UNDERSTANDING RECEPTIVITY TO INTERRUPTIONS IN MOBILE HUMAN-COMPUTER INTERACTION

JOEL E. FISCHER

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

October 2011

Joel E. Fischer: *Understanding Receptivity to Interruptions in Mobile Human-Computer Interaction,* Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy ,

SUPERVISORS:
Steve Benford
Chris Greenhalgh

To my parents.

ABSTRACT

Interruptions have a profound impact on our attentional orientation in everyday life. Recent advances in mobile information technology increase the number of potentially disruptive notifications on mobile devices by an increasing availability of services. Understanding the contextual intricacies that make us receptive to these interruptions is paramount to devising technology that supports interruption management.

This thesis makes a number of contributions to the methodology of studying mobile experiences in situ, understanding receptivity to interruptions, and designing context-sensitive systems.

This thesis presents a series of real-world studies that investigate opportune moments for interruptions in mobile settings. In order to facilitate the study of the multi-faceted ways opportune moments surface from participants' involvement in the world this thesis develops:

- a model of the contextual factors that interact to guide receptivity to interruptions, and
- an adaptation of the Experience-Sampling Method (ESM) to capture behavioural response to interruptions in situ.

In two naturalistic experiments, participants' experiences of being interrupted on a mobile phone are sampled as they go about their everyday lives. In a field study, participants' experiences are observed and recorded as they use a notification-driven mobile application to create photo-stories in a theme park.

Experiment 1 explores the effects of content and time of delivery of the interruption. The results show that receptivity to text messages is significantly affected by message content, while scheduling one's own interruption times in advance does not improve receptivity over randomly timed interruptions. Experiment 2 investigates the hypothesis that opportune moments to deliver notifications are located at the endings of episodes of mobile interaction such as texting and calling. This notification strategy is supported by significant effects in behavioural measures of receptivity, while self-reports and interviews reveal complexities in the subjective experience of the interruption. By employing a mixed methods approach of interviews, observations and an analysis of system logs in the field study, it is shown that participants appreciated location-based notifications as prompts to foreground the application during relative 'downtimes' from other activities. However, an unexpected quantity of redundant notifications meant that visitors soon habituated to and eventually ignored them, which suggests careful, sparing use of notifications in interactive experiences.

Overall, the studies showed that contextual mediation of the timing of interruptions (e.g. by phone activity in Experiment 2 and opportune places in the field study) is more likely to lead to interruptions at opportune moments than when participants schedule their own interruptions. However, momentary receptivity and responsiveness to an interruption is determined by the complex and situated interactions of local and relational contextual factors. These contextual factors are captured in a model of receptivity that underlies the interruption process. The studies highlight implications for the design of systems that seek to manage interruptions by adapting the timing of interruptions to the user's situation. In particular, applications to manage interruptions in personal communication and pervasive experiences are considered.

Past and future exist only too unmistakably in the world,
they exist in the present,
and what being itself lacks in order to be of the temporal order,
is the not-being of elsewhere, formerly and tomorrow.

Maurice Merleau-Ponty (1962)

ACKNOWLEDGMENTS

First and foremost, I would like to thank my supervisors Steve Benford and Chris Greenhalgh at the Mixed Reality Lab for supporting me throughout my time as a PhD student. Thanks for all that help, encouragement, advice, care, and time. Special thanks to Steve, for inviting me to come to Nottingham in the first place, which made all this possible! Thanks to my examiners Alan Dix and Tom Rodden who helped me improve this thesis through their thoughtful and thorough comments. Many thanks to friends and (former) colleagues at the Lab, the Horizon Institute, and the School of Computer Science for help, inspiration, collaboration, and for being mates. In no particular order (probably omitting some): Martin Flintham, Duncan Rowland, Leif Oppermann, Dave Kirk, James Norris, Jo Cranwell, Stuart Reeves, Rachel Jacobs, Dominic Price, Ben Bedwell, Stef Rennick-Egglestone, Hazel Glover, Julie Greensmith, Naisan Benatar, Adriano Galati, Joe Marshall, Mark Paxton, Holger Schnädelbach, Andy Crabtree, Pat Brundell, and Paul Tennent. Special thanks to Abigail Durrant for proof-reading this thesis.

Many thanks to the good people at PARC for supporting me during my internship. Special thanks to Victoria Bellotti for supervising me, and being an inspiration. Many thanks to the STIR team: Nick Yee, Nathan Good, Ellen Isaacs, Dianne Schiano, Nicholas Ducheneaut, and many more at PARC that helped with my research, showed interest and shared exciting stories.

Many thanks to the CHI 2010 Doctoral Consortium student participants and senior panel members for giving advice and feedback on this work, and to the HCI community at various conferences for their valuable comments and questions that moved this work forward.

Last but not least, thanks to my family and friends. Without your love and support this wouldn't have been possible.

This work has partially been funded by an EU Research Scholarship and by the EPSRC.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

ANOVA Analysis of Variance

CSCW Computer-Supported Cooperative Work

DV Dependent Variable

EM Ethnomethodology

EMA Ecological Momentary Assessment

ESM Experience-Sampling Method

FT Free-Text

GPS Global Positioning System

GUI Graphical User Interface

HMD Head-Mounted Display

HCI Human-Computer Interaction

ICT Information and Communication Technology

IM Instant Messaging

IMSM Interruption Management Stage Model

IV Independent Variable

LMM Linear Mixed Models

MC Multiple-Choice

PARC Palo Alto Research Center

PDA Personal Digital Assistant

PH Photo

RSS Really Simple Syndication

SMS Short Messaging Service

TLX Task Load Index

UI User Interface

1

INTRODUCTION

The advent of media has liberated us from the necessity to be at the same place at the same time in order to communicate with one another. With the arrival of Information and Communication Technology (ICT), an almost instantaneous and reciprocal exchange of communication was made possible despite not being in the same place[1]. Today, *Computer-mediated communication* has transformed what used to be the spatial and temporal constraints of communication into mere design choices for the engineering of systems that support communication. Accordingly, applications may be designed to support communication both asynchronously (e.g. email) or synchronously (e.g. Instant Messaging (IM)), and be designed to take place when co-located (e.g. shared whiteboards) or remotely (e.g. video chat) (Ellis et al., 1991).

However, what often remains overlooked when adopting this kind of perspective are the unintended social consequences technological change induces. For example, Meyrowitz (1985) discusses in detail the changed relation to space and place of social situations that comes about with ICT.

