful for exploring a familiar city (P8: *"[PP] information is useful if you are familiar with the city."*), still others thought it redundant (P3: *"Hard to see why I'd use a map in a city I already know."*).

In the web survey, we gained further insight into what digital aids people find useful in helping them explore a city. These are (ordered by count of mentions): POIs along a route (51x), route distance (51x), comments along a route (ranked by highest ratings or recency) (24x), expert travel guides (22x), photos of route segments (17x), no digital aids (13x), number of photographers that took a given path over a time period (9x), number of photos along a route over a time period (9x). While we expected that established aids such as POIs and distance are useful indicators for planning personal routes, it was surprising that there were only few reports of the usefulness of photographer paths to guide exploring a city. This was also evident with photo counts along a route, which is also a novel information aid. This may be because participants are not familiar with such novel aids, especially since in the web survey (to avoid biasing route preference and save time), they were not provided with visualizations of how this information may be visualized. Considering the findings of the lab-study and web survey, it seems that photographer paths as city exploration aids have potential, but stating it without visualizing how it could intelligibly augment an alternative route plan (as was done in the lab-based study) may have influenced its immediate adoption by participants in the web study.

Use of Non-Efficiency Driven Route Planners

Finally, participants were asked what they thought about future route planning applications that generate alternative walkable routes in a city, both locally and abroad, and whether they would use them. Most participants (13/15) were positive about such applications (P7: *"Yes. Helps you to explore more and discover different things, gives you another option."*; P4: *"Yes, especially for a specific corner of a city."*), where the remaining two brought to question why they would use such apps in a city they already knew (P12: *"I would like it more for a city that I don't already know."*; P14: *"Google Maps is enough."*). Responses on what these route planners should be based included: personalized route plans (e.g., in accordance with route travel history) and preset route plans (e.g., museum route, market route, etc.). Together, the responses from participants highlight that a new generation of route planners aimed for exploring cities is desirable.

4.6 Discussion

4.6.1 Study Limitations

In this chapter, we have presented a novel approach based on sequence alignment methods to generate exploration-based route plans in a small-sized city like Amsterdam. Our lab-based user test provided strong evidence that our generated route plans are of potential interest to city dwellers especially by way of hedonic stimulation, under an exploration-based scenario. However, while most participants mentioned they would follow our PP route variation in the CM scenario, this was not the case for the web survey. This brought to question not only whether such a quick survey was able to sufficiently provoke reflection on the given scenarios and routes, but also that alternative city route plans may not be immediately accepted by both locals and visitors to a city. Furthermore, without providing intelligible explanations (cf., Vermeulen (2010)) as was done in the lab-study (visualizing the digital aids on the generated route plans) for why a route was given may have made it difficult for potential users to make an informed judgment.

Another limitation is the real-world evaluation of the generated routes. While we have tried to tease out the differences between each route on participants' route preferences in a lab-based and web-based study, we may not immediately generalize to how participants would react to such routes if used in real-time in an actual situation with a working prototype. A related issue is that the number of route variations presented is limited to two, where participants' reactions may differ in such cases. Nevertheless, the findings from the lab-study provide strong evidence that photographer paths have potential for generating desirable route plans in the city, and importantly highlight the merits of an automatic data-driven approach based on geotagged photographs in a city.

4.6.2 Towards Automatic Exploration-based Route Planners

We started from the question of how human sequential movement patterns can be leveraged as an indicator of interest, and attempted to answer this by relying on a data-driven approach that borrows SA methods from bioinformatics and time geography. Our approach goes against the established human-centered literature on catering for user needs, where typical mobile recommender studies (e.g., Bellotti et al. (2008)) begin with distilling requirements from observing and interacting with users. Here, it can be argued that such a quantitative approach may drastically oversimplify our human needs for exploring a city. However, we wanted to compute routes based on paths taken by photographers in a city *automatically*, without burdening the user to tell us her desires.

Under the foregoing motivation, we have created an opportunity to consider a currently unused information type obtained from geotagged images to guide exploration-based route planning in a city: the number of city photographers that took a given route segment over a certain time period. With this proof-of-concept approach, we have shown it is possible to leverage social geotagged data to cater for the hard problem of automatically generating exploration-based route plans.

4.7 Conclusions

In this chapter, we have presented a proof-of-concept approach that uses sequence alignment methods and the digital footprints of city photographers (obtained through Flickr) to compute exploration-based route plans within a small-sized city like Amsterdam. From our user study with Amsterdam residents, we found that our photographer paths are promising for city exploration, where we believe our findings set the stage for further experimentation with data-driven approaches to tackle the hard problem of automatically generating 'interesting' route plans for exploring both familiar and unfamiliar cities.

The three chapters in Part I of this thesis showed that context-awareness, our first theme of minimal mobile interaction, can contribute to making user interactions in urban settings simpler or more playful. To investigate the second theme of minimal mobile interaction, in Part II of this thesis we focus on non-visual input techniques. In the following chapter (Chapter 5), we investigate the user experience of 3D gesture recognition interaction techniques. Since these 3D gestural interaction techniques are error-prone, we first look at how users deal with errors under varying failed recognition error rates.

Part II

Gestural Input Techniques

5

Effects of Error on Device-based Gesture Interaction

We investigate the usability and user experience of 3D gesture-based interaction. Specifically, what happens when 3D gestures performed by users are not recognized (*Gesture Errors Study*). We present a controlled laboratory study to arrive at which gesture set is more robust to recognition errors under varying error rates. Drawing on experiment logs, video observations, participants' feedback, and a subjective workload assessment questionnaire, results showed that while alphabet gestures are more robust to recognition errors in keeping their signature, mimetic gestures are more robust to recognition errors from a usability and user experience standpoint. Thus, we find that mimetic gestures are better suited for inclusion into mainstream device-based gesture interaction with mobile phones. The work presented in this chapter was published as "Fishing or a Z?: Investigating the Effects of Error on Mimetic and Alphabet Device-based Gesture Interaction" in the Proceedings of the 14th international conference on Multimodal Interaction (El Ali et al., 2012).

5.1 Introduction

Whether we like it or not, errors and failures are an inevitable part of interaction with technology. Device-based gestures (gesturing by moving a device in 3-dimensional space), used in research settings, home environments, or as part of everyday mobile interactions, are gaining potential in becoming a suitable alternative to keyboard or touchscreen-based input, especially when users are encumbered (e.g., manual multitasking (Oulasvirta and Bergstrom-Lehtovirta, 2011)). Yet whether a gesture fails due to poor system design or due to the user's actions, errors impede smooth (multimodal) interaction, as well as the adoption of these novel gesture-based interaction methods. This is critical if gesture-based interaction is included in consumer mobile phones.

These device-based gestures (or motion gestures (Ruiz et al., 2011))[1] are to be distinguished from surface 'touchscreen' gestures (Wobbrock et al., 2009)), which typically

[1]We use the term 'device-based gestures' and not 'motion gestures' to emphasize that these gestures require holding a device, and not solely about motion as in vision-based gesture recognition such as interacting with a Microsoft Kinect©.

involve two-dimensional gesture interaction on a surface, such as a tabletop or mobile touchscreen. In contrast, device-based gestures typically make use of accelerometer and gyroscope sensors to allow users to rotate the device in the air for performing gestures in three-dimensional space. While there has been extensive taxonomy-driven work on classifying gestures by referential function into classes (Ruiz et al., 2011; Rime and Schiaratura, 1991) and recently identifying usable sets of gestures as defined by users (Ruiz et al., 2011; Rico and Brewster, 2010; Kray et al., 2010), it is still an open question which set of gestures are most robust to errors in gesture-based interaction from a performance-centered standpoint.

In this chapter, we look closely at device-based gesture performance using two iconic gesture sets, mimetic (e.g., mimicking a handshaking behavior while holding a mobile device) and alphabet gestures (e.g., drawing the letter ‘N’ in the air using a mobile phone as a brush), and investigate how their referent-independent performance is affected by varying failed recognition rates.

5.2 Research Questions

Our main research question is:

> **RQ 4:** What are the effects of unrecognized 3D gestures on user experience, and how do these affect the design of error-tolerant 3D gesture sets?

Specifically, how do mimetic and alphabet gesture sets evolve in the course of interaction when the performed gesture is not recognized under varying error rates? Here, we investigate how users react when they are less or more familiar with the ideal shape of a gesture, under varying error conditions. Our hypothesis is that since mimetic gestures are less familiar than alphabets, we expect participants to call on their own real-world sensorimotor experience (on how they perform certain activities in daily life) to perform a gesture. We expect the performance of a mimetic gesture to be influenced by this experience, thus morphing the iconic gesture into that previously learned gesture. For example, while a ‘fishing’ gesture might be designed in a particular way to perform some system function (e.g., hold device flat on palm, tilt towards you, and place device back on palm), given unfamiliarity with its designed ideal form, we suspect that this same gesture is more likely to morph into a more natural, learned fishing gesture upon failed recognition attempts. Given the many ways to fish (where variations could be due to cultural or individual differences), we expect the evolved gesture to exhibit more variation in the face of increasing error, especially since participants do not know the ideal shape of the mimetic gesture. Complementarily, this variation arises due to the higher degrees of freedom permitted in performing that gesture.

By contrast, we expect alphabet gestures to exhibit much less variation, instead becoming more rigid and structured after repeated failed recognition attempts – this is because alphabet gestures not only have lower degrees of freedom (namely, 2df), but the set of ideal visual shapes is more familiar to users. Similar to work in speech and handwriting recognition, we expect a hyperarticulation (Read et al., 2002; Oviatt et al., 1998) of gestures as participants begin to gesture more carefully to arrive at the ideal shape required by the recognizer. This we hypothesize will negatively impact the user experience (UX)[2] of performing these gestures.

Investigating the usability differences in a primarily qualitative manner between mimetic and alphabet gestures here yields two main research contributions: first, it aids gesture designers in selecting which gesture set (mimetic or alphabet) is more robust to errors, and hence better suited for inclusion into accelerometer and/or gyroscope equipped mobile devices. This is achieved by providing a deeper understanding of whether some gestures are intrinsically more tolerant to recognition errors than others (e.g., by showing that mimetic gestures, in lacking an ideal shape, can foster more variation in gesture-based interaction). Second, it equips gesture designers with the knowledge of which gesture sets overall induce a lower subjective workload to perform and which increase in duration, especially under conditions of high recognition failure.

Additionally, we provide initial results on how errors can be an impeding factor in performing gestures in public, as well as use-cases for the tested gestures participants reported on. The rest of the chapter is structured as follows: first we provide a review of related work, then we present our study design and methods, give our results and discuss them, provide design recommendations and finally conclude.

5.3 Related Work

5.3.1 Using 3D Gestures in HCI

There is a recent trend in Human-Computer Interaction to provide what are called Natural User Interfaces (NUIs) (Jain et al., 2011). As stated by Jain et al. (2011), this class of interfaces enables users to interact with computers in the way we interact with the world. An important element of such natural interaction is the use of 3D gestures. However gestures alone do not suffice to allow seamless natural interaction, as the user still needs to receive feedback from the system on a performed gesture (Norman, 2010). This is usually complemented by the use of touchscreen buttons, menus, auditory or speech feedback, or some kind of visual feedback from the system. Nevertheless, in allowing the use of 3D gestures for activities such as music composition (e.g., guitar strumming), the primary interaction with a system can be seen as largely natural.

Recently, Grandhi et al. (2011) investigated the naturalness and intuitiveness of gestures, where the goal was to understand how users' mental models are aligned to certain gestures. Relevant finding here is that tasks that suggest the use of a tool should have gestures that pantomime the actual action with the imagined tool in hand. This points to the importance of tool use in gestural interaction where appropriate. Khnel et al. (2011) investigated a user-defined gesture set for gesture-based interaction with smart-home systems (e.g., opening and closing the blinds), which highlighted the merits of this natural interaction method for daily activities in the home. Another line of research by Rico and Brewster (2010) focused on the social acceptability of produced gestures under different settings (e.g., at home, at the pub, etc.), where they also found that some gestural interactions were perceived by users to be enjoyable. The goal of their work was to equip gesture

[2] UX here is based on ISO 9241-210 (1994) definition: "A person's perceptions and responses that result from the use or anticipated use of a product, system or service."

designers with knowledge of which gestures are socially appropriate under which settings and situations. Additionally, they Together, the foregoing studies demonstrate the growing use of 3D gestures as an intuitive and natural interaction alternative to mobile touchscreen interactions, that is generally perceived to be socially acceptable. However, while much research has focused on the naturalness, intuitiveness, and the social consequences of performing certain (surface and device-based) gestures, little research has addressed how issues of failed recognition can transform a produced device-based gesture in the course of interaction.

5.3.2 Dealing with Recognition Errors Across Modalities

Human factors research in multimodal interaction concerned with recognition errors (Mankoff and Abowd, 1999) is a well researched topic in multimodal interfaces (Bourguet, 2006; Oviatt, 2000; Oviatt and VanGent, 1996), where investigations were typically concerned with error handling strategies devised by users in the face of recognition errors (e.g., modality switching to a 'familiar', more efficient modality). In speech-based interfaces, a common finding is that the most intuitive and instinctive way for correcting errors in speech is to repeat the spoken utterance (Suhm et al., 2001) and hyperarticulate it (Oviatt et al., 1998). For multimodal systems, a common error-correction strategy is to repeat a modal action at least once prior to switching to another modality (Oviatt and VanGent, 1996). Oviatt and VanGent (1996) observed a repetition count (called 'spiral depth') of depth 6, before users would switch to another modality. In a follow-up study by Halverson et al. (1999), they tested 3 commercial Automatic Speech Recognition (ASR) systems where they found that a little over 50% of the time, participants would continue to repeat the utterance to a spiral depth of level 3.

However, while recognition errors have been well studied in domains such as speech-based interfaces (Bourguet, 2006; Oviatt et al., 1998), handwriting recognition (Read et al., 2002; Webster and Nakagawa, 1998), and multimodal interfaces (Oviatt, 2000), less attention has been given to usability issues surrounding device-based gesture interaction. An exception is the study by Karam and Schraefel (2006), where they investigated user tolerance for errors in touch-less computer vision-enabled gesture interaction under both desktop (keyboard readily available in front of participants) and ubiquitous computing settings (keyboard not readily available). In this Wizard-of-Oz study, they found that the interaction context played a significant role in how tolerant users were to errors. Specifically, they found that in the ubiquitous computing scenario, users are much more tolerant to errors than in the desktop condition (where recognition error rates can potentially reach 40%) before users will abandon gesture based interaction in favor of traditional, keyboard-based input.

While the results of the study by Karam and Schraefel (2006) are relevant to the current work, our work differs in three main ways: first, their goal was to investigate whether and to what extent participants would switch modalities when confronted with recognition errors, and not study how gestures evolve in response to error. Second, they were concerned with a single gesture, and not different gesture sets and their respective performance by users under different recognition error rates. Finally, their concern was with computer vision-enabled interaction, and not device-based gesture interaction.

5.3.3 Gesture Taxonomies

Rime and Schiaratura (1991) distinguish between symbolic gestures and pantomimic gestures. Symbolic gestures are gestures that have come to take up a single, culturally-specific meaning. Examples include American Sign Language (ASL) gestures and natural language alphabets (e.g., letter ‘C’). Pantomimic gestures on the other hand, are used for showing (through mimicry) the use of movement of some invisible tool or object in the speaker’s hand. For example, when a speaker says “I turned the steering wheel hard to the left”, while mimicking the action of turning a wheel with both hands, they are performing a pantomimic gesture. Here, we are not concerned with the external gesture referent, only its iconic movement in space.

In order to study user-defined motion gestures, Ruiz et al. (2011) devised an initial taxonomy by which to classify gestures. Here, they distinguished between metaphor gestures (metaphor of acting on a physical object other than the phone), physical gestures (direct manipulation), symbolic gestures (visually depicting a symbol by e.g., drawing with the phone), and abstract (gesture-activity mapping is completely arbitrary). In the foregoing taxonomies, the classifications are based on the communicative intent behind performing the gesture (i.e., its representational aspect), and not on the usability and user experience (UX) of manually performing the gesture.

5.4 Methods

5.4.1 Study Design

In this study, we do not to test the communicative intentions of participants while performing gestures (i.e., no referential aspect), but only the manual ’performance’ of the iconic gesture itself. Under this performance-centric view, the mimetic gestures are iconic movements of real-world actions and the alphabet gestures are iconic movements of writing letters.

Mimetic gestures were chosen as they provide a natural means of interaction, making them good candidates for inclusion into mobile technology. Similarly, alphabets can be produced by any literate person, and hence suitable for comparison. Other symbols (e.g., salute or high-five gesture) were not tested to avoid undue cultural bias. Furthermore, like mimetic gestures, alphabets can also be easily learned and recalled, and have practical potential for use in mobile technology (e.g., mid-air writing for user authentication (Ketabdar et al., 2010b)). Given these design considerations, we designed 12 gestures (6 mimetic, 6 alphabetic). Mimetic gestures are: Fishing, Trashing, Shaking, Handshake, Throwing, Glass Filling. These specific gestures were chosen because they represent familiar yet different activities, where all have more than 2 degrees of freedom. Alphabet gestures we chose are English letters (given our participant pools’ language), varied by difficulty (e.g., 2 strokes or 3 to draw the letter). Only 6 different gestures for each set was chosen to avoid participant fatigue, given the high number of trials (200 total) each participant entered. Nevertheless, we believe the chosen gesture samples provide sufficient variation to characterize each gesture set with respect to varying failed recognition rates. Both gesture

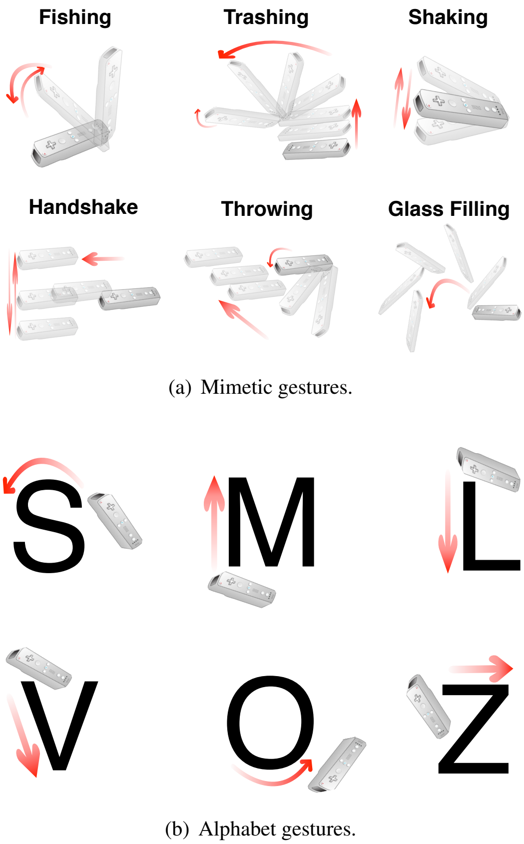

Figure 5.1: The designed gesture sets for (a) mimetic gestures and (b) alphabet gestures.

sets, and their movement in space, are shown in Fig. 5.1.

To investigate how different types of performed gestures respond to varied recognition errors, we used an automated Wizard-of-Oz method (Fabbrizio et al., 2005) where we simulated each of three failed recognition error rate conditions: low (0-20%) error rate, medium error rate (20-40%), high error rate (40-60%). Participants were told that real gesture recognition engines were being tested, where each of the algorithms differs in terms of how sensitive it is to the user's gesture interaction style. When participants performed a gesture, the automated wizard would draw for each gesture block an error rate randomly from the assigned error rate range specific to the condition. When a gesture is performed, the participant receives coarse feedback (recognized / not recognized) on whether or not the performed gesture was successfully recognized.

For this study, testing real recognizers is irrelevant as we are only interested in the usability and user experience (UX) of gesture performance under the chosen error rates. This is both in line with previous work (Karam and Schraefel, 2006), and conveniently allows testing our research questions without the unpredictability of real recognizers. Importantly, here we study gesture performance, not how different feedback improves gesture learnability. Additionally, we tested task-independent gestures for two reasons: first, it allows us to understand the differences between the two gesture sets (mimetic and alphabet) independent of task type. This would eliminate potential participant bias in both workload and expected gesture evolution under error conditions due to the mapping of a given gesture to a task. Second, following the study by Kray et al. (2010), it allows participants to freely speculate about the applied real-world use of these two gesture types.

The conducted experiment was a mixed between- and within-subjects factorial (2 x 3) design. A between-subjects design between the two gesture sets was chosen for two reasons: first, to disallow any transfer effects between the two gesture sets thereby avoiding any contamination between the gesture mental models formed in participants. Second, testing all gestures in one session would excessively lengthen the duration of the experiment, and pose a risk of participant fatigue. There are two independent variables (IVs): gesture type (2 levels: mimetic vs. alphabet) and recognition errors (3 levels: low (0-20%) vs. medium error (20-40%) vs. high error (40-60%), where gesture-type was a between-subjects factor and error rate a within-subjects factor. Each between-subjects condition tested 6 gestures (12 total), randomized across participants. Each gesture occurred in all within-subjects conditions (counterbalanced across participants), in addition to two practice blocks, which resulted in 20 gesture blocks per experimental session. Each block consisted of 10 trials. In a block, participants were asked to perform a given gesture using a Wii Remote® 10 different times (once per trial), where the error rates are randomly distributed within the corresponding recognition error level. In the practice blocks however, the error rate was always low. In total, each participant entered 200 trials (20 practice, 180 test).

The experiment was coded using NBS Presentation®,[3] an experimental control and stimulus delivery software. Interaction and syncing with the Wii Remote was done using GlovePie,[4] a programmable input emulator. Five data sources were collected: modified NASA-TLX workload data (Brewster, 1994; Hart and Wickens, 1990), experiment logs, accelerometer time stamps, gesture video recordings, and post-experiment interviews. The modified NASA-TLX questionnaire assessed participants' subjective workload quantitatively ([0,20] response range) through the index's constituent categories: Mental Workload, Physical Workload, Time Pressure, Effort Expended, Performance Level Achieved and Frustration Experienced (Hart and Wickens, 1990), plus the additional categories of Annoyance and Overall Preference (Brewster, 1994). Given no time pressure imposed on participants in the study, we did not use this category. Additionally, as the Annoyance category is specific to audio interfaces, we only made the additional use of the Overall Preference category. For early signs of changes in the gestures under different error rates (particularly signs of over-articulation), we analyzed the duration of the performed gestures.

5.4.2 Participants

24 participants (16 male, 8 female) aged between 22-41 (M_{age}= 29.6; SD_{age}= 4.5) were recruited. Our participant sample spanned 8 different nationalities, where all but one were right-handed (23/24). Many participants (17/24) had a technical background, and most

[3]http://www.neurobs.com/; last retrieved: 01-08-2013

[4]http://sites.google.com/site/carlkenner/glovepie; last retrieved: 01-08-2013

(19/24) were familiar with gaming consoles that use some form of gesture recognition technology (e.g., Nintendo Wii© or Microsoft Kinect©).

5.4.3 Setup & Procedure

The experiment was carried out at the usability lab at Nokia Research Center in Tampere, Finland. Each experimental session took approximately 1 hour to complete. Participants were seated in front of a monitor, where a tripod-mounted camera was aimed at their gesture interaction space (see Fig. 5.2). They were allowed to define their own interaction space to ensure their comfort during the session, so long as it was still within the camera's view. Prior to the experiment, each participant filled a background information form, signed an informed consent form, and read through detailed instructions for performing the task. After reading the instructions, a tutorial was given on how to perform each gesture. The tutorial involved the experimenter performing each gesture (using the Wii Remote) right next to the participant.

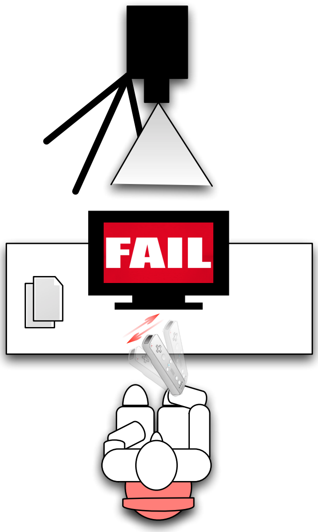

Figure 5.2: Illustration of the experimental setup.

The first two blocks were practice blocks, set always at a low error rate. Before each block, a video of how the gesture to be performed in the next trials was shown on the screen. The performance of the gestures in the videos was identical to how they were performed by the experimenter in the tutorial. The videos were shown to eliminate any failed memory recall effects, where participants' (multimodal) interaction requires a visual input (videos watched) and a translation to somatosensory output (performed gesture). In a trial, a participant would be instructed on screen in text to perform the gesture for that block (e.g., "Perform Fishing gesture."). A participant would have to press and hold the A button on the Wii Remote to start recording the gesture, and release it after completing the gesture. After performing the instructed gesture, if a participant falls into a successful gesture recognition trial, a green checkmark image is flashed on the screen, while in a failed recognition trial, a red image with the word "FAIL" is flashed on the screen. After each block, participants were asked to fill in the modified NASA-TLX questionnaire, where they were provided with an optional 2 min. break after completing it. Participants were allowed to change their responses on the questionnaire at the end of each block, so that their responses per block can be truly relative to one another. After completing the experiment, participants were interviewed for around 10 min. about their experience with the experiment. Afterward, they were thanked for participating, and offered a movie theatre ticket.

5.5 Results

5.5.1 Subjective Workload

The modified NASA-TLX responses were analyzed within groups, per type of gesture. For each modified NASA-TLX category, one-way ANOVA repeated measures tests were conducted comparing results from all error rates. A mixed between- and within-subjects ANOVA revealed no significant differences between the two gesture sets, and therefore not reported. Descriptive statistics (means, standard deviations, 95% confidence intervals, p-values, partial eta squared) of within-subject results for the mimetic and alphabet gestures are shown in Table 5.1 and Table 5.2, respectively. Means and confidence intervals for each category under mimetic and alphabet gesture conditions are shown in Fig. 5.3 and Fig. 5.4, respectively.

For the mimetic gesture condition, a repeated measures ANOVA showed significant differences in the ratings across error rate conditions for Performance Level Achieved ($F(2,22) = 11.19$; $p<0.01$)), Frustration Experienced ($F(2,22) = 23.23$; $p<0.01$)), Overall Preference ($F(2,22) = 5.16$; $p<0.05$)), and Subjective Workload ($F(2,22) = 12.25$; $p<0.01$)). For the alphabet gesture condition, there were significant differences in the ratings across error rate conditions for Effort ($F(2,22) = 6.84$; $p<0.05$)), Performance Level Achieved ($F(2,22) = 23.47$; $p<0.01$)), Frustration Experienced ($F(2,22) = 10.97$; $p<0.01$)), Overall Preference ($F(2,22) = 23.58$; $p<0.01$)), and Subjective Workload ($F(2,22) = 16.63$; $p<0.01$)). Post-hoc pairwise comparisons (with Bonferroni correction[5]) between error conditions (Low-Medium, Low-High, Medium-High) were conducted in every case. Where significant, they are represented in the graphs as bars between low and medium error rate conditions, and as brackets between low and high error rates, with the corresponding significance levels (**, $p < 0.01$; *, $p < 0.05$).

Since there were no significant differences between the two gesture sets, it appears to be that experimentally our wizard-of-Oz recognizer for both gesture conditions had a similar effect on participants, where statistically the two independent groups did not treat the gesture sets differently. However, there were differences between gesture sets with respect to the differences in error conditions. Based on the within-subjects ANOVA results and post-hoc pairwise comparisons on the modified NASA-TLX scores for mimetic gestures (Table 5.1; Fig 5.3) and alphabet gestures (Table 5.2; Fig 5.4), we summarize our findings.

Results showed that while participants in the alphabet gesture condition had to place significantly more effort (Effort Expended) between the low error rate and the medium and

[5] Backward-corrected SPSS© Bonferroni adjusted p-values are reported.

Mimetic Gestures

Category	Error	M	SD	95% CI	P-value	η_p^2
Mental	Low	5.3	3	[3.6, 7]	p= .098	
	Med	5.8	3.3	[4, 7.7]	$F(2,22) = 2.6$	.2
	High	7.2	3.5	[5.3, 9.2]		
Physical	Low	8	4.9	[5.3, 10.8]	p = .365	
	Med	7.7	4.5	[5.1, 10.2]	$F(2,22) = 1$	.1
	High	8.4	5	[5.6, 11.3]		
Effort	Low	7.5	4.7	[4.8, 10.2]	p=.119	
	Med	7.7	4.4	[5.2, 10.3]	$F(2,22) = 2.3$	.2
	High	8.9	4.5	[6.4, 11.5]		
Perform.	Low	15.6	2.6	[14.1, 17.1]	**p<.05**	
	Med	13.7	1.9	[12.6, 14.7]	$F(1.4,15.1) = 11.2$	.5
	High	10.3	4.4	[7.8, 12.8]	(corr. G-G ϵ= .68)	
Frust.	Low	5.5	3.1	[3.7, 7.3]	**p<.001**	
	Med	7.7	3.6	[5.7, 9.7]	$F(2,22) = 23.2$	.7
	High	10.7	4.3	[8.3, 13.2]		
Pref.	Low	13.3	2.7	[11.8, 14.9]	**p<.05**	
	Med	11	2.6	[9.5, 12.5]	$F(2,22) = 5.1$	.3
	High	9.7	4	[7.5, 12]		
Workload	Low	5.1	2.1	[3.9, 6.3]	**p<.001**	
	Med	5.9	2.2	[4.6, 7.2]	$F(2,28) = 13.2$	.5
	High	7.5	2.8	[6, 9.1]		

Table 5.1: Descriptive statistics for mimetic gestures (N=12) under different error rates (Low, Medium, High).

low and high error rate conditions, this was not so for the mimetic gestures. Additionally, for the mimetic gestures, Performance Level Achieved between the low and medium error rates was not significant, while it was significant across all error rates for the alphabet gestures. In the mimetic condition, Frustration Experienced was significant across error conditions, however in alphabet gestures, Frustration was significant only between Low-Medium and Low-High. This shows that for alphabet gestures, frustration is more or less consistently experienced beyond 20% error rates. Interestingly, Overall Preference in the mimetic gesture condition fell significantly only between the low and high error rates and between low and medium, while it significantly dropped for each error rate in the alphabet gesture condition. This hints at a feeling of helplessness in the face of errors for alphabet gestures, possibly because fewer parameters can vary for this gesture set when participants repeatedly try to recover from these errors and still fail.

Finally, Subjective Workload for participants in the mimetic gesture condition significantly differed only between low and high error rate conditions, and between the medium and high error rate conditions, while there were significant workload increases across all error rates for participants in the alphabet gesture condition. Together, these findings suggest that mimetic gestures are better tolerated under error rates of up to 40% (cf., Karam and Schraefel (2006)), compared with error rates of up to only 20% for alphabet gestures. From a usability perspective, our workload results indicate that mimetic gestures are more robust to recognition failure than alphabet gestures, likely due to the higher design space

Alphabet Gestures						
Category	**Error**	M	SD	**95% CI**	P**-value**	η_p^2
	Low	5.8	3.1	[4, 7.6]	p=.872	
Mental	Med	5.5	2.7	[3.9, 7.1]	F(2,22) = .1	.01
	High	6	2.2	[4.8, 7.2]		
	Low	10.2	3.4	[8.3, 12.1]	p = .403	
Physical	Med	9.2	3.4	[7.3, 11.2]	F(1.3,14.1) = .9	.08
	High	10.8	3.2	[9, 12.7]	(corr. G-G ϵ= .69)	
	Low	7.4	2.4	[6, 8.8]	**p<.05**	
Effort	Med	8.8	2.9	[7.1, 10.4]	F(2,22) = 6.8	.4
	High	9.9	3.5	[7.9, 11.9]		
	Low	15.5	2.2	[14.2, 16.8]	**p<.001**	
Perform.	Med	12.3	2.4	[11, 13.7]	F(1.2,13.5) = 23.5	.7
	High	8.7	4.7	[6.1, 11.4]	(corr. G-G ϵ= .61)	
	Low	4.2	2.5	[2.8, 5.7]	**p<.001**	
Frust.	Med	6.3	3.2	[4.5, 8.2]	F(1.1,12.6) = 11	.5
	High	9.6	5.6	[6.4, 12.8]	(corr. G-G ϵ= .57)	
	Low	13.3	2.5	[11.9, 14.8]	**p<.05**	
Pref.	Med	11.2	2.8	[9.7, 12.8]	F(1.3,14.3) = 23.6	.7
	High	8.3	3	[6.6, 10]	(corr. G-G ϵ= .65)	
	Low	5.4	1.1	[4.7, 6]	**p<.001**	
Workload	Med	6.2	1	[5.7, 6.8]	F(2,22) = 16.5	.6
	High	7.9	2.1	[6.7, 9.1]		

Table 5.2: Descriptive statistics for alphabet gestures (N=12) under different error rates (Low, Medium, High).

available for users to experiment with given their unfamiliarity with the ideal shape of the designed gesture.

5.5.2 Gesture Duration

Also within groups, a one-way repeated measures ANOVA was run for each gesture in both the mimetic and alphabet gesture conditions. The means and confidence intervals (based on 12 trials for each gesture per error range) as well as the results from the ANOVA tests for the mimetic and alphabet gesture conditions are summarized in Fig. 4 and Fig. 5, respectively. Post-hoc pairwise comparisons (with Bonferroni correction[6]) between error conditions (Low-Medium, Low-High, Medium-High) were conducted in every case. Where significant, they are represented in the graphs as bars between low and medium error rate conditions, and as brackets between low and high error rates, with the corresponding significance levels (**, $p < 0.01$; *, $p < 0.05$).

In the mimetic gesture condition, there was a significant difference in mean duration due to differences in error rates for only the Fishing gesture, $F(2,22) = 6.24$, $p < 0.01$. Differences between low (2150, 95% CI [1753, 2547]) and high error rates (2616, 95% CI [2154, 3078]) were significant ($p < 0.05$). In contrast, the alphabet gesture condition, there was a significant difference in mean duration in error rates in the L gesture, F(2,22)

[6]Backward-corrected SPSS© Bonferroni adjusted p-values are reported.

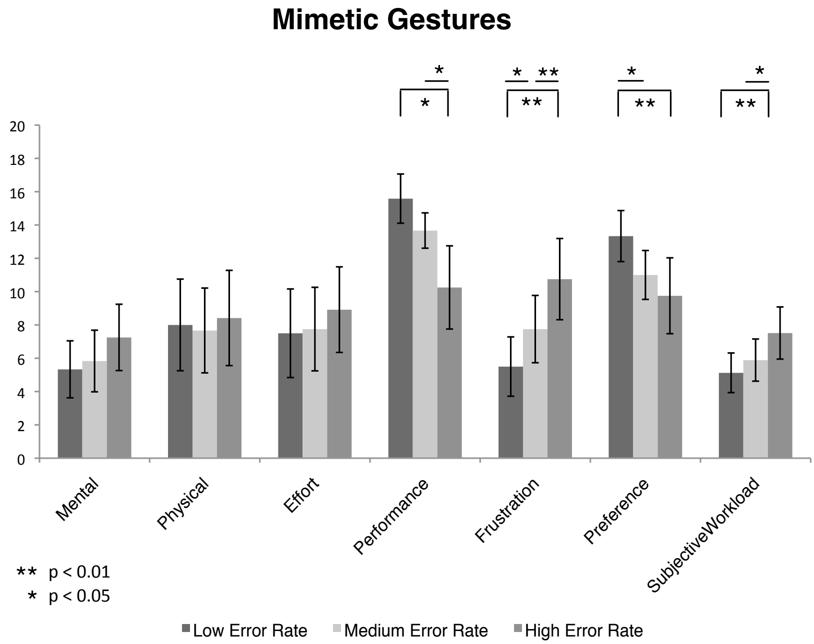

Figure 5.3: Modified NASA-TLX workload measurements for mimetic gestures. Capped error bars represent a 95% confidence interval.

= 5.69, $p < 0.05$, O gesture $F(2, 22) = 4.33$, $p < 0.05$, and V gesture $F(2, 22) = 4.88$, $p < 0.05$. For the L gesture, there were significant differences ($p < 0.05$) only between low (1877, 95% CI [1666, 2088]) and high error rates (2298, 95% CI [1960, 2636]). Also for the O gesture, there were significant differences ($p < 0.05$) only between low (2078, 95% CI [1783, 2372]) and high error rates (2522, 95% CI [2158, 2886]). Likewise for the V gesture, significant differences ($p < 0.05$) were found only between low (1739, 95% CI [1477, 2002]) and high error rates (2517, 95% CI [1841, 2741]).

The results indicate that gesture duration is less affected in the mimetic gesture condition (except for the Fishing gesture), even though duration of performed mimetic gestures is overall higher. Whereas in the alphabet gesture condition, three of the gestures (L, O, V) significantly increased in duration between the low and high error rate conditions (again supporting the 40% error tolerance threshold found by Karam and Schraefel (2006)). However, for both mimetic and alphabet gestures, there appears to be a trend whereby high error rates cause an increase in gesture duration. The increased durations for alphabet gestures suggests that participants performed the gesture more slowly to ensure that they can better control the shape and size of the movement, likely in order to achieve the canonical letter variant. This is in line with work on speech recognition systems, where a common error handling strategy is to hyperarticulate their speech (which results in higher speech duration) (Oviatt et al., 1998). Likewise for handwriting recognition, where users typically rewrite a letter more carefully (Webster and Nakagawa, 1998).

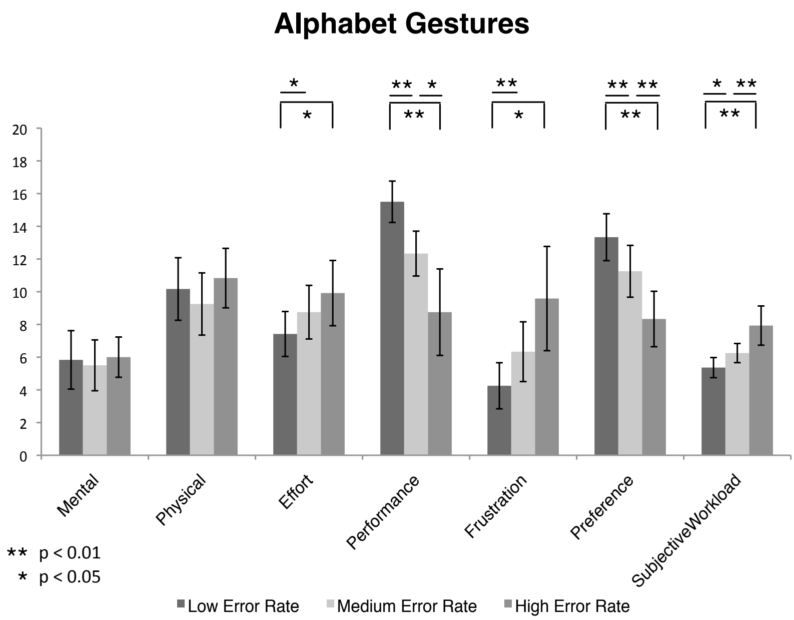

Figure 5.4: Modified NASA-TLX workload measurements for alphabet gestures. Capped error bars represent a 95% confidence interval.

5.5.3 Video Analysis

Observation of participants' behavior from the videos provided early findings on how mimetic and alphabet gestures evolve under varying error conditions. These observations revealed an initial parameter set (shown in Table 5.3) to describe how a gesture evolves under high failed recognition conditions. In our observation analysis below, we used coarse measures to describe how a gesture changed, relative to the subject's style of gesturing (e.g., Fishing gesture in this trial was performed faster than previous trials).

Gesturing Styles. Participants exhibited some variation in gesturing style, which was observed during low error rate trials. In the mimetic condition, some participants were observed to have performed the Fishing gesture as if it was real-world fishing (where participants would experiment with different ways to throw a fishing rod's hook into space), while others a more mechanical, repetitive fishing gesture (varied mainly by speed). The Shaking gesture varied according to the number of swings that a subject performed, and this was done either hurriedly, slow and structured, in a swinging manner, or whether the device was swung vertically or horizontally. The Handshake gesture differed across participants in the frequency of handshakes (which ranged from 2-4), and in terms of how natural the handshaking gesture was (e.g., shaking the hand of someone). In Trashing, participants exhibited both polite and slow trashing, and more angry and intense forms of trashing. For the Glass Filling, the gesture varied in terms of how much the Wii Remote was tilted downward (which maps to how one can vary the rate by which liquid is poured out of a

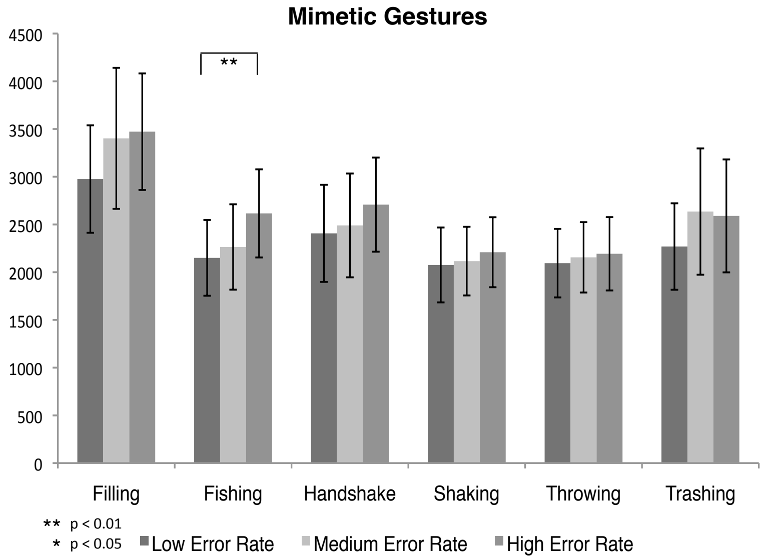

Figure 5.5: Mean durations (in milliseconds) for the mimetic gestures (based on 12 trials per error range). Capped error bars represent a 95% confidence interval.

container). Finally, different throwing gesture styles varied in terms of the direction of the Throwing gesture, and the starting position of the throw. For the alphabet gesture condition, differences in styles of gesturing were due to the speed, breadth, position or completeness of the gesture. Only the O and V gestures varied with respect to direction.