> "Before electronic media, there was ample reason to overlook the difference between physical places and social situations. [...] A given place-situation was spatially and temporally removed from other place-situations. It took time to travel from situation to situation, and disturbance was a measure of social insulation and isolation. Since

1 In this broad definition, ICT includes even early communication technologies such as the telegraph.

> rooms and buildings can be entered only through set doorways, people once could be included in and excluded from situations in clearly observable and predictable ways. Electronic media, however, make significant inroads into the situations once defined by physical location." (Meyrowitz, 1985, p. 116).

With the advent of mobile technology, the degree and nature of dissociation of place-situations and their social situations seem to be ever less predictable. Before people had mobile phones the caller could envision to a certain extent the situation they would intrude upon, for example when calling an acquaintance's landline phone of which one knew the location in the house. The experience of mobile Human-Computer Interaction (HCI) may be characterised by its potential to accompany us almost all the time and wherever we go.

Recent years have seen unprecedented adoption of the mobile phone, with global subscriptions surpassing 5 billion users[2], and the smartphone market share steadily rising[3].

The current advances in "smartphone" technology in particular have expanded the mobile phone's potential to support communication *anywhere*, *anytime* beyond interpersonal communication. A recent study found that the number of people accessing news and information daily on their mobile devices in the U.S. more than doubled between 2008 and 2009[4]. As well as communicating via SMS, email, phone, instant messaging, etc., people on the move may also be engaging with interactive mobile experiences, social networks, and real-time and location-based services, and pervasive games. In short, mobile HCI is becoming increasingly interwoven with our everyday lives.

2 According to the BBC on http://www.bbc.co.uk/news/10569081.
3 According to market analyst Gartner on http://www.gartner.com/it/page.jsp?id=1466313.
4 According to web analyst comScore on http://bit.ly/aYOHkx.

1.1 PROBLEM DEFINITION AND OBJECTIVES

The development of mobile platforms may be characterised by the convergence of software services (e.g. voice call, SMS, email, internet access etc.) and their expandability and customisation through applications (e.g. "app stores"), and the incorporation of hardware features and sensors (e.g. camera, Global Positioning System (GPS), accelerometer and compass).

The expandability of mobile platforms may introduce more and more applications that proactively push information to the user, which increases the device's potential to disrupt many situations inappropriately. For example, as well as the more familiar notifications of direct communication attempts such as an incoming voice call or SMS, the user may also be notified of, for example, a friend nearby using a location-based service, or a status update on a social network service. A side effect is that the mobile device's disruptive potential is increased. This can be a particular problem for mobile users as their context is apt to change radically over time, which increases the possibility of an interruption being inappropriate.

Interruptions are a mundane, yet central part of our everyday lives, which may have both detrimental as well as beneficial effects. As it will be expanded upon, interruptions may have negative effects on task completion and performance, even with potentially life-threatening consequences, for example in complex multitasking environments such as aviation (Latorella, 1998). On the upside, interruptions serve important functions of initiating information exchange or engagement with people (Sproull, 1984), interactive services (Ash et al., 2004) or mobile experiences (Flintham et al., 2007).

At the same time, the extension of mobile platforms to include more and more sensors allows devices to capture aspects of user behaviour. This provides exciting opportunities to tackle a long-standing challenge in *context-aware computing*: to make computing sensitive and responsive to its setting (Dourish, 2004) in order to provide the user with the right thing at the right time (Ho and Intille, 2005).

Building on the study of the phenomenon of interruptions, this thesis also seeks to examine how the space of opportunities in mobile development can be exploited from a designer's perspective to create systems that manage the interruptions in a way that they are more appropriate to current situations.

1.2 MOTIVATION

It has been a long-standing vision and promise that "Machines that fit the human environment instead of forcing humans to enter theirs will make using a computer as refreshing as taking a walk in the woods" (Weiser, 1991, p. 25). Context-aware computing (Schilit et al., 1994) has emerged as an approach to investigate how a computer "can adapt its behaviour in significant ways" to the human context (Weiser, 1991, p. 20). The motivation of this work is to contribute to the design of technology that leads to "true calm and comfort", by incorporating into the design of systems careful consideration of "how they engage our attention" (Weiser and Brown, 1996, p. 8).

Interruptions provide a fascinating phenomenon by which to study human attention, as they have the ability to substantially interfere with and guide our attention. In fact, as shall be shown from the literature (see 2.1), they are a central and important everyday quality of the *conditio humana*. The Janus-faced character of interruptions can cause detrimental effects to the interrupted person in one moment and beneficial effects in the next, which makes interruptions an important phenomenon to study and understand. Interruption may be a well-defined and well-researched phenomenon in the research communities of several disciplines, including Psychology, Computer Science, and

HCI. Yet, it still has many open research questions and opportunities for technical solutions, such as the questions and techniques examined in this thesis.

1.3 RESEARCH QUESTIONS

As a result of the ubiquity of interruptions in everyday live and their Janus-faced character, the need emerges to understand when recipients perceive interruptions as timely, and when they would (or should) rather not be interrupted. This focus on the subjective experience of the interruption is grounded in a literature review in chapter 2. The concept and approach to studying the subjective *receptivity* to interruptions is defined in chapter 4. The primary goal of this thesis is to understand *receptivity* to interruptions. Receptivity to interruptions is explored in a series of three empirical studies, for which the primary research question is:

- What is an opportune moment for an interruption?

In order to understand the ways in which opportune moments surface from the user's current *context*, the effects of different interruption strategies are investigated in the studies. More precisely, the studies investigate systematically the contextual factors that interact to determine people's receptivity. In turn, these studies are grounded in a holistic understanding of *context* that is developed from literature on phenomenology in chapter 2, and literature on interruptions in chapter 3, and synthesised into a research framework in chapter 4 by asking the following question:

- What are the contextual factors that make people receptive to an interruption?

The investigated interruption strategies focus on adapting the *timing* of the interruption. The focus on adapting the timing is not just due to the goal to understand receptivity to interruptions, but also because it is identified as the most practical adaptation done by interruption management systems (see 4.5). This thesis' secondary goal is to learn about the design of systems that manage interruptions by answering the research question:

- What design implications for interruption management systems may be drawn from the empirical studies?

The overall goal of this thesis is to develop a multi-faceted understanding of the ways in which opportune moments surface from the user's context *in situ*, by studying the situated effects of different adaptations of the timing of interruptions.

1.4 RESEARCH AREAS

This thesis adopts a *pragmatic attitude* (see 3.1). A mixed methods approach in order to contribute to knowledge is gained by drawing on, and positioning the thesis within the following research areas:

- *Phenomenological Philosophy* – Phenomenology serves as a backdrop that helps to develop an understanding of mobile HCI and interruptions in terms of the subjective, situated experience.
- *Mobile HCI* – The study of HCI with mobile devices is the overarching research area of this thesis.
- *Interruption management* – The study of the human management of interruptions and their effects, and the design of systems that assist in the management of interruptions frame the application domain of this thesis.

- *Cognitive psychology* – The thesis draws conceptual, terminological, and methodological inspiration from related work in cognitive psychology, especially on attention, information processing, and experimental design and analysis.
- *Ethnography* – The observational method employed in the field study (chapter 8) is inspired by ethnography.
- *Context-aware computing* – The thesis contributes to the discussion and design of context-aware systems through the study of context-sensitive interruption management.

1.5 CONTRIBUTIONS

The diagram in figure 1.1 shows the approximate thematic overlap of the areas and the weight of their influence on this thesis by their size. The key contributions this thesis makes are positioned within the areas of their main contribution. Overall, (A)-(C) contribute to understanding interruptions, while (D) contributes implications for the design of interruption management systems.

(A) *A model of the contextual factors of the interruption process* – A taxonomy that serves the purpose of a sensitising research framework to inform studies of the contextual factors involved in a communication process, such as an interruption. This is developed through a survey of the literature and its veracity is strengthened through the empirical studies (see 4.3).

(B) *An adaptation of the Experience-Sampling Method (ESM) to capture behavioural response to interruptions in situ* – ESM is extended with a temporal model that allows for the study of the interruption as a detailed temporal process (see 3.3.5) to enable the investigation of the *behavioural* effects of aspects of the interruption (see 5.5).

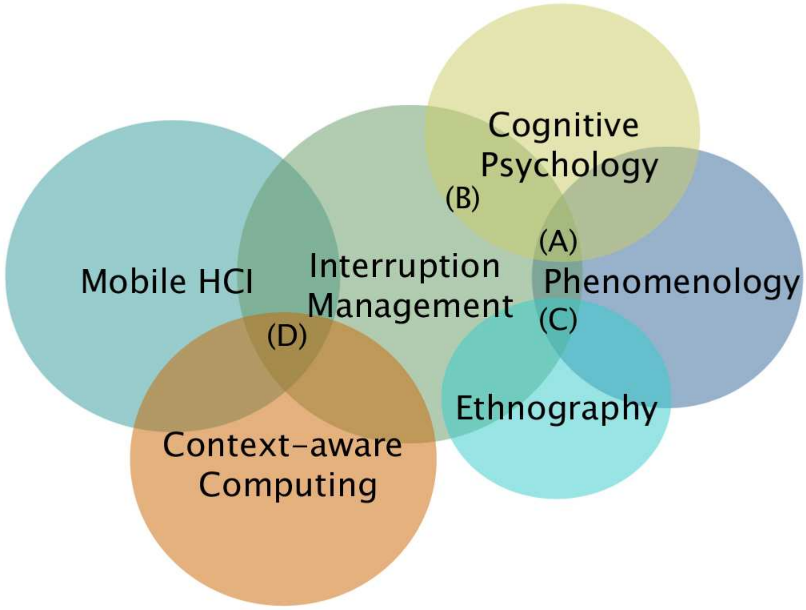


Figure 1.1: Venn diagram of the thematic overlap and influence (size) of relevant research disciplines and the relative positioning of this thesis' contributions (A-D).

(C) *Contextual interaction* – The empirical studies expose several ways in which the situated interplay of local and relational contextual factors determine receptivity to an interruption, such as the interaction of attentional demand in the environment, the social situation, interruption task type and content.

(D) *Temporal adaptation in interruption management* – The studies highlight implications for the design of systems that seek to manage interruptions by adapting the timing of interruptions to the user's situation. The thesis states design recommendations for the application of interruption management to personal communication and to pervasive experiences.

1.6 PUBLICATIONS OF THIS THESIS

Parts of the contents of this thesis have been accepted by peer-review for publication in conference proceedings in HCI or are in submission:

- Chapter 2 expands on the contents in *Fischer, J.E. (2010): Interrupting the Here and Now: Implications and Opportunities.* In *Proceedings of Mensch & Computer 2010, Duisburg, Germany. Oldenbourg.*

- The exploration of opportune moments for interruptions in chapter 3, and the account of Experiment 2 in chapter 7 has been published as *Fischer, J.E., Greenhalgh, C., Benford, S. (2011): Investigating Episodes of Mobile Phone Activity as Indicators of Opportune Moments to Deliver Notifications.* In *Proceedings of the International Conference on Human Computer Interaction with Mobile Devices and Services, MobileHCI '11. New York, NY, USA. ACM.*

- The model of contextual factors of interruptions introduced in chapter 4 has previously appeared in *Fischer, J.E. (2010): Studying and Tackling Temporal Challenges of Mobile HCI.* In *CHI '10 Extended Abstracts on Human Factors in Computing Systems. New York, NY, USA. ACM.*

- An early variant of the temporal model of the interruption process in chapter 5 has been published in *Fischer, J.E. & Benford, S. (2009): Inferring Player Engagement in a Pervasive Experience.* In *Proceedings of SIGCHI conference on Human factors in computing systems, CHI '09. New York, NY, USA. ACM.*

- Some experience-sampling tools discussed in section 5.4.4 have previously been discussed in *Fischer, J.E. (2009): ESM tools - a Critical Review. Position paper. MobileHCI 2009 Workshop on Mobile Living Labs.*

- The account of Experiment 1 in chapter 6 has been published as *Fischer, J.E., Yee, N., Bellotti, V. Good, N., Benford, S., Greenhalgh, C. (2010): Effects of Content and Time on Receptivity to Mobile Interruptions.* In *Proceedings of the International Conference on Human*

Computer Interaction with Mobile Devices and Services, MobileHCI '10. New York, NY, USA. ACM.