Mimetic vs. Alphabet Gesture Evolution. While evolution in gesture performance for both sets was observed in as few as 2 successive error trials (i.e., spiral depth of 2), less variation in gesture performance was observed in the alphabet gesture condition. In both conditions however, we observed that the continued successful recognition of a gesture served as a continuos reinforcer for repeating a gesture in the same manner, irrespective of gesturing style. If however a participant chose to experiment with a variation during those success trials, then the variation was repeated on grounds that it will work on a subsequent trial. We call this repeated gesture variation the ‘canonical variation’. As shown below, these observations are corroborated by participants’ feedback during the interviews.

We observed the following canonical variations of mimetic gestures under high error: S4’s Handshake canonical variant was extending his arm straight and swinging the Wii Remote up and down as if it is a real hand he is reaching out to. S8 exhibited a variation in the Trashing gesture, where the speed of the gesture increased drastically and the end position was raised higher than shown in the video. S10 arrived at the canonical variation of the Glass Filling gesture, which required a slow, calculated twisting of the wrist while the Wii Remote was in the pouring position. S12 exhibited variations on both Fishing and Trashing gestures: Fishing gesture resembled real-world fishing, which involved lightly pulling the Wii Remote diagonally upwards and then with slight force throwing the Wii Remote back

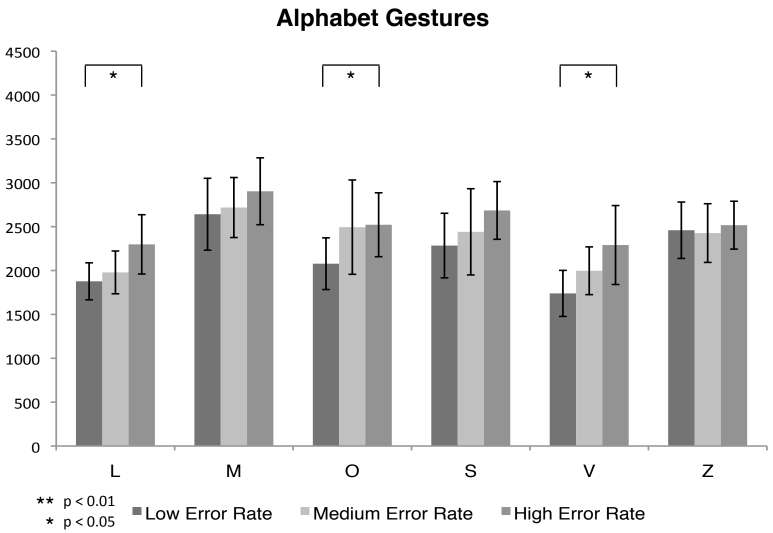

Figure 5.6: Mean duration (in milliseconds) for the alphabet gestures (based on 12 trials per error range). Capped error bars represent a 95% confidence interval.

as if into an imaginary water pool. The Trashing variation was a trashing gesture that was both small in breadth and with almost no movement of the arm and shoulders.

For alphabet gestures, apart from brief intermittent experimentation with breadth, speed, position, completeness and direction of gestures, the main systematic variations observed were attempts to draw the letter symbol more precisely under high error conditions (cf., Webster and Nakagawa (1998)). However, there was one participant who appeared to have arrived at a canonical variation: the letter Z was drawn very quickly with the last stroke (bottom line of a 'Z') more slanted.

Persistence vs. Evolution. When do recognition failures cause participants to persist in repeating the same gesture, and when to push them to explore new gestures? From our observations, certain patterns emerged: first, it seems that if a participant performs a gesture quickly, and it fails, he will experiment with a slower version. If he first performs it slowly, then he will experiment with a faster version. Second, after repeated success trials, the speed of the gesture is the first parameter to vary while the other parameters mostly remain constant, irrespective of gesture class. Usually, though not always, only after failure does exploration take place, where variations come into play. Third, if in a block participants experience a series of successes (4-5), they will be more likely to repeat the same gesture even in the face of repeated errors later in the block, irrespective of gesture class. This successive positive reinforcement suggests that participants have figured out what the recognizer wants in that block. Additionally, if a gesture succeeds too many times in succession ($\geq$4), people seem to apply the principle of least effort and perform incomplete gestures (as witnessed by a downsized version of the Throwing gesture by S6 in the low error condition). This was observed mainly in the alphabet gesture condition, where two

Parameter	Description
Breadth	How big or small (in m^3) was the performed gesture?
Speed	How fast or slow (in m/s) was the gesture performed?
Duration	How long did the gesture take (in milliseconds) to be performed?
Distance	How far or near (in cm) from the subject's body was the gesture performed?
Position	What was the starting position (in x,y,z coordinates) of the performed gesture? And what was the end position?
Power	How much power (in watt) was applied during the performed gesture?
Orientation	What was the initial orientation (heading) of the held device?
Direction	Given fixed orientation, in which direction (x, y, z vector) was the gesture moved?
Rotation	Given changing orientation, in which direction was the gesture rotated to?
Completeness	Was the performed gesture complete or was it incomplete?
Evolution	With respect to the above parameters, how much did a gesture transform across repeated error conditions?

Table 5.3: Gesture parameter descriptions.

participants later mentioned in the interviews that they did not need to complete the gesture for recognition to take place.

5.5.4 Users' Subjective Feedback

Perceived Canonical Variations. Many participants (18/24) across both gesture sets reported that in the face of repeated errors, they would start experimenting with different variations of the gesture, and when feedback was positive, they replicated that variation. This suggests that positive reinforcement after repeated error trials was the driving force behind the step-wise evolution of a given gesture, or put differently, survival of the fittest gesture variation. Supporting our observations, participants reported much less variation in how they performed gestures under high error rate conditions in the alphabet gesture condition than in the mimetic condition. For mimetic gestures, participants reported many novel strategies for how and when a given gesture was recognized (*P9: "The shaking, that was the hardest one because you couldn't just shake freely [gestures in hand], it had to be more precise shaking [swing to the left, swing to the right]... so not just any sort of shaking [shakes hand in many dimensions]"; P11: "What I noticed was that the fishing for example, it was just a rotation [small upward wrist rotation], and when you did it as if you were really fishing, then it didn't work."*).

In line with our hypothesis, the variations under the alphabet gesture condition were perceived to involve a more precise and well structured gesture for recognition to be successful (*P17:"I think within certain blocks I got the pattern of what was working. It was more apparent in the third one [high error rate condition], you have to do it better."*). There were exceptions to this, mainly the drawing of the letter O, which can easily vary (in addition to breadth and speed) along the direction and the start/end position parameters (*P16: "I didn't really change the way I did them [the gestures], except for the O, to see different ways it can be written."*). However, while variations that involved more rigid gesturing of letters was both observed and reported by participants, some participants explicitly expressed arriving at the canonical variations of some letters (*P22: "In Z or S, I made them more rounded, and then it worked better this way."*).

Individual and Cultural Differences. For the mimetic gestures, participants found the Throwing and Glass Filling gestures the most problematic (4 reports each). Throwing in particular appeared not to be as intuitive as the other gestures, possibly because of the many ways people can throw (*P7: "I would have done the throwing gesture differently [shows experimenter how different people throw]"*) The mapping to real-world behaviors was evident when explaining why the Glass Filling gesture is difficult (*P10: "For the glass filling, there are many ways to do it. Sometimes very fast, sometimes slow like beer."*). Interestingly, the Shaking gesture, which is already an input technique in some smart phones (e.g., Apple's iPhone 4®[7]), was not taken favorably by at least one participant (*P4: "Shaking was hard... how long should you shake? I tried here, and it was enough to do two movements... and you have to use your wrist, which can be a problem."*).

Two participants noted that the Fishing gesture was quite easy to perform, however the naturalness of the gesture backfired (*P10: "In the fishing gesture, I was just guessing how to do it. Because I have never done it practically. I cannot really see myself performing well, just simulating it."*). Likewise for the Trashing gesture: (*P10: "I was trying to emulate how I would normally do it. For example trashing, sometimes you're not in the mood, and you do it like this [trashes imaginary object downward softly], very quietly."*) Together, these reports support our hypothesis that mimetic gesture evolve into their real-world counterparts, especially when under high error conditions. Due to the importance of cultural differences in performing certain mimetic gestures, it would be interesting to see whether these cultural forms are the first gestures participants recourse back to under error conditions.

For the alphabet gestures, the O and M alphabet gestures were perceived to be the most difficult (6 and 5 votes, respectively). The O gesture was perceived to be difficult on grounds that there are many ways to write/draw an O (*P17: "The O was a bit funny, because naturally I start from the top, not the bottom."*). This was likely a fault of how the videos in the experiment showed the gesturing of the O, which began from bottom to top. However, the other reason for finding the O difficult was due to the completeness and position parameters (*P13: "For the O, I noticed that if I start here [points in space] and I end here, and there is a gap, then it wouldn't be recognized."*). There were no clear reasons given for why the M gesture was difficult, other than hinting at the number of strokes required (*P14: "The M is more articulated, so you have to spend more time on it."*). V was perceived to be the easiest letter to gesture, due to the ease by which it can be drawn. This

[7] http://www.apple.com/iphone/; last retrieved: 01-08-2013

was so despite that V, like the O, varied along the direction parameter.

Perceived Performance. In general, participants reported they were pleased with their overall performance in both mimetic (7/12) and alphabet (7/12) gesture conditions. Surprisingly, while all participants noticed the difference in difficulty across blocks, some participants in the mimetic gesture condition (6/12) and some in the alphabet gesture condition (7/12) seemed to treat their performance on the low error rate and the medium error rate conditions as more or less the same (*S18: "Between the first and second blocks [low and medium error rate conditions], it was the same...10-15%"*), while attributing very poor performance to the high error rate condition. The reason for this was likely due to the extremely high error rate range (40-60%). This relates to the question of how poor can performance of gesture recognition technology get before the technology is abandoned in favor of well-established methods such as keyboard-based input (cf., the 40% error threshold set by Karam and Schraefel (2006)).

Additionally, participants showed an incredible ability to justify their performance. If participants fell into the high error rate condition first, they attributed their poor performance to the fact that they are still learning how to perform the gestures (*P22: "In the beginning, it was less because I was still learning."*). By contrast, if the high error condition comes later in the block, they attribute their poor performance later due to not putting the kind of effort and attention they did on the first block, which supposedly explains their better performance earlier (*P7: "For the third block [high error rate condition]...I had done many already, I was like I'll just do it and see what happens."*). This is in line with our video observations, where gesture performance was different in high error conditions where gestures tended to evolve more for mimetic gestures (poor observed performance) and get more rigid for alphabet gestures (acceptable observed performance if canonical variation is right).

Use of Gestures on Mobile Phones. For mimetic gestures, when asked whether they saw any real-world use of the tested gestures on mobile phones, most participants (10/12) reported at least one use-case. For the Throwing gesture, a mapping to sending a message was identified as a reasonable interaction method. Another example of Throwing was mobile payment, where one could throw money to a cashier. Similarly for the Handshake gesture, where the handshake could be a form of digital content transaction or a form of business e-card exchange. Trashing was implicated in hanging up a call, deleting content, or what participants did not favor, to turn off an alarm clock by turning over the device. Participants reported that shaking was already available in some mobile phones used for switching to the next song. However, some participants expressed that some of the gestures simply had no use (*P4: "Trashing at a conference is quite natural to turn off a call, but handshake, I can't think of a use."*).

When asked about using alphabet gestures on mobile phones (e.g., gesturing 'S' for sending a message), only half of participants said they would use such gestures (6/12) (*P24: "If there was a phone with this, I would really like that!"*), and some thought it depended on the situation (3/12) (*P19: "If you're on your bike, you could do that."*). From those who would use such gestures, they explicitly mentioned that they have to necessarily be free of error (*P20: "Yeah [I would use such a gesture], if it's really foolproof."*). Reported use cases included gesturing P to pause a game, C for calling or checking a calendar, or most reported, for traversing an interface's menu structure (*P18: "If you need to access a feature that is hidden away in the menu, like starting to write an SMS, then gesture S"*). Some participants mentioned that the gestures need to be explicitly different (*S22: "Those gestures need to be so different so you cannot do wrong gestures...V is part of M."*) and others required a distinct gesture-to-task mapping for each application on your device (*"S23: "The applications should not clash for the same gesture."*). Finally, one participant remarked that while these alphabet gestures might work for the roman alphabet, it would be radically different for chinese characters.

Social Acceptability. For both the mimetic and alphabet gestures, some participants (8/24) expressed concern over performing these gestures in public. This is in line with previous work, which explicitly addressed the question of what kinds of gestures people would perform while in public (Rico and Brewster, 2010). It was surprising to find that there were very few concerns about performing mimetic gestures in public (2/12) (*P9: "I'm not sure [about Fishing], because I'm not fishing myself, why am I doing this?"*), as opposed to performing alphabet gestures (6/12) (*P13: "If I am not in a public place, then yes. Otherwise, you would think I'm a Harry Potter in disguise!"*).

While recognition errors play a clear role in preventing gesturing in public settings (*P16: "When it doesn't take your C, you keep doing it, and it looks ridiculous."*), the breadth of a gesture was also perceived as an important factor (*P18: "Maybe if I do it small, if I don't look very freaky, then it's okay."*). Still, others thought it fun to try novel things in public (*P14: "I never thought about that [alphabet gestures], but why not? I would not find it embarrassing."*).

5.6 Discussion

There are three potential limitations to the present study: first, since our study was conducted in a laboratory, it had less ecological validity. While participants explicitly stated that they would not be comfortable performing some gestures (especially alphabet gestures) in public, from the videos it was evident that all participants performed all the instructed gestures freely and without hesitation. It is interesting to consider here whether the desire to perform all gestures successfully in each block could have overruled whatever embarrassment might come about from merely performing the gesture to execute some device function. At least one participant explicitly mentioned the importance of performance (*P9: "I eventually got the hang of it, and yeah, I really, really wanted to improve my performance!"*). Additionally, gesture performance under failed recognition rates may not reflect performance when a user is mobile (e.g., walking or in public transport).

Second, while testing how task-independent gestures are affected under varying error rates was an explicit design choice, it could be that a gesture to task mapping is necessary for unraveling the usability of a given gesture. While we agree that the task being performed is an important variable in gesture-based interaction, we nevertheless argue that there are differences between individual gestures and importantly between gesture sets that can be unraveled only by leaving out the task. This is to eliminate any potential bias that the task might evoke. For example, calling someone might be considered more urgent in some situations than skipping to the next song, and that might influence the performance and workload required from a gesture. Finally, our design choice in simulating a gesture recognition engine as realistically as possible meant that the errors had to be randomly distributed in each block. In future work, experimentation with different error distributions would better help understand different types of gesture evolution. Additionally, with a real recognition engine, the precise evolution of the gesturing behavior may differ than what was observed in this study.

5.7 Design Recommendations

Implications for Gesture Recognition

Our observations, duration analysis and participant reports showed that mimetic gestures and alphabet gestures do indeed differ under increasing recognition error conditions. While our observations and participant reports showed mimetic gestures tend to vary more into their real-world counterparts when they are repeatedly not recognized, alphabet gestures tend to both increase in duration become more rigid and well structured. This is in line with work on other modalities like speech and handwriting recognition (Read et al., 2002; Oviatt et al., 1998). This suggests that for gesture-based interaction to be accepted in the consumer market, accurate recognition from the first attempt appears to be quite important for mimetic gestures. If a gesture is not recognized from the first instance, there is a risk that the subsequent gesture differs radically from the first, which would be beyond the scope of the recognition algorithm. This is in contrast to alphabet gestures, which in having lower degrees of freedom vary in fewer parameters (mainly speed, breadth, and start/end position) under error conditions. This suggests that recognition engines (in uni- or multimodal systems) can more easily deal with post-failure recognition when this set of gestures is used.

Additionally, participants had no real means to understand the cause of the errors, to avoid errors, or to improve recognition rates. However, they came up with interesting explanations (e.g., canonical variations) why there were more errors in different blocks and what might have caused them (e.g., fatigue). Nevertheless, we observed that they were also active in adapting their gesturing behavior in order to improve recognition errors and to understand the workings of the recognition engine. This seems to suggest that transparency in the gesture recognizer may better support users in their error handling strategies during situations of failed recognition.

Implications for Gesture-based Interaction

It was evident from our results (modified NASA-TLX workload data, mean durations, video observations as well as participants' feedback) that not only do mimetic gestures vary differently than alphabet gestures under error conditions, but also there were differences between individual gestures under each class. We found that while mimetic gestures yield significant increases in overall subjective workload between low and high error rates, and between medium and high error rates, overall workload for alphabet gestures significantly increased across all error rate conditions. This indicates that mimetic gestures are better tolerated under error rates of up to 40%, while alphabet gestures incur significant overall workload with up to only 20% error rates. This is in line with previous work on computer vision-based gesture interaction (Karam and Schraefel, 2006), where our workload results suggested user error tolerance of up to 40% for the mimetic gestures only. Further support that indicates mimetic gestures may be better tolerated can be drawn from the significant increase in mean gesture duration under high error rates for three of the alphabet gestures (L, O, V), but only one of the mimetic gestures (Fishing). Believing that one has to gesture slowly for successful recognition can negatively affect the usability and user experience of a performed gesture, despite that alphabet gestures were performed slightly faster than mimetic gestures.

The two gesture sets also differed in potential use and social acceptability. For mimetic gestures, interesting use cases (e.g., handshake for digital content exchange, throwing for mobile payment) were given, while limited use cases (e.g., interface menu structure traversal) were offered for alphabet gestures. Moreover, alphabet gestures were seen as more embarrassing to perform in public, especially if they are not recognized. From a usability perspective, these findings suggest that mimetic gestures are more promising than alphabet gestures for use during device-based gesture interaction, even under conditions of medium recognition error.

5.8 Conclusions

In this chapter, we described the results of an automated Wizard-of-Oz study to qualitatively investigate how mimetic and alphabet gestures are affected under varying recognition error rates. In line with our hypothesis, it was shown that mimetic gestures, which have a less familiar ideal shape, tend to evolve into diverse real-world variations under high error conditions, while alphabet gestures tend to become more rigid and structured. Moreover, while gesture duration is only one parameter in gesture evolution, we saw that for three of the alphabet gestures, duration increased under high error conditions. In contrast, only one mimetic gesture exhibited such significant increase. Additionally, the interaction videos provided an initial parameter set that can be useful for describing how gestures change under error conditions.

Furthermore, we showed that mimetic gestures seem to be tolerated under error rates of up to 40% (cf., Karam and Schraefel (2006)), while alphabet gestures incur significant overall workload with up to only 20% error rates. From this, we drew usability implications showing the importance of immediate accurate recognition of mimetic gestures (as a way of taming the tendency of these gestures to evolve) and suggested they are better suited than alphabet gestures for inclusion into mainstream device-based gesture interaction with mobile phones.

While this chapter focused on the usability of 3D gesture-based interaction, as mentioned in our study limitations, we mainly looked at task-independent interaction. Looking at task-independent interaction was necessary to investigate the usability issues associated with performing 3D gestures, as pairing with tasks may have influenced gesture performance and preference. Given the promise of gesture-based interaction, we revisited how this form of interaction can be applied in an actual domain. Additionally, given the problematic nature of supporting playful experiences uncovered in Chapter 3, in the next chapter (Chapter 6) we revisit the domain of playful mobile interactions.

6 Playful 3D Gestural Interaction

We revisit the issue of supporting playful urban interactions presented in Chapter 3. Here, we make use of a novel gesture-based paradigm in Mobile HCI to provide such support (*Playful Gestural Interaction Study*). We introduce magnet-based Around Device Interaction (ADI), and its applied use in a playful, music-related context. Using three musical applications developed under the magnet-based ADI paradigm (Air Disc-Jockey, Air Guitar, Air GuitaRhythm), this chapter investigates whether the magnet-based ADI paradigm can be effectively used to support playful music composition and gaming on mobile devices. Based on results from a controlled user study (usability and user experience questionnaire responses, users' direct feedback, and video observations), we show how 3D gestural input can be effectively used to create natural, playful and creative mobile music interactions amongst both musically-trained and non-musically trained users. Additionally, we distill design considerations to optimize playful and creative music interactions using 3D gestural input on today's smartphones. The work presented in this chapter will be published as "Magnet-based Around Device Interaction for Playful Music Composition and Gaming" in the International Journal of Mobile Human Computer Interaction (El Ali and Ketabdar, 2013).

6.1 Introduction

The recent advent of Around Device Interaction (ADI) (Butler et al., 2008) has expanded the interaction space on mobile devices to allow 3D motion gesture interaction around the device, with opportunities for playful music composition and gaming only now taking shape. Using sensors embedded in mobile devices (e.g., (magnetic) compass (Ketabdar et al., 2010a), IR distance sensors (Kratz and Rohs, 2009)), users can now take advantage of the extra interaction space that their mobile device affords, for leisure and entertainment (Davenport et al., 1998).

ADI can be useful for small tangible/wearable mobile or controller devices (e.g., mobile phones or wrist watches) (Ketabdar et al., 2010a). In such devices, it is extremely difficult to operate small buttons and touch screens. By expanding the interaction space around the device, ADI can aid the user in such cases, alongside situations when the device screen is not in line of the user's sight. The ADI paradigm can allow coarse movement-based gestures made in the 3D space around the device to be used for sending different interaction commands such as controlling a portable music player (changing sound volume or music track), zooming, rotation, etc. For mobile phones, it can be also used for dealing with incoming calls (e.g., accepting or rejecting a call). However, ADI need not be limited to use-cases comprising user situational impairments (Ashbrook et al., 2011) or substituting for basic touchscreen tasks (Baudisch and Chu, 2009), but can complement touchscreen interactions with 3D gestures to allow natural, playful interactions in music composition (Ketabdar et al., 2012, 2011) and gaming.

Magnet-based ADI is a novel interaction technique for mobile devices allowing gestural interaction in the whole 3D space around the device.[1] Here, moving a properly shaped magnetic material in hand (e.g. bar shaped, pen, ring) is used to influence the internally embedded compass (magnetometer) sensor in mobile devices by different 3D gestures, hence allowing for touchless interaction around the device. Since the interaction here is based on magnetic fields (which can pass through the hand or clothes, and not depending on users' line of sight), the space at the back and side of device can also be efficiently used for interaction. This technique does not require extra sensors on current smartphones. For these smartphones, it is only necessary to have a properly shaped magnet as an extra accessory. While this can be seen as a limitation of such systems, as will be shown later the use of a magnet allows for a more natural interaction with music related apps.

In this chapter, we look closely at how magnet-based ADI can be used in a playful context, to facilitate natural interaction for music composition and gaming amongst both musically-trained and non-musically trained users. Using three musical applications developed under the ADI paradigm (Air Disc-Jockey, Air Guitar, Air GuitaRhythm), we investigate the potential of ADI for playful interaction, in order to gain insight into the acceptability and naturalness of ADI by users who wish to casually engage in playful mobile music composition. Under this investigation, our primary goal is to explore novel methods using mobile technology to entertain users. The rest of the chapter is structured as follows: first we provide our research questions, followed by a review of related work and our magnet-based ADI framework. We then present our study design and methods, give our results and discuss them, provide design recommendations and finally conclude.

6.2 Research Questions

In this chapter, our main research question is:

> **RQ 5:** How can 3D gestural interaction support playful music composition and gaming on smartphones?

Specifically, we investigate usability and user experience[2] issues, and user acceptance of the ADI paradigm to support playful and creative interactions (Davenport et al., 1998), using three mobile music-related high-fidelity prototype apps (Air Disc-Jockey, Air Guitar, Air GuitaRhythm) that allow natural interaction (Grandhi et al., 2011) using a magnet.

[1] http://www.youtube.com/watch?v=WrVIO-0ak44; last retrieved: 01-08-2013. Anonymized promotional video that illustrates the magnet-based ADI concept in music-related applications.

[2] UX here is based on ISO 9241-210 (ISO DIS 9241-210:1994, 1994) definition: A person's perceptions and responses that result from the use or anticipated use of a product, system or service."

By allowing natural (cf., Grandhi et al. (2011); Khnel et al. (2011)) and in some cases enjoyable (cf., Rico and Brewster (2010)) gesture-based interaction with mobile devices, our hypothesis is that the ADI paradigm would be perceived as a fun and natural means of mobile interaction, in the context of playful gaming and music composition.

Two of the developed prototype apps (Air DJ, Air Guitar) allow free creative expression in composing music, and the last designed as a music game (Air GuitaRhythm). Our target user group are casual gamers and users who wish to compose music non-professionally, as part of everyday playful interactions with mobile devices. However, since two of the prototype apps developed lend themselves to creative music composition, testing users with previous musical training was required to provide insights on whether the ADI concept is perceived differently by those who can and those who cannot compose music. Given this distinction, we expected that users who were musically trained would perceive the creative music apps (Air DJ, Air Guitar) more favorably than those who did not have such training, whereas the gaming app should be perceived similarly by all user groups. This is because the musical game, which is easy to learn and play with a challenging score-based system, is easily accessible by all user groups. By contrast, the musical applications could perhaps be seen as having a higher barrier of accessibility if a user lacks the necessary music skills to compose music.

Investigating usability, user experience and acceptance afforded by the three magnet-based ADI-based musical prototype apps here yields two main research contributions: first, it provides a user-driven concept validation of whether the ADI paradigm, in allowing natural gestural interaction around mobile devices, can support use-cases for playful and creative music interaction (Section *Supporting Playful Music Composition and Gaming*. Second, it equips future interaction designers wishing to make use of magnet-based ADI with design considerations when designing playful and creative mobile ADI interactions (Section *Design Considerations for Applied Magnet-based ADI*). Additionally, we provide initial results on the social acceptability of ADI when interactions take place in public settings, as well as additional use-cases participants reported on.

6.3 Related Work

In addition to the related work listed below, this chapter builds on the related work on using 3D gestures in HCI discussed in Chapter 5 in Section 5.3.1 and partly on the related work on measuring playful experiences discussed in Chapter 3 in Section 3.6.1 .

6.3.1 Around Device Interaction

Several approaches to ADI have been proposed, which focus mainly on solving the occlusion problem (where the user's fingers cover the touch display during interaction). Baudisch and Chu (2009) show the effectiveness of a pointing input method for very small devices (down to 1" display size), where they use a touchscreen on the back of a device to handle occlusion. Butler et al. (2008) use infrared (IR) sensors on the edges of small mobile devices which allow capturing multitouch gestures around the device, specifically on either side of the long-edge of the device. Relatedly, Kratz and Rohs (2009) use six IR distance sensors to allow coarse movement-based hand gestures, in addition to static position-based gestures, which were also shown to be effective to solving the occlusion problem on small mobile devices. Recently, Kratz et al. (2012) demonstrated using depth imaging cameras to allow for back-of and side-of mobile device interaction, specifically focusing on tracking and recognizing gestures in a virtual object rotation task.

Closely related to the present work, Han et al. (2009) tracked a finger-mounted magnet for handwriting input. Relatedly, Harrison and Hudson (2009) used a magnet attached to the user's finger to allow radial and 2D input for a watch device. Ashbrook et al. (2011) also used a magnetically-tracked finger ring, and showed its effectiveness for sighted and eyes-free use for a pointing target-selection task where users could select from up to 8 targets on a menu display. Finally, Ketabdar et al. (2010a) also demonstrated that tracking a magnet in the space around the device can allow high gesture classification accuracy for coarse gestures performed by users around the device.

6.3.2 Gestural Interfaces for Music Interaction

The earliest example of a gesture-based music instrument is the Theremin (Theremin, 1996), which used electronic field sensing of hand positions in space to allow music composition. Since then, there has been a range of digital music instruments that use IR distance (e.g., Airstick free-gesture music controller (Franco, 2005) or vision sensing that go beyond keyboard interactions and allow bodily gesture-based interactions). Related to the present work, Gillian et al. (2009) proposed a gesture-based musical DJ game that uses the mobile device's 3-axis accelerometer and touchscreen interactions where users have to scratch at a specified musical beat (when cued by different multimodal feedback).

Kayali et al. (2008) developed three tangible mobile interfaces as gestural instruments, one of which is relevant here is a simplified guitar prototype that allows strumming frets using the Nintendo DS stylus. Gillian and Paradiso (2012) demonstrated how 3D depth sensors can be used for discrete and continuous control of a gesturally-controlled music instrument. While these presented examples exemplify the use of gestural interaction for controlling virtual music instruments, the interaction techniques offered use motion sensors that require moving the device itself, which risks users losing concentration on the task at hand. In our case, we use the motion of the hand, rather than the device, which affords more natural interaction and allows the phone display to be used more efficiently when composing music. Recently, Ketabdar et al. (2012, 2011) presented a demonstration of a guitar application that also uses magnet-based ADI to allow in air strumming, however their work was more focused on the technical infrastructure behind the guitar app, and therefore lacked any user evaluation. We add to the body of work presented there, by presenting two additional apps (Air DJ and Air GuitaRythm), with a focus on evaluating the apps from a user-centered standpoint.

6.3.3 Evaluating Playful Interactions

In this chapter, we do not deal with console or desktop games, and so the metrics proposed in Section 3.6.1 such as the GEQ or in-game metrics are not very useful to characterize user entertainment with our mobile music-related apps. Instead, we make use of questionnaires to measure users' experience with interactive systems and the perceived usability of the apps. To do this, we follow the approach by Lucero et al. (2011) by using the AttrakDiff2 ™ (Hassenzahl et al., 2003) questionnaire to evaluate playful aspects of interactive systems.

Additionally, we use the System Usability Scale (Brooke, 1996) questionnaire as it is a robust, industry standard in quickly evaluating a system's usability. Both of these scales are discussed under Section *Study Design*. Importantly, for the early stage of our prototype apps, we make use of qualitative observations and interviews, in order to gain insight from users directly on improving our apps and the interaction methods they support. In this case, looking at usability from a performance standpoint (cf., Gajadhar et al. (2010)) (e.g., high scores, touchscreen presses, or magnetic signal deformations) would not be useful for understanding how users deal with this new interaction method for music composition and gaming.

6.4 Magnet-based ADI

6.4.1 Framework

A piece of magnet when moved close enough to a smartphone can influence the compass sensor. The temporal pattern of such an influence is registered by the compass sensor, and can be interpreted as a gestural command using appropriate machine learning algorithms. Getting useful information from the magnetic sensor is not only algorithmically simpler than implementing computer vision techniques, but this approach also does not suffer from illumination variation and occlusion problems. In other words, it does not require direct line of sight into the camera, which enables covering the whole 360° space around the device for interaction.

The output of the compass sensor consists of 3 signals showing the strength and direction of the magnetic field along x, y and z directions. Each sensor reading composes a vector of 3 elements, where a gesture is presented by a certain pattern in a sequence of these vectors. A time derivative function is applied to sensor readings in order to highlight changes in the pattern of magnetic field, and remove effects of earth's magnetic field (which is almost constant). The sequence of vectors is divided into overlapping windows for gesture recognition. Depending on the type of gesture, different techniques can be applied for interpreting sensor readings as a gesture class. In our case, we used heuristic decision rules and Multi-Layer Perceptrons (MLPs) (Minsky and Papert, 1969), which for our framework have been shown to achieve gesture recognition accuracy of 83.7% and 91.4%, respectively (Ketabdar et al., 2010c).

For simple gestures, such as detecting only a triggering action as in the case of our apps below, average norm of vectors inside a window is compared with a predefined or adjustable threshold. Adjusting the sensor can affect the sensitivity for the triggering action. A triggering action involves a rapid motion of hand (with magnet) which causes rapid changes in the pattern of magnetic field around the device, resulting in a significant change in magnetic signal norm for a limited period of time. Motion of the device itself can also cause changes in the compass sensor output, due to displacement of the sensor with respect to earth's magnetic field. Rapid motion of the device is detected based on embedded accelerometer readings, which allows stopping gesture execution. For recognizing more complicated gestures, the sequence of sensor readings is compared with a pre-recorded sequences using template matching techniques such as Dynamic Time Warping (Sakoe and Chiba, 1978). However since here we deal with simple gestures, heuristic decision rules suffice.

6.4.2 Application Design Process

All the applications were implemented on the Apple iPhone 4® as functional interactive prototypes. Two of the applications described below (Air DJ and Air GuitaRhythm) followed a user-centered design process, where interfaces underwent a round of design iterations thereafter. For these apps, qualitative focus group (user insight) sessions were conducted. There were 5 focus group sessions with 10 participants (5 male, 5 female) aged between 20-30, where half had a background in music, and the rest in design. Sessions were carried out in a collaborative setting (2 participants per session). The collaborative dual-testing of participants (who did not know each other beforehand) was conducted as such to ensure discussion in an interactive manner amongst participants on how to improve the apps. Measures included semistructured interviews, think aloud protocols, and observations as protocoled by two observers.

Participants found the concept of magnet-based interaction behind Air DJ and Air GuitaRhythm very appealing, where main feedback involved usability issues (especially the navigation model in GuitaRythm) model and interface redesign suggestions (icon redesign, GUI element reordering, more transparent labels). Participants were also quite positive about using magnets for playful ADI interaction, and were willing to pay for a good magnet to ensure smooth interaction with the apps. However they explicitly stated that they should be readily available in stores and come in different form factors. These earlier findings have been incorporated into the design of the apps described below.

6.4.3 Prototype Applications

Air Disc-Jockey

Air DJ combines standard functionalities usually found in different electronic music instruments. These comprise playback of a song from the users music library, real time control of a lowpass filter applied to the playback, triggering of drum and effect samples and real time synthesis of sounds based on the user's hand movements. The Air DJ interface is shown in Fig. 6.1. A tutorial including an audio/video demonstration of the app can be found by tapping the question mark (6). A song (from iTunes® library) is loaded into the application by pressing the plus icon (7) on the note symbol, and thereafter transcoded to a pcm file. After the transcoding process, playback is started by pressing the Play/Pause button (1). Song title (2) and play progress time (8) are shown inside the music player area. The progress bar (3) indicates the current playback position with respect to the total length of the song. Tapping the bar allows the user to jump to a certain song position according to the tap location. In Settings (11), the user can adjust the music volume and magnetic sensitivity. For activating drums and effects, real-time audio synthesis, and real-time control of a lowpass filter, the user has to press and hold (5) the corresponding touchscreen button(s).

Air DJ allows transforming the user's hand motion to sound using real time audio synthesis. Sound is only generated when the user is actually moving the magnet in the vicinity of the device. Movement is detected from the change of the absolute value of the total magnetic field strength. A total of 3 synthesizer units (4) are available, named Marktree (1, 2,

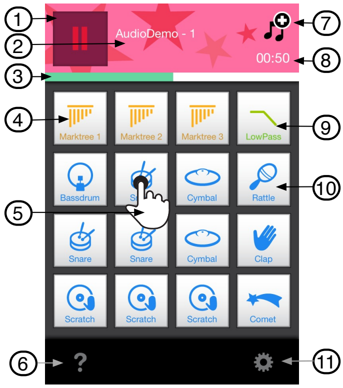

Figure 6.1: Air DJ interface. See text for explanation of labels.

3), which are activated by pressing the corresponding button. When the magnet is moving, Marktree 1 generates short random frequency sinusoids which sum up to a kind of a metallic sound. MarkTree 2 and 3 in contrast generate frequencies corresponding to the notes of two harmonically related minor seven chords (Dmin7, Amin7). During playback of a song, if the LowPass filter (9) is activated, it attenuates high frequencies according to the total strength of the magnetic field. The filter center frequency decreases with decreasing distance between the magnet and the mobile device, hence high frequencies are attenuated stronger when the magnet is close to the device. Air DJ also enables the user to play multiple high quality drum samples (e.g., Snare, Cymbal, Clap) along with the music using hand motion. The drum samples are triggered by pressing one of the blue sample buttons (10) and simultaneously moving the magnet near the mobile device. Main motivation for including these audio synths and effects was that they were determined to be suitable for the magnetic technology used. With respect to the choice of sounds (e.g., MarkTree), these are common sounds that were also deemed suitable as they go well with natural background noise. Furthermore, these audio synths and effects were included as participants from the earlier pilot studies enjoyed playing with them.

Air Guitar

Air Guitar allows playing guitar songs (which are simple lists of chords) by pressing and holding (at least) one fret on a virtual guitar neck while triggering with a magnet. Chords are built from individual guitar samples, where samples cover the visible note range (MIDI: 40...69). The Air Guitar interface is shown in Fig. 6.2. At the time of testing, Air Guitar was still in its early stages and hence the interface and features were still basic (e.g., played songs cannot be saved or shared). The application starts with a virtual guitar neck with six strings (1) and five (I...V) frets (3). The user can play a chord while placing his finger(s) on one or more notes (2) and moving the magnet with the other hand. Open strings (from left to right) correspond to MIDI notes: 40 (E-String), 45 (A-String), 50 (D), 55 (G), 59 (B), 64 (E). The corresponding sample of a selected note is determined using the fret number (I...V), where Sample Number = $\text{MIDI}_{OpenString}$ + Fret Number. At this stage of devel-

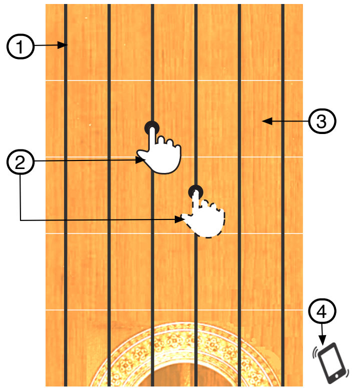

Figure 6.2: Air Guitar interface. See text for explanation of labels.

opment, the Air Guitar app does not support playing open strings. Currently, to get to the settings page where the volume and magnetic sensitivity can be adjusted, the user has to shake the device (4).

Air GuitaRhythm

Air GuitaRhythm introduces an innovative concept for playing music reaction games in the style of Guitar Hero® and Rock Band® on mobile devices using magnet-based touchless 3D gestures. Air GuitaRhythm allows the user to play the lead guitarist of a virtual rock band. In Air GuitaRhythm, songs are delivered with the app and consist of the mp3 file and text file containing information of the game melody (note event time stamps, MIDI note number). The user can choose a tutorial which includes an audio/video demonstration of the app, or select a song. After having selected a song, the play screen is displayed and a counter starts counting backwards from 3 to 1, where then the song playback starts. The Air GuitaRhythm interface is shown in Fig. 6.3. A magnet in hand allows the user to use natural hand gestures similar to real guitar playing to play the notes of the main guitar melody of a song. Air GuitaRhythm challenges the user to move the magnet rhythmically correct (on the dashed line (5)) according to a note pattern shown on the display. Song progress bar (1) and song title (4) are displayed above the note display area, and the three gray icons (3) at the bottom of the screen allow the user to Stop, Pause, or Restart the game. In Settings (10), the user can adjust the music volume and magnetic sensitivity.

After song playback has started, note symbols (6) start moving across the screen from left to right representing notes of the game melody. The user has to move the magnet in the vicinity of the phone to play a note of the melody. Notes can only be played as long as they are in the Play Zone (7). The perfect moment to play a note is when it is aligned with the dashed line (5) of play zone resulting in nice sounding melodies. The user is supposed to move the magnet only if there is a note in the zone. Otherwise an error sound chimes when the score (9) gets lowered. Notes that have been triggered correctly will be displayed in green, and those missed will be displayed in red (as shown). With every missed note, the life bar (2) decreases by one element. Game ends when there are 0 misses left. Triggering

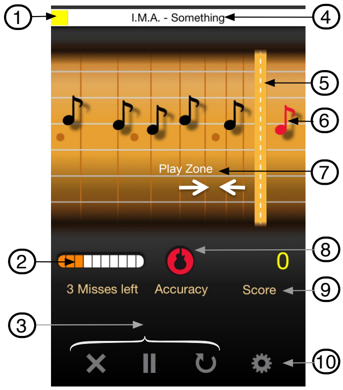

Figure 6.3: Air GuitaRhythm interface. See text for explanation of labels.

a note correctly refills the life bar by one element. Each note that is triggered correctly will increase the score Accuracy (8) (percentage of correctly triggered notes), which indicates how many notes were triggered correctly. The Accuracy score is green when most notes are triggered correctly, and red (as shown) when hardly any notes have been triggered correctly. At the end of a song, the performance of the user is evaluated and compared to the current high-scores of the corresponding song. If the user's score is higher than one of the scores found, the user can add his name to the high-score table.

6.5 Methods

6.5.1 Study Design

To investigate the potential of magnet-based ADI for music composition and gaming, we designed a controlled study to test both the usability and user experience of our mobile music apps. A controlled laboratory study was suitable in this case as it allows drawing rich user insights and concept validation without the unpredictability of in-the-wild testing. While Air DJ and Air GuitaRhythm were high-fidelity prototypes, Air Guitar was still in its early stages. It was nevertheless included in our study as it provided high potential for musical creativity. Together, the applications served as probes into getting users acquainted with the paradigm of magnet-based ADI in general, and for applications to music composition and gaming in particular.

Since only Air GuitaRhythm was a fully developed game (i.e., with a performance scoring system), we expected it to appeal more to the general population of users. The DJ application and the Air Guitar on the other hand, were expected to appeal more to users with at least some musical training. We defined a musically trained participant as a person who plays at least one musical instrument, and has at least 2 years of experience playing it. To deal with the difference in target groups, we tested users who do not have any musical training and users that do. This we hypothesized would provide greater insight into the use of magnet-based ADI for both creative composition (by musically trained users) and for general entertainment and enjoyability (by non-musically trained users).

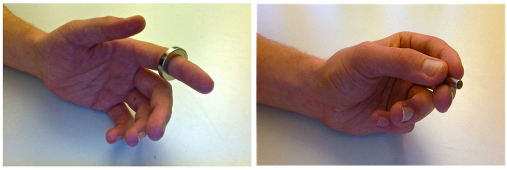

Figure 6.4: Ring magnet (left) and bar magnet (right).

The foregoing design decisions led to a controlled study with a mixed between- and within- subject factorial (2 x 3 x 2) design. There were three independent variables (IVs): music training (2 levels: music training vs. no music training), magnet-based ADI application (3 levels: Air DJ vs. Air Guitar vs. Air GuitaRhythm), and magnet (2 levels: bar-shaped vs. ring magnet). Music training was a between-subjects factor, and ADI application and magnet were within-subjects factors. Each between-subject condition tested all applications and both magnets, counterbalanced and randomized across participants. Participants were given a tutorial on how to use the magnets to interact with each application, and they were allowed to spend as much time as they wanted on each application. To avoid experimental artifacts associated with the form factor of the magnets, participants were asked to play with two different magnets: a bar-shaped magnet (~5cm length, ~0.5cm width) and a ring-shaped magnet (~3cm diameter, ~0.8 cm width), both shown in Fig. 6.4. The ring-shaped magnet had stronger magnetic force due to its thickness. These were calibrated accordingly for use with each app, where users could additionally calibrate the sensitivity if desired. Additionally, testing two different magnets also served as a probe to get participants to imagine later the form factor possibilities for magnet-based ADI. While all users were asked to use both magnets, the usage duration for each were not explicit conditions for this study, so as not to artificially constrain the study setup too much. As mentioned, the use of two magnets was provided primarily to allow users to reflect on the possible form factors that magnets come in, where finally we expect that many shapes and sizes of magnets would be available commercially or at home for users to use.