- Some of the ideas presented in chapter 8 have appeared in *Durrant, A., Rowland, D., Kirk, D., Benford, S., Fischer, J., McAuley, D. (2011): Automics: Souvenir Generating Photoware for Theme Parks.* To appear in *Proceedings of SIGCHI conference on Human factors in computing systems, CHI '11. New York, NY, USA. ACM.*

1.7 STRUCTURE OF THE THESIS

This thesis is structured as four parts. Part I surveys the relevant background literature in two chapters. Chapter 2 introduces phenomenology to develop the situated, subjective perspective from which to study and understand receptivity to interruptions in mobile HCI, and frames and discusses important concepts such as mobility and context. The second part of the chapter provides an overview of the main paradigms of the discourse of mobile HCI in order to position the thesis within its broader HCI field. Chapter 3 surveys the diverse research on interruptions and interruption management. In particular, interruptions in everyday life, research related to opportune moments for interruption, and systems and techniques that support interruption management are introduced. In addition, the chapter defines key terminology and proposes a temporal model of the interruption process that is used to measure behavioural reactions to interruptions in the studies.

Part II develops the approach and methodology employed to study receptivity to interruptions in the participants' everyday environment. Chapter 4 synthesises the model of contextual factors central to the thesis by drawing on communication models and implicit contributions to the model in the literature. The chapter elaborates the focus and the overarching theme of the empirical work alongside the principal research question. Chapter 5 describes the methodology used to study the experience of interruptions and relates this back to its conceptual underpinnings in phenomenology and its key concern to study the subjective experience *in situ*. This chapter then introduces a mixed methods approach to studying interruptions in situ based on Experience-Sampling (ESM) and ethnographic observation, which are supplemented by interviews, questionnaires and diaries. The chapter concludes by detailing how ESM has been adapted to study behavioural responses to interruptions in situ.

Part III covers the empirical work undertaken in this thesis. Chapter 6 reports on a study of effects of content and time of delivery of mobile interruptions on receptivity while participants went about their everyday lives for a period of two weeks. Chapter 7 describes an experiment testing the temporal coordination of interruptions in relation to the participants phone activity as a potential indicator of opportune moments to interrupt. Chapter 8 gives an account of a field study of a real-world application that relied on notification-driven interaction, where notifications were designed to prompt interaction with an application to create photo souvenirs at contextually opportune places.

Part IV reflects the empirical work in this thesis. Chapter 9 discusses the empirical studies in relation to each other, relates them back to broader issues in HCI, interruption management and context-aware computing and draws out implications for design. Chapter 10 concludes this thesis with a brief summary, and by stating the key contributions the thesis makes in terms of methodology, understanding receptivity and implications for the design of systems that employ an interruption management technique. Finally, ideas and implications for future work and the direction of the field are considered.

Part I

BACKGROUND

2

MOBILITY AND HUMAN-COMPUTER INTERACTION

This chapter helps to develop an understanding of the profound impact mobile technology has on our everyday experience by drawing on phenomenological philosophy (see section 2.1). This *phenomenological backdrop* (Dourish, 2001) also frames this thesis' understanding of *context* and its interplay with mobility. Importantly, this backdrop illustrates how interruptions can interfere with our orientation towards the world, and thereby may shape our experience of the world. This will particularly influence the methodology with its focus on the subjective experience and its mixed method approach as outlined in chapter 5.

A subsequent discussion of the paradigms that dominate the discourse of the social and psychological consequences of mobile HCI helps to place the thesis' contribution within the broader context of research in mobile HCI (see section 2.2).

2.1 A PHENOMENOLOGICAL VIEW OF HCI

As HCI is fundamentally concerned with the experience of technology, it is not surprising that ideas of the philosophical tradition most intimately concerned with the human experience have gained currency and influenced widespread adoption of its terminology in HCI. For example, the phenomenological foundation in Winograd and Flores (1985) and Dourish's concept of *Embodied Interaction* (2001) were conceived in order to give a more thorough foundation for thinking and reasoning about the ways in which computer technologies represent the world and "create and manipulate models of reality, people and of action" (Dourish, 2001, in the Preface). Thus, the same foundations in phenomenological philosophy are drawn upon to develop an understanding of the experience of interruptions.

2.1.1 *A short introduction to phenomenology*

According to Merleau-Ponty (1962), every analysis of the *conditio humana* must start with the fact that *existence is prior to essence*, that is, the subject's being in the world is prior to both object perception and self-reflection (Svanaes, 2001).

It is this fundamental structure of existence that Heidegger (1962) described with the notion of *being-in-the-world*. An emphasis on the subjective can be achieved by considering the way in which humans are in the world as opposed to how objects are in the world. Being human does not mean to be passive, but to be actively involved in the world in order to find and create meaning. Husserl introduced the notion of *life-world* to denote the active way in which we experience our mundane living (Fällman, 2003). The life-world differs for each individual. For example,

> "A kitchen affords a different kind of life-world to a chef than to a mechanic, though clearly these two lifeworlds may overlap in some ways as well. A life-world, then, is not just a physical environment, but the patterned ways in which physical environment is functionally meaningful within some activity." (Agre and Horswill, 1997, p. 114)

It is through the process of *involvement* in the world by which we endow meaning to the objects dealt with. This becomes apparent especially when we directly interact with objects in the world, such as cars, doors, tools and computers. However, the active nature of involvement is not restricted to what we mostly understand as direct interaction with objects. The process of perception is already guided by *intentionality*, the fact that experiences are always *experiences of* something (Fällman, 2003). Perception is not the mere process of representing the outer world in the mind; phenomenology after Husserl seeks to break with this Cartesian dualism, which in turn is less extreme than idealism, which completely rejects the existence of an outside world. By seeing the embodied self as part of the world, and as actively involved with it "the body is our general medium for having a world". (Merleau-Ponty, 1962, p. 146). The *body-subject*, as Merleau-Ponty calls it, necessarily looks onto the world from *one* viewpoint. The subjective nature of experience (i.e. being-in-the-world) is a consequence of considering existence and perception of the embodied self as prerequisites of *being-in-the-world*. Merleau-Ponty also acknowledges that this viewpoint is shaped by the kind of body, and by history, memories, culture and beliefs. These give past experience an existential influence on the present, illustrate the active ways of involvement in the world, and explain the individualistic character of life-worlds.

The phenomenological perspective then becomes useful when considering mobile HCI. The involvement of the embodied self changes with spatial and temporal location and orientation, but is guided by our previous lived experiences, shaping the individual experience of mobile interaction.

2.1.2 *Human-technology-world relations*

What is it that the interaction with technology adds to our involvement in the world? In an attempt to "consider what occurs when experience is directed through, with, and among technological artifacts" (Ihde, 1986, p. 139), and what Fällman (2003) refers to as a *phenomenology* *of human-technology relations*, Ihde describes three different forms of relations between a user, the technology, and the world.

First, Ihde introduces the *embodiment relation*, according to which the technology is incorporated into one's experience of bodily engaging in the world. This relation may be seen as a technological extension of the senses of bodily perception, or *vice versa*, as an internalisation of external devices "into" the bodily experience. For example, sight may be extended through a telescope or corrected through spectacles, hearing may be amplified through speakers or hearing aids, and touch may be extended to become a means for orientation in the world, as shall be explicated further in 2.1.5 by the example of a blind man's walking stick. According to this relation, technology mediates the experience of the world – the world is experienced *through* the technology. Ihde (1986, p. 141) formalises this relationship as:

(HUMAN-TECHNOLOGY)⟶WORLD

Secondly, there may be technologies that are positioned between the user and the world, but in contrast to the *embodied relation*, these often *represent* an aspect of the world in some form of abstraction rather than to *mediate* it. Here, the user is required to interpret the state of the world, rather than to perceive the world directly. Thus, the attentional focus is more on the technology itself. Examples of such technologies include the thermometer or speedometer, which provide their users with abstracted information about the world (temperature) or information in relation to the world (speed), but the user's perceptual focus is not on the world but on the technological instrument itself. Ihde (1986, p. 142) thus calls this relationship the *hermeneutic relation* and formalises it in this way:

HUMAN⟶(TECHNOLOGY-WORLD)

Thirdly, Ihde introduces a relationship in which the technology does not mediate or even abstract the world, it represents a relationship that is primarily one between the user and the technology; and the world seems to merely present a background to that interaction. He terms this relationship as an *alterity relation* and it should be very familiar to us as it incorporates the experience of virtual reality, computer games, chat rooms and other virtual representation of the world. It may also include using more traditional software programs, such as word processors or spreadsheet applications. Ihde formalises this relationship as:

$$\textsc{Human} \longrightarrow \textsc{Technology}\ (\text{-}\textsc{World})$$

As the focus of this thesis is on mobile technology, it intuitively appears that the physical *arrangement* of mobile devices lends itself to mediate all three kinds of technology-world relations. This may be due to the fact that they accompany our *mobility* in the world, which is arguably characterised by more frequently changing context[1] than more stationary technology usage settings.

For example, a mobile application for blind people that translates the device's camera's input pointed at the surrounding world into auditory signals[2] may be seen as an example of how the mobile phone mediates the bodily perception of the world in an *embodiment relation*. As an example of a mobile application that facilitates a *hermeneutic relation* consider an application that uses the device's GPS to display to you the speed at which you're travelling, or the microphone as a sound level meter, or a thermometer to display the current temperature. If the user is immersed in a game that is played in a virtual world displayed on the screen of the device, this represents an example of an *alterity relation*.

The interaction with mobile technology may be characterised by frequent *switches* between the different kinds of relations – in fact these

1 We shall see later on in this chapter that it is not actually the context that changes, what changes is our contextual *involvement*.
2 The mobile application vOICe by Peter Meijer is available at http://www.seeingwithsound.com.

switches may even be seen as their defining character (Fällman, 2003, p. 141). For example, (Fällman, 2003) analyses how a user traverses the different relations in the process of using a digital camera to capture a photo. While the user is looking at the world through the viewfinder she is in an *embodiment relation* with the technology, while when she looks at the captured image on the device's screen she may have switched to an *alterity relation*. As it will be illustrated these kinds of switches may be caused by interruptions, a fact which anticipates the role interruptions may play in our experience of the world.

2.1.3 *Context*

In the following, a phenomenological perspective of *context* is developed that frames the understanding of context in this thesis. In this view, context becomes an *interactional problem* rather than a *representational* one, i.e. rather than being context per se, something may or may not become contextually relevant in the moment because it is individually perceived to be (Dourish, 2004). Or as Svanaes (2001) puts it, "context can never be a property of the world, but [...] context rather is the horizon within which the user makes sense of the world." (Svanaes, 2001, p. 380).

The representational view of context is prevalent in the literature on *context-aware computing* (Dey et al., 2001, e.g.), the field that seeks to make computing sensitive and responsive to its settings. However, the representational view is problematic in that it assumes context to be an ontological problem of modelling entities with properties and relations that exist in the world (Dourish, 2004). Although this approach is useful from an engineer's perspective as it helps to break down the complexities of system design, such formalisms may be less helpful in developing an understanding of how such systems will be *perceived* by the user. According to the phenomenological perspective, we encounter the world, and find and create meaning in it through the process of involved perception – involvement. Therefore, throughout the rest of this thesis, the centrality of the subjective experience is implied, even when discussing *contextual factors* (see 4.4) in a seemingly objective and reflective way.

Based on this view of context, it seems important to discuss the concept of *mobility* to inform the ways in which to investigate the phenomenon of interruption in mobile HCI. By exploring mobility further in the following, both the conception of interruptions, and the employed methodology and research approach are motivated.

2.1.4 *Mobility*

As a consequence of the proposed view of context (see 2.1.3), being mobile is understood as a form of being-in-the-world (Fällman, 2003) – mobility and context are seen as inextricably intertwined. Accordingly for mobile HCI, the focus is on the subjective experience whilst being mobile, which entails not only the interaction with the technology, but first and foremost the ongoing involvement with the world as negotiated and enacted in the moment (Dourish, 2004).

For example, in an account of how the experience of using a mobile phone interferes with involvement in the immediate environment, Light (2008) describes some of the contextual complexities of dealing with a phone call whilst being mobile. Mobile phones are becoming smaller, and almost attached to the body, or experienced as hands-free earpieces. Mobile phone practices weave into other experiences of technology, such as driving or integrate other technologies, such as music players and cameras. By ringing to notify us of an incoming call, the mobile phone refocuses attention and triggers a negotiation not just with the surrounding environment, such as a crowd to navigate or a car to drive, but also with the cognitive spaces the recipient finds themself in when being called. Light (2008) illustrates Heidegger's notion that *our concern precedes our whereabouts* by the example of how one of her participant's regularly anticipates the day to come whilst walking to work, and thus already occupies multiple spaces (the environment and the anticipated workplace). She conjectures that

> "Mobility brings with it the likelihood of being in the wrong place, or surrounded by the wrong things, to deal most effectively with the call in hand. [...] calls, especially those on the move, are likely to involve additional transitions, as the local environment changes around the recipient and invites the person to cope with complex physical and social negotiations." (Light, 2008, p. 398-399).