To measure the usability and user experience (our dependent variables) of each musical app, five data sources were collected: a) AttrakDiff2™ (Hassenzahl et al., 2003) questionnaire[3] responses b) System Usability Scale (SUS) (Brooke, 1996) responses c) Likert-scale questions about participants' attitudes toward magnet-based ADI and the given prototype apps tested d) video recordings of participants' gestures, and e) post-experiment interviews, to get direct user feedback on magnet-based ADI.

AttrakDiff2 measures pragmatic and hedonic qualities of interactive systems by allowing participants to provide ratings on a 7-point semantic differential scale (range [-3, 3]) for 28 attributes, resulting in 4 quality dimensions: 1) Pragmatic Quality (PQ), which measures usability of a product (or in our case each application). Here, PQ gives insight into how easy and straightforward it was to use each application 2) Hedonic Quality - Identification (HQ-I), which gives insight into the extent that users can identify with each application 3) Hedonic Quality - Stimulation (HQ-S), which gives insight into the extent

[3]AttrakDiff2™ is a questionnaire originally developed to measure the perceived attractiveness of interactive products based on hedonic and pragmatic qualities.

that each application stimulates users with novelty 4) Attractiveness (ATT), which provides a global appeal value and quality perception of each application.

Despite that the applications were all prototypes, we nevertheless decided to additionally administer the SUS (10-item questionnaire on a 5-point Likert scale) to gain additional insight (aside from the PQ category in AttrakDiff2) into the ease of use, efficiency, and satisfaction of each application. The SUS has been shown to be a robust and reliable standalone tool for measuring perceived usability of interactive systems, where a score of 70 and above indicates an acceptable score (Bangor et al., 2008). While there is some overlap between the SUS and AttrakDiff2, collecting multiple sources of data provides stronger evidence of findings. Additionally, usability is only one dimension of AttrakDiff2 (which is more focused on UX issues of enjoyment and novelty).

Likert-scale questions (4-item; $\alpha = .71$) we gave participants asked about their first impression of the apps, how comfortable it was to play with each app, how easy to learn using each app, and whether or not they enjoyed making music with each app. An additional item asked whether or not they would be willing to carry a magnet around. Additionally, we had a semi-structured interview at the end of each testing session, where users could give their feedback directly on what they thought about ADI using magnets, their expectations about availability of magnets when they download these apps, as well as their preferences for the magnet form factor (shape, size, color). Additionally, they were asked about other application use-cases that could potentially benefit from the magnet-based ADI paradigm. To gain insight into whether this mode of interaction is socially acceptable, they were asked whether or not they would interact with mobile phones using magnets in public places (e.g., in the metro, bus, or on a public street).

6.5.2 Participants

24 participants (15 male, 9 female) aged between 23-39 (M_{age}= 27.2; SD_{age}= 4.1) were recruited. Half had musical training, and the other half no musical training. This was identified through the recruiting process and later through the information forms participants had to fill in before each test session. Our participant sample spanned 13 different nationalities, where all were right-handed. No left-handed participants were tested as this was not an explicit aspect of our research questions. Half (12/24) had a technical background, and nearly half (11/24) were familiar with gaming consoles that use some form of gesture recognition technology (e.g., Nintendo Wii© or Microsoft Kinect©). Most of the musically trained participants played the guitar (9/12) among other instruments (piano or accordion), with the rest having been trained to play only the piano (3/12).

6.5.3 Setup & Procedure

The study was carried out at the usability lab at Telekom Innovation Laboratories (Berlin, Germany). Each experimental session took between 1-1.5 hours. Participants were tested in pairs, and provided each with an iPhone 4 with the prototype apps, as well as two magnets (bar- and ring-shaped). Study participants were guided by two experimenters. They were seated at opposite ends of a table, where a tripod-mounted camera was aimed at their gesture interaction space. They were allowed to define their own interaction space (within the camera angle's view) to ensure their comfort during the session. At start of the session, each participant filled a background information form, signed an informed consent form, and read through instructions for performing the task. Before each condition, they were given a quick tutorial and demo on how to play with each app.

After the tutorial, participants would then play with each app (with no set time limit). They were asked to try out both magnets. After stopping each application, they were asked to fill in the AttrakDiff2, the SUS, and the constructed Likert-scale intermediate questionnaire. All participant responses were set on the same questionnaire, to ensure that responses were relative to one another. After playing with all apps, they were briefly interviewed about their experiences ($\sim$10 min.) of the experimental session and the magnet-based ADI paradigm and given applications. Afterward, they were thanked for participating, signed a receipt form, and offered a monetary reward for participating (which participants knew they would get).

6.6 Results

6.6.1 Perceived Usability & User Experience

AttrakDiff2 Responses

We ran an independent one-way Analysis of Variance (ANOVA) (with all assumptions satisfied) between groups (musically trained vs. non-musically trained) for each AttrakDiff2 dimension (PQ, HQ-I, HQ-S, ATT), however no significant differences between groups were found. Therefore, both groups were treated as a uniform sample. We ran a repeated measures ANOVA comparing mean AttrakDiff2 responses across all participants on each dimension for each of the tested apps (Air DJ, Air Guitar, Air GuitaRhythm). Results showed significant differences in responses across quality dimensions for only PQ, HQ-I, and ATT. Post-hoc pairwise comparisons (with Bonferroni correction[4]) between each app (Air DJ, Air Guitar, Air GuitaRhythm) were conducted in every case. Results (means, standard deviations, confidence intervals, significance (α = .05), and effect size (partial eta-squared values)) for each tested app are shown in Table 6.1. Where significant, dimensions and app names are represented in bold, and where a particular pairwise comparison is not significant, app names are in (additional) italics.

For the PQ dimension, participants perceived clear differences between the Air DJ and the Air Guitar apps, and between the Air Guitar and Air GuitaRhythm apps, but not between the Air DJ and Air GuitaRhythm apps. Lack of a difference in the latter case is not surprising, given that the Air Guitar app was still in the early stages of development, and usability issues associated with its use were expected (which we discuss below). Scores for both Air DJ and Air GuitaRhythm apps in general showed that the current usability of those prototype apps was satisfactory. For the HQ-I dimension, while results showed an overall significant difference between AttrakDiff2 responses, post-hoc pairwise comparisons were not significant. With the scores for all apps all close to zero, we can draw that our participant sample did not clearly identify with these music-related prototype apps. This could be due to the novelty of magnet-based ADI for playful music composition and gaming,

[4]Backward-corrected SPSS© Bonferroni adjusted p-values are reported.

Dimension	App	M	SD	CI	P-value	η_p^2
	DJ	.7	1	[.2,1]	**p=.001**	
PQ	**G**	-.1	1.1	[-.6,.4]	$F(2,46) = 8.3$	.3
	GR	1	.9	[.7,1.4]		
	DJ	.2	1.1	[-.1,.9]	**p=.02**	
HQ-I	G	-.3	1	[-.8,.1]	$F(2,46) = 4.2$	.2
	GR	.3	1.2	[-.2,.8]		
	DJ	.6	.9	[.2,1]	p=.29	
HQ-S	G	.2	1.2	[-.4,.7]	$F(2,46) = 1.3$	.1
	GR	.5	1.2	[-.04,1]		
	DJ	.8	1.2	[.3,1.3]	**p=.03**	
ATT	G	.1	1.4	[-.5,.7]	$F(2,46) = 3.9$	.1
	GR	1	1.4	[.5,1.6]		

Table 6.1: Descriptive and inferential statistics of participant (N=24) responses (range [-3, 3]) on tested apps (DJ: Air DJ, G: Air Guitar, GR: Air GuitaRhythm) for each AttrakDiff2 dimension: PQ = Pragmatic Quality, HQ-I: Hedonic Quality-Identity, HQ-S: Hedonic Quality-Stimulation, ATT: Attractiveness.

which may take time to be accepted by users as an established alternative mode of mobile interaction.

For the HQ-S dimension, there were no statistical differences, where average responses were between 0 and 1. Here, we expected responses on HQ-S to be higher, given the highly positive qualitative responses from participants in the exit interview (described below). This could be due to the limited stimulation categories of the AttrakDiff2 questionnaire, or due to participants hesitant about providing very high ratings in their responses. While this finding may be interpreted with caution, it should nevertheless serve as an indicator that the engagement and stimulation factors associated with interacting with these music-related prototype apps can be improved upon. For the ATT dimension, results showed an overall significant difference, however post-hoc pairwise comparisons did not. Response scores for the Air DJ and Air GuitaRhythm were on or around 1, indicating that both apps generally appealed to participants. For the Air Guitar app, the near-zero response could have been influenced by the poor usability of the app, which may have affected its current attractiveness.

System Usability Scale Responses

Measured SUS responses were calculated according to (Brooke, 1996), and analyzed in terms of average score frequency distributions. Results are shown in Fig. 7.9. For the Air DJ app, half of participants (12/24) gave an acceptable SUS score (70 or above). For Air Guitar, few participants (4/24) gave an acceptable score, and for Air GuitaRhythm, slightly more than half (15/24) gave an acceptable SUS score. For the Air DJ and Air GuitaRhythm apps, these scores reflect that the ADI-based apps using magnets are nearly ready for entering the consumer market with only few issues remaining (as will be discussed later). For the Air Guitar app, the acceptable usability of the current app was quite low, which is not surprising given the early stage of development during time of testing.

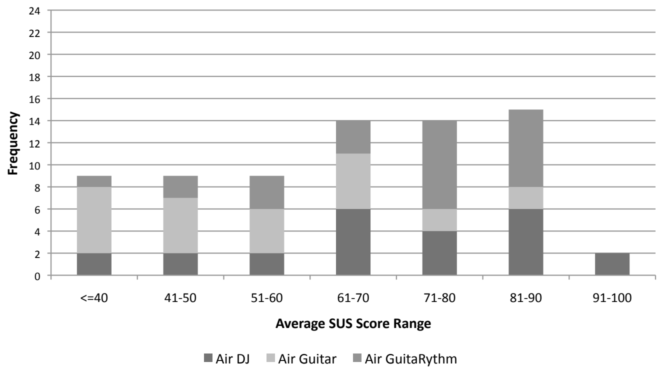

Figure 6.5: Frequency distribution of mean System Usability Scale responses across participants (N=24) for all tested apps.

6.6.2 Users' Subjective Feedback

After interacting with each of the magnet-based ADI apps, participants were given a 7-point Likert scale (1-strongly disagree, 7-strongly agree) exit questionnaire to gather their overall feedback on each of the tested apps (Medians (Md) and Interquartile Ranges (IQR) are reported). This was followed by a semi-structured interview.

Overall User Acceptance of Magnet-based ADI

For all tested apps, participants reported that they had a positive first impression of interacting with the apps (Air DJ: Md=5, IQR=3-6; Air Guitar: Md=5, IQR=2.8-5; Air GuitaRhythm: Md=5, IQR=4-6). This was confirmed during the interviews, were most participants (20/24) had a positive overall impression of composing music and gaming using magnet-based ADI (P6: "*It was really cool with the magnet, I mean I've never even heard of that before!*").

While overall responses were positive, some participants had concerns regarding the originality (P12: "*I have a Nintendo Wii, and I've seen similar technology so I was not so impressed.*"), the use of magnets (P24: *"Sensitivity of the magnet was not good.. especially the DJ app, I think it's much more practical to tap on the instrument [touchscreen] then using a magnet."*), and the limited features of the prototype apps (P17: *"Some of the apps you can make more music, but some were boring (like the Air Guitar app)... maybe would be cooler if you can make more things with the magnet."*). When participants were asked about their expectations of the availability of magnets, all participants stated (as in early user insight sessions) that magnets should be available at electronic and large department stores (e.g., Apple Store®, Woolworths®). Four participants mentioned that if the apps require a particular shape or strength of a magnet, then such magnets should be readily available for purchase.

All participants found the apps easy to use (Air DJ: Md=6, IQR=4.8-7; Air Guitar: Md=6, IQR=3.8-6; Air GuitaRhythm: Md=6, IQR=5-7). Participants were generally comfortable interacting with the Air DJ (Md=5, IQR=5-6) and Air GuitaRhythm apps (Md=6, IQR=4-6) using a magnet, however were neutral with respect to the Air Guitar app (Md=3, IQR=2-5). When asked about whether they are willing to carry a magnet with them to interact with such apps, few participants (8/24) reported they would. Likewise to the mixed adoption of pen-based computing (Kurtenbach, 2010), this was expected as users do not always want to carry an additional accessory. However, participants that stated they would carry a magnet, mentioned (x5) that attaching the magnet to their keychain or as part of the phone's casing were the easiest methods to carry it around.

Participants (18/24) were generally quite positive about using magnets for such interaction (P11: *"I find the idea itself nice, having the interaction outside in your own personal area around the phone.. these small gestures in your personal space feel very natural, it feels good."*). Main issues concerning the interaction included the sensitivity of the detection (P10: *"I thought magnets weren't sensitive enough [especially for Air Guitar App], so that was a bit annoying."*), having to carry the magnet (P14: *"The magnet is small and maybe you can lose it in your pocket."*), and using magnets near electronic devices (P12: *"I was feeling a bit uncomfortable because I think the magnet affects the hard disk, so I was a bit scared."*). The concern of the sensitivity of the magnet can be adjusted accordingly through user-defined calibration. Regarding the possible damage to the smartphone from the magnet, the magnets used for these applications are not strong enough to interfere with the smartphone's hard disk.[5]

Magnet Form Factor

Participants were asked about their preferences for a magnet size or shape for ADI. Around half (13/24) preferred the magnet ring, some (9/24) the bar-shaped magnet, and the remaining (2/24) had other preferences. The ring was preferred due to the ease of carrying it (P3: *"I liked the ring more because you can slip it on and you can have a free hand."*), the natural interaction it affords (P4: *"For the Air Guitar app, a bar makes more sense, but I liked the ring. I thought the ring felt more natural."*), or aesthetic reasons (P13: *"I wouldn't mind wearing a (colorful) ring, as a fashion accessory."*). The strength of the magnetic signal from the ring was perceived to be both good (P18: *"It's [the Ring] stronger and easier to use."*) and bad (P3: *"The ring was stronger and there was many double strokes and I didn't like that."*). Despite that participants could caliberate the magnet sensitivity, this was not always done correctly. This indicates that the calibration procedure should be more transparent to users, so as to avoid very weak or very strong magnetic signals. Other issue reported with the ring-shaped magnet is whether or not it fit the participant's finger (P16: *"At first I thought I liked the ring better, but then it was a bit awkward to use – it fell off at times. So maybe one that fits me better."*).

The bar-shaped magnet was preferred by some participants because it resembled a stylus/pen, which was easy to grasp. One participant stated that it resembles an instrument, which is suitable for these apps. Two participants mentioned it was smaller (and therefore easier to carry). The main concern over the bar-shaped magnet was that the magnetic signal was weaker, and so lacked sensitivity during music composition and gaming, despite calibration. Some participants (4/24) mentioned that for the Air Guitar app, the magnet should probably be shaped like a guitar pick (P5: *"As long as we're playing guitar, best*

[5]However, all magnets pose risks to magnetic strips (e.g., on credit cards) at close range. With the rise of new card readers such as Near-Field Communication (NFC) however, this problem is avoided.

to be shaped as a guitar pick. More realistic."). Likewise with a drumming application, where the magnet should be shaped like a drum stick.

Interaction Methods and Styles

Based on our video observations and users' feedback, there arose a number of issues concerning the supported interaction method of interaction using a magnet. First, participants differed in how they held the smartphone during interaction, where some preferred to lay the device on the table, and others held the device in one hand and the magnet in the other. For putting the device on the table, this was the case for the Air Guitar and Air DJ apps, where both touchscreen interactions (holding an instrument button or holding string(s) with fingers) and moving the magnet required simultaneous hand actions (P3: *"It's a good idea to use magnets, but hard to press buttons while holding a magnet and a phone."*). This was especially of concern for the Air Guitar app, as holding both the smartphone and magnet could pose risks in dropping items (P7: *"Air Guitar here is a bit complicated because you have to use more than one finger, to put the device down and use a magnet. And I'm afraid of dropping the phone so I don't hold it."*). These concerns are in line with previous work that investigated the effects of encumbrance (manual multitasking) on mobile interactions like pointing and typing on a keyboard (Oulasvirta and Bergstrom-Lehtovirta, 2011), which may negatively impact task performance.

Another concern was the form factor of the mobile device for playing Air Guitar, where a physical guitar body extension would allow easier grasping of the smartphone as if it were a real guitar (P5: *"Holding the iPhone [for Air Guitar] was kind of uncomfortable, you would need a physical extension."*). Given that the smartphone's display may not always be visible to the user, this brought up the question of whether enabling vibrotactile feedback (Marshall and Wanderley, 2006) (varied by different parameters such as rhythm and waveform) on the strings could better allow for eyes-free air guitar interaction. For the Air DJ app, one participant found it more practical to not gesture altogether using a magnet (P24: *"For the DJ app, I think it's much more practical to tap on the instrument [touchscreen] then using a magnet."*). Together, these findings show that using the magnet-based ADI paradigm for everyday playful interactions like music composition and gaming requires further design considerations (effects of encumbrance, form factor, vibrotactile feedback) when merging physical (gesture-based) and digital (touchscreen) interactions.

Playfulness & Professional Music Performance

All participants reported enjoying composing music and playing with the apps (Air DJ: Md=5, IQR=3-5; Air Guitar: Md=4, IQR=2-5; Air GuitaRhythm: Md=5, IQR=3.8-6). When participants were asked to rank their favorite app, Air GuitaRhythm ranked the highest (12x), followed by Air Guitar (7x) and the Air DJ app (5x). Given that Air GuitaRhythm was the only app developed with full gaming elements, it was perceived to be overall the most engaging and fun (P23: *"The first one [Air GuitaRhythm] you had a goal, and it was really fun, but the other two I didn't know what to do."*). This is in line with previous work on flow experiences where challenge (here in the form of a game score) strongly influences fun and engagement (Csikszentmihalyi, 1990). However, amongst those who could play music, the Air Guitar app was perceived to be the most creative of the apps (P13: *"I liked the guitar app because it was very creative."*).

Even though the developed applications were not targeted towards professional musicians, a few participants with musical training expressed concern over the apps. For the Air DJ app, one participant found it too simple (P6: *"The DJ game was too simple, there should have been more functions. Like a keyboard or something. Or maybe even track mixing."*). For the Air Guitar app, one participant (who is a bass guitarist) mentioned the problem of not supporting open strings (P5: *"If there were open strings, the guitar app would be much better."*). Another participant (who is a sound production engineer and musician) mentioned that generally he would prefer the magnet-based ADI if it allowed for continuous control over the magnet signal (P21: *"For me, it would be good to find a way to measure the magnet signal so that it is not only a trigger, but a continuous signal. So basically a controller."*). Despite his concerns, this same participant was able to easily compose music with the Air Guitar app, where he composed the Jingle Bells tune upon request (see attached Video).

Social Acceptability

Participants were asked about how socially acceptable interacting with the ADI-based apps using a magnet is in public places, and whether they would do it in public (e.g., metro, café). Most (18/24) stated they would (P17: *"Yeah, I wouldn't care, people can do what they want. It actually feels quite natural to play with this."*), however two of those participants mentioned that they would do this only if they had headphones on. From the six participants who stated they would not, three mentioned they would if they were with a group of friends (P19: *"If I'm alone, no. But with friends, yeah why not."*). Another participant mentioned he would engage in such interaction if the magnet form factor was more appropriate (P5: *"Yes, as long as it looks cooler than this. If it was a magnet pick, maybe yeah."*). Together, these findings provide early indicators that magnet-based ADI may become part of people's daily lives, even in public settings. However, to fully verify this would require a longitudinal test of these apps in users' daily environment.

Other Magnet-based ADI Application Areas

When participants were asked about other potential use-cases for magnet-based ADI, many (16/24) had immediate ideas of other application areas. These included gesturing in the air for text entry and drawing/painting, substituting or extending basic mobile touchscreen interactions (answering call, rejecting call) when the device is occluded (e.g., in one's pocket), rhythmic skill practice (for musical training), and especially gaming. For gaming, this included first person shooters, sports games like tennis, and even an Angry Birds© adaptation. Some mentioned the potential for multiplayer collaborative gaming, especially for games like Air GuitaRhythm (however they did express that the magnets might interfere with the other player's smartphone). Together, these suggestions provide further evidence on the potential of the magnet-based ADI paradigm for supporting user activities, even outside of the music composition and gaming domains.

6.7 Discussion

6.7.1 Supporting Playful Music Composition and Gaming

Despite that the music-related applications we developed under the magnet-based ADI paradigm were not targeted towards professional musicians, we still expected to see a significant difference in mean AttrakDiff2 dimension scores between the musically-trained and non-musically trained groups, especially for the Air Guitar app. While there were differences for the Air Guitar in usability (PQ: Mean musically-trained: -.4; Mean non-musically trained: .2) and perceived novelty and stimulation (HQ-S: Mean musically-trained: 0.5; Mean non-musically trained: -.1), these differences were not statistically significant. However, they do partially indicate that musically-trained users were more critical of the usability of the Air Guitar app, as well as perceived it as more novel given the creativity it affords from them.

From our AttrakDiff2 scores and SUS scores, we showed that the apps based on the magnet-based ADI paradigm were generally positively perceived, and aside from the Air Guitar app, were perceived to be usable. From our user subjective reports and observations, we showed that the magnet-based ADI paradigm can indeed support playful music composition and gaming on mobile devices, and that this mode of interaction is a fun method of musical interaction. Based on these findings, we can confidently state that the creative apps (Air DJ, Air Guitar) can be used for music composition on the go, by amateur musicians and musically-affine users alike. The Air GuitaRhythm app, already at a quite usable stage, established a novel form of musical gaming experience for mobile devices. Taken together, our findings confirm our hypothesis that the magnet-based ADI paradigm can go beyond HCI work focused on user situation impairments or improving user performance when using a given ADI interaction technique (e.g., pointing and target selection (Ashbrook et al., 2011)), but be effectively applied to support playful music composition and gaming in mobile interaction.

6.7.2 Study Limitations

There are three potential limitations to the present study. First, since our study was conducted in a laboratory, it had less ecological validity. However, since participants were tested in pairs (where in most cases they did not know each other beforehand) and given the presence of the two experimenters, the experimental setting closely resembled natural situations. Moreover, given that participants mostly found the magnet-based ADI to be socially acceptable (amidst present strangers) and their positive responses on the naturalness of this mode of interaction, our results can likely be generalized to outside of the laboratory usage scenarios. However, at this stage it is difficult to predict whether long-term usage of these apps would provide the same level of entertainment for the everyday user. To address this, we propose to include gamification (Deterding et al., 2011) elements to all the music apps, to ensure long-term user engagement (discussed under Section *Design Considerations for Magnet-based ADI*).

A second limitation was that two of the tested apps (Air DJ, Air Guitar) were still in an earlier stage of development, where participants explicitly mentioned that the Air DJ app could benefit from more features, and the Air Guitar app was buggy at times (either too sensitive or not sensitive resulting in double or no strokes). Indeed, while these two apps could have been improved upon further, our findings indicate that they were nevertheless useful probes into the validation and suggestion of design improvements for the magnet-based ADI paradigm. Related to this point, at this stage we have not measured the time spent by users on each app. While such a measure might provide useful insight into user engagement, this would be more useful at a stage when the apps are more fully developed, and deployed to users across smartphone game stores (e.g., Apple App Store© or Google Play Market©). With a higher sample of users, and observing play time interaction over a longer period, this measure would be more useful.

Lastly, since the present work had dealt with only magnet-based ADI, we may not immediately generalize our findings to other ADI paradigms (especially in cases where optic techniques are used instead of magnets). In such cases, the interaction of using or not not using a magnet may sufficiently differ to warrant a redesign of the tested apps. Furthermore, in other variants of the ADI paradigm, the space of interaction around the device may be more complex, for example by switching the mode of interaction depending on where around the mobile device the user interacts (e.g., front, back, or side of mobile device (cf., Baudisch and Chu (2009))). While we have not made a comparison of our magnet-based ADI for playful music composition and gaming with other ADI variants, our findings nevertheless showed that participants understood the concept behind ADI (P11: *"This kind of personal space interaction felt very natural, would it be possible to do this without the magnet?"*) and its application areas.

6.8 Design Recommendations

Based on our findings, we draw design considerations that improve the user experience of using magnet-based ADI applications, specifically in the context of music-related (gaming) applications.

1. **Designing Natural Interactions:** As revealed in the interview responses (Sections *AttrakDiff2 Responses*, *Overall User Acceptance of Magnet-based ADI*, and *Playfulness & Professional Music Performance*), users enjoy interacting with a magnet, and they find this mode of interaction quite natural. For interaction designers, this means that making use of gestural input techniques alongside touchscreen interaction (cf., Norman (2010)) for domains such as music gaming and composition is a worthwhile design goal to support entertainment.

2. **Supporting Magnet Availability & Use:** Based on participant responses (Section *Overall User Acceptance of Magnet-based ADI*), it is a valid consideration that magnets should perhaps be readily available in stores, and come in different form factors (shape, size, color) (cf., Section *Magnet Form Factor*), especially for applications that afford a direct form mapping between the magnet and the instrument/tool (e.g., guitar pick, sword) it uses (cf., Grandhi et al. (2011)). Additionally, based on some participant responses, support for carrying the magnet (keychains, phone casings) should also perhaps be readily available.

3. **Transparency in Application Calibration**: Based on user's remarks concerning magnet sensitivity (Section *Overall User Acceptance of Magnet-based ADI* and Sec-

tion *Magnet Form Factor*), should have immediate access and information on calibrating the magnet sensitivity, given the shape, size and strength of the magnet with respect to the application.

4. **Effective Use of Multimodal Feedback:** Given user concerns over interaction methods with the music-related apps (Section *Interaction Methods and Styles*), the apps should be augmented with multimodal feedback where necessary (e.g., vibrotactile alongside visual feedback on digital guitar strings), so as to allow smooth simultaneous physical and touchscreen interaction. Additionally, physical form extensions to a mobile device would allow for more natural musical interactions (e.g., guitar casing).

5. **Gamification for User Engagement:** To ensure engagement with creative musical apps by non-musically trained users, game-like elements (gamification) (Deterding et al., 2011) can augment the user experience (cf., Section *Interaction Methods and Styles*).

6. **Addition of User-requested Features:** To extend these musical apps to support professional musicians, extra features need to be added (e.g., magnet as controller, open strings, DJ mixing) (cf., Section *Playfulness & Professional Music Performance*) alongside other application areas for magnet-based ADI (cf., Section *Other Magnet-based ADI Application Areas*), such as rhythmic skill practice and a wider variety of games.

7. **Requirement of Longitudinal In-the-Wild Testing:** Given that most of our participants found magnet-based ADI to be socially acceptable when used in public settings (cf., Section *Social Acceptability*), designing ADI for social settings, while still requires further longitudinal real-world testing, provides an early indicator that this form of interaction may be socially acceptable.

6.9 Conclusions

In this chapter, we presented the results of a user study to investigate whether 3D gestural input using the magnet-based ADI technique can be effectively used to support playful music composition and gaming on mobile devices. In line with our hypothesis and the goal of this chapter, we were able to show that 3D gestural input does offer a playful and natural interaction method for composing music and playing music-related games, and is entertaining to users. We believe we have set the stage for further experimentation with applied use-cases across domains for 3D gestural input in general, and magnet-based ADI in particular.

In this chapter we looked at how magnet-based ADI can be applied in a playful interaction context. As an additional example of 3D gesture-based interaction in daily mobile interactions, in the next chapter (Chapter 7) we look at a common task performed by mobile users: authentication.

7 Usability and Security Tradeoff in 3D Gestural Authentication

Using the magnet-based ADI paradigm introduced in Chapter 6, we investigate the usability and security trade-off in 3D gestural authentication (*Gestural Authentication Study*). We replicated a controlled security study to assess the vulnerability of this authentication method against video-based shoulder surfing attacks. Additionally, in a separate controlled study, we measured the usability and user experience issues associated with performing air gestural signatures for smartphone security access. For this, we measured user experience using experience, subjective workload, and trust questionnaires as well as analyzed performed signature durations. Our security analysis provided further validation of the security of this authentication method, and with our user experience research, we arrived at a set of design recommendations for optimizing the user experience of using this magnetic gestural authentication method. The work presented in this chapter is currently under peer-review as "Investigating the Usability and Security Trade-off in Magnetic Gestural Authentication" in the International Journal of Human Computer Studies (El Ali et al., 2013a).

7.1 Introduction

As presented in Chapter 6, magnet-based ADI is a subset of ADI, which also allows gestural interaction in the whole 3D space around the device by deforming an embedded compass sensor's magnetic field. Previous work (Sieger et al., 2010) has shown a need for additional layers of security for different settings, including entering a pin code to access ATMs or unlock smartphones. Indeed, in a survey with 465 participants asking about security methods on mobile phones, Ben-Asher et al. (2011) found that only 26.7% of respondents perceived PIN-based entry methods to be a secure method of user authentication.

ADI opens up new forms of user authentication, where users can perform mid-air 3D gestures (e.g., their signature) to gain access to a system (e.g., Patel et al. (2004)). This promises a fast, secure and natural method for user authentication. User authentication is achieved here by allowing a user to gesture in mid-air and verifying whether this signature matches that user's recorded template signature (Guse et al., 2011). Despite that 3D gesture authentication is not generally perceived (as assessed by a web survey) as providing a high level of security by users (Sieger et al., 2010), recent work has shown that 3D gestures are in fact quite secure against video-based shoulder surfing attacks (Sahami Shirazi et al., 2012). However, to fully understand whether this kind of authentication method would be culturally and commercially adopted, the trade-off between security and usability issues (Weir et al., 2010; Cranor and Garfinkel, 2005; Yee, 2004; Schultz et al., 2001) behind this form of interaction needs to be further investigated.

In this chapter, we look closely at the security and usability issues associated with magnet-based ADI and its applied use for user authentication purposes. While this work focuses on smartphone security access, this method of authentication is applicable to security access of any device (e.g., laptops, doors or ATMs) if embedded with a magnetometer. To investigate the usability issues associated with this method of authentication, a controlled laboratory experiment was set up where users had to define magnetic 3D signatures. Usability and user experience data was collected using a a mixed-methods approach, including measuring signature performance duration, usability and user experience Likert-scale questionnaires, and interviews. To assess the vulnerability of this authentication method, a separate video-based shoulder surfing attack scenario (cf., Sahami Shirazi et al. (2012)) was set up in a controlled setting where users had to try to forge some of the signatures defined in the usability study.

The rest of the chapter is structured as follows: first we provide our research questions followed by a review of related work and a description of our magnet-based ADI framework. We then introduce the usability study design and methods, followed by the security study design and methods. Afterward, we present our results and discuss them, provide design recommendations and conclude.

7.2 Research Questions

In this chapter, our main research question is:

> **RQ 6:** How does 3D gestural interaction affect the usability and security of mobile gestural user authentication?

Specifically, what is the security of this method (as measured by recognition accuracy under a video-based shoulder surfing attack scenario), and what are the usability issues (as measured by overall system acceptance, perceived security, gesture recall, and gesture duration) associated with using this kind of user authentication method?

This chapter builds directly on previous work, where under a video-based shoulder surfing attack scenario, Sahami Shirazi et al. (2012) showed that this authentication method is indeed quite secure against visual-based forgeries. The present study attempts to replicate the findings by Sahami Shirazi et al. (2012), and additionally investigate the usability and user experience issues associated with using this method. Guse et al. (2011) have shown that accelerometer and gyroscope-based gestural authentication for predefined simple gestures (e.g., circle gesture, infinity gesture) under different forgery conditions (Naive Forgery, Semi-naive Forgery, and Visual Forgery) and for different recognition algorithms (Dynamic Time Warping (DTW) (Sakoe and Chiba, 1978), Hidden Markov Models (HMMs) (Rabiner, 1989)) is a secure method for user authentication. These previous findings led to our first hypothesis, that magnetic gestural authentication is a secure method of user authentication under a video-based shoulder surfing attack scenario.

Previous work has shown that users perceive performing 3D motion gestures for HCI-related tasks (e.g., smart-home control (Khnel et al., 2011)) as a natural mode of interaction (Ruiz et al., 2011; Grandhi et al., 2011), which depending on the gesture performed can also be enjoyable (Rico and Brewster, 2010). Given that magnetic gestural signatures allow a natural mode of interaction, and given the prevalence and acceptance of handwritten paper-based (or wet ink) signatures, our second hypothesis was that magnetic gestural authentication would be perceived to be a usable method for authentication, and provide a positive user experience (UX)[1] amongst participants.

Assessing the usability and user experience issues associated with this method includes measuring responses on whether users are able to define their own unique signature (characteristic-based component), recall their own signature (knowledge-based component), and are willing to carry an external accessory such as a magnet for authentication purposes (token-based component). Additionally, investigating the usability and user experience issues associated with this method also requires investigating how long on average it takes participants to perform and recall a signature (i.e., signature duration), the perceived level of trust by users towards this type of authentication system, as well as how users perceive the difficulty in forging a signature with full video evidence.

Investigating the security, usability and user experience afforded by this authentication method here yields two main research contributions. First, it provides further evidence on the vulnerability of the magnetic gestural authentication method towards video-based shoulder surfing attacks. Second, it provides insight into whether this authentication method (even if shown to be secure) is a usable method for user authentication, and in cases where it falls short, how the usability and UX issues can be addressed.

7.3 Related Work

In addition to the related work listed below, this chapter builds on the related work on using 3D gestures in HCI discussed in Chapter 5 in Section 5.3.1 as well as the related work on Around Device Interaction discussed in Chapter 6 in Section 6.3.1.

7.3.1 Protecting Against Shoulder Surfing Attacks

The degree of a shoulder surfing attack generally depends on the situation in which it occurs (e.g., on the street or at the cashier in a supermarket). Keypads or touch screens in alphanumeric or graphical passwords (Biddle et al., 2012; Suo et al., 2005; Renaud and Angeli, 2004) are particularly vulnerable, since an adversary can easily obtain a direct view of the interaction with the authentication interface. Examples of graphical passwords that use graphics or pictorial representations include PassFaces (Tari et al., 2006), Jiminy (Renaud and Smith, 2001), VIP (De Angeli et al., 2002), Déjà Vu (Dhamija and Perrig, 2000), and Passpoints (Wiedenbeck et al., 2005).

[1]UX here is based on ISO 9241-210 (1994) definition: A person's perceptions and responses that result from the use or anticipated use of a product, system or service."

Magnet-based gesture authentication intersects with 3 authentication schemes available (Shanmugapriya and Padmavathi, 2009; Farella et al., 2006): 1) knowledge based, where users have to recall their own defined gestural signatures (cf., password recall) 2) token-based, where users have to possess a magnet (cf., an RFID tag) and 3) Characteristic or biometric-based, which can either be physiological (e.g., fingerprints, face or iris recognition) or behavioral (e.g., keystroke dynamics (Shanmugapriya and Padmavathi, 2009), gait, or hand motion). Since users' signatures are (largely) unique and characteristic of their hand and arm motion (cf., paper-based or 2D touchscreen signatures), magnetic gestural authentication falls under a biometric behavioral scheme.

While physiological biometric methods provide a high level of security given that they are based on an individual's unique biological markers, they typically involve costly hardware, and even then are not completely risk free (cf., the review by Faundez-Zanuy (2004) on the vulnerability of biometric security systems, or a specific instance applied to attacking fingerprint-based authentication (Uludag and Jain, 2004)). Likewise, in a typical password authentication scheme that involves alphanumeric passwords, dictionary attacks (Morris and Thompson, 1979) can easily succeed, especially since people are used to choosing easy to recall passwords. Indeed, in a case study of 14,000 UNIX passwords, searching from a dictionary of merely 3x106 words revealed almost 25% of the passwords (Klein, 1990). Preventing dictionary attacks for many techniques leads to either heavy computational load (Lin et al., 2001) or user requirements that reduce overall system acceptance by users. Other techniques, such as adopting strong password policies (Computer Security Center U.S., 1986), graphical passwords (Biddle et al., 2012), or designing cognitive games (Roth et al., 2004) again illustrate the classic trade-off between usability and security of user authentication methods (Tari et al., 2006).

A suitable alternative for dealing with shoulder surfing attacks include gaze-based authentication schemes (Kumar et al., 2007) and iris recognition (Daugman, 2004), however these methods pose their own challenges. In the case of gaze-based authentication, it is difficult for users to select a secure password. While users' recall strategies can be augmented by different cognitive mechanisms such as cued-recall (Bulling et al., 2012), this provides further processing costs on the user. Iris recognition, on the other hand, can be vulnerable to attacks that mimic the vitality of the iris or live tissue verification (Prabhakar et al., 2003). Likewise with fingerprint-based authentication schemes, where research has shown that such systems are vulnerable to attacks using artificial 'gummy' fingers made of molds where the live finger was pressed (Tari et al., 2006). The foregoing studies highlight the high tension between different user authentication methods that are both secure and usable.

7.4 Magnetic Gestural Authentication Framework

The idea behind our magnetic gestural authentication framework is to use the embedded magnetic sensor (or magnetometer) of a smartphone as a means of authenticating users. A piece of magnet when moved close enough to a smartphone can influence the compass sensor. A typical magnetic sensor contains a 3D Hall effect sensor that registers the strength of the magnetic field along different dimensions. A Hall effect sensor produces a voltage (Hall potential VH) proportional to the magnetic flux density (B in Tesla), due in par to this so-called Hall effect. The output from the sensor is provided in the x, y, and z coordinates of the smartphone. This output can have different ranges depending on the device (e.g., the iPhone 4® has a value range between 128 μT). Sliding a magnet around the device changes the original magnetic field around the device. The temporal pattern of such an influence is registered by the compass sensor, and can be interpreted as a gestural command using appropriate machine learning algorithms.

A time derivative function is applied to the sensor readings in order to highlight changes in the pattern of magnetic field, and remove effects of earth's magnetic field (which is almost constant). The sequence of vectors is divided into overlapping windows for gesture recognition. In order to match templates, we adapt a template matching algorithm called multi-dimensional Dynamic Time Warping (DTW) (ten Holt et al., 2007) to analyze different 3D magnetic signatures. DTW is suitable for measuring similarities between two signal sequences that may vary in time or speed, and can operate with a limited number of templates and still achieve very accurate results. Getting useful information from the magnetic sensor here is not only algorithmically simpler than implementing computer vision techniques, but this approach also does not suffer from illumination variation and occlusion problems. In other words, it does not require direct line of sight into the camera, which enables covering the whole 360° space around the device for interaction.

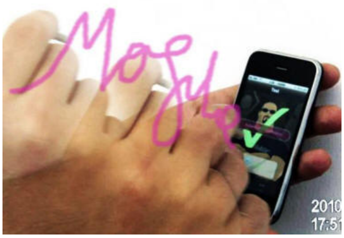

Figure 7.1: An illustration of a user defining a magnetic air signature.

Whenever a user performs a new signature around the device (illustrated in Fig. E.10 on a iPhone 3GS®), the compass sensor registers the temporal patterns of magnetic field along its three axes. Then, DTW is used to compare this multi-dimensional time series signal with pre-recorded templates of the user's signature for authentication. If the distance of a new input gesture with respect to the prerecorded signature is less than some threshold, the person is considered as a legitimate user and granted access to a smartphone or protected device. In the tested prototype (described in Section 7.5.1), in order to define an authentication gesture or magnetic signature, the user arbitrarily moves an appropriate permanent magnet (e.g., a magnetic token/stylus or a magnet in a finger ring) around the device along 3D trajectories.

7.5 Usability Study

7.5.1 Study Design

To investigate the potential of this magnetic gestural authentication system, we designed a controlled study to collect unique signatures from participants and to test the usability

(a) Foam mobile device model

(b) Magnet form factor

Figure 7.2: (a) The foam mobile device mockup with embedded SHAKE sensor box used for recording signatures (b) the magnet used for making a gestural signature, with one pole labelled to ensure correct signature reproduction.

and user experience of this authentication method. A controlled laboratory study was suitable in this case as it parallels real-world situations of user authentication (e.g., unlocking a mobile device or laptop in one's own room, or verification at an ATM in an enclosed space). Additionally, it allows drawing rich user insights and concept validation without the unpredictability of field testing. We adopted a mixed-methods approach that was largely qualitative and based on user insights, in order to test how practical this method would be if it were to be adopted and integrated in the lives of users (Dourish et al., 2004). The collected signatures here were used as material for the shoulder surfing attack scenario (described in the Security Study). Given these design decisions of collecting participant signatures and testing the usability of the method, we set up a single factor experimental design where participants were required to make a unique gestural signature, and later recall this signature.

Apparatus

For recording signatures, we used our magnetic gestural authentication prototype. To obtain precise magnetometer signal information, the SHAKE SK6 (Hughes, 2006) sensor was used instead of an iPhone 3GS/4®. The SHAKE sensor is able to sense magnetic fields from proximate magnetic material and transmit the data to a PC over a Bluetooth connection. We used an authentication software on the PC to capture the magnetometer signals. The SHAKE sensor however has a small form factor (almost the size of a matchbox), which might influence how participants viewed this interaction method. To ensure that there was no such influence, the SHAKE sensor was embedded in a foam model (shown in Fig. 7.2(a)) that replicated the height (115 mm) and width (59 mm) dimensions of an iPhone 4®, however kept the thickness of the SHAKE sensor box (∼15 mm). To perform magnetic gestural signatures, participants were provided with a bar-shaped magnet (∼5cm height, ∼0.5cm width) with one magnetic pole marked (shown in Fig. 7.2(b)).