Light's account suggests a multi-faceted view of mobility. In the following, spatial and temporal properties of mobility are considered in more detail.

2.1.4.1 *Spatial mobility*

Merleau-Ponty (1962) introduces the dual concept in which the body can both be spatially perceived *among* other objects and *directly* as an experiencing or living body distinct from the spatiality of objects. Whereas objects are spatial in their position, the body is spatial in its situation. Whereas the objects in the world are organised with reference in Cartesian space, the bodily space is organised in reference to potential affordances in the world (Svanaes, 2001).

Similarly, Heidegger dismisses Cartesian distance as a measure of closeness, as there would be far more intrinsic experiences of spatiality. It is rather the attention to something or someone, or as Heidegger calls it "the circumspection of concern" that "brings something close by" (Heidegger, 1962, p. 142). For example, we may feel closer to someone that we are talking to on the mobile phone than the person opposite of us on the bus; and we may perceive the picture on the wall as closer than the spectacles on our nose (cf. Mulhall, 1996).

Thus, Heidegger posits that "circumspective being in the world is spatial" (Heidegger, 1962, p. 145), in that it shows the character of *de-severance* and *directionality*. *De-severance* is the process by which closeness is achieved, which presupposes *directionality* towards a region out of what is de-severed brings itself close by, so that one can come across it with regard to its place (cf. Heidegger, 1962, p. 143). Hereby, *region* and *place*, again are understood not as measureable position in space, but equipment is understood as 'having its place', e.g. fitted and installed, where the region denotes the 'whither' to which the totality of places for a context of equipment gets alloted (cf. Heidegger, 1962, p. 136).

In a similar vein, Harrison and Dourish (1996) stress how physical spaces become meaningful through the occasioning of space, through activities and social interactions that take place in them, transforming them into *places*; consider for example the difference between a "house" (space) and a "home" (place). Harrison and Dourish (1996) illustrate how *space* refers to the geometrical position and orientation of physical arrangements, and *place* arises out of space through interaction, appropriation, and meaning-making in social practices. "Space is the opportunity; place the understood reality" (ibid. p., 69). As a result, places are invested with understanding of behavioural appropriateness, cultural expectations towards its role, function and nature, social meaning, and convention (cf. ibid.).

In the face of people's mobility through those places, Dourish (2006b) also points out how the appropriation of space is dependent on one's access to and legitimacy within the space. Accordingly, mobility may also be a result of avoiding problematic encounters with the authorities (e.g. for a homeless person) or a form of labour (e.g. for a taxi driver) (Dourish, 2006b).

For the designer of mobile technology, it is difficult to cater for the scenarios it is used in, as its usage settings are inherently uncertain. We are all familiar with the consequences of context-insensitive mobile phones, ringing, buzzing and beeping at the wrong time in the wrong place. Spatial properties may not just render the mobile device's "behaviour" inappropriate, they may also diminish its usability, or even make it dangerous; for instance think of reading or typing on your phone while driving or walking at the side of a road. Thus, context-aware systems need not only be aware of location, e.g. GPS coordinates, but need to have a notion of the semantic meaning of that space for the individual.

2.1.4.2 *Temporal mobility*

The temporality of experience is related to the temporality of being in the world. Merleau-Ponty (1962) cites Husserl in explaining how temporal experience unfolds in a continuous current view onto the world, where the past and the future are present in a continuum of *retentions* and *protensions*. A protension is an anticipation of a future event, and a retention is any past moment. The current view is influenced by retentions and protensions, and retentions and protensions are influenced by the current view through which one experiences now as well as experiences the protensions and retentions. Merleau-Ponty (1962) rejects the notion that a memory refers to a past event: retentions would be present events, even if they were the shadow of past experiences.

> "If my brain stores up traces of the bodily process which accompanied one of my perceptions, and if the appropriate nervous influx passes once more through these already fretted channels, my perception will reappear, but it will be a fresh perception, weakened and unreal perhaps, but in no case will this perception, which is

present, be capable of pointing to a past event [...]" (Merleau-Ponty, 1962, p. 480)

Following Merleau-Ponty, the current viewpoint onto the world is always shaped by all the previous viewpoints one has had onto the world. Previous experience gives the lens through which we look onto the world a unique tint. Following from the above, this is not just limited to the experience of *now*, but also applies to the experience of the past (retentions) and the anticipation of future events (protentions). For example, as mentioned in section 2.1.3, Light (2008) describes the way her participants anticipated experience (i.e. protensions) and thereby the experience of multiple temporal spaces as a routine element of mobility.

The notion of a relationship between temporal properties of mobility and experience anticipates the next section where phenomenology's key concept to understanding the interactional achievement of context is looked at: that is, involvement.

2.1.5 *Involvement and readiness-to-hand*

Heidegger has posited, "one must not understand a human being's existence (being-in-the-world) as simply a matter of spatial and temporal location with respect to other objects" (Fällman, 2003, p. 157). "Human mobility is a matter of shifting contexts; of changing involvements" (ibid.). Merleau-Ponty describes this involvement with the world as being directed by an embodied intentionality towards the world (Svanaes, 2001). Our orientation towards the world changes according to the direction of our intentionality.

This is often illustrated by the experience of tool use. Both Heidegger and Merleau-Ponty give compelling examples of how the experience of interacting with and through objects changes by adapting and extending the bodily experience through external devices (Svanaes, 2001). Merleau-Ponty (1962) uses the example of a blind man's walking stick to illustrate this. When the skill is learned to perceive the world through the stick, the stick no longer exists as an object, it is internalised by the body, it has become a part of "me". Learning is the process by which external objects are internalised. Ihde's adaptation for when these relations include technology (1986), and Fällman's extension (2003) for mobile devices have already been introduced (see section 2.1.2).