Defining & Recalling Gestural Signatures

For defining a signature, participants held the foam device model in one hand, and the magnet in the other hand with the marked end always facing upward (to ensure consistency across recorded templates). All users used the same magnet to define the signatures. To ensure that the recorded signature was unique, they were allowed to practice performing this signature as long as they wanted. There was no limitation on the shape or length of the signature, nor on whether they used one or two hands to perform the signature, so long as the magnetometer was able to detect the interaction. Once they were done practicing, they could then record this signature using the authentication software. To ensure that the recorded signatures for participants were consistent, they had to record their signature five times, resulting in five signature templates. The average distance between the five templates (Sign_d) was used as the target template for attacks in the security study. To collect video evidence of the performed signatures for later use in the shoulder surfing attack scenario, participants were recorded from four different angles (front, left, right, rear) while they performed their initial signature.

Once participants recorded their signature, they had to fill in questionnaires prompting them about the usability and user experience of this security method. In order to assess whether they could correctly recall their own signature, they were asked to perform the same signature again after filling in the questionnaires. For the recall signature, participants had three chances to perform the correct signature, resulting in three signature templates. They were however not given feedback (discussed in Section 7.8) as to whether or not they performed the recalled signature correctly, as signal comparisons were done offline. To ensure that the recall scenario was natural, participants were not instructed that they would have to recall the signature they performed, as it would have risked excessively deliberate efforts to memorize the exact signature gesture movement. The lag period between the original signature and the recall signature lasted between 7-10 min (depending on how fast a participant filled out the forms).

Measuring Usability & User Experience

Aside from collecting signal data, to measure the usability and user experience (our dependent variables) of interaction using the magnetic gestural authentication system, seven data sources were collected: a) signing duration of original and recall signature templates b) System Usability Scale (SUS) (Brooke, 1996) responses c) NASA-TLX questionnaire (Hart and Wickens, 1990) responses d) System Trust Scale (STS) (Jian et al., 2000) responses e) Likert-scale questions about participants' attitudes toward using magnetic gestural signatures as an authentication scheme f) video recordings of participants' signatures from four angles g) post-experiment interviews, to get direct user feedback on this authentication method.

Durations of signatures and recalled signatures were measured as they provide knowledge into how long a given signature takes in comparison to existing security methods (e.g., PIN or graphical passwords), as well as provide validation whether original signatures differ from recalled signatures, and from attempted forgery attacks. We administered the SUS (10-item questionnaire on a 5-point Likert scale) to gain insight into the ease of use, efficiency, and satisfaction of this authentication system. The SUS has been shown to be a robust and reliable tool for measuring perceived usability of interactive systems, where a score of 70 and above indicates an acceptable score (Bangor et al., 2008). We gave participants the NASA-TLX questionnaire (Hart and Wickens, 1990) to assess their perceived subjective workload quantitatively ([0,20] response range) through the index's constituent categories: Mental Workload, Physical Workload, Time Pressure, Effort Expended, Performance Level Achieved and Frustration Experienced. Subjective Workload was calculated as follows: (Mental + Physical + Time + Frustration + (20 - Performance Level)) / 5). The STS was given to gain insight into whether participants trusted or distrusted this authentication system. Finally, Likert-scale questions (Cronbach's α=.78) and post-session interviews were given to gain additional insight into how participants perceived the magnetic gestural authentication method.

Participants

20 participants (14 male, 6 female) aged between 20-38 (M_{age}= 29.7; SD_{age}= 5) were recruited. Our participant sample spanned 13 different nationalities, where most were right-handed (18/20). Half (10/20) had a technical background, and more than half (13/20) were familiar with gaming consoles that use some form of gesture recognition technology (e.g., Nintendo Wii© or Microsoft Kinect©). Nearly half (8/20) use some form of security scheme to secure their mobile device or laptop, where 4-digit PIN passwords were the most common. However, all were familiar with password and PIN-based security schemes.

Setup & Procedure

The study was carried out at the usability lab at Telekom Innovation Laboratories (Berlin, Germany). Each session took between 45-60 min. Participants were tested in pairs, where each was provided with a foam model with embedded SHAKE sensor and a pole labeled magnet (shown in Fig. 7.2). Participants were guided by two experimenters. They were seated at opposite ends of a table. Webcams captured the front and side angles of their gesture interaction space, and a tripod-mounted camera was mounted behind them to capture the rear angle (see illustration in Fig. 7.3). Recorded video streams were all synchronized using EvoCam 4©. They were allowed to define their own interaction space (within the cameras' angle views) to ensure their comfort during the session. At the start of the session, each participant filled a background information form, signed an informed consent form, and read through instructions for performing the task. Before starting, they were given a quick tutorial and demo on how to hold the magnet and foam-based sensor, and how to record a signature. To record a signature, they had to press and hold the button on the SHAKE sensor (see Fig. 7.2(a)), where the LED turns blue when the button is held. To stop recording, the button had to be released.

After the tutorial, participants would record five signature templates. After recording their signature, they were asked to fill in the SUS, NASA-TLX, STS, and the Likert-scale questionnaire. Afterwards, they were asked to reproduce their original signatures, and record those recalled signatures three times. Once they finished recording their recalled signature templates, they were briefly interviewed about their experiences ($\sim$10 min.) of the experimental session and the magnetic gestural authentication system. Finally, they were thanked for participating, signed a receipt form, and offered a monetary reward for

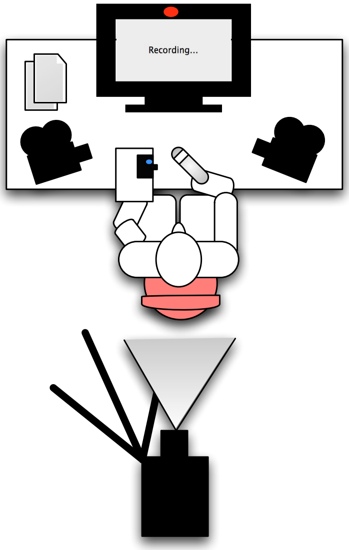

Figure 7.3: An illustration of the standard laboratory setup for collecting signature data from participants. Note: in the actual setup, two participants were tested per session.

participating.

7.6 Security Study

7.6.1 Study Design

Video-based Shoulder Surfing Attacks

To investigate the security of the magnetic gestural authentication method, we built on previous work (Sahami Shirazi et al., 2012) and designed a follow-up controlled experiment to assess the vulnerability of this method against video-based shoulder surfing attacks. Under this scenario, we assume the worst case scenario where the adversary has full access to HD video evidence of the different angles of performed signatures. Since 2D cameras are widely available to attackers, these were used instead of depth cameras (e.g., Kinect©). To ensure that the adversary has sufficient information on the performed gestural signatures, the video recordings of signatures from the usability study were provided to attackers to try and forge the targeted signatures. These videos captured defined signatures from four different angles: front, left, right, rear. To make this scenario realistic, we put a restriction on the number of security attacks, where an adversary was allowed only a total of three attempts. It should be noted that our study is based on adversarial attacks of average, everyday persons, and not skilled forgers.

Measuring Security

As in ink-based 2D signatures, some signatures are easier to forge than others (e.g., few strokes or clear letters of a person's name). If our gestural authentication method is to provide security across all signature types, then it has to be secure enough against both easy and difficult signatures, and against a variation in signature styles (e.g., using one or two hands for performing a signature). Therefore, we decided to additionally test the vulnerability of this method when videos for attack showed both easy and difficult signatures, and signatures that used one or two hands to perform. To determine which signatures from the usability study were easy or difficult, two independent coders were recruited and asked to make a checklist amongst the resulting (post data cleaning) 15 signatures (see Section 7.7.1) made and to determine which are easy and difficult. The easy and difficult signatures were varied amongst those with a variety of styles (such as using one or two hands to perform). Their lists were subsequently cross-checked, and matching candidates for easy and difficult signatures were nominated. The experimenter then selected two easy signatures and two difficult signatures for forgers to target in this study.

Figure 7.4: Participant using paper aids for targeting a signature.

The foregoing design decisions led to a within-subject factorial (2 x 1) design, where all participants had to forge gestural signatures, and signature difficulty (2 levels: easy vs. difficult) was a within-subjects factor. Participant assignment was randomized, and order of presented videos was counterbalanced. As in the usability study, participants were given a short training and all relevant hints for forging, such as grasping the foam SHAKE device with the correct position and orientation and how the magnet was held (marker on magnet always up). The four videos (corresponding to different view angles of each signature) recorded in the usability study were shown to each participant, who were then asked to forge the targeted signatures. Participants used the same foam SHAKE device and magnet as the one used in the video. There was no restriction to the study duration, where participants could watch the videos as many times as desired. Additionally, they could speed/slow down the videos, as well as step through each frame individually. Aside from being allowed to practice the signature motion as long as desired, they were also given a notepad and pen to draw the gestured signature motion if they wanted (e.g., Fig. 7.4).

As in the usability study, duration of forged signatures were measured. This was done to make duration comparisons between the original signature and mean forged signature duration. After forging each signature, participants were given a short Likert-scale questionnaire (Cronbach's α=.64) that asked about their experiences in forging that particular signature. Responses for each signature were recorded on the same questionnaire, taking into account the counterbalancing of the presented videos for forgery. After all forgery attempts were made, participants were given another short Likert-scale questionnaire (Cronbach's α=.7) and an exit semi-structured interview to inquire about the difficulty in forging signatures in general, the clarity of the videos and their video angle preferences, and what their general attitudes are to this gestural authentication scheme.

Participants

20 participants (11 male, 9 female) aged between 20-34 (M_{age}= 27.1; SD_{age}= 3.6) were recruited. Participant sample spanned 10 different nationalities, where all were right-handed. Nearly half (19/20) had a technical background, and half (10/20) were familiar with gaming consoles that use some form of gesture recognition technology. More than half (13/20) used some form of security scheme to secure their mobile device or laptop, where again 4-digit PIN passwords were the most common, and all were familiar with password and PIN-based security schemes.

Setup & Procedure

The study was also carried out at the usability lab at Telekom Innovation Laboratories (Berlin, Germany). Each session took around 60 min. Participants were again tested in pairs, where each was provided with a foam model with embedded SHAKE sensor and the same pole labeled magnet used in the usability study. Participants were guided by two experimenters. They were seated at opposite ends of a table. Each participant was seated in front of a PC, where they could inspect the video footage of the gestural signatures to be forged. At start of the session, each participant filled a background information form, signed an informed consent form, and read through instructions for performing the task.

Before each condition (easy or difficult signature), they were given a tutorial and instructed how to hold the magnet and foam-based sensor as was done in the usability study. They had as long as they wanted to practice the signature. Once they were ready, they were allowed to record three signature attacks. After recording the signatures, they were asked to fill in a Likert-scale questionnaire, and were briefly interviewed about their experiences ($\sim$10 min.) of the experimental session and their forgery attempts. Finally, they were thanked for participating, signed a receipt form, and offered a monetary reward for participating.

7.7 Results

7.7.1 Security

Data Cleaning

To validate signature templates recorded by participants in the usability study for consistency, a data cleaning procedure using 5-fold cross-validation was used. A cutoff point on similarity between signature templates (Sign$_t$, t = $\{1,...,5\}$) was set. If the ratio was greater than 1.5, then the signature data for that participant was considered inconsistent, and had to be rejected. If it was lower than 1.5, then the signature was shown to be consistent. This cleaning process resulted in removal of data of five participants, leaving consistent signature template data for 15 participants. For the security study, one participant did not correctly record the forgery templates, and had to be excluded. This left data from 19 forgers to be considered.

Recall

As in 2D ink-based signatures, a person's signature varies each time it is performed to some degree. In order to define a 3D magnetic signature and check the repeatability, the user is required to enter a signature template five times (Sign_t, $t = \{1,...,5\}$). Once the user successfully registers his own personalized 3D signature, the system can be used. The average DTW distance of all templates is then calculated and used as the main signature (Sign_d). For each signature (Sign_t), a ratio was calculated by dividing by the main signature ($\text{Ratio}_t = \text{Sign}_t/\text{Sign}_d$).

To show whether a participant was able to recall his own signature, the recall distance score (Recall_t, $t = \{1,2,3\}$) for each signature was compared with the ratio of the main signature (Ratio_t). If the ratio value for a recalled signature was above this value, then it can be shown that the participant was unable to recall his signature. To find the acceptable threshold, we found the minimum and maximum thresholds for accepting a signature as a participant's own, in addition to the ratio across all participants. The lower the threshold (θ) value, the higher the acceptance rate. For original signatures made, the thresholds across 15 participants for each recall attempt are shown in Fig. 7.5, with threshold values summarized in Table 7.1. At $\bar{\theta} = 2$, the percentage of successful login attempts is 84.4%. However, as will be shown below (Section 7.7.1), this threshold value need not be this low to protect against security attacks, while still allowing eligible logins.

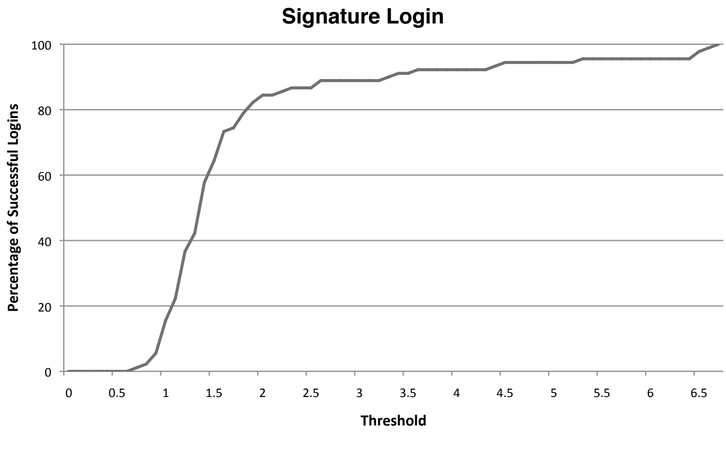

Figure 7.5: For login attempts across users ($N = 15$), with $\theta = 6.5$, the percentage of successful logins is 100% successful.

	$t1$	$t2$	$t3$	Range	$\bar{\theta}$	SD
Recalled Signature	1.8	1.5	1.6	.7-3.6	1.6	.2

Table 7.1: Recall attempt thresholds (θ) and mean threshold $\bar{\theta}$ for recalled signatures (N = 15).

Forgery of Original Signatures

For forgery attempts, a similar procedure was applied to identify the corresponding ratio of threshold values across trials. In order to forge a login, an attacker has three attempts (Forge$_t$, $t = \{1,2,3\}$). For each attack attempt, Ratio$_t$ = Forge$_t$/Sign$_d$ is calculated. If Ratio$_t$ is smaller than a given threshold (θ), the forgery is successful. Despite that the first successful attempt is enough to authenticate the forger, we take the average forgery threshold value in our analysis. For attack of easy and difficult signatures, the thresholds across 19 participants for easy and difficult signature attacks are shown in Fig. 7.6. Threshold values are summarized in Table 7.2.

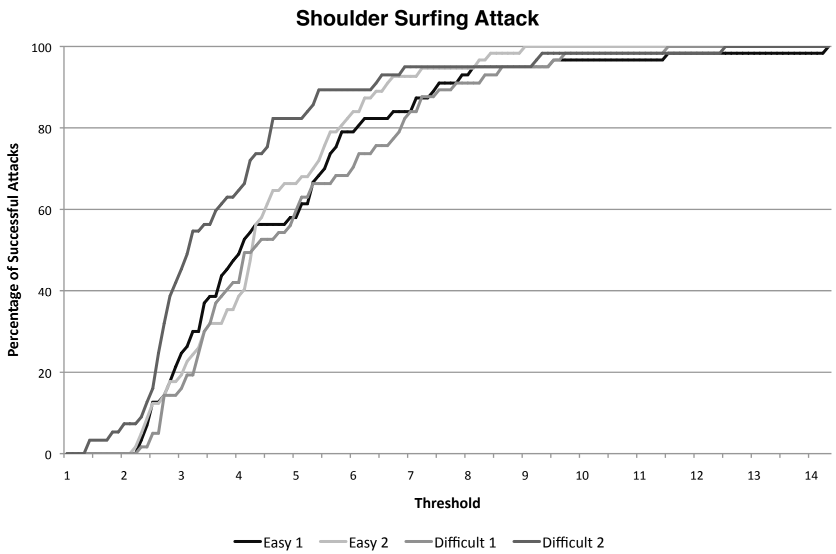

Figure 7.6: For easy and difficult signature attacks across forgers (N = 19), with θ = 2.4, the percentage of successful attacks is less than 10%.

Based on these results, with θ = 2.4, the percentage of successful attacks on both easy and difficult signatures is less than 10%. To evaluate the accuracy of authenticating eligible users for these signatures, we compared the original main signature made by that participant (Sign$_d$) with each recall attempt made (Recall$_t$, $t = \{1,2,3\}$). For recall attempts across 15 participants (Section 7.7.1), with $\bar{\theta}$ = 2.4, the percentage of successful logins is 86.7%. This shows that tuning the threshold value for signature recognition around 2.4 provides a good balance between false acceptance (allowing attackers to login) and true acceptance (permitting eligible logins) in protecting against video-based shoulder-surfing

Condition	$t1$	$t2$	$t3$	**Range**	$\bar{\theta}$	SD
Easy Signature$_1$ Forgery	5.1	4.7	4.5	2.3-14.3	4.8	.3
Easy Signature$_1$ Recall	3.4	1.2	2.5		2.4	1.1
Easy Signature$_2$ Forgery	4.4	4.5	4.5	2.2-9	4.5	.06
Easy Signature$_2$ Recall	2	1.7	1.6		1.8	.21
Difficult Signature$_1$ Forgery	5	4.9	4.8	2.3-11.5	4.9	.1
Difficult Signature$_1$ Recall	1.5	1.4	1.4		1.4	.1
Difficult Signature$_2$ Forgery	3.6	3.9	3.9	1.4-12.5	3.8	.2
Difficult Signature$_2$ Recall	1.5	.9	1		1.1	.3

Table 7.2: Attack and recall attempt thresholds (θ) and mean thresholds ($\bar{\theta}$) for recalled (N = 4) and forged signatures (N = 19).

attacks. The average ratios for attacks and recall attempts across both easy and difficult signatures are shown in Table 7.2.

ROC Analysis of Magnetic Signature Data

To gain further insight into the security of the gestural authentication method, we used Equal Error Rate (EER) to measure accuracy. This is the rate at which False Acceptance Rate (FAR) and False Rejection Rate (FRR) are equal. FAR is the probability that a non-authorized person is identified and FRR is the probability that an authorized person is not identified. We use this measure because typically the number of genuine cases in a verification system are much smaller than the number of forgery cases. In this case, we made use of all recalled signatures defined in the usability study, as well as forgeries made in the security study. We have 15 (signers) x 3 (recall samples) = 45 cases for genuine recall signatures, and 19 (forgers) x 3 (attack samples) x 4 (signatures) = 228 attacks. FAR and FRR were calculated as follows:

$$FAR = \frac{\#\, of\; verified\, forgery\, cases}{\#\, of\;\, forgery\, cases}$$

$$FRR = \frac{\#\, of\; rejected\, genuine\, cases}{\#\, of\;\, genuine\, cases}$$

To calculate EER, for each threshold value (0-14.3) the corresponding FAR and FRR were derived. The value of the point at which FAR and FRR are equal is the EER. As shown in Table 7.3, the The EER is 13%, at threshold value of 2.4, which shows that the magnetic gestural authentication system provides both security (even in high-risk situations), as well as usable access to users of the system. To plot the Receiver Operating Characteristic (ROC) curve, the True Acceptance Rate (TAR) was calculated (100 - FRR). All the (FAR,TAR) pairs were used to plot the ROC curve (shown in Fig. 7.7).

To further investigate whether the EER rate was an artifact of the 2-handed signatures, which might have tipped the balance between having a strong signature and being able to recall it easily, we removed those signatures from our analysis. In this case, we had 15 (signers) x 3 (recall samples) = 45 cases for genuine recall signatures, and 19 (forgers) x 3 (attack samples) x 2 (signature difficulty) = 114 attacks. The EER rate for these two signatures drops to 4% at threshold value of 2.4 (Table 7.3), with the corresponding ROC curve (shown in Fig. 7.8). This highlights the need for users to consider making a strong signature one the one hand, but one that is also easy to recall (discussed in Section 7.9).

	Threshold	EER (%)
All signatures	2.4	13
1-handed signatures	2.4	4

Table 7.3: Threshold and EER (%) for detection error rates.

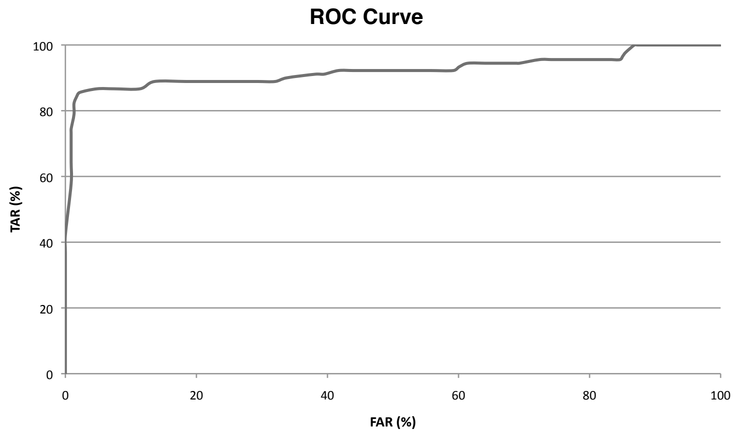

Figure 7.7: ROC curve for magnetic gestural authentication system.

Original & Forged Signature Duration

To further establish the difference between original signatures defined and forged signatures, the mean duration of the attacks on easy and difficult signatures were compared with the mean duration of the corresponding original signature. Means, standard deviations, and one-sample t test statistics of easy and difficult signatures for both original and forged signatures are summarized in Table 7.4. The one-sample t test (where the T value was set as the original mean signature duration for each signature respectively) revealed no significance differences for all but the second easy signature. This shows that attackers were on average able to closely reproduce the duration of the attacked signatures, except for a single signature. However as the forgery results (Section 7.7.1) showed, attackers were largely unable to successfully break into the system.

7.7.2 Usability & User Experience

Gestural Signature Duration

To assess whether the durations of recalled signatures were similar to the original signatures made, the mean duration of signatures were analyzed for the 20 participants in the usability

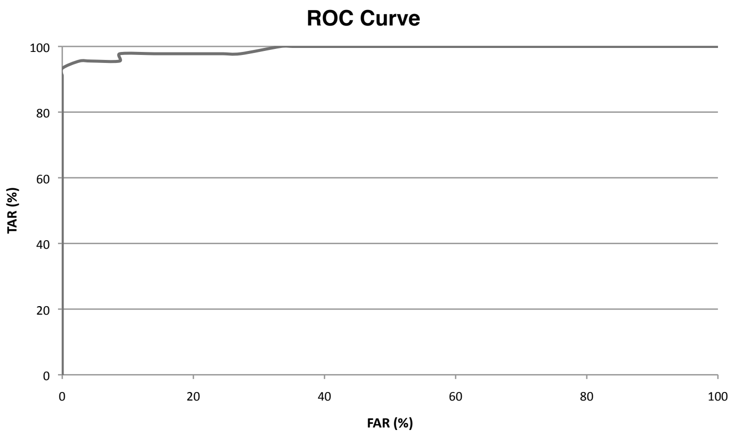

Figure 7.8: ROC curve for magnetic gestural authentication system based on two (1-handed) signatures only.

Signature	**Mean Original**	SD	**Mean Forgery**	SD	***One-sample T-test***
Easy$_1$	3412	211	3253	557	t(18)= -1.2, p= .23
Easy$_2$	2884	277	3782	708	t(18)= 5.5, **p= .00**
Difficult$_1$	4417	497	4593	1574	t(18)= .5, p= .63
Difficult$_2$	2475	527	3382	1235	t(18)= 3.2, p= .01

Table 7.4: Mean duration (m/s), standard deviations, and one-sample t test statistics of original signatures based on 5 templates each and mean durations and standard deviations of forged signatures based on 3 attacks across forgers (N=19).

study. Despite that for security analysis 5 participants were removed due to inconsistent signature templates, here we chose to analyze the mean durations of defined and recalled signatures for all participants. This was done to ensure that the duration analysis paralleled real-world situations, where a signature may differ in shape and movement, but still comply with the overall duration across different recorded templates. A paired-samples t-test was run, however revealed no significant differences between mean durations of original and recalled signatures. These findings are summarized in Table 7.5. This is in line with our expectations, where signatures that belong to participants are not easily forgotten, at least with respect to duration of the performed signature. To further assess the user experience of this authentication method, participants in the exit questionnaire for the usability study where asked whether they perceived magnetic gestural authentication to be a fast method for user authentication. Participants were quite positive, stating that this is indeed a fast method for authentication (Md=5.5, IQR=3-6).

Signature	M	SD	***Paired-samples T-test***
Original	3329	1379	t(39) = 2.1, p = .04
Recall	3143	1232	

Table 7.5: Mean duration (m/s) of original 5 signature templates made by users (N=20) and later recall duration (based on 3 templates) with paired-samples t test statistic.

System Usability Scale Responses

Measured SUS responses across participants (N=20) were calculated according to Brooke (1996), and analyzed in terms of average score frequency distributions. Results are shown in Fig. 7.9. Only few participants (6/20) gave a score greater than 70, which indicates that the tested magnetic gestural authentication prototype is not yet ready for use in the consumer market. This is not surprising, given that the system was still a prototype (involving bulky and light foam models with embedded SHAKE sensors and a complicated toolkit interface for recording gestures on a PC). Nevertheless, it does reveal that if this kind of authentication system is to be perceived as usable, it will have to provide (as discussed in Section 7.9) for smoother interaction without burdening the user with the details of template recording and allow for actual mobile devices to be used for recording magnetic signals instead of foam models.

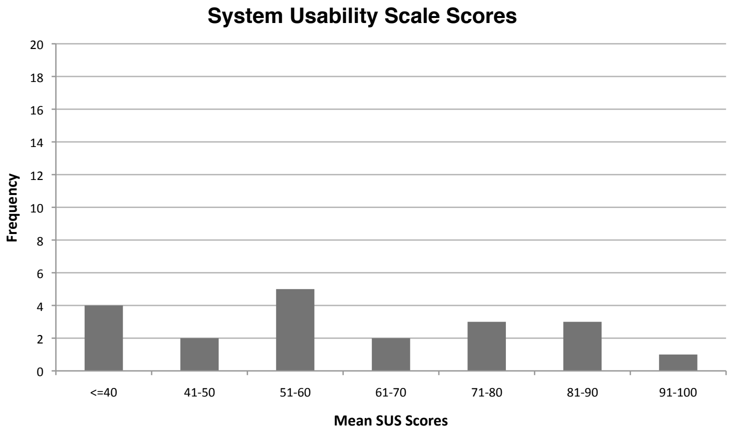

Figure 7.9: Frequency distribution of mean System Usability Scale responses across participants (N= 20) to the magnetic gestural authentication system.

NASA-TLX Responses

To investigate the overall subjective workload incurred on participants using the magnetic gestural authentication system, the NASA-TLX questionnaire (Hart and Stavenland, 1988) was administered after participants defined their original signature. Mean responses and confidence intervals are shown in Fig. 7.10. From these results, it can be seen that the mean Subjective Workload is 6.7, which provides evidence that this authentication method does not impose high subjective workload on participants. This provides additional support to our hypothesis that magnetic gestural authentication provides a natural and fast method of user authentication.

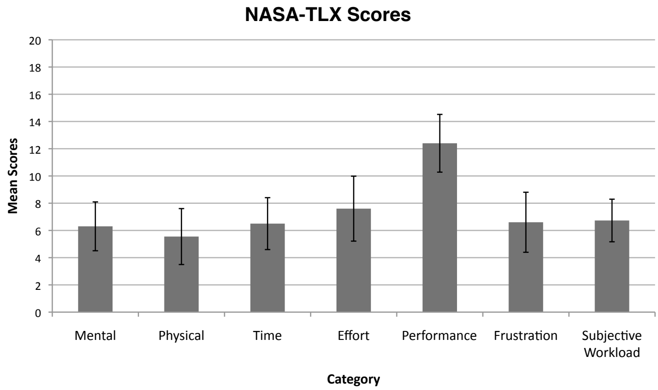

Figure 7.10: Mean NASA-TLX workload responses [range 0-20] across participants (N= 20) to the magnetic gestural authentication system. Capped error bars represent 95% confidence intervals.

System Trust Scale Responses

To investigate whether participants trusted the magnetic gestural authentication system they interacted with, we administered the System Trust Scale (Jian et al., 2000) after participants defined their signature. Participant responses were split into separate categories for trust and distrust of the system, where responses on both categories followed a normal distribution (Shapiro-Wilk test for Trust ($t(20) = .97$; $p = .80$) and Distrust ($t(20) = .97$, $p = .82$)). A paired-samples t-test was conducted to assess the differences in mean scores between perceived trust and distrust, however no significant differences were found. Means, standard deviations, and paired-samples t-test statistic are shown in Table 7.6.

Analyzing the responses in terms of number of participants who trusted the system versus those who distrusted the system, we found that half of participants (10/20) trusted the system. As is common with novel user authentication schemes (cf., Ben-Asher et al. (2011), it takes time as well evidence for an authentication system to gain the trust of users. In our case, it seems that only half of our participant sample trusted the system. This makes it difficult to state clearly at this early stage of the system whether participants would truly trust the system in the future, however our security results suggest that if participants were informed about the actual security of the system, an increase in system trust over time seems likely.

STS Category	M	SD	*Paired-samples T-test*
Trust	3.9	1	t(19) = .5, p = .15
Distrust	3.4	.7	

Table 7.6: Means, standard deviations, and paired-samples t-test statistics for Trust and Distrust categories of the System trust Scale responses across participants (N=20) to the magnetic gestural authentication system.

7.7.3 Users' Subjective Feedback

Overall Gestural Authentication System Acceptance

To assess the overall acceptance of the proposed magnetic gestural authentication system, users' subjective feedback was gathered via Likert-scale questionnaire responses and interview responses. This data was collected for both participants in the usability study (N=20) and the security study (N=20).

Usability Study Participant Responses

For the usability study, when asked whether they would use this gestural authentication method on a daily basis (e.g., to unlock their smartphone), participants found this method to be quite suitable for daily use (Md=5, IQR=3-6). When asked whether this method was secure enough for their devices, participants responded positively (Md=4, IQR=3-5.25). However, when asked whether this gestural authentication method was better than PINs or passwords, participants did not think so (Md=2.5, IQR=2-5.25). With respect to whether they found the defined air signatures easy to recall, participants were confident that these signatures are easy to recall (Md=6, IQR=4.75-7). When asked whether making air signatures caused them any fatigue, participants did not find these air signatures tiring (Md=5.5, IQR=3-6.25). Together, these questionnaire responses highlight that while the magnetic gestural authentication system is usable, and is perceived as an adequate scheme to secure devices, it is still perceived as less secure than traditional security methods such as PINs and passwords.

After participants completed the usability study, they were asked for their own subjective feedback concerning the presented gestural authentication system. When asked about their overall impression concerning security access using magnet-based air signatures, very few participants were positive (3/20), nearly half were not sure (9/20), with the remaining participants negative about the security such a system provides (8/20). Amongst those who were positive, one participant brought up the point that one needs to recall only one signature, akin to having one handwritten signature associated with one's bank account (*P7: "It's secure because normally a person uses only one signature, for example in the bank for your bank account. Yes, I find it secure."*).

Amongst those who were not sure about the system, they raised valid concerns, including whether or not the system would be able to recognize their own signature later if they forgot it (*P3: "Sometimes I wasn't sure what I was writing, so if it is to replace passwords, I wonder whether my device can recognize my own signature."*), whether performing their signature in a public place is safe (*P14: "My problem is, if someone is looking at you, and your signature is not complex enough, and you do it in public, then someone can steal it."*), whereas others stated they need study results to determine whether this method is secure enough (*P15: "I'd have to wait and see study results to see how secure it was."*). With respect to the first concern, performing air signatures is not very different than performing a handwritten ink-based signature, where with sufficient practice (and recorded signature templates), the system would be able to recognize their signature. Concerning the second main concern raised, the performed signature gesture would have to be small and perhaps performed just slightly above the palm of one's hand or the smartphone so as to provide cover from others who might be watching. This point will be elaborated on in the discussion (Section 7.8). For the last concern raised, it is not surprising that people need strong evidence about the security of a given authentication scheme before trusting it (which is where the present work fares into), especially given the already widespread adoption of password and PIN-based authentication methods (which have a low bar of acceptance due to their ease of use).

Participants who were negative about this authentication system shared similar concerns to those were unsure, including the fear that they may forget their own signatures (*P8: "I think I wouldn't be able to get into my own phone again, because I would forget it."*) but also the belief that PINs and passwords are already secure enough (*P5: "I would still prefer to use PIN combinations, because I think it is secure enough already."*). For the latter concern, it appears that this notion that PINs and passwords are secure enough stems from the ubiquity of this authentication scheme, however as mentioned before (see Section 7.3.1), this is not only a misconception (Morris and Thompson, 1979), but also is contrary to the findings by (Ben-Asher et al., 2011) who found that participants generally do not perceive PIN-based entry methods to be a secure method of mobile authentication.

Security Study Participant Responses

In the security study, when asked whether gestural magnetic signatures provided strong security for their devices, participants were again positive about this authentication method (Md=5, IQR=4-6). In the post-session interviews, participants were again asked about their overall impression concerning security access using magnet-based air signatures. By contrast to the usability study, here most participants (14/20) found this method to be secure, a few (4/20) who were unsure, and the remaining (2/20) negative about this authentication method. The high discrepancy between the responses in this study likely stems from the fact that in the security study, participants were asked to forge other people's signatures, which was apparently a difficult task.

Those who had an overall positive impression found this to be a secure method despite knowing that someone may have full video access to the signature under attack (*P1: "I think it provides good security, if you do something really difficult, then it is secure. Even if you record it, it is still secure."*). At the same time, the issue of security can also depend on which device one uses this authentication scheme for, and how many other additional layers of security there are besides air signatures (*P3: "For a business setting, like my business phone, it is fine. For ATMs, I'm not sure, but if research says it is quite secure, then okay. Probably a good idea to have it besides the PIN."*). From those few that were negative, the main reason was a general distrust of this kind of system and the nature of the interaction method, where the performed signature may take too long to perform (*P8: "I don't like the system. I think it is very useless. Lots of time to do this."*).

The positive impressions given by participants in this study notwithstanding, participant responses were not as positive when asked directly whether they thought this authentication method is more secure than established methods like PINs and (graphical) passwords. Here, less than half of participants (6/20) found magnetic gestural authentication more secure, a few (4/20) stated clearly that it is less secure, where the rest (10/20) stated either that the security level was the same or were unsure (*P5: "I'm generally skeptical to new things, probably just as safe or unsafe as the paper signatures."*). Finally, those participants were asked whether they themselves would use this authentication method in their everyday lives (for both smartphone access and ATMs) if this method was shown to be secure. Here, less than half (7/20) stated they would in general (*P4: "If it was safe enough, I would use it. I like the idea, I like the concept."*), a few (4/20) said they would not (*P10: "For myself, I don't think so. At this point, no. Maybe if I tried to do my own elaborated signature, then maybe yeah."*), and the rest (9/20) stated they would use this method however it is contingent on which device they are using this security protocol for (*P2: "Not for daily use. But for ATMs, I would use it."*; *P9: "For a mobile yes. But not for my bank account."*). The latter statements concerning using this method depending on what device the authentication scheme is used for highlights the usability and security tradeoff, where as stated earlier, participants find entering a 4-digit PIN easier and quicker than performing a signature in the air, but at the cost of breakable security.

Perceived Forgery Difficulty

In the security study, participants were given several Likert-scale questions concerning the difficulty of forging air signatures and using this authentication method. In general, participants found forging any air signature to be difficult (Md=5, IQR=2-3.25) and certainly more difficult to forge than handwritten wet ink-based signatures. With respect to the different signatures tested, participants found it difficult to forge the easy signature (Md=4, IQR=3-4.25) and very difficult to forge the difficult signature (Md=6, IQR=6-7). Participants also agreed that both forging easy signatures (Md=4.5, IQR=4-6) and difficult signatures (Md=6, IQR=3-6) were more difficult to forge than forging handwritten signatures.

With respect to the difficulty of following hand movements of users performing a signature in the videos, participants were neutral concerning the easy signatures (Md=3, IQR=2-4) but found it very difficult to follow hand movements of the difficult signatures being performed (Md=6.5, IQR=4.75-7). While these results provide strong evidence concerning the level of perceived forgery difficulty, the vulnerability of a given signature does depend on how simple a given signature is (*P6: "Depends if the movements are complicated, then it is very difficult to forge it."*). To gain insight into which video angles were most helpful for participants in their forgery attempts, the back (posterior) view was mentioned to be most helpful (17x), followed by the frontal (anterior) view (9x), the left (3x) and right view (2x).

Gesturing Using Magnets

To investigate how participants in the usability study perceived the gestural interaction afforded by using magnets, they were asked about their attitudes towards gesturing using a magnet, carrying a magnet, and paying for a personal magnet. Participants found the gestural interaction using magnets to be intuitive (Md=4, IQR=3-5), which lends support that this form of interaction is easy to learn and use (*P2: "For me, it's very intuitive, I think it is very easy to learn and use."*). Nevertheless, one concern raised was fine motor coordination in certain situations such as when a person is intoxicated (*P16: "I found the magnets very easy to use. Only thing I worry about is how well it would work if you're drunk, that could be an issue. Just your coordination, goes out the window when you're drunk."*). This is a valid concern, and indeed unlike a PIN code or password where the code can be stored away as a backup plan for such situations, authentication using magnetic signatures requires a replication of a predefined level of human motor performance. Nevertheless, with sufficient practice of one's signature, this problem may be alleviated. Such cases however do point back to the finding raised earlier (see Section 7.7.3) that there should always be a backup plan in case one forgets his own signature (discussed in Section 7.9).

When the usability study participants were asked about whether they would be willing to carry a magnet with them, responses were in general positive (Md=3.5, IQR=2-6), however with clear exceptions expressed by some participants (e.g., *P17: "It was alright [the experiment session], but I wouldn't carry a magnet with me."*). When participants were asked whether they would be willing to pay for a good magnet to ensure that authentication proceeds smoothly (i.e., without any potential weak signals resulting from a poor choice of magnet), participant responses were neutral (Md=3, IQR=2-6). This is not unexpected, as it does bring to question why participants would pay for an additional accessory when existing security methods such as PINs and passwords do not require them. However, this same question was posed to participants in the security study, and their responses were slightly more positive concerning paying for a magnet (Md=3.5, IQR=2.75-5.25). This difference might reflect the fact that participants faced with the forgery situation came to appreciate that they were provided with the very same magnet that the attacked users in the video were using. In such a case, paying for a personalized magnet would indeed provide an additional layer of security. In other words, just as one pays for an extra pair of house keys that one carries around, likewise can be said of magnets for security access. This issue will be elaborated on below (Section 7.7.3).

Form Factors

Magnet Form Factor: Given that the proposed authentication method requires making use of 3D gestures to perform a signature, we expected the form factor of both the magnet and the SHAKE foam model to have an influence on users' preferences and attitudes towards acceptance of this mode of interaction. To investigate this, participants were asked about whether they faced any difficulties performing the signature gestures, and whether the SHAKE foam model was a well designed probe into the everyday use of this method.

Participants had no problems at all performing the gestures for their defined signatures (Md=6, IQR=4.75-7), however their level of comfort in performing them was borderline acceptable (Md=3.5, IQR=2-6). A few participants mentioned explicitly that they did not like the bar-shaped magnet, because it was too small and can easily slip from the fingers (*P10: "Well, I like the experiment. But the magnet is quite uncomfortable to hold, it was quite slippery. If it were more ergonomic, and had a different shape, then it would be good. Definitely better than entering a password."*). Additionally, while others had a problem with the size and thickness of the magnet (*P9: "I was not too comfortable with the magnet and device, because magnet was too thick. I prefer paper based, because it is finer, the* *shape is finer."*), still others requested a bigger sized magnet (*P11: "Also you need a bigger magnet. If it is more like a pen, would feel more natural."*).

These forgoing concerns raise the importance of ensuring an optimal balance between the usability of a given tool (such as a magnet) with a particular form factor and its ergonomic design to ensure comfort in performing air signatures. Additionally, this brings up the issue whether users should in fact be provided with a personalized magnet to ensure greater security (e.g., *P5: "If you have a personalized magnet, then it would be quite secure."*). However, if users are provided with a personalized magnet, then the risk of losing this magnet would require additional effort from the user to regain the exact same method, where an additional protocol to allow this has to be implemented (e.g., by banks or smartphone and laptop manufacturers).

SHAKE Foam Model Form Factor

To gain insight into how the participants in the usability study found the form factor of the SHAKE foam model, we explicitly asked participants about this in the exit interview. More than half (13/20) of participants found the designed foam model acceptable as a probe into the actual use of this authentication method (*P5: "The foam model was realistic. At one point I imagined it would be a smartphone."*), while the rest (7/20) found it too bulky and/or too light (*P12: "Maybe it was a little too big. It was also too light."*). Also, one participant mentioned explicitly that he did not like the button on the SHAKE sensor (*P11: "I guess it is okay, but I didn't like the button, too pointy."*).

While the height and width factor of the SHAKE model was designed with the same dimensions as an iPhone 4®, the thickness (depth) of the model was dictated by the thickness of the SHAKE sensor used. Nevertheless, even those participants who found it too thick or too light conceded that they served as believable surrogates for smartphones (*P9: "Shape is too big, it is long and thick. But it was fine."*), which meant that we did have external validity, given the prototype stage this authentication system is currently at.

Social Acceptability

The aim of an authentication system is to allow users access to the secured system or device, at any place or time. However, for the proposed gestural authentication system, there is a performative aspect where making gestures in public spaces may bring to question social acceptability issues in performing 3D signatures in the air while out in public. To gain an idea of how the users in the usability study perceived such situations, they were asked whether they would feel awkward performing air signatures in public places (e.g., supermarket, mall, etc.). While participants found this mode of interaction to be very socially acceptable (Md=5.5, IQR=3-6), this may boil down to individual differences (e.g., *P16: "I think it can be kind of embarrassing if you stand in front of ATMs and write your signature in the air."*).

Moreover, as was shown in previous work (Rico and Brewster, 2010), the willingness to perform a given gesture in public places depends on both the given context and which gesture is being performed. Also, if a given signature gesture is not recognized, and the user has to repeat it, this may also cause reluctance in performing the gesture in a public setting (El Ali et al., 2012). These social acceptability results, which were obtained through a controlled setting (discussed in Section 7.8), provide an early indication that the social performative aspect of magnetic gestural authentication would have to be further researched longitudinally if a system based on gestural interaction is to be adopted for cultural and commercial use.

7.8 Discussion

There are four potential limitations to the present study. First, in the current magnetic gestural authentication prototype, participants were not given feedback on their performance when both defining their own signature, recalling their own signature, or upon attempted forgery attacks. This was due to the way the data was analyzed, where the magnetometer signal data was processed offline. This limitation was raised by one participant in the security study (*P4: "I mean it is hard to tell about my performance, because I don't know if I made it. No feedback, no percentage of success."*).

Indeed, in a real system, the difference between a performed signature and the recorded templates this signature should match would be visualized to the user in some representation (e.g., a percentage bar indicating a confidence score of how similar the performed signature was to the original). However, if one considers established methods of authentication such as PINs or passwords, then users of these authentication schemes are likewise not given feedback on their accuracy of an entered PIN or password. An exception to this are paper-based handwritten signatures, where for example in some situations, the employees at a person's bank may choose to disclose that person's original signature and ask him to prove that he is able to perform that very same signature again in front of them. Since showing the person's signature is not an option in the case of the proposed authentication system (due to security issues), a visualization of the difference between the performed and original signature would appear to be the most effective method to handle feedback on user performance.

Second, in the usability study, while we investigated the performance of participants to recall their own defined signature, we have only tested short term memory recall (Tulving, 2002, 1993). As has been shown in the cognitive sciences, there are different kinds of episodic memory, including working memory, short-term memory, and long-term memory (Baddeley, 2003; Ericsson and Kintsch, 1995; Baddeley et al., 1974). To assess whether this kind of authentication system can be used for longer periods of time, one would have to investigate whether participants are able to recall their own signature across days, weeks, and possibly months. While this kind of memory test for recall signatures is beyond the scope of the current study, this issue was raised by participants who expressed doubts whether they would be able to access their own device if they were to forget their own signature. In such situations, the recommendation to include a backup authentication scheme (e.g., a security question or PIN) in case one forgets his own signature should be implemented. This will be discussed further in Section 7.9.

Third, in this study, we did not make a direct security comparison between the proposed gestural authentication system and other security methods such as PINs, passwords, graphical passwords (Biddle et al., 2012), keystroke dynamics (Shanmugapriya and Padmavathi, 2009; Monrose and Rubin, 2000), or touch movement dynamics (Sae-Bae et al., 2012). While we did ask participants in the questionnaires and interviews whether they perceived this method to be better or worse than established methods, the comparisons were limited to what participants were already familiar with (namely, PINs and passwords), and not more advanced (gestural) security methods currently being developed and tested (e.g., behavioral biometric methods based on keystroke pattern dynamics (Shanmugapriya and Padmavathi, 2009)). Nevertheless, in this study we demonstrated the security level that the proposed system provides, as well as the usability issues associated with performing air signatures using a magnet, which we believe provides a sufficient initial assessment of this novel authentication method.

Fourth, while our goal was to investigate the usability and security tradeoff in magnetic gestural authentication across all devices, the current stage of the prototype (in using the SHAKE sensor) was geared towards smartphone security access. In other words, the design of our study was restricted to form factors associated with mobile (smartphone) devices. Nevertheless, we took care in our questionnaire and interview questions to ensure that participants were able to identify that the tested authentication system is applicable to other devices, including laptops, desktops, or ATMs.

7.9 Design Recommendations

In this chapter, our goal was to answer our research question of whether magnetic gestural authentication is both a usable and secure method of authentication for different contexts (e.g., daily use on mobile devices or for security access at bank institutions). In line with our first hypothesis and previous work, our results showed that the gestural authentication system is indeed secure against visual-based forgery attacks. Additionally, our results showed that this system also allows authentication of eligible users, especially when it makes use of a person's biometric signature (which may be transferred from wet ink-based signatures). With respect to our second hypothesis, our usability results showed that this method of authentication, in allowing for natural 3D gestures, provides a good user experience for participants (namely, low subjective workload, natural gestural interaction, easy to recall biometric signatures). However, given the early stage of the tested authentication prototype, a number of design considerations arose that can aid improving the usability, acceptance and eventual adoption of this kind of security system for personal and institutional use:

1. **Ensuring Transparency of Security Results**: As is common in the introduction of novel authentication systems (Weir et al., 2010), it will take time for users to fully trust the system (cf., Section 7.7.3 & 7.7.2). In order to gain user trust and acceptance, empirical data demonstrating the security of the system should be readily available.

2. **Best Practice Guidelines on Usable and Secure Signatures**: While participants mentioned that making these air signatures was quick, this does depend on the choice of a given signature (cf., Section 7.7.3 & 7.7.3). Currently, there are no best practices for what constitutes a secure signature – however for this system (insofar as the DTW algorithm is used to match signature templates), a long and complicated signature is an indicator for how secure a signature is. In such cases, this may come at the cost of speed of performing a signature (cf., speed accuracy tradeoff (Wickelgren, 1977)), which may negatively impact the usability of the system. Therefore,

arriving at a minimum signature duration could help optimize the balance between security and usability. Furthermore, given the drop from a 13% to 4% EER in our ROC analysis when we factored two handed signatures out provides further evidence on the tradeoff between a usable (specifically easy to recall) signature and a secure signature. Finally, given participants' concern about public stealing of their signature, guidelines on making a signature in a clandestine manner should be provided (e.g., signing on the palm of one's hand so as to provide cover from the prying eyes of nearby strangers).

3. **Improving Usability vs. Workload**: While the usability scores on the SUS were low, using the system did not incur high subjective workload for participants (cf., Section 7.7.2 & 7.7.2). As the proposed system is still in the prototype stage, it is not surprising that the general usability of the system is low. Nevertheless, despite this early stage of the system, our NASA-TLX workload results gave clear indications that performing air signatures does not provide high subjective workload, and when additionally considering users' subjective responses related to gesturing using magnets (Section 7.7.3), this adds to the body of evidence that this authentication scheme is a natural and easy to use authentication method. Therefore, usability of the system, not subjective workload issues, should be addressed.

4. **Complement Standard Authentication Methods**: As was mentioned by some participants, this kind of authentication method would strongly benefit (at least in the earlier stages of use) to be used alongside standard methods (cf., Sections 7.7.3 & 7.7.3). This is to ensure that a backup plan is available in case one forgets his own signature, in addition to providing an additional layer of security.

5. **Designing for Form Factor**: Form factors of the mobile device and magnet appear to play an important role in adoption of this new security method (cf., Section 7.7.3). While some participants found the shape and size of the magnet and foam model to be acceptable, others did not. This raises the issue of whether personalized magnets, in providing more security, should be provided.

6. **Restricting Context of Use Can Facilitate Adoption**: While gesturing using a magnet was perceived to be intuitive, users may not always want to carry a magnet around (cf., Section 7.7.3). This can be solved by limiting the use cases and contexts of use in which magnets are used (e.g., gestural authentication only for security access at bank institutions), or embedding magnets in devices. Relatedly, standardized magnets should perhaps be readily available for purchase, and personalized magnets (if provided) should additionally be provided on demand by the service provider (e.g., mobile device manufacturer or bank institute).

7. **Longitudinal Analysis Requirement For Social Acceptability**: Our early social acceptability results provide initial clues that this method of authentication may indeed be socially acceptable (cf., Section 7.7.3). However, to conclusively state this would require further longitudinal research in users' natural settings.

7.10 Conclusions

In this chapter, we have presented a user authentication method based on magnetic gestural interaction. In line with our hypothesis and prior work, our results showed that this authentication method is secure against visual-based shoulder surfing attacks. With respect to usability, participants found the authentication method to be quite natural, easy to recall biometric signatures, and providing overall low subjective workload. However, the current stage of the system raised doubts about the general usability (as measured by the SUS) and trust of the system. From these results, design considerations were derived that should serve as a starting point if this kind of authentication method is to be accepted as a standard authentication method on today's smartphones, but also for ATMs and desktops/laptops. Taken together, we hope that our findings have provided a solid overview of the security and usability tradeoff in magnetic gestural authentication, and can guide future authentication methods that draw on the natural gestural mode of interaction afforded by (magnet-based) ADI.

The three chapters in Part II of this thesis showed that gestural input techniques, like context-aware solutions, can contribute to making user interactions simpler and provide a positive user experience across different domains; namely, task-independent, playfulness, and user authentication. In the following chapter (Chapter 8), we provide a summary of the research carried out in this thesis, discuss how this ties into minimal mobile HCI, and provide future directions.

8 Conclusions

This thesis set out to investigate the usability and user experience (UX) issues faced when designing mobile interactions that require minimal visual attention from users when they interact with smartphones, namely *minimal* mobile human computer interaction. The two parts of this thesis address two main themes:

- Context aware solutions to improve usability and UX when users interact with smartphones (Part I)
- The usability and UX of using gestural input techniques alongside screen-based interaction (Part II)

For achieving minimal mobile interaction using context-aware solutions (Part I), we investigated the usability and user experience issues associated with using location-aware multimedia messaging (*LMM Study* and *Playful LMM Study*) and exploration-based route planners (*Route Planner Study*) to support urban exploration and playfulness. For achieving minimal mobile interaction using gestural input techniques (Part II), we investigated the usability and user experience issues associated with discrete mobile 3D gestural input in a task-independent setting (*Gesture Errors Study*), and used 3D gestural interaction to support playfulness (*Playful Gestural Interaction Study*) and user authentication (*Gestural Authentication Study*). Below, we revisit the research questions we raised in Section 1.2.2, and summarize the findings.

8.1 Summary

Context-aware solutions, drawing from context-aware computing research, have the goal of making mobile devices sense, learn from, and adapt to the user and his/her changing context. In addition, non-visual interaction techniques, drawing from multimodal interaction research, allow users to interact with a system through non-visual means, such as gestures. In both, the goal is to free users' attentional resources. In this thesis, we show that both approaches are suitable for designing and developing usable minimal mobile interactions.

8.1.1 Context-awareness

In Part I of this thesis, we began our investigation into minimal mobile interaction by first focusing on the contextual factors associated with location-based media production and consumption (Chapter 2). This allowed us to understand the interaction issues users face in an urban environment. We asked the following question:

> **RQ 1:** How do users create and interact with urban location-based content, and how should this impact the design of future Location-aware Multimedia Messaging (LMM) systems?

To answer this, we conducted a pilot ethnographic study with participants using an LMM system, and followed up with another longitudinal study using a multimodal diary method (*LMM Study*). In the pilot study, while users positively perceived the LMM prototype, they stated explicitly that they needed to spend more time with the system to test its usefulness. In this followup multimodal diary study, we found that users were indeed willing to use a system that supports creating and consuming location-based content, specifically to document their daily experiences. Upon analysis of the content they created, we found that the most common tasks our users were interested in were reporting activities (a form of citizen journalism using mobile devices (Kramer et al., 2008)) and expressing appreciation over their surroundings, for primarily entertainment and aesthetic purposes. This gave us data into how users, if they were to use such messaging systems, would use the supported functionality. Our analysis, based on an episodic memory framework grounded in Cognitive Science, provided us with insight into the contextual factors governing what, where, how and why people create and consume location-based user generated content (see Section 2.7). This provided the initial groundwork for the need of minimal mobile interactions, given the high information load that is expected to be incurred on users when viewing these messages in urban settings.

Looking more closely at the user interaction with our LMM prototype, we see that it provided a suitable initial use case for investigating minimal mobile interaction. This was because the multimedia messages users created were automatically anchored to the location at which they were created. This allowed for a more restricted, but simpler, form of interaction with user generated multimedia content. Additionally, the Augmented Reality (AR) presentation of messages made interaction easier for users. This was achieved through the designed input techniques, which involved both touchscreen and gestural movement of the smartphone to view messages. Together, the findings point to the importance of designing systems that do not require heavy interaction from users. This allowed us to provide concrete design recommendations (see Section 2.8) for the functionality and interaction methods that future LMM systems should support. Additionally, it allowed us to arrive at the concept of minimal mobile interaction, and how it can ease processing costs from users.

While we found that the ease of use of the LMM prototype was positively perceived by users (particularly when in an urban setting where user attentional resources are limited), our users also found the system too limited. The primary limitation for our users was that using the system was no longer enjoyable and fun to play with after the initial phases of interaction. Given that our users valued that this system should be more engaging and fun, it posed a challenge: how our LMM prototype, with its designed minimal interaction, can better support playful interactions among users. This lead to our second research question:

> **RQ 2:** How can location-aware multimedia messaging (LMM) systems be used to support playful urban interactions?

Using LMM systems as a case study, we answered this question (in Chapter 3) by presenting an analysis of how fun or playfulness can be studied and designed for in urban environments (*Playful LMM Study*). Drawing on the findings of Chapter 2 and an envisioned tourism-based scenario illustrating how LMM can be used (see Section 3.1), we discussed in detail what playful experiences are and three problems that arise in realizing the scenario:

1. How playful experiences can be inferred (the inference problem)
2. How interacting with the system can be motivated and maintained (the maintenance problem)
3. How playful experiences can be measured (the measurement problem)

We responded to each of these problems by distilling three design considerations (see Section 3.7) for playful, minimal mobile HCI:

1. User experiences can be approached as information-rich representations or as arising from the interaction between a user and system
2. Incentive mechanisms can be mediators of fun and engagement under minimal mobile interaction settings
3. Measuring experiences requires a balance in the choice of testing methodology

For playful experiences in mobile interaction, we provided guidelines that facilitate system designers and developers in integrating playful elements into systems that may not be otherwise initially designed to support playful interactions. In this sense, the experience we gained in running multiple user studies investigating LMM systems that require minimal interaction from users, has allowed us to identify the necessary ingredients for a system to be playful for users. While we did not run another user test iteration on the LMM prototype, we revisited some of these issues in Chapter 6.

While in the initial two studies (Chapters 2 and 3) we focused on the overall user experience and elicited playfulness of multimedia messaging behavior at urban locations, these studies also showed that urban interactions are connected across locations. To study this (Chapter 4), we designed a system to allow pedestrians to explore a city's different locations (*Route Planner Study*). To maintain the requirement of designing minimal mobile interactions, we wanted to additionally avoid burdening users in supplying lengthy user preferences. In this chapter, we asked:

> **RQ 3:** How can we automatically generate routes to support pedestrians in exploring a city?

To fulfill the minimal mobile interaction requirement, we made use of a smartphone's context-aware capabilities (in this case, location sensing using GPS). To do this, we made use of geotagged data provided by the photo sharing website Flickr.[1] Geotagged photos provided a unique window into city photographers' experiences, under the assumption that

[1] http://www.flickr.com/; last retrieved: 01-08-2013

there must have been something of interest for a photographer to take a photo at a given location. By borrowing sequence alignment methods from bioinformatics, we designed, built, and evaluated a simple context-aware city route planner to facilitate city residents and visitors in exploring a city (where we used the city of Amsterdam, the Netherlands, as a testbed for experimentation). Applying sequence alignment techniques on 5 years of geotagged photos allowed us to create walkable city routes based on paths traversed by multiple photographers (see Section 4.4). Understanding movement behavior of city pedestrians allowed us to generate off-beat exploration-based routes. We believe our work opens up opportunities for future data science researchers to make use of quick sequence alignment methods for analyzing geotagged data, despite that the intended purpose of these methods applies to a different domain (namely, aligning protein and DNA sequences). Furthermore, our work adds to the body of methods (e.g., Hidden Markov Models) that can account for the sequential character of human behavior, especially at a coarse level as is the case with GPS coordinates.

To test whether our generated routes would be desired by users, we conducted a user study with Amsterdam residents to compare our routes with the most efficient (i.e., shortest distance between two locations) and popular route variations (where we defined popular as the volume of geotagged activity at a location in a given a time period). We collected experience questionnaire data, web survey responses, and user interviews. Our results showed that our generated routes (which were based on aligning sequences of the locations where photographers took photos at in a given time period) were perceived as indeed more stimulating and more suitable for city exploration than the efficient and popular route variations. Moreover, while digital aids based on photographer paths can potentially aid city exploration, we found that their acceptance in mainstream route planners depends on their visualization. These results provided us with insights into which digital information aids users would like to make use of when exploring a city (see Section 4.5.3), even under a minimal mobile interaction setting that relies on a data-driven approach. In short, our user-driven findings into what kinds of digital aids best support users' city exploration needs can be used to improve future digital information aids for city exploration.

We believe future researchers and designers can now consider a currently unused information type obtained from geotagged images to guide exploration-based route planning in a city, namely, the number of city photographers that took a given route segment over a certain time period. Furthermore, we show that it is possible to leverage social geotagged data to cater for the hard problem of automatically generating exploration-based route plans, and that these route plans are desired by users.

8.1.2 Gestural Input Techniques

In Part II of this thesis, we investigated how non-visual input techniques such as 3D gestural input can aid minimal mobile interaction. We focused on the use of 3D gestural interaction in three different domains: Task-independent (Chapter 5), Playfulness (Chapter 6), and User Authentication (Chapter 7). Below, we revisit the research questions we posed for each of these chapters.

We began our investigation by focusing on the usability and user experience issues associated with 3D mobile gestural interaction when recognition errors occur. This was necessary to establish whether users would be willing to adopt such minimal mobile interaction techniques even in the face of errors. In such situations, if a performed gesture does not get recognized, this would become even more frustrating for users, forcing them to switch back to standard touchscreen interactions (e.g., Karam and Schraefel (2005)). Moreover, this would aid in identifying which gesture sets are most robust to recognition errors. In Chapter 5, we posed the following question:

> **RQ 4:** What are the effects of unrecognized 3D gestures on user experience, and how do these affect the design of error-tolerant 3D gesture sets?

To answer this question, we ran a primarily qualitative study (*Gesture Errors Study*) to investigate how two sets of iconic gestures, mimetic and alphabetic, that vary in familiarity are affected under varying failed recognition error rates (0-20%, 20-40%, 40-60%). For our investigation, we developed and used an automated Wizard-of-Oz method to emulate 3D gesture recognition. As in Chapter 4, we made use of methods in a seemingly different field and applied them to evaluate the usability and UX of 3D gestural interaction. This allows interaction designers to quickly and easily test user frustration and tolerance to gesture recognition errors. Moreover, given the low cost and rapid development of this method, it provides a quick and cost effective means for mobile gesture designers to identify immediately which gesture sets provide the best user experience.

Drawing on experiment logs, video observations, participants feedback, and a subjective workload assessment questionnaire, our empirical results revealed two main findings (see Section 5.7): first, mimetic gestures tend to evolve into diverse variations (within the activities they mimic) under high error rates, while alphabet gestures tend to become more rigid and structured. Second, mimetic gestures were tolerated (with respect to performance duration and workload) under recognition error rates of up to 40%, while alphabet gestures by contrast tend to increase in duration and incur significant overall workload with up to only 20% error rates. These results show that while alphabet gestures are more robust to recognition errors with respect to gestural performance (i.e., their signature), mimetic gestures are more robust to recognition errors from a usability and user experience standpoint. This makes mimetic gestures better suited for inclusion into small, handheld devices that support 3D gestural interaction. While we have shown that mimetic gestures are better tolerated by users under recognition errors than alphabet gestures during 3D gesture-based interaction, our results also demonstrate the need to account for human factors when designing novel interaction techniques. Specifically, to account for errors when interacting with a system.

The findings in Chapter 5 provide support for using a minimal mobile interaction technique such as 3D gestural interaction. In our experimental setup, users were given feedback on a desktop display after performing a gesture, making our results generalizable for minimal mobile interaction, although not eyes-free interaction. Despite that users were willing to tolerate up to 40% errors, we had so far focused on the usability issues of 3D gestural interaction in a task-independent interaction. Looking at task-independent interaction was necessary to avoid any potential confounds with other variables (e.g., urgency of a task) that may influence gesture performance and preference. Given these promising findings, we revisited how this form of interaction can be applied in an actual domain. Given the problematic nature of supporting playful interactions discussed in Chapter 3, in our next investigation (Chapter 6) we revisited the domain of playful mobile interactions. In this next chapter, we asked:

RQ 5: How can 3D gestural interaction be used to support playful music composition and gaming on smartphones?

To answer our question, we made use of a recent interaction techniques in HCI known as magnet-based Around Device Interaction (ADI). This paradigm allows expanding the interaction space on mobile devices to allow 3D gesture interaction using magnets around the device. Since the interaction here is based on magnetic fields (which can pass through the hand or clothes, and not depending on users line of sight), the space at the back and side of device can also be efficiently used for interaction. This technique does not require extra sensors on current smartphones – it is only necessary to have a properly shaped magnet as an extra accessory.

For this study, we focused on the applied use of 3D gestural interaction (employing magnet-based ADI) in a playful, music-related context. This study (*Playful Gestural Interaction Study*) allowed us to look closely at how 3D gesture-based interaction can be synergistically coupled with minimal touchscreen interaction to facilitate playfulness. Using three musical applications developed under this magnet-based ADI paradigm (Air Disc-Jockey, Air Guitar, Air GuitaRhythm), we investigated whether this kind of minimal mobile interaction can be effectively used to support playful music composition and gaming on smartphones. To test this, we ran a controlled user study, where we collected usability and user experience questionnaire responses, users' direct feedback, and video observations. Our results showed that magnet-based ADI can be effectively used to create natural, playful and creative music interactions. From this, we distilled magnet-based ADI design recommendations to optimize the user experience for playful and creative music interactions in today's smartphones.

By revisiting the domain of playfulness in this applied ADI setting, some recurring issues surfaced. First, the inference problem (the difficulty in assessing whether or not a given interaction is playful; Section 3.4, p. 41) applies in this study, as we were able to infer playfulness from the responses participants gave in the questionnaires and interviews. However, it is difficult to quantify the amount of excitement and playfulness they experience without resorting to physiological measures, which were unavailable at the time, costly, and limited in what we can infer from the collected signals (van der Zwaag et al., 2010). Nevertheless, our mixed methods approach of assessing quantitatively through questionnaires and gathering subjective responses provided a lens into the desired user interaction. Second, the maintenance problem (the problem of ensuring a fun and enjoyable experience over time; Section 3.5, p. 43) applied to two of the music apps (Air DJ and Air Guitar), since in both there was no in-built incentive mechanism to continuously gauge users' interests. For the Air GuitaRhythm app, embedded gaming elements allowed even the non-musically trained user to get engaged and experience challenge and flow. This not only provides support to our design consideration of providing incentive mechanisms (Section 3.7, p. 28) presented in Chapter 3, but also underscores the need to keep aware of these issues when designing games in a minimal mobile interaction setting.

Despite that dealing with playful interactions poses a challenge, in this thesis we have dealt with it in two distinct contexts (namely playfulness of a context-aware system and how gestural input can support casual play). From this applied use of 3D gestural input, we provided design recommendations (Section 6.8) for interaction designers who wish to use 3D gesture-based interaction for creating playful user experiences in the music production domain, and for designing engaging 3D gesture-based games under a minimal mobile interaction setting where screen estate is limited (e.g., on a mobile screen).

In our final chapter (Chapter 7), we presented an additional example of how minimal mobile interaction can be used, by focusing on a common, daily mobile task: user authentication. Unlike play and enjoyability, authentication should be performed quickly and effortlessly without comprising security. In user authentication, a fundamental challenge is designing and implementing a security method that ensures that the method is both usable by users, while at the same time providing sufficiently strong security against any kind of adversarial attack. To stay in line with our minimal mobile interaction requirement, we also made use of magnet-based ADI, given that this 3D gestural input technique is quick to perform and achieves high recognition accuracy (Ketabdar et al., 2010b). Here, we looked at how 3D gestural input using magnet-based ADI can enable usable and secure user authentication on small, handheld devices. In this chapter, we asked:

> **RQ 6:** How does 3D gestural interaction affect the usability and security of mobile gestural user authentication?

While previous work provided an early assessment of the vulnerability of magnetic 3D gestural authentication under a video-based shoulder surfing attack scenario, the usability and user experience issues were not addressed. To answer our question, we verified the security of this method in a controlled lab study, and tested the usability and user experience issues in another controlled study (*Gestural Authentication Study*). Our results showed that this authentication method is indeed secure against video-based shoulder-surfing attacks. Given the novelty of this user authentication method (i.e., currently not used as a standard authentication scheme on smartphones), there is little data on whether this authentication method is secure enough for use in daily mobile interaction. To address this limitation, we largely replicated prior research (Sahami Shirazi et al., 2012), despite that it is uncommon in the HCI community to replicate previous research (see for example the CHI 2011 SIG panel RepliCHI (Wilson et al., 2011)). In our replicated experimental setup, we additionally accounted for a previously untested variable that could influence the security results, namely, accounting for variation amongst air signatures performed using this gestural authentication method. In so doing, we were able to retrace the steps and rerun the security analysis presented in prior work, and were able to verify the previous security results for this magnetic gestural authentication technique.

From a user experience standpoint, our participants found the method to be quite natural, with easy to recall gestural signatures, and providing low subjective workload. This allowed us to draw design recommendations (see Section 7.9) that should serve as a starting point if this kind of authentication method is to be accepted as a standard authentication scheme on small, handheld devices. This provides support that non-visual input techniques such as 3D gesture-based interaction can be synergistically coupled with screen-based interaction to improve the UX of interaction. Moreover, given that we dealt with both the security and usability and UX issues associated with this method that combined both quantitative (usability Likert-scale questionnaires, duration analysis) and qualitative methods (semi-structured interviews), we were able to arrive at a more complete account of how users perceive and interact with this authentication method. Another main finding here is that while users found 3D gestural authentication to be natural, easy to perform, and with high memorability for performed signatures, around half of our participants did not trust the security of the system. This was expected given that this form of authentication is not common on today's smartphone devices. To address this, one of our design recommendations suggests that interaction designers ensure transparency of the security results (e.g., through summarized fact sheets in the application) to gain user trust, especially for those users that are early adopters.

In this investigation, the larger lesson is that a combination of both security analysis and HCI methods to investigate the usability and user experience of a novel authentication method is necessary to have a better understanding of whether an authentication method would be adopted. In this sense, security researchers would benefit from adopting a mixed-methods HCI-driven approach to assess the usability and larger user experience of a given method.

8.2 Towards Minimal Mobile HCI

We started this thesis by looking at the current state in the fields of mobile HCI, context-aware computing, and multimodal interaction. We reviewed the benefits that eyes-free interaction can provide users, particularly when in a crowded, urban setting. There, we also highlighted that while there has been much work designing and developing such eyes-free interaction techniques, the usability and user experience (UX) issues associated with these techniques is an ongoing research effort across HCI communities. Specifically, we asked whether we can reduce user attentional costs associated with smartphone use, and ultimately improve the usability and user experience of such interaction. In Table 8.1, we provide a summary of all the research chapters and how they tie in with minimal mobile interaction, where we show the different systems and interaction techniques presented, their domain, input modalities, output modalities, and how minimal interaction is achieved.

System / Technique	Domain / Chapter	Input Modality	Output Modality	Minimal Interaction?
Location-aware Multimedia Messaging	Urban Exploration (Ch. 2), Playfulness (Ch. 3)	Touchscreen, 3D Gestures	Visual (Augmented Reality)	Context-awareness, Visuo-gestural Input
Exploration-based Route Planner	Urban Exploration (Ch. 4)	Touchscreen	Visual (Display)	Context-awareness
3D Gestural Interaction	Task-independent (Ch. 5)	3D gestures	Visual (Screen Display)	Gestural Input
Magnet-based ADI	Playfulness (Ch. 6), User Authentication (Ch. 7)	3D gestures, Touchscreen	Visual (Display), Non-speech Audio	Visuo-gestural Input, Gestural Input

Table 8.1: Summary of systems/interaction techniques and their domain, input modalities, output modalities, and how minimal interaction is achieved.

If we take a step back, we can assess the problem in the following way: people are using their small, handheld personal devices (i.e., their smartphones) for many different tasks while in an urban setting. This kind of setting is complex and demands high information processing efforts (cognitive, perceptual, motor, and social load) from users. This processing load incurred on users additionally comes largely independent of the mobile task at hand. Examples of research efforts dedicated to address this range from improving urban wireless network access and speed to ensure continuous internet access, to making smartphones more responsive to their users' needs through presenting personalized content to minimize search behavior.

In the HCI community, large research efforts have focused on designing and developing so-called eyes-free interaction techniques, so that users need not interact with their smartphones using their visual sense. However, these efforts have largely remained within the confines of research laboratories, the adoption of complete eyes-free techniques are slow, and the techniques demonstrated are limited in use. To address this, this thesis presented an alternative approach for designing mobile interactions, where we introduced the concept of *minimal* mobile HCI. In Chapter 1 (p. 7), we defined *minimal* mobile HCI as a subset of eyes-free mobile interaction that allows minimal combination of the visual modality with other sensory modalities to minimize attentional demand, frustration, and situational impairments when users interact with smartphones. **This minimal interaction approach focuses on improving the usability and user experience of interacting with smartphones using a simple heuristic: we should keep designing for (mobile) screens, but augment the screen-based interaction by either making the system more intelligent using context-awareness solutions, or complementing user interaction with non-visual interaction techniques such as 3D gestural input.** Interaction designers need not abandon screen-based interaction, nor stop designing for users' visual modality.

To achieve this kind of minimal mobile interaction, we made use of two developments in computing that influence mobile HCI: context-aware computing (where contextual information can be used to make systems more intelligent) and non-visual input techniques, specifically 3D gestural interaction. For minimal mobile HCI using context-awareness, we can use context-aware solutions so that a user's mobile device can sense, learn from, and adapt to the user to free his or her attentional resources while in an urban setting. In our research, we have primarily made use of a smartphone's ability to sense location (through the device's embedded GPS sensor). We have studied this in two domains: urban exploration and playfulness. In the urban exploration domain, we found that location information can be useful for supporting the creation and consumption of location-based multimedia content (*LMM Study*). Additionally, it can be used for automatically generating exploration-based routes in small, touristic cities like Amsterdam (*Route Planner Study*). In the domain of playfulness, by further interpreting our findings on how people use LMM systems, we found that automatic location information can be used to create fun and enjoyable user experiences. In each of those studies, the automatically inferred location information was used as a primary source to minimize users' interaction with the tested system, in particular minimizing their reliance on their visual modality. While this minimal interaction sometimes came at the cost of user engagement (e.g., in the *LMM study*), it provided us with clues on how to increase it by embedding playful elements (*Playful LMM Study*). It also raised awareness that other contextual sources can be used, by for example making extra inferences on what locations may be interesting to visit in a city based on the sequence of detected photo geotags made by photographers in a city (*Route Planner Study*).

Taken together, the three context-awareness studies we carried out (in Part I) raised the question of how to balance minimal mobile interaction while still fulfilling users' needs in domains like urban exploration and playfulness. This gives rise to a tradeoff in user interaction that takes place in an urban setting: on the one hand, visual interaction needs to be kept at a minimum to ensure safety and minimum distraction when in crowded areas, and on the other hand, the designed system features to support a desired user task should be complete and cater for the best user experience. Under this tradeoff, we believe minimal mobile interaction provides a good compromise, as designers and developers need not abandon designing for the visual modality, which allows for a bigger design space when designing system features.

For minimal mobile HCI using gestural input, we can make use of non-visual interaction techniques such as 3D gesture-based interaction to allow the user to interact with his or her smartphone without relying on the visual sense. In our research, we have mainly looked at the usability of non-task specific discrete 3D gesture interaction, and at two applied use cases of how 3D gestural interaction can support user tasks. We have studied this in three domains: task-independent, playfulness, and user authentication. In the task independent gestural input domain, we found that 3D gestural interaction is a usable method for users and still provides an overall positive user experience for failed gesture recognition error rates up to 40% (*Gesture Errors Study*). In the playfulness domain, we found that a specific 3D gestural interaction technique using magnets is fun and enjoyable for users who wish to compose music and play music-related games on their smartphone (*Playful Gestural Interaction Study*). In the user authentication domain, we found that the security of the magnet-based 3D gestural interaction method is strong enough to protect users, and while users found the method highly usable, they were reluctant to immediately adopt this form of authentication without full transparency on its security (*User Authentication Study*).

In each of those studies, the supported 3D gestural input was used as a primary means to minimize users' interaction with the tested system, and largely free up their visual modality. However, in each of those studies, there were recurring issues that needed addressing:

- Performing 3D gestures caused fatigue after continued performance ($\sim$30 min.)
- While some 3D gestures were easy to learn and perform, some were difficult to recall
- Performing 3D gestures was sometimes perceived to be a slower method of input than touchscreen-based interaction
- Performing some 3D gestures in public places was not perceived to be socially acceptable

Each of the issues raised are current topics of research in the HCI, Ergonomics, and Interaction Design communities. The above issues indeed raise the question of to what extent 3D gestural interaction can become a standard on small, handheld devices. However, the findings from our studies show that this type of interaction (namely, 3D gestural input), if these limitations are addressed, generally provide a positive user experience. Concretely, designers should ensure that the supported smartphone 3D gesture sets are ergonomic, easy to perform and recall, and can be performed quickly. With respect to social acceptability, designers can collect data on which gesture sets are more socially acceptable in a given culture. However, we believe this latter finding to be a minor issue if this mode of interaction becomes a standard – what was previously socially unacceptable can become acceptable if adopted by enough users.

Taken together, the three 3D gestural input studies we carried out in the thesis (Part II) point to the benefits that this kind of non-visual interaction method can reap when augmented with touchscreen interaction. Augmenting touchscreen interaction with 3D gestural input not only lowers users' visual attention costs, but can also be used to support playfulness (*Playful gestural Interaction Study*), as well as give rise to a novel gesture-based security method (*Gestural Authentication Study*). As we showed in our work on context-aware solutions, the synergistic coupling of 3D gestural input with touchscreen-based interaction expands the design space for mobile systems and applications, while still minimizing the processing costs (particularly visual) associated with interaction in an urban setting.

System / Technique	Environmental	Social	Device	Personal
Location-aware Multimedia Messaging	-	+	+	+
Exploration-based Route Planner	-	+	-	+
3D Gestural Interaction	+	+	+	+
Magnet-based ADI	+	+	+	+

Table 8.2: Summary of how systems/techniques presented in this thesis fulfill different factors affecting user motivations for using eyes-free mobile interaction (based on Yi et al. (2012)). The '+' sign denotes that a factor is fulfilled, and the '-' sign denotes that a factor is not fulfilled.

In our introduction (Chapter 1, we provided a review of user motivations for why users would want to use eyes-free interaction (Section 1.1.3, p. 6). All the studies dealt with in this thesis have made use of minimal mobile interaction in some form, but not completely eyes-free interaction. However, minimal mobile interaction fulfills many of the same motivations for why users would want to use eyes-free interaction techniques. The usability and user experience evaluations of the systems and techniques presented in this thesis, and how they relate to user motivations for the different factors (Environmental, Social, Device, Personal) associated with using those techniques, are shown in Table 8.2. As can be seen, the Social and Personal factors are satisfied for all the systems and techniques presented. As we have not dealt directly with environmental issues in user interaction, nor safety critical domains, the Environmental factor does not apply for the LMM prototype and the Exploration-based Route Planner. For Device factors, our route planner work does not apply as the focus was on generating routes, and not on visualizing routes on smartphone screens. Together, our studies provide support that the reasons motivating users to use completely eyes-free techniques can to a large extent be addressed by minimal mobile HCI.

To conclude, in our research chapters, we have studied the usability and user experience issues of exemplary systems and techniques. While there are clear usability and user experience issues that surround these techniques, we have provided recommendations to overcome them. This allows improving users' experience when interacting with these mobile technologies. Taken together, this thesis provides evidence that minimal mobile interaction, as a design approach, provides a modest, yet effective goal for designing mobile interactions that are suitable for use in urban settings. Even with limited screen estate, the use of context-aware solutions or 3D gestural input techniques allows more flexility and a larger design space for interaction designers to simplify mobile user interfaces and the resulting user interaction. Specifically:

Minimal mobile HCI, using context-aware solutions (such as automatic location sensing) or 3D gestural input techniques, expands the design space for designing and developing systems/applications used in urban settings, while keeping user interaction costs at a minimum.

The work presented on minimal mobile HCI opens opportunities for future research. These future directions are presented below (Section 8.3).

8.3 Future Directions

Quantitative Modeling of Minimal Mobile Interaction
The minimal mobile interaction concept we introduced was geared towards designing simple context-aware and multimodal mobile interactions that make minimal use of touchscreen interactions. The work presented in this thesis has laid the initial groundwork for validating that a middle ground between full touchscreen interaction and eyes-free interaction is possible.

A clear extension of our work is to provide methods and tools to precisely assess in a quantitative manner the upper and lower bounds for minimal mobile interaction. In other words, determining the number of microinteractions (cf., Saffer (2013); Wolf et al. (2011)) a mobile interaction session should have to reduce interaction from users. This would provide a quantitative basis for such minimal interaction. A relevant step in this direction is the work by Oulasvirta and Bergstrom-Lehtovirta (2010), where their goal was to arrive at what they call a multimodal flexibility index (MFI), calculated from changes in users' performance induced by blocking of sensory modalities. In our case, the aim is not to block sensory modalities and model the changes in performance, but to arrive at a precise assessment of the number of microinteractions required from each modality, and to achieve a balance where drawing resources from the visual modality are kept at a minimum.

Tools that make use of such quantitive measurements can aid HCI researchers and interaction designers in developing new mobile interaction design methods with the goal of minimizing user attentional costs, especially when in an urban setting. Such a tool could be realized as a minimal mobile interaction pattern library (cf., design patterns[2]), where contributions to this library state precisely how methods were combined, under what setting, and what the resulting effects on usability and user experience were.

Testing Other Interaction Methods
In our work on context-awareness, we have looked closely at how location-awareness (accessed through location sensors such as GPS) can help minimize interaction from users. However, there is more to context than location sensing. It is possible to detect nearby people, devices, lighting, noise level, network availability, and even social situations (Dey et al., 2001; Dey and Abowd, 1999). For example, with today's smartphones, we could detect how close a person is to a device through proximity sensors. A step in this direction can be found in recent work by Kostakos et al. (2013), where they use proximity sensing

[2]For example, see the best practices interaction design pattern library offered by Welie.com (http://www.welie.com/patterns/index.php; last retrieved: 01-08-2013) or by UI Patterns (http://ui-patterns.com; last retrieved: 01-08-2013).

using bluetooth sensors to model and understand passenger flow during public transport bus transits. Another example is detecting whether or not a person is busy by detecting how many other persons s/he is currently collocated with. For example, again using low-cost bluetooth sensors, it becomes possible to infer whether a user is in a crowded context or not, which can be used to minimize interruptions; for example, smartphone push notifications for less urgent activities (cf., McFarlane and Latorella (2002)).

In our work on gestural interaction techniques, we have looked closely at 3D gestural input. However, other input techniques could also aid in minimizing interaction. The most prominent example here is voice-based interfaces (Jain et al., 2011), where a user can issue a command to his or her device with little to no interaction with the smartphone device. Other ways to support minimal interaction includes making use of multimodal output, an area we have briefly touched upon here. How to design and develop effective notification cues (Mcfarlane and Mcfarlane, 1997) that minimize interruption to a user's current activity are also highly relevant to ensuring minimal mobile interaction. These notification cues can also be tested under different modalities, such as using haptic feedback, speech or non-speech auditory feedback, and under minimal visual feedback settings (cf., Google Glass).

Considered in light of minimal mobile HCI, (multimodal) output or feedback can be synergistically coupled with screen-based interaction, without requiring complete removal of screen-based interaction (c.f., EarPod, an eyes-free menu selection technique for smartphones (Zhao et al., 2007)). A relevant step in this direction is to make use of crossmodal feedback cues (e.g., Hoggan et al. (2009)), where visual information is mapped to other modalities (such as tactile or auditory cues). In crossmodal interaction, the physical parameters of feedback cues are to a large extent amodal (e.g., duration, rhythm), so that the same information can be communicated across modalities, during daily user interaction (Hoggan and Brewster, 2010). From a minimal mobile interaction perspective, the user may still draw on her visual modality during interaction, however the visual information processed is augmented with tactile and/or auditory cues. Testing whether these cues can effectively reduce attentional costs and improve users' experience in interacting across a range of tasks is a promising direction for further work.

Longitudinal, In-the-Wild Testing

We have presented several studies to show how minimal mobile interaction is possible. In all but the *Playful LMM Study*, we ran user evaluations to assess usability and user experience issues that arise from user interaction. While in the *LMM Study* in Chapter 2 we studied user's multimedia messaging behavior in the wild by using a multimodal diary method (El Ali et al., 2010), this kind of in the wild testing was not currently possible for the other user studies. The breadth of the research carried out in this thesis, using early prototypes and systems, meant that the systems and techniques required first an evaluation in the laboratory. This was both a limitation and a necessary aspect of our work. On the one hand, the early systems and techniques tested (e.g., exploration-based route planner system, and all the 3D gestural techniques studied) were still in the earlier stages of development, and therefore not production ready. On the other hand, precisely due to the earlier stages of these systems, from an experimentation perspective we required maximum control over the different usability and user experience variables.

Given that our approach in this thesis was to study multiple systems and techniques, it also meant we were unable to thoroughly follow through with each system and technique to the stage of deployment where users can use them in their natural, everyday settings. In some cases this was an enforced limitation (e.g., the magnet-based ADI music apps and authentication techniques are patented technologies). In other cases, it was a side effect of studying multiple systems and techniques at the cost of a full redesign and (re-)evaluation of a smaller set of systems or techniques in the wild. Our design choices notwithstanding, to truly validate the use of minimal mobile interaction methods, whether through context-awareness or gestural interaction techniques, users should be allowed to use these systems or techniques in their natural, everyday setting. For future researchers interested in taking the minimal mobile HCI further, testing users in such in-the-wild settings should be a priority, as it would provide stronger ecological validity for minimal mobile interaction.

8.4 Looking Ahead

The concept of minimal mobile HCI we introduced in this thesis parallels recent debates on the web amongst HCI researchers and interaction designers on whether the future of interaction with interfaces rests with Graphical User Interfaces (GUIs) or not. Some argue that we should do away with GUIs completely to improve users' experience when interacting with technology,[3] others argue that GUIs are here to stay,[4] while others find the debate meaningless as design is about solving problems for a particular user group.[5] With minimal mobile HCI, while we have also taken a middle ground on whether we should rely on screen-based interactions or not, we believe the question of including a GUI (which heavily relies on users' visual sense) this is highly dependent on the user task and importantly *where* the user interaction takes place.