Heidegger (1962) famously uses the example of hammering to show how the experience of the hammer is *ready-to-hand* (*zuhanden*). That is, Heidegger's way of saying that the object is internalised and not present any more in the hammerer's consciousness than are the tendons of the hammerer's arm (cf. Winograd and Flores, 1985). It is through the process of *breakdown* that the hammer becomes *present-at-hand*, i.e. the hammerer becomes conscious of the hammer as an object again when it slips from the grasp of their hand. However, as Heidegger points out, in this state it is not completely devoid of its readiness-to-hand. Rather, it withraws itself from the readiness-to-hand for something with which one *concerns* oneself with (cf. Heidegger, 1962, p. 103) – it is this notion of concern, or *care* that Heidegger understands to be essential for Being (cf. Mulhall, 1996, p. 36) – which in this thesis is referred to as *involvement* (cf. Merleau-Ponty).

The notion of readiness-to-hand brings with it a fairly complex conceptual background that Heidegger aims to explicate as an 'equipmental totality' (Heidegger, 1962, p. 97), within which it finds its place. For example, a mobile phone only exists as a mobile phone in relation to its usability in order to call someone, the electricity saved in the battery that powers it and that has been produced, in relation to the mobile phone network and so on. Heidegger calls this the various ways of the 'in-order-to' of an equipment, such as "serviceability, conduciveness, usability, manipulability." (Heidegger, 1962, p. 97). The *readiness-to-hand* of a totality of equipment may be summarised as exhibiting the following characteristics (cf. Mulhall, 1996):

- A *for-which* the equipment is usable – an end product, for example a pen may be used for writing a letter, a phone to call someone and so on. This directedness is the *towards-which* of equipment.
- A *whereof* the equipment itself is made or constituted, the raw material and implicitly the labour and the intentions of the producers of the equipment.
- The *towards-which* of the work typically is destined towards *recipients*, who are encountered as part of the work.

These characteristics show that involvement in the world is always social, or in other words, "Being of Dasein is essentially Being-with-Others" (Mulhall, 1996, p. 64). The 'Others' are typically implicated the 'whereof' equipment is made and produced, and in the 'towards-which' equipment is used for. In this sense, *readiness-to-hand* is inherently intersubjective (ibid., p. 65).

Arguably, Heidegger's concern with equipment has made phenomenological thought popular in HCI. For example, Dourish (2001) applies Heidegger's notion of how the orientation towards things in the world may change from *ready-to-hand* to *present-at-hand* to the computer mouse. Dourish illustrates that as long as "I act through the mouse, the mouse is an extension of my hand [...]." (ibid., p. 109), and is in Heidegger's terms ready-to-hand. Then, when the mouse cannot be moved further at the edge of the mousepad, the orientation towards the mouse changes. "I become conscious of the mouse mediating my action, precisely because of the fact that it had been interrupted. The mouse becomes the object of my attention as I pick it up and move it back to the center of the mousepad." (ibid.), and becomes in Heidegger's terms present-at-hand.

It is because of the unreflected character of involvement that the readiness-to-hand of equipment (seemingly paradoxically) only becomes visible when it becomes *unready-to-hand* for our "concernful dealings" (e.g., the task at hand) in various ways (cf. Heidegger, 1962, p. 102). Heidegger introduces three ways in which equipment may become *unhandy*: through absence, damage, and through obstacles. These are addressed in turn.

First, Heidegger claims that something that is missing in our concernful dealings amounts to coming across something unready-to-hand. The ready-to-hand enters the mode of *obtrusiveness*, and reveals itself as something just present-at-hand (precisely because of its absence) (cf. Heidegger, 1962, p. 103).

Secondly, *breakdowns*, as illustrated by the example of the mouse may cause annoyance in the use of mobile technology, because they essentially create a barrier to successful interaction *through* the technology. Breakdowns may be caused by various flaws, such as software bugs, hardware failures, network congestion or unavailability, or human error. Consider the common example and cause of annoyance of "loosing signal" on a train, which causes the audio quality to diminish, or even the call to be dropped; with the phone this is likely to divert the attention from the conversational partner being *ready-to-hand* to the mobile phone being *present-at-hand*, i.e. *unready-to-hand* for the current concern of having a conversation on the phone.

Thirdly, and what appears most relevant in the context of this thesis is what Heidegger phrases as "something *unready-to-hand* which is not missing at all and not unusable, but which 'stands in the way' of our concern." (Heidegger, 1962, p. 103). Keeping to the example of the mobile phone, it is not just the interaction with the technology that may be interrupted by a breakdown, for the mobile phone has the potential to 'get in the way' of involvement in any kind of activity in the world. It effectively becomes an obstacle for which we are concerned with, as "the Being of that which still lies before us and calls for our attending to it." (Heidegger, 1962, p. 104). Unknowingly, Heidegger has provided a valuable definition of *interruption* for this thesis as 'an obstacle for what is *ready-to-hand*'.

Technology breakdowns and obstacles illustrate how interruption of our involvement in the world is an inherent feature of HCI. In fact, Dourish used the same terminology of *interrupted* experience in the example above. As Dourish describes, breakdowns in the interaction are caused "because it had been interrupted". In addition to breakdowns, more obvious examples of obstacles include interruptions such as notifications of system alerts, calendar events, or incoming communication. In effect, both explicit interruptions such as notifications and implicit interruptions such as breakdowns are technology-mediated interruptions, which have the potential to interfere with our orientation towards the world by directing and guiding our attention. This discovery is precisely why the discussion of the phenomenological perspective is fruitful in developing an understanding of the potential impact that technology-mediated interruptions have in everyday life. In this broad view, interruptions become a mundane, yet essential constituent of the human experience in that they initiate change in our orientation towards the world.

2.2 MOBILE HUMAN-COMPUTER INTERACTION

In this section, paradigms and particularities of mobile HCI are presented and discussed in order to frame the research area of this thesis. Whereas the previous section looked at mobility in a broader sense, this section discusses the social and psychological consequences that arise when interacting with technology whilst being mobile. Particular attention is paid to the challenges the mobile phone causes in light of the human's natural physical limitations, the idea that the mobile device is with us *anywhere, anytime*, and how as a result of mobility, interaction becomes fragmented across time and space. The concept of *context-awareness* and its relevant approaches to tackle some of the challenges in mobile HCI are introduced subsequently; in particular, techniques to coordinate remote communication are reviewed, as they represent the family of techniques that are used to support interruption management.

2.2.1 *Limitation of attention*

The term "human factors" describes the social and psychological consequences of and prerequisites for socio-technical systems in terms of the "limitations" the human physique and cognition bring to our interactions with technology. For instance, *Moore's law* may predict the doubling of the number of transistors on an integrated circuit approximately every two years, and consequently influence the miniaturisation of devices in approximately the same rate (Moore, 1965). However, the human physique does not undergo the same process. Hence, there is a limit to how small physically operated devices can get, for example touch screens or buttons on mobile devices.