Unlike these GUI vs. no GUI debates, we have proposed two solutions (namely context-awareness and gestural input techniques) to facilitate the advancement of the HCI and Interaction Design fields. We have dealt with how each, under a minimal mobile interaction setting, can improve usability and user experience of the studied systems and techniques in a scientific manner. The proposed minimal mobile HCI does not require eliminating the GUI, nor keeping it there constantly, it requires first and foremost an assessment of where the anticipated interaction will take place. If the envisioned interaction with the smartphone is to be used in an urban setting, then the designer should consider solutions whereby screen-based interaction is *augmented*, not replaced. While the place of interaction is paramount, the user task can additionally influence the choice of what augmentation is used.

We have presented two such augmentations: context-awareness and gestural input techniques. We studied these in a small set of domains, namely urban exploration (Chapters 2 and 4), playfulness (Chapter 3 and 6), task-independent (Chapter 5), and user authentication (Chapter 7). Through our studies, we have provided evidence to persuade designers designers, developers, as well as entrepreneurs to think in terms of reducing user attentional costs – that a synergestic combination with either context-awareness, 3D gestural input, or even both, can provide users with a better overall user experience. In short, at least for the next 10 years, we strongly believe screen-based interaction is here to stay, but to improve

[3] http://www.cooper.com/2012/08/29/the-best-interface-is-no-interface; last retrieved: 01-08-2013
[4] http://www.elasticspace.com/2013/03/no-to-no-ui; last retrieved: 01-08-2013
[5] http://scottberkun.com/2013/the-no-ui-debate-is-rubbish/; last retrieved: 01-08-2013

the usability and UX of this interaction, we should focus on methods and design approaches that aim to augment this interaction, not replace it.

Appendices

A
Appendix Chapter 2

LMM Study Interview Questions

Warmup

- **How is it going?**
- **How did your day go so far?**

Intro

- **Could you please describe the environment around us?**
- **How are you feeling right now? Sad? Happy? Anxious?**
- **What do you normally do here? What kinds of activities do you do in such an area?**
- **Do you normally see your friends here?**

  What about people that you've seen before but don't really know?
- **When you think about this area, what do you remember or associate it with? Feel free to walk around.**
- **Can you think about one memorable thing you did here, and if possible explain it to me?**

Into the Cafe

1. Explanation of LMM prototype
2. Quick Demo
3. First Tryout as Image

- **After using LMM prototype, what do you think is the purpose of this application?**

  Can you please explain?

1. In-depth Explanation of anchoring & sharing in prototype

Figure A.1: Semi-Structured interview questions for the pilot LMM study. Form continues on next page (1/3).

2. Topic of expression: drinking coffee/beverage at a cafe

- **Could you please explain the drawing you made?**

  Why did you choose to do it here?
- **Could you please explain the text you wrote?**

  Why did you choose to write this text at this spot?
- **Could you please explain the photograph you made?**

  Why did you take the photo here?

2-MIN. BREAK

- **Which did you prefer best: drawing an image, writing text, or taking photographs?**

  Why this?
- **Was there something specific in the environment that you directed your *(text, drawing, photo)* on?**
- **When you were scanning the environment to make a *(text, drawing, photo)*, what caught your attention most? Why?**
- **Let me now ask you about a few questions about your sharing preferences. Who would you like to have see the expressions you've made here? Friends, family, etc.?**
- **For a text, image, or photo, would you like the expression you made to stay here for anyone to see?**
- **Suppose that the expression you made is visible to others on their mobile devices only when you're close to it, that is in the area. Would do you prefer that?**

- **If yes, would you like others to approach you about an expression you left behind?**

  If no, then why not?

General Social & Urban Context Questions

- **Does the fact that people are around here affect the kinds of expressions you would make?**

  If yes, then why? Do you feel affected by this more when writing text, drawing, or taking photographs?

  If no, then why not? Is this so for writing text, drawing, and taking photographs?

Figure A.2: Semi-Structured interview questions for the pilot LMM study. Form continues on next page (2/3).

- **Do you feel that using such an application makes you more aware of the environment?**

  If yes, then why yes?

  If no, then why no?

- **Do you feel that using such an application makes you more aware of the people around you?**

  If yes, then why yes?

  If no, then why no?

About the Application Itself

- **Was it easy to use the application?**

  Explain

- **Was is it easy enough to make an expression while you were observing the environment?**

  Explain

- **What about if you were involved in the event, that is, doing something like drinking coffee?**

  Explain

- **What would you improve in this application?**

  Extra features? Musical soundtrack? Notifications?

- **Finally, what do you think the purpose of LMM prototype is now that you're quite familiar with it?**

Figure A.3: Category Attribution Task instruction form. Form continued from previous page (3/3).

How Do You Augment Your Reality in Public Places?

Instructions:

When you make a multimedia expression, please answer all the questions below (in the Question Template). Please make a multimedia expression 2 times per day. The questions should be answered only when you make an expression at a public place. The public place can be the same across days.

Each time you want to express something, consider what kind of media you would like to make the expression in. You can make it in 'one or more' of the following expressions (you are not however restricted to these):

a) draw an image
b) write a short text
c) take a photograph
d) shoot a video
e) leave a voice audio recording (i.e, a podcast)
f) upload a song

- If you choose to draw an image or write text, then please enclose your drawing or text within a box of the same size as the screen of the mobile device shown below.

- If you choose to take a photograph(s), shoot a video(s), leave a voice audio recording(s), or upload a song(s), then please email these files to abdallah.elali@gmail.com after the week has ended.

For every expression that you make, you can choose to make it:

a) **Public:** anyone can see/hear the expression you made
b) **Private:** only your friends or a network you choose can see/hear the expression you made

Figure A.4: LMM multimodal diary study questions. Form continues on next page (1/2).

Question Template:

Date:

Time:

Expression type: a) drawing **b)** text **c)** photograph **d)** video **e)** voice recording **f)** song **g)** other:

Title/Name of the expression:

Public or Private?

1) Where are you right now?
2) Please explain why you made the expression at this place.
3) Please describe how you are feeling right now. (e.g., happy, sad, anxious, excited, lazy, etc.)
4) Please describe the environment around you.
5) Who are you with right now?
6) What were you doing before you made the expression?
7) Is there an event going on where you are (e.g., sunset, festival, live band, market, dinner, etc.)? If yes, please describe the event.
8) If yes to question 7), are you participating in this event, or did you only observe it?
9) If yes to question 7), is this the first time you participate/see such an event?
10) Were you able to express what you wanted? If not, please state why you couldn't.
11) Was there something specific in the environment that you directed this expression at? If yes, please state what it is.

Figure A.5: LMM multimodal diary study questions. Form continued from previous page (2/2).

LMM Study Exit Interview

1. **Did you face any difficulties in filling in the diary during the last week?**

   If yes, can you please explain?

2. **Where there days where you felt more inspired to post something (e.g., over the weekend)?**

3. **Where there particular locations that made you feel more inspired to post an expression?**

4. **Which media type did you make use of most during the last week?**

   Was there a reason for using this type over the others?

5. **Did you feel the diary made you more aware of your everyday environment?**

   Please explain.

6. **Was your overall experience during the last week different because you had to fill in the diary?**

   Please explain.

7. **Would you use a mobile application that supports posting various types of multimedia content at locations?**

   Please explain.

8. **If you were viewing expressions made by others at a location, would you be interested in seeing information similar to what was asked from you in the diary?**

   At what level of detail would you like to see it? Would you want the information to adapt to your current situation and/or desires?

9. **If a mobile application asked you questions similar to the ones in the diary (e.g., who you were with, what event is happening, etc.), would you fill in this information?**

   At what level of detail would you be willing to do so?

10. **Would you like to add anything further?**

Figure A.6: LMM study exit interview questions.

LMM Study - Category Attribution Task

Age: ______

Sex: M / F

Education level (Bachelor's/Master's/Doctorate): ___________________

Listed below are explanations provided by people about why they made a multimedia expression (e.g., a song, a photo, a video, text, etc.). Your task is to provide TASK and DOMAIN categories that you believe best classify an explanation.

For each explanation listed below, please assign TASK and DOMAIN categories according to your first impression. If you believe more than one task or domain is needed, then please fill them in. Below are lists of TASKS and DOMAINS you can use. You can list the number of the TASK and DOMAIN category in each row.

Note: if you choose *Other,* then please fill in the category you think best classifies the explanation.

Tasks:

1) Appreciation (i.e., enjoying the qualities of something)
2) Recommendation (i.e., recommending something to others)
3) Criticism (i.e., being critical of something)
4) Altruism (i.e., actively seeking to help)
5) Self-reflection (i.e., reflecting on one's own actions or feelings)
6) Activity reporting (i.e., reporting to people what you have done)
7) *Other (please state what you think)*

Domains:

1) Aesthetics (e.g., a beautiful scene)
2) Entertainment (e.g., a film)
3) Transport (e.g., train ride)
4) Products & Services (e.g., eating food or putting money in the bank)
5) Health & Well-being (e.g., feeling sick)
6) Education (e.g., studying)
7) Architecture (e.g., a building)
8) *Other (please state what you think)*

Figure A.7: Category Attribution Task instruction form. Form continues on next page (1/2).

#	Media Type	Explanation	Task	Domain
1	drawing	Wanted to try out my first multimedia expression		
2	photo	Beauty of the sun shining over the bldg. opposite the icy water		
3	song	We were studying for a few hours, and now we are listening to music instead. The song I chose fits well the work atmosphere		
4	audio recording	I was struck by the sudden intensity of the voices		
5	photo	OV-chipkaart organization messed up because my picture is on it only half way		
6	photo	I wanted to make a picture of a black mountain bike disappearing in the distance, but the picture had a bad focus so it did not work out as well as I expected		
7	video	The scenery is just beautiful; the snow is still fresh on the trees		
8	photo	I needed to make an expression		
9	photo	Most interesting location of the day		
10	text	I'm in a private area so no one should be able to copy credit card data		
11	text	Because I liked the movie		
12	text	Because my hands were freezing!		
13	photo	The boat has an actual stove that I made my coffee on - very old fashioned!		
14	photo	Wow, nice view		
15	photo	Excessive amount of people on the train		
16	photo	I liked the shot		
17	song	We're almost as good as the original		
18	photo	A sign of good morning world we're there again		
19	text	I saw a bird eating snow		
20	photo	As a real dutchy, when there is ice, you need to go there		

[…]

Figure A.8: Category Attribution Task instruction form. Only a sample of 20 questions shown. Form continued from previous page (2/2).

B
Appendix Chapter 4

AttrakDiff ™ #: _______

Amsterdam Central Station → Museumplein

Consider the following scenario:
"You are at Amsterdam Central Station with three friends who are visiting town for the weekend, and they would now like to go and see the Museumplein. It is around 2 o'clock in the afternoon on a Saturday, and the skies are clear. You all have free time on your hands, so you decide to walk there. Before embark on your journey, you reflect on which route to take."

OR
(scenario & questionnaire presented depending on participant condition)

Waterlooplein → Westerkerk

Consider the following scenario:
"You are at the Waterlooplein with a good friend who has just come back from a vacation. You would like to catch up, and decide to go have coffee somewhere near the Westerkerk. It is around 7 o'clock in the evening on a Sunday, with some clouds in the sky. You two decide to walk there. Before you embark on your journey, you reflect on how you want to get there."

Here, you are asked to evaluate 3 routes that take you from Central Station to Museumplein. Following are pairs of words to assist you in your evaluation. Each pair represents extreme contrasts. The possibilities between the extremes enable you to describe the intensity of the quality you choose. Do not spend time thinking about the word-pairs. Try to give a spontaneous response. You may feel that some pairs of terms do not adequately describe the entire route. In this case, give an answer that most closely relates to the highest segment of a route. Keep in mind that there is no right or wrong answer. Your personal opinion is what counts!

Figure B.1: Scenarios and modified AttrakDiff2 questionnaire for the Route Planner study. Form continues on next page (1/3).

AttrakDiff ™

#: _______

1: first route variation 2: second route variation 3: third route variation
(Mark spaces, not lines. You can change ratings you've done earlier, if you wish)

1	Faster		Slower
2	Isolating		Connective
3	Pleasant		Unpleasant
4	Inventive		Conventional
5	Simple		Complicated
6	Professional		Unprofessional
7	Practical		Impractical
8	Ugly		Attractive
9	Likeable		Disagreeable
10	Cumbersome		Straightforward
11	Stylish		Tacky
12	Predictable		Unpredictable
13	Cheap		Premium
14	Alienating		Integrating
15	Brings me closer to people		Separates me from people
16	Unpresentable		Presentable
17	Rejecting		Inviting
18	Unimaginative		Creative
19	Good		Bad

Figure B.2: Modified AttrakDiff2 questionnaire for the Route Planner study. Form continues on next page (2/3).

AttrakDiff ™

#: ________

20	Confusing		Clearly Structured
21	Repelling		Appealing
22	Bold		Cautious
23	Innovative		Conservative
24	Dull		Captivating
25	Undemanding		Challenging
26	Motivating		Discouraging
27	Novel		Ordinary
28	Unruly		Manageable

Complete only after filling the above questionnaire 3 times (once for each route variation):

1) Given the scenario above, I would initially follow:
   a. Route 1
   b. Route 2
   c. Route 3
   d. Either one will do
   e. None of them

2) I picked [Route 1 / Route 2 / Route 3 / __________] because:
   a. It is more scenic
   b. It is more convenient
   c. Routes are too similar to make a difference
   d. Other: _____________

Figure B.3: Modified AttrakDiff2 questionnaire and short questions for the Route Planner study. Form continued from previous page (3/3).

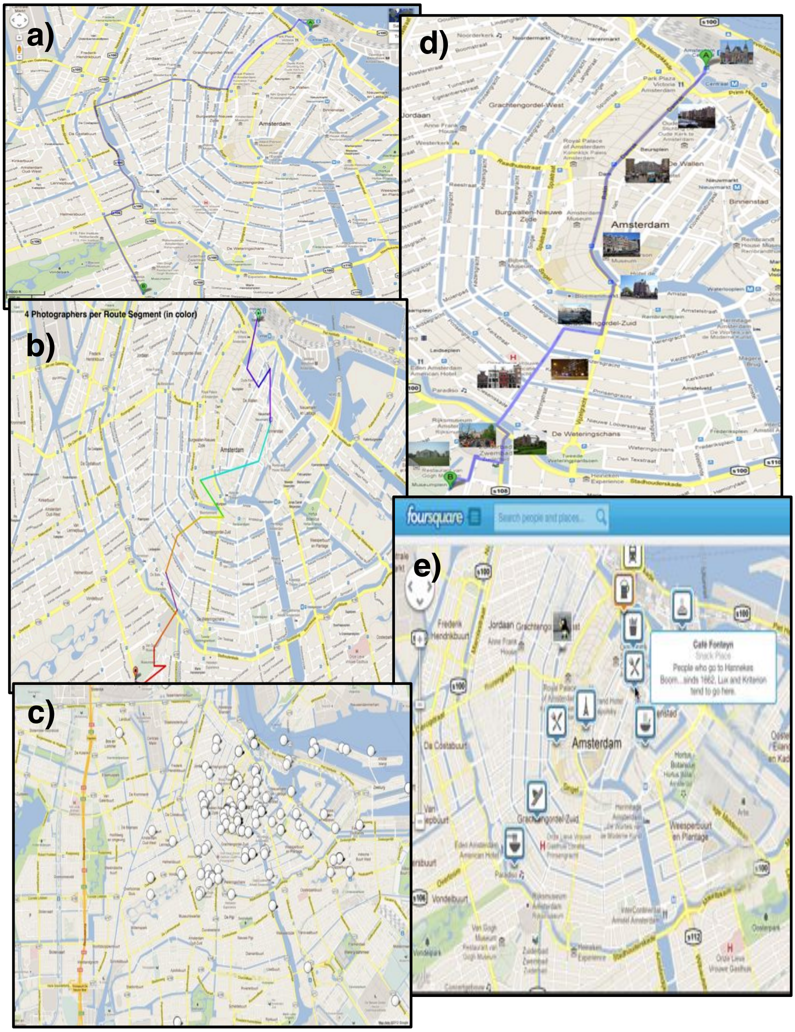

Figure B.4: The different digital information aids presented to participants. a) Google Maps shortest route b) Color-coded Photographer Paths c) Photo density geopoints d) Photo thumbnail geopoints e) Foursquare POIs

“Which route would you take?”
Interview Questions

1) What was your overall impression of the experiment session?

2) What did you think of the scenarios? Were they realistic?

3) For the first route (Centraal → Museumplein), what in your opinion were the differences in the three route variations you saw? Please tell in detail which part of the routes you liked, disliked, etc. Rank them.

4) For the second route (Waterlooplein → Westerkerk), what in your opinion were the differences in the two route variations you saw? Please tell in detail which part of the routes you liked, disliked, etc. Rank them.

5) What kind of information would you like to have in deciding to take a route when you have time (i.e., for exploration)?

6) What do you think about this kind of information: specific paths that were taken by a certain number of photographers over a period of 5 years in a city (like Amsterdam).

Explanation of the different routes (random, google maps, algorithm). Hand out different sheets (plain interesting route, photo route, photographer segment saliency & counts, combination photos + photographic paths).

7) Which of these information types do you think helps you most to decide how to plan your route in a city for exploration? How important is each kind of information for you? Combination?

8) If there was an app that recommends alternative scenic routes in a city based on your start and end location, is this something you would use? Locally and abroad?

Figure B.5: Route Planner study interview questions.

“Which route would you take?”
Web Survey

NOTE: The original web survey, in addition to the below questions, additionally asked basic demographic information (age, gender, familiarity with route planners) and showed static images of our generated routes.

I find the following kinds of information helpful when planning a route (by foot) for exploring a city with friends: *
You can check more than one.

[These are checkboxes, and last has a free form text box]

-Established Points of Interest (e.g., restaurants, cafes, etc.) along a route,
-Photos of the different parts of a route,
-Comments (most recent first) left by others about the different parts of a route,
-Number of photos taken along a route over a specified time period (e.g., 1 year),
-Route segments that a number of city photographers took over a specified time period (e.g., 1 year),
-Walking distance of a route
-Walking time of a route,
-Asking strangers on the street
-None, I like getting lost!
-Other (e.g.,)

What kind of information do you feel helps you most when planning a route in a city you want to explore?
E.g., popular points of interest, tourist guides, asking locals, recommendations from friends, etc.

[Free form text box]

Figure B.6: Route Planner study web survey questions.

C
Appendix Chapter 5

Rating Scale Definitions		
Title	**Endpoints**	**Description**
Mental Demand	*Low/High*	How much mental, visual and auditory activity was required? (e.g. thinking, deciding, calculating, looking, listening, scanning, searching)
Physical Demand	*Low/High*	How much physical activity was required? (e.g. pushing, pulling, turning, controlling)
Time Pressure	*Low/High*	How much time pressure did you feel because of the rate at which things occurred? (e.g. slow, leisurely, rapid, frantic)
Effort Expended	*Low/High*	How hard did you work (mentally and physically) to accomplish your level of performance?
Performance Level Achieved	*Poor/Good*	How successful do you think you were in doing the task set by the experimenter? How satisfied were you with your performance? Don't just think of your score, but how you felt you performed.
Frustration Experienced	*Low/High*	How much frustration did you experience? (e.g. were you relaxed, content, stressed, irritated, discouraged)
Overall Preference Rating	*Low/High*	Rate your overall preference for each of the recognition algorithms. Which one made the task the easiest? The first block, the second, or the third?

Figure C.1: Modified NASA-TLX questionnaire category explanations. Form continues on next page (1/2).

#: ____

MENTAL DEMAND

Low — High

PHYSICAL DEMAND

Low — High

TIME PRESSURE

Low — High

EFFORT EXPENDED

Low — High

PERFORMANCE LEVEL ACHIEVED

Poor — Good

FRUSTRATION EXPERIENCED

Low — High

OVERALL PREFERENCE RATING

Low — High

Figure C.2: Modified NASA-TLX questionnaire. Form continued from previous page (2/2).

"How well are mobile gestures recognized?" Interview Questions

1) What is your overall impression of the experiment session?

2) How did you feel your performance was?

3) Did you notice any change in how you performed the gestures?

4) Did you notice any difference between each of the individual gestures?

5) Applied use of gestures in mobile phones?

6) How socially acceptable are the gestures when performed in public? Mimetic gestures? Alphabet gestures?

Figure C.3: Gesture Errors study exit interview questions.

D

Appendix Chapter 6

AttrakDiff ™

#: _______

MagiMusic AttrakDiff2 Evaluation

Following are pairs of words to assist you in your evaluation of each MagiMusic app. Each pair represents extreme contrasts. The possibilities between the extremes enable you to describe the intensity of the quality you choose.

Do not spend time thinking about the word-pairs. Try to give a spontaneous response. You may feel that some pairs of terms do not adequately describe the iPhone® app. In this case, please still be sure to give an answer. Keep in mind that there is no right or wrong answer. Your personal opinion is what counts!

1: first app 2: second app 3: third app
(Mark spaces, not lines. You can change ratings you've done earlier, if you wish)

1	Human		Technical
2	Isolating		Connective
3	Pleasant		Unpleasant
4	Inventive		Conventional
5	Simple		Complicated
6	Professional		Unprofessional
7	Practical		Impractical
8	Ugly		Attractive
9	Likeable		Disagreeable
10	Cumbersome		Straightforward
11	Stylish		Tacky
12	Predictable		Unpredictable
13	Cheap		Premium
14	Alienating		Integrating

Figure D.1: AttrakDiff2 questionnaire for the Playful ADI study. Form continues on next page (1/2).

AttrakDiff ™

15	Brings me closer to people	☐☐☐☐☐☐☐	Separates me from people
16	Unpresentable	☐☐☐☐☐☐☐	Presentable
17	Rejecting	☐☐☐☐☐☐☐	Inviting
18	Unimaginative	☐☐☐☐☐☐☐	Creative
19	Good	☐☐☐☐☐☐☐	Bad
20	Confusing	☐☐☐☐☐☐☐	Clearly Structured
21	Repelling	☐☐☐☐☐☐☐	Appealing
22	Bold	☐☐☐☐☐☐☐	Cautious
23	Innovative	☐☐☐☐☐☐☐	Conservative
24	Dull	☐☐☐☐☐☐☐	Captivating
25	Undemanding	☐☐☐☐☐☐☐	Challenging
26	Motivating	☐☐☐☐☐☐☐	Discouraging
27	Novel	☐☐☐☐☐☐☐	Ordinary
28	Unruly	☐☐☐☐☐☐☐	Manageable

Figure D.2: AttrakDiff2 questionnaire for the Playful ADI study. Form continued from previous page (2/2).

#: _____

System Usability Scale

Instructions: After playing with each MagiMusic app, please tick each box below with the number of the app, to best describe your reaction to it.

1: First app **2**: Second app **3**: Third app

		Strongly Disagree				Strongly Agree
1.	I think that I would like to use this app frequently.	☐	☐	☐	☐	☐
2.	I found this app unnecessarily complex.	☐	☐	☐	☐	☐
3.	I thought this app was easy to use.	☐	☐	☐	☐	☐
4.	I think that I would need assistance to be able to use this app.	☐	☐	☐	☐	☐
5.	I found the various features in app were well integrated.	☐	☐	☐	☐	☐
6.	I thought there was too much inconsistency in this app.	☐	☐	☐	☐	☐
7.	I would imagine that most people would learn to use this app very quickly.	☐	☐	☐	☐	☐
8.	I found this app very cumbersome/awkward to use.	☐	☐	☐	☐	☐
9.	I felt very confident using this app.	☐	☐	☐	☐	☐
10.	I needed to learn a lot of things before I could get going with this app.	☐	☐	☐	☐	☐

This questionnaire is based on the System Usability Scale (SUS), which was developed by John Brooke while working at Digital Equipment Corporation. © Digital Equipment Corporation, 1986.

Figure D.3: System Usability Scale.

#:_____

“MagiMusic” Exit Questionnaire

Please answer each question based on your experience with the iPhone® apps. For most questions, you have to fill in a number (between 1-7; 1-very bad & 7-very good) that best represents your experience for each app.

1) What was your first impression of the app?

1	2	3	4	5	6	7
Very bad						Very good

App 1: ______________

App 2: ______________

App 3: ______________

2) Was it easy to learn?

1	2	3	4	5	6	7
Very easy						Very difficult

App 1: ______________

App 2: ______________

App 3: ______________

3) Was it comfortable to play with the app?

1	2	3	4	5	6	7
Very comfortable						Extremely uncomfortable

App 1: ______________

App 2: ______________

App 3: ______________

Figure D.4: Playful ADI study exit questionnaire. Form continues on next page (1/2).

4) Did you enjoy making music with the app? What was good or bad?

1 2 3 4 5 6 7
Not at all Very much so

App 1: ______________ Good/bad: _________________________

App 2: ______________ Good/bad: _________________________

App 3: ______________ Good/bad: _________________________

5) I would be willing to carry a magnet with me to play with these apps.

1 2 3 4 5 6 7
Not at all Definitely

Figure D.5: Playful ADI study exit questionnaire. Form continued from previous page (2/2).

1) What is your overall impression of the experiment session? *[warm up]*

2) How do you feel about interacting with the apps using a magnet? How would you use it?

3) What are your expectations about the availability of magnets when you download a MagiMusic app?

4) Preference for magnet size or shape? What for you is the perfect magnet?

5) What other apps *[typical mobile tasks]* can you think of where you can use magnets to interact with your phone?

6) Would you play with these apps in public places (e.g., bus, metro, mall, street, etc.)? Would you find it socially acceptable?

Figure D.6: Playful ADI study exit interview questions.

E
Appendix Chapter 7

#:_____

"MagiSign" Gesture Questionnaire

Please answer each question based on your experience with performing each set of gestures (4 total). Choose a number (between 1 & 7) that best describes your experience with gesture signatures using a magnet.

1) **These gestures take too long to perform.**

1 2 3 4 5 6 7
Strongly Disagree *Strongly Agree*

2) **These signature gestures are suitable for daily use.**

1 2 3 4 5 6 7
Strongly Disagree *Strongly Agree*

3) **These gesture signatures are secure enough.**

1 2 3 4 5 6 7
Strongly Disagree *Strongly Agree*

4) **I found it comfortable performing these signature gestures.**

1 2 3 4 5 6 7
Strongly Disagree *Strongly Agree*

5) **Performing these gestures in front of strangers is awkward.**

1 2 3 4 5 6 7
Strongly Disagree *Strongly Agree*

Figure E.1: Authentication ADI usability study questionnaire. Form continues on next page (1/2).

6) I would prefer performing these gesture signatures over passwords.

1 2 3 4 5 6 7
Strongly Disagree *Strongly Agree*

7) I have problems performing these gestures.

1 2 3 4 5 6 7
Strongly Disagree *Strongly Agree*

8) I can easily remember the signature gestures I made.

1 2 3 4 5 6 7
Strongly Disagree *Strongly Agree*

Figure E.2: Authentication ADI usability study questionnaire. Form continued from previous page (2/2).

#:_____

"MagiSign" Exit Questionnaire

Please answer each question based on your experience in this study with performing air signatures using a magnet for security access.

1) **Performing gestures with a magnet takes too much time.**

1 2 3 4 5 6 7
Strongly Disagree *Strongly Agree*

2) **Gesture input using a magnet is intuitive.**

1 2 3 4 5 6 7
Strongly Disagree *Strongly Agree*

3) **Entering a password (PIN or alphanumeric password) is easier than gesture signing with a magnet.**

1 2 3 4 5 6 7
Strongly Disagree *Strongly Agree*

4) **One handed gestures with a magnet are enough for securing my mobile device.**

1 2 3 4 5 6 7
Strongly Disagree *Strongly Agree*

5) **Gesturing with a magnet is exhausting.**

1 2 3 4 5 6 7
Strongly Disagree *Strongly Agree*

6) **These two handed gestures I made with a magnet can be easily forged.**

1 2 3 4 5 6 7
Strongly Disagree *Strongly Agree*

7) **I would prefer performing these gesture signatures over passwords.**

1 2 3 4 5 6 7

Figure E.3: Authentication ADI usability study exit questionnaire. Form continues on next page (1/2).

Strongly Disagree *Strongly Agree*

8) I would have problems performing these gestures in public.

1 2 3 4 5 6 7
Strongly Disagree *Strongly Agree*

9) I can easily remember the signature gestures I made.

1 2 3 4 5 6 7
Strongly Disagree *Strongly Agree*

10) I found gesturing using two hands to be a secure method for authentication.

1 2 3 4 5 6 7
Strongly Disagree *Strongly Agree*

11) I would be willing to carry a magnet with me for secure mobile access authentication.

1 2 3 4 5 6 7
Strongly Disagree *Strongly Agree*

12) I would be willing to pay for a good magnet to ensure secure authentication.

1 2 3 4 5 6 7
Strongly Disagree *Strongly Agree*

13) How would you like to carry a magnet with you (e.g., as a ring, necklace, on keychain, separate magnet, etc.)?

By: __________________

14) For the two handed signatures, I prefer to perform the signatures:

a) Serially (move one hand first, then the other hand)
b) Simultaneously (use both hands at the same time)
c) Both are fine

Figure E.4: Authentication ADI usability study exit questionnaire. Form continued from previous page (2/2).

#:_____

System Trust Scale

Please answer each question based on your experience with the air signature security authentication system using a magnet. Please circle a number (not at all=1; extremely=7) that best describes your experience.

1	2	3	4	5	6	7
Not at all						*Exteremely*

1) **The air signature security system is reliable.**

   1 2 3 4 5 6 7

2) **The air signature security system is deceptive.**

   1 2 3 4 5 6 7

3) **The air signature security system's actions will have a harmful or injurious outcome.**

   1 2 3 4 5 6 7

4) **I am confident in the air signature security system.**

   1 2 3 4 5 6 7

5) **The air signature security system provides security.**

   1 2 3 4 5 6 7

6) **The air signature security system has integrity.**

   1 2 3 4 5 6 7

7) **I am suspicious of the air signature security system's intent, action, or outputs.**

   1 2 3 4 5 6 7

Figure E.5: System Trust Scale for the Authentication ADI usability study. Form continues on next page (1/2).

8) The air signature security system is dependable.

1 2 3 4 5 6 7

9) I can trust the air signature security system.

1 2 3 4 5 6 7

10) I am familiar with the air signature security system.

1 2 3 4 5 6 7

11) The air signature security system behaves in an underhanded/dishonest manner.

1 2 3 4 5 6 7

12) I am wary of / cautious about the air signature security system.

1 2 3 4 5 6 7

Figure E.6: System Trust Scale for the Authentication ADI usability study. Form continued from previous page (2/2).

“MagiSign”
Interview Questions

1) What is your overall impression of the experiment session?

2) How do you feel about mobile security access using magnet-based signatures? Would you use it?

Figure E.7: Authentication ADI usability study exit interview questions.

#:_____

“MagiForge” Video Set Questionnaire

Please answer each question based on your experience with forging the signatures from the videos. For each video set, fill in a number (between 1 & 7) that best describes your experience with forging those signatures.

1) It was easy to forge the signatures made with a magnet.

1 2 3 4 5 6 7
Strongly Disagree *Strongly Agree*

Video set 1: ______________

Video set 2: ______________

2) It was hard to follow the hand movements of the person from the videos.

1 2 3 4 5 6 7
Strongly Disagree *Strongly Agree*

Video set 1: ______________

Video set 2: ______________

3) Forging these signatures is easier than forging signatures on paper.

1 2 3 4 5 6 7
Strongly Disagree *Strongly Agree*

Video set 1: ______________

Video set 2: ______________

4) Select preferred video angles for forgery:

a) Front b) Back c) Left d) Right e) All f) Combination: _________

5) General Comments:

____

Figure E.8: Category Attribution Task instruction form. Form continued from previous page (3/2).

#:_____

"MagiForge" Exit Questionnaire

Please answer the questions based on your experience of forging air signatures.

1) **Overall, I found forging one handed signatures to be easy.**

1 2 3 4 5 6 7
Strongly Disagree *Strongly Agree*

2) **Overall, I found forging two handed signatures to be easy.**

1 2 3 4 5 6 7
Strongly Disagree *Strongly Agree*

3) **Forging someone's air signature is more difficult than forging a signature on paper.**

1 2 3 4 5 6 7
Strongly Disagree *Strongly Agree*

4) **Overall, one handed gestures with a magnet provides strong security for mobile devices.**

1 2 3 4 5 6 7
Strongly Disagree *Strongly Agree*

5) **Overall, two handed gestures with a magnet provides strong security for mobile devices.**

1 2 3 4 5 6 7
Strongly Disagree *Strongly Agree*

6) **I would be willing to pay for a good magnet to ensure secure and usable authentication when making air signatures.**

1 2 3 4 5 6 7
Strongly Disagree *Strongly Agree*

Figure E.9: Authentication ADI security study intermediate questionnaire.

“MagiForge”
Interview Questions

1) What is your overall impression of the experiment session?

2) How do you feel about mobile security access using magnet-based signatures? In comparison with other security methods (PIN, graphical passwords, etc.).

3) What was you impression about forging one-handed vs. two-handed signatures?

4) Would you use this method of authentication?

Figure E.10: Authentication ADI security study exit questionnaire.

Bibliography

Amin, A., Townsend, S., Ossenbruggen, J., and Hardman, L. (2009). Fancy a drink in canary wharf?: A user study on location-based mobile search. In *INTERACT '09: Proceedings of the 12th IFIP TC 13 International Conference on Human-Computer Interaction*, pages 736–749. Springer-Verlag. 23

Appan, P., Sundaram, H., and Birchfield, D. (2004). Communicating everyday experiences. *SRMC*, pages 17–24. 20

Arase, Y., Xie, X., Hara, T., and Nishio, S. (2010). Mining people's trips from large scale geo-tagged photos. In *Proc. MM '10*, pages 133–142. ACM. 55

Ark, W. S. and Selker, T. (1999). A look at human interaction with pervasive computers. *IBM Syst. J.*, 38(4):504–507. 4

Ashbrook, D., Baudisch, P., and White, S. (2011). Nenya: subtle and eyes-free mobile input with a magnetically-tracked finger ring. In *Proc. CHI '11*, pages 2043–2046. ACM. 6, 96, 98, 112

Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, 4(10):829–839. 3, 138

Baddeley, A., Hitch, G., et al. (1974). Working memory. *The psychology of learning and motivation*, 8:47–89. 3, 138

Bangor, A., Kortum, P. T., and Miller, J. T. (2008). An empirical evaluation of the system usability scale. *IJHCI*, 24(6):574–594. 105, 122

Bardzell, J., Bardzell, S., Pace, T., and Karnell, J. (2008). Making user engagement visible: a multimodal strategy for interactive media experience research. In *CHI '08: CHI '08 extended abstracts on Human factors in computing systems*, pages 3663–3668, New York, NY, USA. ACM. 45

Baudisch, P. and Chu, G. (2009). Back-of-device interaction allows creating very small touch devices. In *Proc. CHI '09*, pages 1923–1932, NY, USA. ACM. 5, 6, 96, 97, 113

Bellotti, V., Begole, B., Chi, E. H., Ducheneaut, N., Fang, J., Isaacs, E., King, T., Newman, M. W., Partridge, K., Price, B., Rasmussen, P., Roberts, M., Schiano, D. J., and Walendowski, A. (2008). Activity-based serendipitous recommendations with the magitti mobile leisure guide. In *Proc. CHI '08*, pages 1157–1166. ACM. 68

Bellotti, V. and Edwards, K. (2001). Intelligibility and accountability: human considerations in context-aware systems. *Hum.-Comput. Interact.*, 16(2):193–212. 38

Ben-Asher, N., Kirschnick, N., Sieger, H., Meyer, J., Ben-Oved, A., and Möller, S. (2011). On the need for different security methods on mobile phones. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*, MobileHCI '11, pages 465–473, New York, NY, USA. ACM. 115, 132, 134

Benford, S., Rowland, D., Flintham, M., Drozd, A., Hull, R., Reid, J., Morrison, J., and Facer, K. (2005). Life on the edge: supporting collaboration in location-based experiences. In van der Veer, G. C. and Gale, C., editors, *CHI*, pages 721–730. ACM. 41

Bentley, F. and Barrett, E. (2012). *Building Mobile Experiences*. The MIT Press. 4

Bentley, F. and Metcalf, C. J. (2009). The use of mobile social presence. *IEEE Pervasive Computing*, 8:35–41. 20

Bernhaupt, R., IJsselsteijn, W., Mueller, F., Tscheligi, M., and Wixon, D. R. (2008). Evaluating user experiences in games. In Czerwinski, M., Lund, A. M., and Tan, D. S., editors, *CHI Extended Abstracts*, pages 3905–3908. ACM. 44

Biddle, R., Chiasson, S., and Van Oorschot, P. (2012). Graphical passwords: Learning from the first twelve years. *ACM Comput. Surv.*, 44(4):19:1–19:41. 117, 118, 138

Bourguet, M.-L. (2006). Towards a taxonomy of error-handling strategies in recognition-based multi-modal human-computer interfaces. *Signal Process.*, 86:3625–3643. 76

Brewster, S. (1994). *Providing a structured method for integrating non-speech audio into human-computer interfaces*. PhD thesis, University of York. 79

Brewster, S. (2002). Overcoming the lack of screen space on mobile computers. *Personal Ubiquitous Comput.*, 6(3):188–205. 7

Brewster, S., Lumsden, J., Bell, M., Hall, M., and Tasker, S. (2003). Multimodal 'eyes-free' interaction techniques for wearable devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '03, pages 473–480, New York, NY, USA. ACM. 6, 7

Brooke, J. (1996). SUS: A Quick and Dirty Usability Scale. In Jordan, P. W., Thomas, B., Weerdmeester, B. A., and McClelland, I. L., editors, *Usability Evaluation in Industry*. Taylor & Francis., London. 99, 104, 107, 121, 131

Brown, B. and Chalmers, M. (2003). Tourism and mobile technology. In *Proc. ECSCW '03*, pages 335–355. Kluwer Acad. 51

Bulling, A., Alt, F., and Schmidt, A. (2012). Increasing the security of gaze-based cued-recall graphical passwords using saliency masks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 3011–3020, New York, NY, USA. ACM. 118

Bulling, A., Roggen, D., and Tröster, G. (2009). Wearable eog goggles: eye-based interaction in everyday environments. In *CHI '09: Proceedings of the 27th international conference extended abstracts on Human factors in computing systems*, pages 3259–3264, New York, NY, USA. ACM. 45

Burrell, J. and Gay, G. K. (2002). E-graffiti: Evaluating real-world use of a context-aware system. *Interacting with Computers*, 14(4):301–312. 21, 25, 27, 30, 44

Butler, A., Izadi, S., and Hodges, S. (2008). Sidesight: multi-"touch" interaction around small devices. In *Proc. UIST '08*, pages 201–204, NY, USA. ACM. 95, 97

Cakmakci, O. and Rolland, J. (2006). Head-worn displays: A review. *Journal of Display Technology*, 2(3). 5

Casson, A. J., Smith, S., Duncan, J. S., and Rodriguez-Villegas, E. (2008). Wearable eeg: what is it, why is it needed and what does it entail? In *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, pages 5867–5870. 45

Chapman, C. N. and Lahav, M. (2008). International ethnographic observation of social networking sites. In *CHI '08: CHI '08 extended abstracts on Human factors in computing systems*, pages 3123–3128, New York, USA. ACM. 31

Cheng, A.-J., Chen, Y.-Y., Huang, Y.-T., Hsu, W. H., and Liao, H.-Y. M. (2011). Personalized travel recommendation by mining people attributes from community-contributed photos. In *Proc. MM '11*, pages 83–92. ACM. 54

Cherubini, M., Gutierrez, A., de Oliveira, R., and Oliver, N. (2010). Social tagging revamped: supporting the users' need of self-promotion through persuasive techniques. In *CHI '10: Proceedings of the 28th international conference on Human factors in computing systems*, pages 985–994, New York, NY, USA. ACM. 44

Cheverst, K., Davies, N., Mitchell, K., and Friday, A. (2000). Experiences of developing and deploying a context-aware tourist guide: the guide project. In *MobiCom '00: Proceedings of the 6th annual international conference on Mobile computing and networking*, pages 20–31, New York, NY, USA. ACM. 41

Chittaro, L. (2006). Visualizing information on mobile devices. *Computer*, 39(3):40–45. 7

Chittaro, L. (2009). Distinctive aspects of mobile interaction and their implications for the design of multimodal interfaces. *Journal on Multimodal User Interfaces*, 3(3):157–165. 3, 29, 48

Clements, M., Serdyukov, P., de Vries, A., and Reinders, M. (2010a). Finding wormholes with flickr geotags. In Gurrin, C., He, Y., Kazai, G., Kruschwitz, U., Little, S., Roelleke, T., Rüger, S., and van Rijsbergen, K., editors, *Advances in Information Retrieval*, volume 5993 of *LNCS*, pages 658–661. Springer. 55

Clements, M., Serdyukov, P., de Vries, A. P., and Reinders, M. J. (2010b). Using flickr geotags to predict user travel behaviour. In *Proc. SIGIR '10*, pages 851–852. ACM. 54

Computer Security Center U.S. (1986). *Department of Defense password management guidelines*. Department of Defense Computer Security Center, [Fort George G. Meade, Md. :. 118

Comte, A. (1880). *A general view of positivism / translated from the French of Auguste Comte by J.H. Bridges*. Reeves & Turner, London, 2nd ed. edition. 42

Consolvo, S., Harrison, B. L., Smith, I. E., Chen, M. Y., Everitt, K., Froehlich, J., and Landay, J. A. (2007). Conducting in situ evaluations for and with ubiquitous computing technologies. *Int. J. Hum. Comput. Interaction*, 22(1&2):103–118. 22, 43

Consolvo, S. and Walker, M. (2003). Using the experience sampling method to evaluate ubicomp applications. *IEEE Pervasive Computing*, 2(2):24–31. 46, 47

Cramer, H., Mentis, H., and Fernaeus, Y. (2010). Serious work on playful experiences: a preliminary set of challenges. In *'Fun, seriously?' Workshop at CSCW*. Savannah, GA, USA. 38, 39, 41, 48

Cranor, L. F. and Garfinkel, S. (2005). *Security and Usability: Designing Secure Systems That People Can Use*. O'Reilly Media, illustrated edition edition. 11, 116

Csikszentmihalyi, M. (1990). *Flow: The Psychology of Optimal Experience*. Harper Perennial. 41, 110

Daugman, J. (2004). How iris recognition works. *Circuits and Systems for Video Technology, IEEE Transactions on*, 14(1):21 – 30. 118

Davenport, G., Holmquist, L. E., and Thomas, M. (1998). Fun: A condition of creative research. *IEEE MultiMedia*, 5(3):10–15. 95, 96

De Angeli, A., Coutts, M., Coventry, L., Johnson, G. I., Cameron, D., and Fischer, M. H. (2002). Vip: a visual approach to user authentication. In *Proceedings of the Working Conference on Advanced Visual Interfaces*, AVI '02, pages 316–323, New York, NY, USA. ACM. 117

De Choudhury, M., Feldman, M., Amer-Yahia, S., Golbandi, N., Lempel, R., and Yu, C. (2010). Automatic construction of travel itineraries using social breadcrumbs. In *Proc. HT '10*, pages 35–44. ACM. 54

de Leon, M. P., Balasubramaniam, S., and Donnelly, W. (2006). Creating a distributed mobile networking testbed environment-through the living labs approach. In *Proceeding on Testbeds and Research Infrastructures for the Development of Networks and Communities*, pages 134–139. 2nd International IEEE/Create-Net Conference. 45

Deterding, S., Sicart, M., Nacke, L., O'Hara, K., and Dixon, D. (2011). Gamification. using game-design elements in non-gaming contexts. In *Proc. CHI EA '11*, pages 2425–2428, NY, USA. ACM. 34, 112, 114

Dey, A. K. (2001). Understanding and using context. *Personal & Ubiquitous Computing*, 5(1):4–7. 38

Dey, A. K. and Abowd, G. D. (1999). Towards a better understanding of context and context-awareness. In *In HUC '99: Proceedings of the 1st international symposium on Handheld and Ubiquitous Computing*, pages 304–307. Springer-Verlag. 154

Dey, A. K., Abowd, G. D., and Salber, D. (2001). A conceptual framework and a toolkit for supporting the rapid prototyping of context-aware applications. *Human-Computer Interaction*, 16(2-4):97–166. 4, 38, 154

Dhamija, R. and Perrig, A. (2000). Déjà vu: a user study using images for authentication. In *Proceedings of the 9th conference on USENIX Security Symposium - Volume 9*, SSYM'00, pages 4–4, Berkeley, CA, USA. USENIX Association. 117

Dourish, P. (2001). *Where the action is: the foundations of embodied interaction*. MIT Press, Cambridge, MA, USA. 43

Dourish, P. (2004). What we talk about when we talk about context. *Personal and Ubiquitous Computing*, 8(1):19–30. 39, 42

Dourish, P., Grinter, E., Delgado de la Flor, J., and Joseph, M. (2004). Security in the wild: user strategies for managing security as an everyday, practical problem. *Personal Ubiquitous Comput.*, 8(6):391–401. 120

El Ali, A. and Ketabdar, H. (2013). Magnet-based around device interaction for playful music composition and gaming. To be published in IJMHCI. 8, 15, 95

El Ali, A., Ketabdar, H., and Nack, F. (2013a). Investigating the usability and security trade-off in magnetic gestural authentication. Under Review; IJHCS. 8, 15, 115

El Ali, A., Kildal, J., and Lantz, V. (2012). Fishing or a z?: investigating the effects of error on mimetic and alphabet device-based gesture interaction. In *Proceedings of the 14th ACM international conference on Multimodal interaction*, ICMI '12, pages 93–100, New York, NY, USA. ACM. 8, 14, 73, 137

El Ali, A. and Nack, F. (2009). Touring in a living lab: some methodological considerations. In *Proceedings of the Mobile Living Labs 09*, pages 23–26, Enschede, NL. MobileHCI '09, Novay. 45, 46

El Ali, A., Nack, F., and Hardman, L. (2010). Understanding contextual factors in location-aware multimedia messaging. In *Proceedings of the 12th international conference on Multimodal Interfaces*. 8, 14, 19, 43, 155

El Ali, A., Nack, F., and Hardman, L. (2011). Good times?! 3 problems and design considerations for playful hci. *International Journal of Mobile Human-Computer Interaction*, 3(3):50–65. 8, 14, 37

El Ali, A., van Sas, S. N., and Nack, F. (2013b). Photographer paths: sequence alignment of geotagged photos for exploration-based route planning. In *Proceedings of the 2013 conference on Computer supported cooperative work*, CSCW '13, pages 985–994, New York, NY, USA. ACM. 8, 14, 51

Ericsson, A. K. and Kintsch, W. (1995). Long-term working memory. *Psychological Review*, 102(2):211–245. 138

Eriksson, M., Niitamo, V., and Kulkki, S. (2005). State-of-the-art in utilizing living labs approach to user-centric ict innovation – a european approach. Technical report, Luleå University of Technology, Sweden. 45

Fabbrizio, G. D., Tur, G., and Hakkani-Tur, D. (2005). Automated wizard-of-oz for spoken dialogue systems. In *Proc. INTERSPEECH 2005*, pages 1857–1860. 78

Farella, E., O'Modhrain, S., Benini, L., and Riccó, B. (2006). Gesture signature for ambient intelligence applications: a feasibility study. In *Proceedings of the 4th international conference on Pervasive Computing*, PERVASIVE'06, pages 288–304, Berlin, Heidelberg. Springer-Verlag. 118

Faundez-Zanuy, M. (2004). On the vulnerability of biometric security systems. *Aerospace and Electronic Systems Magazine, IEEE*, 19(6):3 – 8. 118

Fischer, P. T. and Hornecker, E. (2012). Urban hci: spatial aspects in the design of shared encounters for media facades. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 307–316, New York, NY, USA. ACM. 4

Forman, G. and Zahorjan, J. (1994). The challenges of mobile computing. *Computer*, 27(4):38–47. 7

Franco, I. (2005). The airstick: a free-gesture controller using infrared sensing. In *Proc. NIME '05*, pages 248–249, Singapore. NUS. 98

Froehlich, J., Chen, M. Y., Consolvo, S., Harrison, B., and Landay, J. A. (2007). Myexperience: a system for in situ tracing and capturing of user feedback on mobile phones. In *MobiSys '07: Proceedings of the 5th international conference on Mobile systems, applications and services*, pages 57–70, New York, NY, USA. ACM. 46

Gajadhar, B., Nap, H., de Kort, Y., and IJsselsteijn, W. (2010). 'let's sweep some mines together': social interaction & competition in casual games. In Calvi, L. and Nuijten, K., editors, *Playability and player experience: Proc. Fun and Games 2010 Workshop, 15-17, Leuven (B)*, pages 19 – 30, Leuven. NHTV Press, NHTV Press. 99

Gillian, N., O'Modhrain, S., and Essl, G. (2009). Scratch-off: A gesture based mobile music game with tactile feedback. In *Proc. NIME '09*. Pittsburgh, USA. 98

Gillian, N. and Paradiso, J. A. (2012). Digito: A fine-grain gesturally controlled virtual musical instrument. In *Proc. NIME '12*. Ann Arbor, USA. 98

Girardin, F., Calabrese, F., Fiore, F. D., Ratti, C., and Blat, J. (2008). Digital footprinting: Uncovering tourists with user-generated content. *IEEE Pervasive Computing*, 7:36–43. 54

Grandhi, S. A., Joue, G., and Mittelberg, I. (2011). Understanding naturalness and intuitiveness in gesture production: insights for touchless gestural interfaces. In *Proc. CHI '11*, pages 821–824, NY, USA. ACM. 75, 96, 97, 113, 117

Greenberg, S. and Buxton, B. (2008). Usability evaluation considered harmful (some of the time). In *CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, pages 111–120, New York, NY, USA. ACM. 44, 45

Greenfield, A. (2006). *Everyware: The Dawning Age of Ubiquitous Computing*. New Riders Publishing, Berkeley, CA. 4, 39, 50

Griswold, W. G., Shanahan, P., Brown, S. W., Boyer, R., Ratto, M., and et al. (2003). Activecampus – experiments in community-oriented ubiquitous computing. *IEEE Computer*, 37:73–81. 21, 23

Guse, D., Kirschnick, N., Kratz, S., and Möller, S. (2011). Gesture-based user authentication for mobile devices. 115, 116

Hägerstrand, T. (1970). What about people in Regional Science? *Papers in Regional Science*, 24(1):6–21. 53

Halverson, C., Horn, D., Karat, C., and Karat, J. (1999). The beauty of errors: patterns of error correction in desktop speech systems. In *Proc. INTERACT '99*, pages 133–140. IOS Press. 76

Han, X., Seki, H., Kamiya, Y., and Hikizu, M. (2009). Wearable handwriting input device using magnetic field: Geomagnetism cancellation in position calculation. *Precision Engineering*, 33(1):37 – 43. 98

Harrison, C. and Hudson, S. E. (2009). Abracadabra: wireless, high-precision, and unpowered finger input for very small mobile devices. In *Proc. UIST '09*, pages 121–124, NY, USA. ACM. 98

Hart, S. and Wickens, C. (1990). *Manprint: an Approach to Systems Integration*, chapter Workload Assessment and Prediction, pages 257–292. Van Nostrand Reinhold. 79, 121, 122

Hart, S. G. and Stavenland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In Hancock, P. A. and Meshkati, N., editors, *Human Mental Workload*, chapter 7, pages 139–183. Elsevier. 131

Hassenzahl, M., Burmester, M., and Koller, F. (2003). AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. *Mensch & Computer 2003. Interaktion in Bewegung*, pages 187–196. 61, 98, 104

Hassenzahl, M. and Sandweg, N. (2004). From mental effort to perceived usability: transforming experiences into summary assessments. In *CHI '04: CHI '04 extended abstracts on Human factors in computing systems*, pages 1283–1286, New York, USA. ACM. 40

Hassenzahl, M. and Tractinsky, N. (2006). User experience - a research agenda. *Behaviour & Information Technology*, 25(2):91–97. 40

Hecht, B., Schöning, J., Erickson, T., and Priedhorsky, R. (2011). Geographic human-computer interaction. In *Proc. CHI '11 EA*, pages 447–450. ACM. 52

Hochmair, H. (2004). Towards a classification of route selection criteria for route planning tools. *Developments in Spatial Data Handling*, pages 481–492. 51, 55

Hoggan, E. and Brewster, S. A. (2010). Crosstrainer: testing the use of multimodal interfaces in situ. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 333–342, New York, NY, USA. ACM. 155

Hoggan, E., Raisamo, R., and Brewster, S. A. (2009). Mapping information to audio and tactile icons. In *Proceedings of the 2009 international conference on Multimodal interfaces*, ICMI-MLMI '09, pages 327–334, New York, NY, USA. ACM. 155

Hughes, S. (2006). Shake - sensing hardware accessory for kinaesthetic expression model sk6. In *SAMH Engineering Services*. Blackrock, Ireland. 120

Husserl, E. (1893-1917). *On the Phenomenology of the Consciousness of Internal Time*. Springer. 42

Iso, I. O. F. S. (1998). International ergonomic requirements for office work with visual display terminals (vdts) - part 11 : Guidance on usability. *International Organization*, 1998(2):28. 7

ISO DIS 9241-210:1994 (1994). *Ergonomics of human system interaction - Part 210: Human-centred design for interactive systems (formerly known as 13407)*. International Organization for Standardization (ISO), Switzerland. 7, 40, 75, 96, 117

Jacucci, G., Oulasvirta, A., and Salovaara, A. (2007). Active construction of experience through mobile media: a field study with implications for recording and sharing. *Personal and Ubiquitous Computing*, 11(4):215–234. 21

Jain, J., Lund, A., and Wixon, D. (2011). The future of natural user interfaces. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '11, pages 211–214, New York, NY, USA. ACM. 6, 75, 155

Jian, J.-Y., Bisantz, A. M., and Drury, C. G. (2000). Foundations for an empirically determined scale of trust in automated systems. *International Journal of Cognitive Ergonomics*, 4(1):53–71. 121, 132

Jiang, X., Chen, N. Y., Hong, J. I., Wang, K., Takayama, L., and Landay, J. A. (2004). Siren: Context-aware computing for firefighting. In *Pervasive*, volume 3001, pages 87–105. 41

Kane, S. K., Wobbrock, J. O., and Smith, I. E. (2008). Getting off the treadmill: evaluating walking user interfaces for mobile devices in public spaces. In *Proceedings of the 10th international conference on Human computer interaction with mobile devices and services*, MobileHCI '08, pages 109–118, New York, NY, USA. ACM. 3

Kapoor, A. and Horvitz, E. (2008). Experience sampling for building predictive user models: a comparative study. In *CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, pages 657–666, New York, NY, USA. ACM. 29

Karam, M. and Schraefel, M. C. (2005). A taxonomy of gestures in human computer interactions. Technical report, University of Southampton. 147

Karam, M. and Schraefel, M. C. (2006). Investigating user tolerance for errors in vision-enabled gesture-based interactions. In *Proc. AVI '06*, pages 225–232, NY, USA. ACM. 76, 78, 82, 84, 90, 93

Kärkkäinen, L. and Laarni, J. (2002). Designing for small display screens. In *Proceedings of the second Nordic conference on Human-computer interaction*, NordiCHI '02, pages 227–230, New York, NY, USA. ACM. 3

Kayali, F., Pichlmair, M., and Kotik, P. (2008). Mobile tangible interfaces as gestural instruments. In *Mobile Music Workshop*. 98

Kellar, M., Reilly, D., Hawkey, K., Rodgers, M., MacKay, B., Dearman, D., Ha, V., MacInnes, W. J., Nunes, M., Parker, K., Whalen, T., and Inkpen, K. M. (2005). It's a jungle out there: practical considerations for evaluation in the city. In *CHI '05: CHI '05 extended abstracts on Human factors in computing systems*, pages 1533–1536, New York, NY, USA. ACM. 45

Kenteris, M., Gavalas, D., and Economou, D. (2011). Electronic mobile guides: a survey. *Personal and Ubiquitous Computing*, 15:97–111. 51

Ketabdar, H., Jahanbekam, A., Yuksel, K. A., Hirsch, T., and Haji Abolhassani, A. (2011). Magimusic: using embedded compass (magnetic) sensor for touch-less gesture based interaction with digital music instruments in mobile devices. In *Proc. TEI '11*, pages 241–244. ACM. 96, 98

Ketabdar, H., Moghadam, P., Naderi, B., and Roshandel, M. (2012). Magnetic signatures in air for mobile devices. In *Proceedings of the 14th international conference on Human-computer interaction with mobile devices and services companion*, MobileHCI '12, pages 185–188, New York, NY, USA. ACM. 96, 98

Ketabdar, H., Roshandel, M., and Yüksel, K. A. (2010a). Towards using embedded magnetic field sensor for around mobile device 3d interaction. In *Proc. MobileHCI '10*, pages 153–156, NY, USA. ACM. 95, 98

Ketabdar, H., Yuksel, K. A., JahanBekam, A., Roshandel, M., and Skripko, D. (2010b). Magisign: User identification/authentication based on 3d around device magnetic signatures. In *Proc. UBICOMM '10*. 77, 149

Ketabdar, H., Yüksel, K. A., and Roshandel, M. (2010c). Magitact: interaction with mobile devices based on compass (magnetic) sensor. In *Proc. IUI '10*, pages 413–414, NY, USA. ACM. 6, 99

Kjeldskov, J., Graham, C., Pedell, S., Vetere, F., Howard, S., Balbo, S., and Davies, J. (2005). Evaluating the usability of a mobile guide: the influence of location, participants and resources. *Behaviour and Information Technology*, 24(1):51–65. 60

Klein, D. V. (1990). "foiling the cracker": A survey of, and improvements to, password security. Technical report, Carnegie Mellon University. 118

Kostakos, V., Camacho, T., and Mantero, C. (2013). Towards proximity-based passenger sensing on public transport buses. *Personal and Ubiquitous Computing*, pages 1–10. 154

Kostakos, V., Nicolai, T., Yoneki, E., O'Neill, E., Kenn, H., and Crowcroft, J. (2009). Understanding and measuring the urban pervasive infrastructure. *Personal Ubiquitous Comput.*, 13(5):355–364. 46, 49, 52

Kostakos, V., O'Neill, E., Penn, A., Roussos, G., and Papadongonas, D. (2010). Brief encounters: Sensing, modeling and visualizing urban mobility and copresence networks. *ACM Trans. Comput.-Hum. Interact.*, 17(1):1–38. 46

Kramer, M. A. M., Reponen, E., and Obrist, M. (2008). Mobimundi: exploring the impact of user-generated mobile content – the participatory panopticon. In *Proceedings of the 10th international conference on Human computer interaction with mobile devices and services*, MobileHCI '08, pages 575–577, New York, NY, USA. ACM. 144

Kratz, S., Guse, D., Rohs, M., Müller, J., Bailly, G., and Nischt, M. (2012). Palmspace: Continuous around-device gestures vs. multitouch for 3d rotation tasks on mobile devices. In *Proc. AVI '12*, NY, USA. ACM. 98

Kratz, S. and Rohs, M. (2009). Hoverflow: exploring around-device interaction with ir distance sensors. In *Proc. MobileHCI '09*, pages 1–4, NY, USA. ACM. 95, 97

Kray, C., Nesbitt, D., Dawson, J., and Rohs, M. (2010). User-defined gestures for connecting mobile phones, public displays, and tabletops. In *Proc. MobileHCI '10*, pages 239–248, NY, USA. ACM. 74, 79

Kumar, M., Garfinkel, T., Boneh, D., and Winograd, T. (2007). Reducing shoulder-surfing by using gaze-based password entry. In *Proceedings of the 3rd symposium on Usable privacy and security*, SOUPS '07, pages 13–19, New York, NY, USA. ACM. 118

Kuniavsky, M. (2003). *Observing the User Experience: A Practitioner's Guide to User Research (Morgan Kaufmann Series in Interactive Technologies) (The Morgan Kaufmann Series in Interactive Technologies)*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA. 45

Kurashima, T., Iwata, T., Irie, G., and Fujimura, K. (2010). Travel route recommendation using geotags in photo sharing sites. In *Proc. CIKM '10*, pages 579–588. ACM. 54, 61

Kurtenbach, G. (2010). Pen-based computing. *XRDS*, 16(4):14–20. 109

Khnel, C., Westermann, T., Hemmert, F., Kratz, S., Mller, A., and Mller, S. (2011). I'm home: Defining and evaluating a gesture set for smart-home control. *International Journal of Human-Computer Studies*, 69(11):693 – 704. 75, 97, 117

Law, E. L.-C., Roto, V., Hassenzahl, M., Vermeeren, A. P., and Kort, J. (2009). Understanding, scoping and defining user experience: a survey approach. In *CHI '09: Proceedings of the 27th international conference on Human factors in computing systems*, pages 719–728, New York, NY, USA. ACM. 39, 40

Li, K. A., Baudisch, P., and Hinckley, K. (2008). Blindsight: eyes-free access to mobile phones. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '08, pages 1389–1398, New York, NY, USA. ACM. 6

Lim, M. Y. and Aylett, R. (2009). An emergent emotion model for an affective mobile guide with attitude. *Applied Artificial Intelligence*, 23(9):835–854. 41

Lin, C. L., Sun, H. M., and Hwang, T. (2001). Attacks and solutions on strong-password authentication. *IEICE Transactions on Communications*, E84-B(9):2622–2627. 118

Love, S. (2005). *Understanding Mobile Human-Computer Interaction*. Information Systems Series (ISS). Elsevier Science. 3

Lu, X., Wang, C., Yang, J.-M., Pang, Y., and Zhang, L. (2010). Photo2trip: generating travel routes from geo-tagged photos for trip planning. In *MM '10*, pages 143–152. ACM. 54

Lucero, A., Holopainen, J., and Jokela, T. (2011). Pass-them-around: collaborative use of mobile phones for photo sharing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 1787–1796, New York, NY, USA. ACM. 98

Lucero, A., Lyons, K., Vetek, A., Järvenpää, T., White, S., and Salmimaa, M. (2013). Exploring the interaction design space for interactive glasses. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '13, pages 1341–1346, New York, NY, USA. ACM. 5

Mandryk, R., Inkpen, K., and Calvert, T. (2006). Using psychophysiological techniques to measure user experience with entertainment technologies. *Behaviour and Information Technology*, 25(2):141–158. 44

Mankoff, J. and Abowd, G. D. (1999). Error correction techniques for handwriting, speech, and other ambiguous or error prone systems. Technical report, GIT-GVU. 76

Mann, S. (1997). Wearable computing: A first step toward personal imaging. *Computer*, 30(2):25–32. 5

Marshall, M. T. and Wanderley, M. M. (2006). Vibrotactile feedback in digital musical instruments. In *NIME '06*, pages 226–229. 110

McFarlane, D. C. and Latorella, K. A. (2002). The scope and importance of human interruption in human-computer interaction design. *Hum.-Comput. Interact.*, 17(1):1–61. 155

Mcfarlane, D. C. and Mcfarlane, D. C. (1997). *Interruption of People in Human-Computer Interaction: A General Unifying Definition of Human Interruption and Taxonomy*. PhD thesis, George Washington University. 3, 155

Minsky, M. L. and Papert, S. A. (1969). *Perceptrons*. Cambridge, MA: MIT Press. 99

Monrose, F. and Rubin, A. D. (2000). Keystroke dynamics as a biometric for authentication. *Future Generation Computer Systems*, 16(4):351–359. 138

Morris, R. and Thompson, K. (1979). Password security: a case history. *Commun. ACM*, 22(11):594–597. 118, 134

Müller, J., Alt, F., Michelis, D., and Schmidt, A. (2010). Requirements and design space for interactive public displays. In *Proceedings of the international conference on Multimedia*, MM '10, pages 1285–1294, New York, NY, USA. ACM. 5

Nack, F. (2003). Capturing experience - a matter of contextualising events. In *Proceedings of the 2003 ACM SIGMM workshop on Experiential telepresence*, New York, NY, USA. ACM. 20, 40

Nacke, L. E., Drachen, A., Kuikkaniemi, K., Niesenhaus, J., Korhonen, H. J., van den Hoogen, W. M., Poels, K., IJsselsteijn, W. A., and de Kort, Y. A. W. (2009). Playability and player experience research. In *Proceedings of DiGRA 2009: Breaking New Ground: Innovation in Games, Play, Practice and Theory*. DIGRA, London, UK. 41, 44, 45

Norman, D. A. (2010). Natural user interfaces are not natural. *interactions*, 17(3):6–10. 6, 75, 113

Obrenovic, Z., Abascal, J., and Starcevic, D. (2007). Universal accessibility as a multimodal design issue. *Commun. ACM*, 50(5):83–88. 3

Okuyama, K. and Yanai, K. (2011). A travel planning system based on travel trajectories extracted from a large number of geotagged photos on the web. In *PCM*. 54

Oulasvirta, A. and Bergstrom-Lehtovirta, J. (2010). A simple index for multimodal flexibility. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 1475–1484, New York, NY, USA. ACM. 154

Oulasvirta, A. and Bergstrom-Lehtovirta, J. (2011). Ease of juggling: studying the effects of manual multitasking. In *Proc. CHI '11*, pages 3103–3112, NY, USA. ACM. 73, 110

Oulasvirta, A., Tamminen, S., Roto, V., and Kuorelahti, J. (2005). Interaction in 4-second bursts: the fragmented nature of attentional resources in mobile hci. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '05, pages 919–928, New York, NY, USA. ACM. 3, 6

Oviatt, S. (2000). Taming recognition errors with a multimodal interface. *Commun. ACM*, 43:45–51. 76

Oviatt, S. (2003). *The human-computer interaction handbook*, chapter Multimodal interfaces, pages 286–304. L. Erlbaum Associates Inc., Hillsdale, NJ, USA. 5

Oviatt, S., Maceachern, M., and anne Levow, G. (1998). Predicting hyperarticulate speech during human-computer error resolution. *Speech Communication*, 24:87–110. 74, 76, 84, 92

Oviatt, S. and VanGent, R. (1996). Error resolution during multimodal human-computer interaction. In *Proc. ICSLP '96*, volume 1, pages 204–207. 76

Papliatseyeu, A. and Mayora Ibarra, O. (2008). Nailing the reality with GeoMedia: location-aware multimedia tags. In *Proceedings of MobiMedia'08 Conference*, Oulu, Finland. ACM. (to appear). 21

Parhi, P., Karlson, A. K., and Bederson, B. B. (2006). Target size study for one-handed thumb use on small touchscreen devices. In *Proceedings of the 8th conference on Human-computer interaction with mobile devices and services*, MobileHCI '06, pages 203–210, New York, NY, USA. ACM. 3

Patel, S. N., Pierce, J. S., and Abowd, G. D. (2004). A gesture-based authentication scheme for untrusted public terminals. In *Proc. UIST '04*, pages 157–160, NY, USA. ACM. 115

Paulos, E. and Jenkins, T. (2005). Urban probes: Encountering our emerging urban atmospheres. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 341–350. CHI'05. 4

Persson, P. and Fagerberg, P. (2002). Geonotes: a real-use study of a public location-aware community system. Technical Report. 21, 23, 32, 44

Poels, K., de Kort, Y., and Ijsselsteijn, W. (2007). "It is always a lot of fun!": exploring dimensions of digital game experience using focus group methodology. In *Future Play '07: Proceedings of the 2007 conference on Future Play*, pages 83–89, New York, NY, USA. ACM. 41

Prabhakar, S., Pankanti, S., and Jain, A. (2003). Biometric recognition: Security and privacy concerns. *Security & Privacy, IEEE*, 1(2):33–42. 118

Preece, J., Rogers, Y., and Sharp, H. (2002). *Interaction Design*. John Wiley & Sons, Inc., New York, NY, USA, 1st edition. 1

Rabiner, L. (1989). A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2):257 –286. 116

Raghunath, M. T. and Narayanaswami, C. (2002). User interfaces for applications on a wrist watch. *Personal Ubiquitous Comput.*, 6(1):17–30. 5

Read, J. C., MacFarlane, S., and Casey, C. (2002). Oops! silly me! errors in a handwriting recognition-based text entry interface for children. In *Proc. NordiCHI '02*, pages 35–40, NY, USA. ACM. 74, 76, 92

Renaud, K. and Angeli, A. D. (2004). My password is here! an investigation into visuo-spatial authentication mechanisms. *Interacting with Computers*, 16(6):1017–1041. 117

Renaud, K. and Smith, E. (2001). Jiminy: "helping users to remember their passwords. In *Annual Conference of the South African Institute of Computer Scientists and Information Technologists*. 117

Rico, J. and Brewster, S. (2009). Gestures all around us: user differences in social acceptability perceptions of gesture based interfaces. In *Proc. MobileHCI '09*, pages 64:1–64:2, NY, USA. ACM. 3

Rico, J. and Brewster, S. (2010). Usable gestures for mobile interfaces: evaluating social acceptability. In *Proc. CHI '10*, pages 887–896, NY, USA. ACM. 74, 75, 91, 97, 117, 137

Rime, B. and Schiaratura, L. (1991). *Fundamentals of Nonverbal Behavior*, chapter Gesture and speech, pages 239–281. Cambridge University Press. 74, 77

Roth, V., Richter, K., and Freidinger, R. (2004). A pin-entry method resilient against shoulder surfing. In *Proceedings of the 11th ACM conference on Computer and communications security*, CCS '04, pages 236–245, New York, NY, USA. ACM. 118

Ruiz, J., Li, Y., and Lank, E. (2011). User-defined motion gestures for mobile interaction. In *Proc. CHI '11*, pages 197–206, NY, USA. ACM. 73, 74, 77, 117

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39:1161–1178. 31

Sae-Bae, N., Ahmed, K., Isbister, K., and Memon, N. (2012). Biometric-rich gestures: a novel approach to authentication on multi-touch devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 977–986, New York, NY, USA. ACM. 138

Saffer, D. (2013). *Microinteractions: Designing With Details*. O'Reilly Media, Incorporated. 154

Sahami Shirazi, A., Moghadam, P., Ketabdar, H., and Schmidt, A. (2012). Assessing the vulnerability of magnetic gestural authentication to video-based shoulder surfing attacks. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems*, CHI '12, pages 2045–2048, New York, NY, USA. ACM. 116, 123, 149

Sakoe, H. and Chiba, S. (1978). Dynamic programming algorithm optimization for spoken word recognition. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 26(1):43 – 49. 99, 116

Schöning, J., Hecht, B., and Starosielski, N. (2008). Evaluating automatically generated location-based stories for tourists. In *CHI EA '08*, pages 2937–2942. ACM. 60

Schultz, E. E., Proctor, R. W., Lien, M. C., and Salvendy, G. (2001). Usability and Security An Appraisal of Usability Issues in Information Security Methods. *Computers & Security*, 20(7):620–634. 116

Shanmugapriya, D. and Padmavathi, G. (2009). A survey of biometric keystroke dynamics: Approaches, security and challenges. *CoRR*, abs/0910.0817. 118, 138, 139

Shoval, N. and Isaacson, M. (2007). Sequence alignment as a method for human activity analysis in space and time. *Annals of the Association of American Geographers*, 97(2):282–297. 53, 54

Sieger, H., Kirschnick, N., and Möller, S. (2010). User preferences for biometric authentication methods and graded security on smartphones. In *SOUPS '10 Symposium on Usable Privacy and Security*. ACM. 115, 116

Singh, V. K., Jain, R., and Kankanhalli, M. S. (2009). Motivating contributors in social media networks. In *WSM '09: Proceedings of the first SIGMM workshop on Social media*, pages 11–18, New York, USA. ACM. 31, 44, 48

Snavely, N., Garg, R., Seitz, S. M., and Szeliski, R. (2008). Finding paths through the world's photos. In *SIGGRAPH '08*, pages 15:1–15:11. ACM. 54

Sohn, T., Li, K. A., Lee, G., Smith, I., Scott, J., and Griswold, W. G. (2005). Place-its: A study of location-based reminders on mobile phones. In *Ubicomp*, pages 232–250. Springer. 21, 30

Ståhl, A., Höök, K., Svensson, M., Taylor, A. S., and Combetto, M. (2009). Experiencing the affective diary. *Personal and Ubiquitous Computing*, 13:365–378. 21, 23, 34

Suh, Y., Shin, C., Woo, W., Dow, S., and Macintyre, B. (2011). Enhancing and evaluating users' social experience with a mobile phone guide applied to cultural heritage. *Personal Ubiquitous Comput.*, 15:649–665. 61

Suhm, B., Myers, B., and Waibel, A. (2001). Multimodal error correction for speech user interfaces. *ACM Trans. Comput.-Hum. Interact.*, 8:60–98. 76

Suo, X., Zhu, Y., and Owen, G. (2005). Graphical passwords: a survey. In *Computer Security Applications Conference, 21st Annual*, pages 10 pp. –472. 117

Tamminen, S., Oulasvirta, A., Toiskallio, K., and Kankainen, A. (2004). Understanding mobile contexts. *Personal Ubiquitous Comput.*, 8(2):135–143. 3, 6

Tari, F., Ozok, A. A., and Holden, S. H. (2006). A comparison of perceived and real shoulder-surfing risks between alphanumeric and graphical passwords. In *Proceedings of the second symposium on Usable privacy and security*, SOUPS '06, pages 56–66, New York, NY, USA. ACM. 117, 118

ten Holt, G. A., Reinders, M. J., and Hendriks, E. A. (2007). Multi-dimensional dynamic time warping for gesture recognition. In *Thirteenth annual conference of the Advanced School for Computing and Imaging*. 119

Theremin, L. S. (1996). The design of a musical instrument based on cathode relays. *Leonardo Music J.*, 6:49–50. 98

Tuite, K., Snavely, N., Hsiao, D.-y., Tabing, N., and Popovic, Z. (2011). Photocity: training experts at large-scale image acquisition through a competitive game. In *CHI '11*, CHI '11, pages 1383–1392. ACM. 54

Tulving, E. (1993). What is episodic memory? *Current Directions in Psychological Science*, pages 67–70. 20, 40, 138

Tulving, E. (2002). Episodic memory: From mind to brain. *Annu. Rev. Psychol.*, 53:1–25. 40, 138

Turk, M. and Robertson, G. (2000). Perceptual user interfaces (introduction). *Commun. ACM*, 43(3):32–34. 5

Uludag, U. and Jain, A. K. (2004). Attacks on biometric systems: A case study in fingerprints. In *Proc. SPIE-EI 2004, Security, Seganography and Watermarking of Multimedia Contents VI*, pages 622–633. 118

van der Zwaag, M. D., van den Broek, E. L., and Janssen, J. H. (2010). Guidelines for biosignal driven hci. In Girouard, A., Mandryk, R., Nacke, L., Solovey, E. T., Tan, D., and Jacob, R. J. K., editors, *Proceedings of ACM CHI2010 workshop - Brain, Body, and Bytes: Physiological user interaction, Atlanta, GA, USA*, pages 77–80, New York. ACM. 148

Vazquez-Alvarez, Y. and Brewster, S. A. (2011). Eyes-free multitasking: the effect of cognitive load on mobile spatial audio interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 2173–2176, New York, NY, USA. ACM. 6, 7

Vermeulen, J. (2010). Improving intelligibility and control in ubicomp. In *Proc. Ubicomp'10*, pages 485–488. ACM. 62, 68

Webster, R. and Nakagawa, N. (1998). An interface-oriented approach to character recognition based on a dynamic model. *Pattern Recognition*, 31(2):193–203. 76, 84, 87

Weir, C. S., Douglas, G., Richardson, T., and Jack, M. (2010). Usable security: User preferences for authentication methods in ebanking and the effects of experience. *Interact. Comput.*, 22(3):153–164. 116, 139

Weiser, M. (1991). The computer for the 21st century. *Scientific American*, 265(3):66–75. 4

Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta Psychologica*, 41(1):67 – 85. 139

Wiedenbeck, S., Waters, J., Birget, J.-C., Brodskiy, A., and Memon, N. (2005). Passpoints: design and longitudinal evaluation of a graphical password system. *Int. J. Hum.-Comput. Stud.*, 63(1-2):102–127. 117

Wigelius, H. and Väätäjä, H. (2009). Dimensions of context affecting user experience in mobile work. In *INTERACT '09: Proceedings of the 12th IFIP TC 13 International Conference on Human-Computer Interaction*, pages 604–617, Berlin, Heidelberg. Springer-Verlag. 40, 41

Williams, A. and Dourish, P. (2006). Imagining the city: The cultural dimensions of urban computing. *IEEE Pervasive Computing*, pages 38–43. 19, 38

Williamson, J., Murray-Smith, R., and Hughes, S. (2007). Shoogle: excitatory multimodal interaction on mobile devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '07, pages 121–124, New York, NY, USA. ACM. 7

Wilson, C. (2008). Activity patterns in space and time: calculating representative hagerstrand trajectories. *Transportation*, 35:485–499. 55

Wilson, M. L., Mackay, W., Chi, E., Bernstein, M., Russell, D., and Thimbleby, H. (2011). Replichi - chi should be replicating and validating results more: discuss. In *PART 2 — —— Proceedings of the 2011 annual conference extended abstracts on Human factors in computing systems*, CHI EA '11, pages 463–466, New York, NY, USA. ACM. 149

Wobbrock, J. O., Morris, M. R., and Wilson, A. D. (2009). User-defined gestures for surface computing. In *Proceedings of the 27th international conference on Human factors in computing systems*, CHI '09, pages 1083–1092, New York, NY, USA. ACM. 73

Wolf, K., Naumann, A., Rohs, M., and Müller, J. (2011). Taxonomy of microinteractions: defining microgestures based on ergonomic and scenario-dependent requirements. In *Proceedings of the 13th IFIP TC 13 international conference on Human-computer interaction - Volume Part I*, INTERACT'11, pages 559–575, Berlin, Heidelberg. Springer-Verlag. 154

Yee, K.-P. (2004). Aligning security and usability. *Security Privacy, IEEE*, 2(5):48 –55. 116

Yi, B., Cao, X., Fjeld, M., and Zhao, S. (2012). Exploring user motivations for eyes-free interaction on mobile devices. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems*, CHI '12, pages 2789–2792, New York, NY, USA. ACM. 6, 153

Zhang, H., Law, E., Miller, R. C., Gajos, K. Z., Parkes, D. C., and Horvitz, E. (2012). Human computation tasks with global constraints. In *CHI'12*, pages 217–226. ACM. 55

Zhao, S., Dragicevic, P., Chignell, M., Balakrishnan, R., and Baudisch, P. (2007). Earpod: eyes-free menu selection using touch input and reactive audio feedback. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '07, pages 1395–1404, New York, NY, USA. ACM. 6, 7, 155

Zwick, C., Schmitz, B., Kühl, K., and Dewsbury, V. (2005). *Designing for small screens: mobile phones, smart phones, PDAs, pocket PCs, navigation systems, MP3 players, game consoles.* AVA Academia, Lausanne. 3

Summary

In the last 20 years, the widespread adoption of personal, mobile computing devices in everyday life, has allowed entry into a new technological era in Human Computer Interaction (HCI). The constant change of the physical and social context in a user's situation made possible by the portability of mobile devices means that the user's attention becomes limited. This can negatively impact the user experience. To deal with this problem, this thesis draws from two developments in HCI, context-awareness and 3D gestural input. From these developments, we introduce the concept of minimal mobile HCI, a subset of eyes-free mobile interaction that allows minimal combination of the visual modality with other sensory modalities to improve the user experience of interacting with smartphones.

In the first part, we look closely at the design and evaluation of location-aware multimedia messaging systems, and how they can lead to playfulness in a minimal interaction setting. We then look at how urban interactions are connected across locations by designing and evaluating an exploration-based route planner that makes use of large amounts of geotagged data. In the second part, we look closely at the usability and user experience issues associated with 3D mobile gestural input when recognition errors occur. Upon showing that this form of interaction is usable in the face of errors, we then investigate two example applications. We revisit the playfulness domain, and investigate how 3D gestural input can be applied to music composition and gaming. Finally, we look at how minimal mobile interaction can be used to support mobile user authentication using 3D gestural input.

Together, our user studies show that interaction designers need not abandon screen-based interaction, nor stop designing for users' visual modality, only complementing it with context-awareness or 3D gestural input solutions. This can expand the design space for designing and developing mobile systems/applications that keep user interaction costs at a minimum.

Samenvatting

In de afgelopen twintig jaar heeft de wijdverspreide acceptatie van persoonlijke, mobiele computerapparatuur in het dagelijks leven, een nieuw technologisch tijdperk ingeluid op het gebied van mens-computerinteractie (MCI). Constante veranderingen van de fysieke- en sociale context van de gebruikerssituatie mogelijk gemaakt door de inherente draagbaarheid van mobiele apparatuur hebben geresulteerd in een slechts beperkte aandacht van de gebruiker voor de mobiele apparatuur en haar applicaties. Dit kan de gebruikerservaring negatief benvloeden. Om met dit probleem om te gaan, richt het onderzoek beschreven in dit proefschrift zich op twee ontwikkelingen binnen het MCI domein, namelijk context awareness en 3D gestural input. Ontleent aan deze ontwikkelingen, introduceren wij het concept van minimaal mobiele MCI; een onderdeel van zogenaamde eyes-free mobiele interactie welke een minimale combinatie van visuele modaliteit mogelijk maakt met andere sensormodaliteiten, waardoor de gebruikerservaring van het smartphonegebruik verbeterd kan worden.

In het eerste gedeelte van dit proefschrift richten we ons op het ontwerp en de validatie van systemen voor locatiebewuste multimediale berichten, en hoe deze kunnen leiden tot speelse interactie in een minimale setting. Vervolgens kijken we naar hoe stedelijke interacties met elkaar verbonden zijn over verschillende locaties, middels het ontwerpen en valideren van een exploration-based route planner welke gebruik maakt van een grote hoeveelheid geotagged data.

In het tweede gedeelte van dit proefschrift onderzoeken we nauwlettend de bruikbaarheid en de algemene gebruikerservaring welke is geassocieerd aan het gebruik van 3D mobile gestural interactie, specifiek op momenten dat fouten in de herkenning van bewegingen optreden. We laten zien dat 3D mobile gestural interactie zelfs bruikbaar is bij het optreden van detectiefouten, en onderzoeken vervolgens twee voorbeeldapplicaties. Daarna richten we ons nogmaals op hoe 3D gestural interactie kan worden toegepast in een speelse context, door te kijken naar bewegingsgestuurde muziekcompositie en het spelen van computerspelen. Tenslotte onderzoeken we hoe minimale mobiele interactie kan worden gebruikt voor het ondersteunen van mobiele gebruikersauthenticatie, gebruikmakende van 3D gestural input.

Tezamen laten onze gebruikerstudies zien dat interactieontwerpers schermgebaseerde interactiemethoden niet per direct op hoeven te geven, noch dat ze zouden moeten stoppen met het ontwerpen voor de visuele modaliteit van de gebruikers; ze moeten deze juist aanvullen met context-awareness of 3D gestural input oplossingen. Hiermee kunnen de mogelijkheden voor het ontwerpen en het ontwikkelen van mobiele systemen en applicaties worden verhoogd zodanig dat de werkelijke interactie die is vereist van de gebruiker tot een minimum wordt beperkt.

SIKS Dissertation Series

1998

1 Johan van den Akker (CWI) *DEGAS - An Active, Temporal Database of Autonomous Objects*
2 Floris Wiesman (UM) *Information Retrieval by Graphically Browsing Meta-Information*
3 Ans Steuten (TUD) *A Contribution to the Linguistic Analysis of Business Conversations within the Language/Action Perspective*
4 Dennis Breuker (UM) *Memory versus Search in Games*
5 Eduard W. Oskamp R(UL) *Computerondersteuning bij Straftoemeting*

1999

1 Mark Sloof (VU) *Physiology of Quality Change Modelling; Automated Modelling of Quality Change of Agricultural Products*
2 Rob Potharst (EUR) *Classification using Decision Trees and Neural Nets*
3 Don Beal (UM) *The Nature of Minimax Search*
4 Jacques Penders (UM) *The Practical Art of Moving Physical Objects*
5 Aldo de Moor (KUB) *Empowering Communities: A Method for the Legitimate User-Driven Specification of Network Information Systems*
6 Niek J.E. Wijngaards (VU) *Re-Design of Compositional Systems*
7 David Spelt (UT) *Verification Support for Object Database Design*
8 Jacques H.J. Lenting (UM) *Informed Gambling: Conception and Analysis of a Multi-Agent Mechanism for Discrete Reallocation*

2000

1 Frank Niessink (VU) *Perspectives on Improving Software Maintenance*
2 Koen Holtman (TU/e) *Prototyping of CMS Storage Management*
3 Carolien M.T. Metselaar (UvA) *Sociaal-organisatorische Gevolgen van Kennistechnologie; een Procesbenadering en Actorperspectief*
4 Geert de Haan (VU) *ETAG, A Formal Model of Competence Knowledge for User Interface Design*
5 Ruud van der Pol (UM) *Knowledge-Based Query Formulation in Information Retrieval*
6 Rogier van Eijk (UU) *Programming Languages for Agent Communication*
7 Niels Peek (UU) *Decision-Theoretic Planning of Clinical Patient Management*
8 Veerle Coupé (EUR) *Sensitivity Analysis of Decision-Theoretic Networks*
9 Florian Waas (CWI) *Principles of Probabilistic Query Optimization*
10 Niels Nes (CWI) *Image Database Management System Design Considerations, Algorithms and Architecture*
11 Jonas Karlsson (CWI) *Scalable Distributed Data Structures for Database Management*

2001

1 Silja Renooij (UU) *Qualitative Approaches to Quantifying Probabilistic Networks*
2 Koen Hindriks (UU) *Agent Programming Languages: Programming with Mental Models*
3 Maarten van Someren (UvA) *Learning as Problem Solving*
4 Evgueni Smirnov (UM) *Conjunctive and Disjunctive Version Spaces with Instance-Based Boundary Sets*
5 Jacco van Ossenbruggen (VU) *Processing Structured Hypermedia: A Matter of Style*
6 Martijn van Welie (VU) *Task-Based User Interface Design*
7 Bastiaan Schonhage (VU) *Diva: Architectural Perspectives on Information Visualization*
8 Pascal van Eck (VU) *A Compositional Semantic Structure for Multi-Agent Systems Dynamics*
9 Pieter Jan 't Hoen R(UL) *Towards Distributed Development of Large Object-Oriented Models, Views of Packages as Classes*
10 Maarten Sierhuis (UvA) *Modeling and Simulating Work Practice BRAHMS: a Multiagent Modeling and Simulation Language for Work Practice Analysis and Design*
11 Tom M. van Engers (VU) *Knowledge Management: The Role of Mental Models in Business Systems Design*

2002

1 Nico Lassing (VU) *Architecture-Level Modifiability Analysis*
2 Roelof van Zwol (UT) *Modelling and Searching Web-based Document Collections*
3 Henk Ernst Blok (UT) *Database Optimization Aspects for Information Retrieval*
4 Juan Roberto Castelo Valdueza (UU) *The Discrete Acyclic Digraph Markov Model in Data Mining*
5 Radu Serban (VU) *The Private Cyberspace Modeling Electronic Environments Inhabited by Privacy-Concerned Agents*
6 Laurens Mommers (UL) *Applied Legal Epistemology; Building a Knowledge-based Ontology of the Legal Domain*
7 Peter Boncz (CWI) *Monet: A Next-Generation DBMS Kernel For Query-Intensive Applications*

8 Jaap Gordijn (VU) *Value Based Requirements Engineering: Exploring Innovative E-Commerce Ideas*
9 Willem-Jan van den Heuvel (KUB) *Integrating Modern Business Applications with Objectified Legacy Systems*
10 Brian Sheppard (UM) *Towards Perfect Play of Scrabble*
11 Wouter C.A. Wijngaards (VU) *Agent Based Modelling of Dynamics: Biological and Organisational Applications*
12 Albrecht Schmidt (UvA) *Processing XML in Database Systems*
13 Hongjing Wu (TU/e) *A Reference Architecture for Adaptive Hypermedia Applications*
14 Wieke de Vries (UU) *Agent Interaction: Abstract Approaches to Modelling, Programming and Verifying Multi-Agent Systems*
15 Rik Eshuis (UT) *Semantics and Verification of (UM)L Activity Diagrams for Workflow Modelling*
16 Pieter van Langen (VU) *The Anatomy of Design: Foundations, Models and Applications*
17 Stefan Manegold (UvA) *Understanding, Modeling, and Improving Main-Memory Database Performance*

2003

1 Heiner Stuckenschmidt (VU) *Ontology-Based Information Sharing in Weakly Structured Environments*
2 Jan Broersen (VU) *Modal Action Logics for Reasoning About Reactive Systems*
3 Martijn Schuemie (TUD) *Human-Computer Interaction and Presence in Virtual Reality Exposure Therapy*
4 Milan Petkovic (UT) *Content-Based Video Retrieval Supported by Database Technology*
5 Jos Lehmann (UvA) *Causation in Artificial Intelligence and Law – A Modelling Approach*
6 Boris van Schooten (UT) *Development and Specification of Virtual Environments*
7 Machiel Jansen (UvA) *Formal Explorations of Knowledge Intensive Tasks*
8 Yong-Ping Ran (UM) *Repair-Based Scheduling*
9 Rens Kortmann (UM) *The Resolution of Visually Guided Behaviour*
10 Andreas Lincke (UT) *Electronic Business Negotiation: Some Experimental Studies on the Interaction between Medium, Innovation Context and Cult*
11 Simon Keizer (UT) *Reasoning under Uncertainty in Natural Language Dialogue using Bayesian Networks*
12 Roeland Ordelman (UT) *Dutch Speech Recognition in Multimedia Information Retrieval*
13 Jeroen Donkers (UM) *Nosce Hostem – Searching with Opponent Models*
14 Stijn Hoppenbrouwers KUN *Freezing Language: Conceptualisation Processes across ICT-Supported Organisations*
15 Mathijs de Weerdt (TUD) *Plan Merging in Multi-Agent Systems*
16 Menzo Windhouwer (CWI) *Feature Grammar Systems - Incremental Maintenance of Indexes to Digital Media Warehouse*
17 David Jansen (UT) *Extensions of Statecharts with Probability, Time, and Stochastic Timing*
18 Levente Kocsis (UM) *Learning Search Decisions*

2004

1 Virginia Dignum (UU) *A Model for Organizational Interaction: Based on Agents, Founded in Logic*
2 Lai Xu (UvT) *Monitoring Multi-party Contracts for E-business*
3 Perry Groot (VU) *A Theoretical and Empirical Analysis of Approximation in Symbolic Problem Solving*
4 Chris van Aart (UvA) *Organizational Principles for Multi-Agent Architectures*
5 Viara Popova (EUR) *Knowledge Discovery and Monotonicity*
6 Bart-Jan Hommes (TUD) *The Evaluation of Business Process Modeling Techniques*
7 Elise Boltjes (UM) *Voorbeeld$_{IG}$ Onderwijs; Voorbeeldgestuurd Onderwijs, een Opstap naar Abstract Denken, vooral voor Meisjes*
8 Joop Verbeek (UM) *Politie en de Nieuwe Internationale Informatiemarkt, Grensregionale Politiële Gegevensuitwisseling en Digitale Expertise*
9 Martin Caminada (VU) *For the Sake of the Argument; Explorations into Argument-based Reasoning*
10 Suzanne Kabel (UvA) *Knowledge-rich Indexing of Learning-objects*
11 Michel Klein (VU) *Change Management for Distributed Ontologies*
12 The Duy Bui (UT) *Creating Emotions and Facial Expressions for Embodied Agents*
13 Wojciech Jamroga (UT) *Using Multiple Models of Reality: On Agents who Know how to Play*
14 Paul Harrenstein (UU) *Logic in Conflict. Logical Explorations in Strategic Equilibrium*
15 Arno Knobbe (UU) *Multi-Relational Data Mining*
16 Federico Divina (VU) *Hybrid Genetic Relational Search for Inductive Learning*
17 Mark Winands (UM) *Informed Search in Complex Games*
18 Vania Bessa Machado (UvA) *Supporting the Construction of Qualitative Knowledge Models*
19 Thijs Westerveld (UT) *Using generative probabilistic models for multimedia retrieval*
20 Madelon Evers (Nyenrode) *Learning from Design: facilitating multidisciplinary design teams*

2005

1 Floor Verdenius (UvA) *Methodological Aspects of Designing Induction-Based Applications*
2 Erik van der Werf (UM) *AI techniques for the game of Go*
3 Franc Grootjen (RUN) *A Pragmatic Approach to the Conceptualisation of Language*
4 Nirvana Meratnia (UT) *Towards Database Support for Moving Object data*
5 Gabriel Infante-Lopez (UvA) *Two-Level Probabilistic Grammars for Natural Language Parsing*
6 Pieter Spronck (UM) *Adaptive Game AI*
7 Flavius Frasincar (TU/e) *Hypermedia Presentation Generation for Semantic Web Information Systems*
8 Richard Vdovjak (TU/e) *A Model-driven Approach for Building Distributed Ontology-based Web Applications*
9 Jeen Broekstra (VU) *Storage, Querying and Inferencing for Semantic Web Languages*
10 Anders Bouwer (UvA) *Explaining Behaviour: Using Qualitative Simulation in Interactive Learning Environments*
11 Elth Ogston (VU) *Agent Based Matchmaking and Clustering - A Decentralized Approach to Search*
12 Csaba Boer (EUR) *Distributed Simulation in Industry*
13 Fred Hamburg (UL) *Een Computermodel voor het Ondersteunen van Euthanasiebeslissingen*
14 Borys Omelayenko (VU) *Web-Service configuration on the Semantic Web; Exploring how semantics meets pragmatics*
15 Tibor Bosse (VU) *Analysis of the Dynamics of Cognitive Processes*
16 Joris Graaumans (UU) *Usability of XML Query Languages*
17 Boris Shishkov (TUD) *Software Specification Based on Re-usable Business Components*
18 Danielle Sent (UU) *Test-selection strategies for probabilistic networks*
19 Michel van Dartel (UM) *Situated Representation*
20 Cristina Coteanu (UL) *Cyber Consumer Law, State of the Art and Perspectives*
21 Wijnand Derks (UT) *Improving Concurrency and Recovery in Database Systems by Exploiting Application Semantics*

2006

1 Samuil Angelov (TU/e) *Foundations of B2B Electronic Contracting*
2 Cristina Chisalita (VU) *Contextual issues in the design and use of information technology in organizations*
3 Noor Christoph (UvA) *The role of metacognitive skills in learning to solve problems*
4 Marta Sabou (VU) *Building Web Service Ontologies*
5 Cees Pierik (UU) *Validation Techniques for Object-Oriented Proof Outlines*
6 Ziv Baida (VU) *Software-aided Service Bundling - Intelligent Methods & Tools for Graphical Service Modeling*
7 Marko Smiljanic (UT) *XML schema matching – balancing efficiency and effectiveness by means of clustering*
8 Eelco Herder (UT) *Forward, Back and Home Again - Analyzing User Behavior on the Web*
9 Mohamed Wahdan (UM) *Automatic Formulation of the Auditor's Opinion*
10 Ronny Siebes (VU) *Semantic Routing in Peer-to-Peer Systems*
11 Joeri van Ruth (UT) *Flattening Queries over Nested Data Types*
12 Bert Bongers (VU) *Interactivation - Towards an e-cology of people, our technological environment, and the arts*
13 Henk-Jan Lebbink (UU) *Dialogue and Decision Games for Information Exchanging Agents*
14 Johan Hoorn (VU) *Software Requirements: Update, Upgrade, Redesign - towards a Theory of Requirements Change*
15 Rainer Malik (UU) *CONAN: Text Mining in the Biomedical Domain*
16 Carsten Riggelsen (UU) *Approximation Methods for Efficient Learning of Bayesian Networks*
17 Stacey Nagata (UU) *User Assistance for Multitasking with Interruptions on a Mobile Device*
18 Valentin Zhizhkun (UvA) *Graph transformation for Natural Language Processing*
19 Birna van Riemsdijk (UU) *Cognitive Agent Programming: A Semantic Approach*
20 Marina Velikova (UvT) *Monotone models for prediction in data mining*
21 Bas van Gils (RUN) *Aptness on the Web*
22 Paul de Vrieze (RUN) *Fundaments of Adaptive Personalisation*
23 Ion Juvina (UU) *Development of Cognitive Model for Navigating on the Web*
24 Laura Hollink (VU) *Semantic Annotation for Retrieval of Visual Resources*
25 Madalina Drugan (UU) *Conditional log-likelihood MDL and Evolutionary MCMC*
26 Vojkan Mihajlovic (UT) *Score Region Algebra: A Flexible Framework for Structured Information Retrieval*
27 Stefano Bocconi (CWI) *Vox Populi: generating video documentaries from semantically annotated media repositories*
28 Borkur Sigurbjornsson (UvA) *Focused Information Access using XML Element Retrieval*

2007

1 Kees Leune (UvT) *Access Control and Service-Oriented Architectures*
2 Wouter Teepe (RUG) *Reconciling Information Exchange and Confidentiality: A Formal Approach*
3 Peter Mika (VU) *Social Networks and the Semantic Web*

4 Jurriaan van Diggelen (UU) *Achieving Semantic Interoperability in Multi-agent Systems: a dialogue-based approach*
5 Bart Schermer (UL) *Software Agents, Surveillance, and the Right to Privacy: a Legislative Framework for Agent-enabled Surveillance*
6 Gilad Mishne (UvA) *Applied Text Analytics for Blogs*
7 Natasa Jovanovic' (UT) *To Whom It May Concern - Addressee Identification in Face-to-Face Meetings*
8 Mark Hoogendoorn (VU) *Modeling of Change in Multi-Agent Organizations*
9 David Mobach (VU) *Agent-Based Mediated Service Negotiation*
10 Huib Aldewereld (UU) *Autonomy vs. Conformity: an Institutional Perspective on Norms and Protocols*
11 Natalia Stash (TU/e) *Incorporating Cognitive/Learning Styles in a General-Purpose Adaptive Hypermedia System*
12 Marcel van Gerven (RUN) *Bayesian Networks for Clinical Decision Support: A Rational Approach to Dynamic Decision-Making under Uncertainty*
13 Rutger Rienks (UT) *Meetings in Smart Environments; Implications of Progressing Technology*
14 Niek Bergboer (UM) *Context-Based Image Analysis*
15 Joyca Lacroix (UM) *NIM: a Situated Computational Memory Model*
16 Davide Grossi (UU) *Designing Invisible Handcuffs. Formal investigations in Institutions and Organizations for Multi-agent Systems*
17 Theodore Charitos (UU) *Reasoning with Dynamic Networks in Practice*
18 Bart Orriens (UvT) *On the development and management of adaptive business collaborations*
19 David Levy (UM) *Intimate relationships with artificial partners*
20 Slinger Jansen (UU) *Customer Configuration Updating in a Software Supply Network*
21 Karianne Vermaas (UU) *Fast diffusion and broadening use: A research on residential adoption and usage of broadband internet in the Netherlands between 2001 and 2005*
22 Zlatko Zlatev (UT) *Goal-oriented design of value and process models from patterns*
23 Peter Barna (TU/e) *Specification of Application Logic in Web Information Systems*
24 Georgina Ramírez Camps (CWI) *Structural Features in XML Retrieval*
25 Joost Schalken (VU) *Empirical Investigations in Software Process Improvement*

2008

1 Katalin Boer-Sorbán (EUR) *Agent-Based Simulation of Financial Markets: A modular, continuous-time approach*
2 Alexei Sharpanskykh (VU) *On Computer-Aided Methods for Modeling and Analysis of Organizations*
3 Vera Hollink (UvA) *Optimizing hierarchical menus: a usage-based approach*
4 Ander de Keijzer (UT) *Management of Uncertain Data - towards unattended integration*
5 Bela Mutschler (UT) *Modeling and simulating causal dependencies on process-aware information systems from a cost perspective*
6 Arjen Hommersom (RUN) *On the Application of Formal Methods to Clinical Guidelines, an Artificial Intelligence Perspective*
7 Peter van Rosmalen (OU) *Supporting the tutor in the design and support of adaptive e-learning*
8 Janneke Bolt (UU) *Bayesian Networks: Aspects of Approximate Inference*
9 Christof van Nimwegen (UU) *The paradox of the guided user: assistance can be counter-effective*
10 Wauter Bosma (UT) *Discourse oriented Summarization*
11 Vera Kartseva (VU) *Designing Controls for Network Organizations: a Value-Based Approach*
12 Jozsef Farkas (RUN) *A Semiotically oriented Cognitive Model of Knowledge Representation*
13 Caterina Carraciolo (UvA) *Topic Driven Access to Scientific Handbooks*
14 Arthur van Bunningen (UT) *Context-Aware Querying; Better Answers with Less Effort*
15 Martijn van Otterlo (UT) *The Logic of Adaptive Behavior: Knowledge Representation and Algorithms for the Markov Decision Process Framework in First-Order Domains*
16 Henriette van Vugt (VU) *Embodied Agents from a User's Perspective*
17 Martin Op't Land (TUD) *Applying Architecture and Ontology to the Splitting and Allying of Enterprises*
18 Guido de Croon (UM) *Adaptive Active Vision*
19 Henning Rode (UT) *From document to entity retrieval: improving precision and performance of focused text search*
20 Rex Arendsen (UvA) *Geen bericht, goed bericht. Een onderzoek naar de effecten van de introductie van elektronisch berichtenverkeer met een overheid op de administratieve lasten van bedrijven*
21 Krisztian Balog (UvA) *People search in the enterprise*
22 Henk Koning (UU) *Communication of IT-architecture*
23 Stefan Visscher (UU) *Bayesian network models for the management of ventilator-associated pneumonia*
24 Zharko Aleksovski (VU) *Using background knowledge in ontology matching*
25 Geert Jonker (UU) *Efficient and Equitable exchange in air traffic management plan repair using spender-signed currency*
26 Marijn Huijbregts (UT) *Segmentation, diarization and speech transcription: surprise data unraveled*

27 Hubert Vogten (OU) *Design and implementation strategies for IMS learning design*
28 Ildiko Flesh (RUN) *On the use of independence relations in Bayesian networks*
29 Dennis Reidsma (UT) *Annotations and subjective machines- Of annotators, embodied agents, users, and other humans*
30 Wouter van Atteveldt (VU) *Semantic network analysis: techniques for extracting, representing and querying media content*
31 Loes Braun (UM) *Pro-active medical information retrieval*
32 Trung B. Hui (UT) *Toward affective dialogue management using partially observable markov decision processes*
33 Frank Terpstra (UvA) *Scientific workflow design; theoretical and practical issues*
34 Jeroen de Knijf (UU) *Studies in Frequent Tree Mining*
35 Benjamin Torben-Nielsen (UvT) *Dendritic morphology: function shapes structure*

2009

1 Rasa Jurgelenaite (RUN) *Symmetric Causal Independence Models*
2 Willem Robert van Hage (VU) *Evaluating Ontology-Alignment Techniques*
3 Hans Stol (UvT) *A Framework for Evidence-based Policy Making Using IT*
4 Josephine Nabukenya (RUN) *Improving the Quality of Organisational Policy Making using Collaboration Engineering*
5 Sietse Overbeek (RUN) *Bridging Supply and Demand for Knowledge Intensive Tasks - Based on Knowledge, Cognition, and Quality*
6 Muhammad Subianto (UU) *Understanding Classification*
7 Ronald Poppe (UT) *Discriminative Vision-Based Recovery and Recognition of Human Motion*
8 Volker Nannen (VU) *Evolutionary Agent-Based Policy Analysis in Dynamic Environments*
9 Benjamin Kanagwa (RUN) *Design, Discovery and Construction of Service-oriented Systems*
10 Jan Wielemaker (UvA) *Logic programming for knowledge-intensive interactive applications*
11 Alexander Boer (UvA) *Legal Theory, Sources of Law & the Semantic Web*
12 Peter Massuthe (TU/e), Humboldt-Universtät zu Berlin *Operating Guidelines for Services*
13 Steven de Jong (UM) *Fairness in Multi-Agent Systems*
14 Maksym Korotkiy (VU) *From ontology-enabled services to service-enabled ontologies* making ontologies work in e-science with ONTO-SOA
15 Rinke Hoekstra (UvA) *Ontology Representation - Design Patterns and Ontologies that Make Sense*
16 Fritz Reul (UvT) *New Architectures in Computer Chess*
17 Laurens van der Maaten (UvT) *Feature Extraction from Visual Data*
18 Fabian Groffen (CWI) *Armada, An Evolving Database System*
19 Valentin Robu (CWI) *Modeling Preferences, Strategic Reasoning and Collaboration in Agent-Mediated Electronic Markets*
20 Bob van der Vecht (UU) *Adjustable Autonomy: Controling Influences on Decision Making*
21 Stijn Vanderlooy (UM) *Ranking and Reliable Classification*
22 Pavel Serdyukov (UT) *Search For Expertise: Going beyond direct evidence*
23 Peter Hofgesang (VU) *Modelling Web Usage in a Changing Environment*
24 Annerieke Heuvelink (VU) *Cognitive Models for Training Simulations*
25 Alex van Ballegooij (CWI) *"RAM: Array Database Management through Relational Mapping"*
26 Fernando Koch (UU) *An Agent-Based Model for the Development of Intelligent Mobile Services*
27 Christian Glahn (OU) *Contextual Support of social Engagement and Reflection on the Web*
28 Sander Evers (UT) *Sensor Data Management with Probabilistic Models*
29 Stanislav Pokraev (UT) *Model-Driven Semantic Integration of Service-Oriented Applications*
30 Marcin Zukowski (CWI) *Balancing vectorized query execution with bandwidth-optimized storage*
31 Sofiya Katrenko (UvA) *A Closer Look at Learning Relations from Text*
32 Rik Farenhorst and Remco de Boer (VU) *Architectural Knowledge Management: Supporting Architects and Auditors*
33 Khiet Truong (UT) *How Does Real Affect Affect Affect Recognition In Speech?*
34 Inge van de Weerd (UU) *Advancing in Software Product Management: An Incremental Method Engineering Approach*
35 Wouter Koelewijn (UL) *Privacy en Politiegegevens; Over geautomatiseerde normatieve informatie-uitwisseling*
36 Marco Kalz (OUN) *Placement Support for Learners in Learning Networks*
37 Hendrik Drachsler (OUN) *Navigation Support for Learners in Informal Learning Networks*
38 Riina Vuorikari (OU) *Tags and self-organisation: a metadata ecology for learning resources in a multilingual context*
39 Christian Stahl (TUE), Humboldt-Universität zu Berlin *Service Substitution – A Behavioral Approach Based on Petri Nets*
40 Stephan Raaijmakers (UvT) *Multinomial Language Learning: Investigations into the Geometry of Language*
41 Igor Berezhnyy (UvT) *Digital Analysis of Paintings*
42 Toine Bogers (UvT) *Recommender Systems for Social Bookmarking*

43 Virginia Nunes Leal Franqueira (UT) *Finding Multi-step Attacks in Computer Networks using Heuristic Search and Mobile Ambients*
44 Roberto Santana Tapia (UT) *Assessing Business-IT Alignment in Networked Organizations*
45 Jilles Vreeken (UU) *Making Pattern Mining Useful*
46 Loredana Afanasiev (UvA) *Querying XML: Benchmarks and Recursion*

2010

1 Matthijs van Leeuwen (UU) *Patterns that Matter*
2 Ingo Wassink (UT) *Work flows in Life Science*
3 Joost Geurts (CWI) *A Document Engineering Model and Processing Framework for Multimedia documents*
4 Olga Kulyk (UT) *Do You Know What I Know? Situational Awareness of Co-located Teams in Multidisplay Environments*
5 Claudia Hauff (UT) *Predicting the Effectiveness of Queries and Retrieval Systems*
6 Sander Bakkes (UvT) *Rapid Adaptation of Video Game AI*
7 Wim Fikkert (UT) *Gesture interaction at a Distance*
8 Krzysztof Siewicz (UL) *Towards an Improved Regulatory Framework of Free Software. Protecting user freedoms in a world of software communities and eGovernments*
9 Hugo Kielman (UL) *A Politiele gegevensverwerking en Privacy, Naar een effectieve waarborging*
10 Rebecca Ong (UL) *Mobile Communication and Protection of Children*
11 Adriaan Ter Mors (TUD) *The world according to MARP: Multi-Agent Route Planning*
12 Susan van den Braak (UU) *Sensemaking software for crime analysis*
13 Gianluigi Folino (RUN) *High Performance Data Mining using Bio-inspired techniques*
14 Sander van Splunter (VU) *Automated Web Service Reconfiguration*
15 Lianne Bodenstaff (UT) *Managing Dependency Relations in Inter-Organizational Models*
16 Sicco Verwer (TUD) *Efficient Identification of Timed Automata, theory and practice*
17 Spyros Kotoulas (VU) *Scalable Discovery of Networked Resources: Algorithms, Infrastructure, Applications*
18 Charlotte Gerritsen (VU) *Caught in the Act: Investigating Crime by Agent-Based Simulation*
19 Henriette Cramer (UvA) *People's Responses to Autonomous and Adaptive Systems*
20 Ivo Swartjes (UT) *Whose Story Is It Anyway? How Improv Informs Agency and Authorship of Emergent Narrative*
21 Harold van Heerde (UT) *Privacy-aware data management by means of data degradation*
22 Michiel Hildebrand (CWI) *End-user Support for Access to Heterogeneous Linked Data*
23 Bas Steunebrink (UU) *The Logical Structure of Emotions*
24 Dmytro Tykhonov (TUD) *Designing Generic and Efficient Negotiation Strategies*
25 Zulfiqar Ali Memon (VU) *Modelling Human-Awareness for Ambient Agents: A Human Mindreading Perspective*
26 Ying Zhang (CWI) *XRPC: Efficient Distributed Query Processing on Heterogeneous XQuery Engines*
27 Marten Voulon (UL) *Automatisch contracteren*
28 Arne Koopman (UU) *Characteristic Relational Patterns*
29 Stratos Idreos (CWI) *Database Cracking: Towards Auto-tuning Database Kernels*
30 Marieke van Erp (UvT) *Accessing Natural History - Discoveries in data cleaning, structuring, and retrieval*
31 Victor de Boer (UvA) *Ontology Enrichment from Heterogeneous Sources on the Web*
32 Marcel Hiel (UvT) *An Adaptive Service Oriented Architecture: Automatically solving Interoperability Problems*
33 Robin Aly (UT) *Modeling Representation Uncertainty in Concept-Based Multimedia Retrieval*
34 Teduh Dirgahayu (UT) *Interaction Design in Service Compositions*
35 Dolf Trieschnigg (UT) *Proof of Concept: Concept-based Biomedical Information Retrieval*
36 Jose Janssen (OU) *Paving the Way for Lifelong Learning; Facilitating competence development through a learning path specification*
37 Niels Lohmann (TUE) *Correctness of services and their composition*
38 Dirk Fahland (TUE) *From Scenarios to components*
39 Ghazanfar Farooq Siddiqui (VU) *Integrative modeling of emotions in virtual agents*
40 Mark van Assem (VU) *Converting and Integrating Vocabularies for the Semantic Web*
41 Guillaume Chaslot (UM) *Monte-Carlo Tree Search*
42 Sybren de Kinderen (VU) *Needs-driven service bundling in a multi-supplier setting - the computational e3-service approach*
43 Peter van Kranenburg (UU) *A Computational Approach to Content-Based Retrieval of Folk Song Melodies*
44 Pieter Bellekens (TUE) *An Approach towards Context-sensitive and User-adapted Access to Heterogeneous Data Sources, Illustrated in the Television Domain*
45 Vasilios Andrikopoulos (UvT) *A theory and model for the evolution of software services*
46 Vincent Pijpers (VU) *e3alignment: Exploring Inter-Organizational Business-ICT Alignment*
47 Chen Li (UT) *Mining Process Model Variants: Challenges, Techniques, Examples*

48 Milan Lovric (EUR) *Behavioral Finance and Agent-Based Artificial Markets*
49 Jahn-Takeshi Saito (UM) *Solving difficult game positions*
50 Bouke Huurnink (UvA) *Search in Audiovisual Broadcast Archives*
51 Alia Khairia Amin (CWI) *Understanding and supporting information seeking tasks in multiple sources*
52 Peter-Paul van Maanen (VU) *Adaptive Support for Human-Computer Teams: Exploring the Use of Cognitive Models of Trust and Attention*
53 Edgar J. Meij (UvA) *Combining Concepts and Language Models for Information Access*

2011

1 Botond Cseke (RUN) *Variational Algorithms for Bayesian Inference in Latent Gaussian Models*
2 Nick Tinnemeier (UU) *Work flows in Life Science*
3 Jan Martijn van der Werf (TUE) *Compositional Design and Verification of Component-Based Information Systems*
4 Hado van Hasselt (UU) *Insights in Reinforcement Learning; Formal analysis and empirical evaluation of temporal-difference learning algorithms*
5 Base van der Raadt (VU) *Enterprise Architecture Coming of Age - Increasing the Performance of an Emerging Discipline*
6 Yiwen Wang (TUE) *Semantically-Enhanced Recommendations in Cultural Heritage*
7 Yujia Cao (UT) *Multimodal Information Presentation for High Load Human Computer Interaction*
8 Nieske Vergunst (UU) *BDI-based Generation of Robust Task-Oriented Dialogues*
9 Tim de Jong (OU) *Contextualised Mobile Media for Learning*
10 Bart Bogaert (UvT) *Cloud Content Contention*
11 Dhaval Vyas (UT) *Designing for Awareness: An Experience-focused HCI Perspective*
12 Carmen Bratosin (TUE) *Grid Architecture for Distributed Process Mining*
13 Xiaoyu Mao (UvT) *Airport under Control. Multiagent Scheduling for Airport Ground Handling*
14 Milan Lovric (EUR) *Behavioral Finance and Agent-Based Artificial Markets*
15 Marijn Koolen (UvA) *The Meaning of Structure: the Value of Link Evidence for Information Retrieval*
16 Maarten Schadd (UM) *Selective Search in Games of Different Complexity*
17 Jiyin He (UvA) *Exploring Topic Structure: Coherence, Diversity and Relatedness*
18 Mark Ponsen (UM) *Strategic Decision-Making in complex games*
19 Ellen Rusman (OU) *The Mind ' s Eye on Personal Profiles*
20 Qing Gu (VU) *Guiding service-oriented software engineering - A view-based approach*
21 Linda Terlouw (TUD) *Modularization and Specification of Service-Oriented System*
22 Junte Zhang (UvA) *System Evaluation of Archival Description and Access*
23 Wouter Weerkamp (UvA) *Finding People and their Utterances in Social Media*
24 Herwin van Welbergen (UT) *Behavior Generation for Interpersonal Coordination with Virtual Humans On Specifying, Scheduling and Realizing Multimodal Virtual Human Behavior*
25 Syed Waqar ul Qounain Jaffry (VU) Analysis and Validation of Models for Trust Dynamics
26 Matthijs Aart Pontier (VU) *Virtual Agents for Human Communication - Emotion Regulation and Involvement-Distance Trade-Offs in Embodied Conversational Agents and Robots*
27 Aniel Bhulai (VU) *Dynamic website optimization through autonomous management of design patterns*
28 Rianne Kaptein (UvA) *Effective Focused Retrieval by Exploiting Query Context and Document Structure*
29 Faisal Kamiran (TUE) *Discrimination-aware Classification*
30 Egon van den Broek (UT) *Affective Signal Processing* ASP: Unraveling the mystery of emotions
31 Ludo Waltman (EUR) *Computational and Game-Theoretic Approaches for Modeling Bounded Rationality*
32 Nees-Jan van Eck (EUR) *Methodological Advances in Bibliometric Mapping of Science*
33 Tom van der Weide (UU) *Arguing to Motivate Decisions*
34 Paolo Turrini (UU) *Strategic Reasoning in Interdependence: Logical and Game-theoretical Investigations*
35 Maaike Harbers (UU) *Explaining Agent Behavior in Virtual Training*
36 Erik van der Spek (UU) *Experiments in serious game design: a cognitive approach*
37 Adriana Burlutiu (RUN) *Machine Learning for Pairwise Data, Applications for Preference Learning and Supervised Network Inference*
38 Nyree Lemmens (UM) *Bee-inspired Distributed Optimization*
39 Joost Westra (UU) *Organizing Adaptation using Agents in Serious Games*
40 Viktor Clerc (VU) *Architectural Knowledge Management in Global Software Development*
41 Luan Ibraimi (UT) *Cryptographically Enforced Distributed Data Access Control*
42 Michal Sindlar (UU) *Explaining Behavior through Mental State Attribution*
43 Henk van der Schuur (UU) *Process Improvement through Software Operation Knowledge*
44 Boris Reuderink (UT) *Robust Brain-Computer Interfaces*
45 Herman Stehouwer (UvT) *Statistical Language Models for Alternative Sequence Selection*

46 Beibei Hu (TUD) *Towards Contextualized Information Delivery: A Rule-based Architecture for the Domain of Mobile Police Work*
47 Azizi Bin Ab Aziz (VU) *Exploring Computational Models for Intelligent Support of Persons with Depression*
48 Mark Ter Maat (UT) *Response Selection and Turn-taking for a Sensitive Artificial Listening Agent*
49 Andreea Niculescu (UT) *Conversational interfaces for task-oriented spoken dialogues: design aspects influencing interaction quality*

2012

1 Terry Kakeeto (UvT) *Relationship Marketing for SMEs in Uganda*
2 Muhammad Umair (VU) *Adaptivity, emotion, and Rationality in Human and Ambient Agent Models*
3 Adam Vanya (VU) *Supporting Architecture Evolution by Mining Software Repositories*
4 Jurriaan Souer (UU) *Development of Content Management System-based Web Applications*
5 Marijn Plomp (UU) *Maturing Interorganisational Information Systems*
6 Wolfgang Reinhardt (OU) *Awareness Support for Knowledge Workers in Research Networks*
7 Rianne van Lambalgen (VU) *When the Going Gets Tough: Exploring Agent-based Models of Human Performance under Demanding Conditions*
8 Gerben de Vries (UvA) *Kernel Methods for Vessel Trajectories*
9 Ricardo Neisse (UT) *Trust and Privacy Management Support for Context-Aware Service Platforms*
10 David Smits (TUE) *Towards a Generic Distributed Adaptive Hypermedia Environment*
11 J.C.B. Rantham Prabhakara (TUE) *Process Mining in the Large: Preprocessing, Discovery, and Diagnostics*
12 Kees van der Sluijs (TUE) *Model Driven Design and Data Integration in Semantic Web Information Systems*
13 Suleman Shahid (UvT) *Fun and Face: Exploring non-verbal expressions of emotion during playful interactions*
14 Evgeny Knutov (TUE) *Generic Adaptation Framework for Unifying Adaptive Web-based Systems*
15 Natalie van der Wal (VU) *Social Agents. Agent-Based Modelling of Integrated Internal and Social Dynamics of Cognitive and Affective Processes*
16 Fiemke Both (VU) *Helping people by understanding them - Ambient Agents supporting task execution and depression treatment*
17 Amal Elgammal (UvT) *Towards a Comprehensive Framework for Business Process Compliance*
18 Eltjo Poort (VU) *Improving Solution Architecting Practices*
19 Helen Schonenberg (TUE) *What's Next? Operational Support for Business Process Execution*
20 Ali Bahramisharif (RUN) *Covert Visual Spatial Attention, a Robust Paradigm for Brain-Computer Interfacing*
21 Roberto Cornacchia (TUD) *Querying Sparse Matrices for Information Retrieval*
22 Thijs Vis (UvT) *Intelligence, politie en veiligheidsdienst: verenigbare grootheden?*
23 Christian Muehl (UT) *Toward Affective Brain-Computer Interfaces: Exploring the Neurophysiology of Affect during Human Media Interaction*
24 Laurens van der Werff (UT) *Evaluation of Noisy Transcripts for Spoken Document Retrieval*
25 Silja Eckartz (UT) *Managing the Business Case Development in Inter-Organizational IT Projects: A Methodology and its Application*
26 Emile de Maat (UvA) *Making Sense of Legal Text*
27 Hayrettin Gurkok (UT) *Mind the Sheep! User Experience Evaluation & Brain-Computer Interface Games*
28 Nancy Pascall (UvT) *Engendering Technology Empowering Women*
29 Almer Tigelaar (UT) *Peer-to-Peer Information Retrieval*
30 Alina Pommeranz (TUD) *Designing Human-Centered Systems for Reflective Decision Making*
31 Emily Bagarukayo (RUN) *A Learning by Construction Approach for Higher Order Cognitive Skills Improvement, Building Capacity and Infrastructure*
32 Wietske Visser (TUD) *Qualitative multi-criteria preference representation and reasoning*
33 Rory Sie (OUN) *Coalitions in Cooperation Networks (COCOON)*
34 Pavol Jancura (RUN) *Evolutionary analysis in PPI networks and applications*
35 Evert Haasdijk (VU) *Never Too Old To Learn – On-line Evolution of Controllers in Swarm- and Modular Robotics*
36 Denis Ssebugwawo (RUN) *Analysis and Evaluation of Collaborative Modeling Processes*
37 Agnes Nakakawa (RUN) *A Collaboration Process for Enterprise Architecture Creation*
38 Selmar Smit (VU) *Parameter Tuning and Scientific Testing in Evolutionary Algorithms*
39 Hassan Fatemi (UT) *Risk-aware design of value and coordination networks*
40 Agus Gunawan (UvT) *Information Access for SMEs in Indonesia*
41 Sebastian Kelle (OU) *Game Design Patterns for Learning*
42 Dominique Verpoorten (OU) *Reflection Amplifiers in self-regulated Learning*
43 Withdrawn
44 Anna Tordai (VU) *On Combining Alignment Techniques*
45 Benedikt Kratz (UvT) *A Model and Language for Business-aware Transactions*
46 Simon Carter (UvA) *Exploration and Exploitation of Multilingual Data for Statistical Machine Translation*

47 Manos Tsagkias (UvA) *Mining Social Media: Tracking Content and Predicting Behavior*
48 Jorn Bakker (TUE) *Handling Abrupt Changes in Evolving Time-series Data*
49 Michael Kaisers (UM) *Learning against Learning - Evolutionary dynamics of reinforcement learning algorithms in strategic interactions*
50 Steven van Kervel (TUD) *Ontology driven Enterprise Information Systems Engineering*
51 Jeroen de Jong (TUD) *Heuristics in Dynamic Sceduling; a practical framework with a case study in elevator dispatching*

2013

1 Viorel Milea (EUR) *News Analytics for Financial Decision Support*
2 Erietta Liarou (CWI) *MonetDB/DataCell: Leveraging the Column-store Database Technology for Efficient and Scalable Stream Processing*
3 Szymon Klarman (VU) *Reasoning with Contexts in Description Logics*
4 Chetan Yadati (TUD) *Coordinating autonomous planning and scheduling*
5 Dulce Pumareja (UT) *Groupware Requirements Evolutions Patterns*
6 Romulo Goncalves (CWI) *The Data Cyclotron: Juggling Data and Queries for a Data Warehouse Audience*
7 Giel van Lankveld (UT) *Quantifying Individual Player Differences*
8 Robbert-Jan Merk (VU) *Making enemies: cognitive modeling for opponent agents in fighter pilot simulators*
9 Fabio Gori (RUN) *Metagenomic Data Analysis: Computational Methods and Applications*
10 Jeewanie Jayasinghe Arachchige (UvT) *A Unified Modeling Framework for Service Design*
11 Evangelos Pournaras (TUD) *Multi-level Reconfigurable Self-organization in Overlay Services*
12 Marian Razavian (VU) *Knowledge-driven Migration to Services*
13 Mohammad Safiri (UT) *Service Tailoring: User-centric creation of integrated IT-based homecare services to support independent living of elderly*
14 Jafar Tanha (UvA) *Ensemble Approaches to Semi-Supervised Learning Learning*
15 Daniel Hennes (UM) *Multiagent Learning - Dynamic Games and Applications*
16 Eric Kok (UU) *Exploring the practical benefits of argumentation in multi-agent deliberation*
17 Koen Kok (VU) *The PowerMatcher: Smart Coordination for the Smart Electricity Grid*
18 Jeroen Janssens (UvT) *Outlier Selection and One-Class Classification*
19 Renze Steenhuizen (TUD) *Coordinated Multi-Agent Planning and Scheduling*
20 Katja Hofmann (UvA) *Fast and Reliable Online Learning to Rank for Information Retrieval*
21 Sander Wubben (UvT) *Text-to-text generation by monolingual machine translation*
22 Tom Claassen (RUN) *Causal Discovery and Logic*
23 Patricio de Alencar Silva (UvT) *Value Activity Monitoring*
24 Haitham Bou Ammar (UM) *Automated Transfer in Reinforcement Learning*
25 Agnieszka Anna Latoszek-Berendsen (UM) *Intention-based Decision Support. A new way of representing and implementing clinical guidelines in a Decision Support System*
26 Alireza Zarghami (UT) *Architectural Support for Dynamic Homecare Service Provisioning*
27 Mohammad Huq (UT) *Inference-based Framework Managing Data Provenance*
28 Frans van der Sluis (UT) *When Complexity becomes Interesting: An Inquiry into the Information eXperience*

In the last 20 years, the widespread adoption of personal, mobile computing devices in everyday life, has allowed entry into a new technological era in Human Computer Interaction (HCI). The constant change of the physical and social context in a user's situation made possible by the portability of mobile devices means that the user's attention becomes limited. This can negatively impact the user experience. To deal with this problem, this thesis draws from two developments in HCI, context-awareness and 3D gestural input. From these developments, we introduce the concept of minimal mobile HCI, a subset of eyes-free mobile interaction that allows minimal combination of the visual modality with other sensory modalities to improve the user experience of interacting with smartphones.

In the first part, we look closely at the design and evaluation of location-aware multimedia messaging systems, and how they can lead to playfulness in a minimal interaction setting. We then look at how urban interactions are connected across locations by designing and evaluating an exploration-based route planner that makes use of large amounts of geotagged data. In the second part, we look closely at the usability and user experience issues associated with 3D mobile gestural input when recognition errors occur. Upon showing that this form of interaction is usable in the face of errors, we then investigate two example applications. We revisit the playfulness domain, and investigate how 3D gestural input can be applied to music composition and gaming. Finally, we look at how minimal mobile interaction can be used to support mobile user authentication using 3D gestural input.

Together, our user studies show that interaction designers need not abandon screen-based interaction, nor stop designing for users' visual modality, only complementing it with context-awareness or 3D gestural input solutions. This can expand the design space for designing and developing mobile systems/applications that keep user interaction costs at a minimum.

ISBN: 978-94-6182-364-9