Another important example of the limitations imposed by human nature is related to *attention*. "Information overload" is one of the more prominent and popular examples of problems associated with the notion of attention as a sparse resource or commodity that is challenged by the amount of sensory stimulation available from ICT. The term *attention economy* has been introduced (Davenport and Beck, 2001) to cater for the idea that information consumes attention (and vice versa) and that a wealth of information creates a poverty of attention that needs to be allocated efficiently.

It seems that much of the wealth of information in the world is now delivered to us through ICT, making it part of the problem. Hudson et al. (2002) and others suggest that technology may also be part of the solution, for example by filtering content that is relevant and important to us (discussed in more detail below in 2.2.4). Furthermore, it seems that the problems of limited attention are exacerbated when the user context is mobile and cognitive resources are needed to negotiate moment-by-moment orientation in the world. Consider the limitation of not being able to read information on a device's screen whilst driving (Perry et al., 2001) – at least safely.

In an effort to aid the designer of mobile technology, Oulasvirta et al. (2005b) have devised the *Resource Competition Framework*. It is introduced as an analytical tool to describe and capture a mapping of cognitive faculties in mobility tasks such as walking, way finding, sidestepping, and avoiding collisions to tasks in mobile HCI such as typing information and searching for information on a mobile device's display (Oulasvirta et al., 2005b). The framework makes assumptions about the available cognitive resources and their competition, that they are *functionally modular*, that they have a *limited capacity*, and that the limited resources may be *pooled* (reserved) by certain tasks.

This framework may be at odds with a holistic view such as embodied by the concept of involvement (see 2.1.5). However, it is useful in that it may be applied early in the design process of mobile technology to reflect on the challenges it may pose to the user's attention whilst being mobile. In the context of this thesis, the framework helped to select the places in a theme park where users were prompted to complete interactive tasks in a mobile application (see chapter 8). In more detail, an estimation of the attentional demands of the interactive tasks and the attentional demands to navigate the environment helped to hypothesise which places might be opportune to engage with the mobile application (see section 8.1.1.2, p. 218).

2.2.2 *Anytime, anywhere?*

The characterisation of mobile phone technology as an enabler of "access anytime, anywhere" has been described as one of the major promises of mobile technologies as the removal of the "binding between a fixed space and a person's information and communication resources" (Perry et al., 2001, p. 324). Nevertheless, assuming such an idealistic view often falls short of discussing social consequences for its users. As Sadler et al. (2006) observe in their study of freelancers' use of mobile phones, the device induces a potential for *perpetual availability* that creates a necessity to manage availability throughout the day in accordance with the mobile worker's shifting contexts. Availability was observed to be negotiated based on the contingencies surfacing from the task at hand, stressing the fluid and dynamic character of availability.

A study investigating the physical proximity of a user to their mobile phone by Patel et al. (2006) does show that 14 out of 16 participants kept their phones switched on 85% of the time on average (including nights), but only 58% of the time on average did they also have the phone within arm's reach. The actual availability is likely to be even lower, considering that actual availability is influenced by the multi-faceted ways in which we are involved in the world, rather than physical proximity to the phone alone. Perhaps unsurprisingly, when users were away from home, they carried their mobile phone with them significantly more often than when at home. Studies that take a critical look at the *anywhere-anytime* paradigm then often reveal insights that are contrary to the rhetoric of *always on*.

2.2.3 *Fragmentation of interaction*

Due to limited attentional resources whilst being mobile, mobile interaction necessarily becomes *fragmented* into multiple dimensions that are explored in the following subsections. Social interaction as mediated through mobile devices needs to be managed across domains. This in turn may exacerbate the difficulty in maintaining a desired ambiguity of social interaction such as *plausible deniability*, the phenomenon that the recipient relies on the sender's lack of information to excuse a lack of responsiveness (Nardi et al., 2000). With increasing mutual awareness, e.g. of location and activity through such techniques as discussed below in 2.2.5, the social acceptability of interaction management techniques such as *plausible deniability* are at risk.

2.2.3.1 *Spatial fragmentation*

Whereas one might argue that the simplistic term "anywhere" assumes a geometric notion of space where all places are functionally equivalent, Harrison and Dourish (1996) posit that space becomes meaningful through occasioning, through the activities that take place in them. One may also consider the social aspects of "anywhere" (Perry et al., 2001) to understand how interaction may be spatially fragmented and differentiated by its surroundings. For example, Perry et al. (2001) point out that anywhere access may be *possible*, but not *acceptable*, for example considering the kind of mobile phone conversation one might have when surrounded by others as opposed to when there is no one else around. Mobile HCI may cross boundaries (e.g. home and office) that one may have previously tried to keep separated (Aoki, 2005).

In addition, mobility and its accompanying anticipated restriction of access to information may lead to planning strategies such as "planful opportunism" to compensate for the anticipated lack of access (Perry et al., 2001), e.g. by printing out timetables and reservation codes needed on a trip.

2.2.3.2 *Temporal fragmentation*

Mobile HCI is not just subject to spatial fragmentation, but also to temporal fragmentation. The intrusive character of mobile phones and their attachment to an individual makes their temporal properties more evident, as Prasopoulou et al. (2006) observe. For example, consider the notion of *work time* and *private time*. Whereas in the age before mobile communication one would simply not be available at work outside of *work time*, mobile technology may cause disruption by channelling all communication to one device (Mazmanian et al., 2006), increasing the burden on the user to manage and account for availability to the social *cohorts* she is engaged with (Tolmie et al., 2008).

Many people may simply consider it inappropriate to have work-related conversations outside of *work time* (Perry et al., 2001). Among the unintended consequences of this blurring of the boundary between work and private time is that professionals may fear that not answering their mobile phone for work related phone calls after normal office hours would be seen as an evasion of responsibility by their employers (Prasopoulou et al., 2006).

Another consequence of the ways in which mobility fragments interaction temporally is illustrated by the use of *dead time* to resume information work tasks which have been stalled, or handle ongoing communication tasks such as email correspondence, often in order to keep on top of one's work (Perry et al., 2001). *Dead time* often occurs when mobile; it is time spent in transit, e.g. on public transportation, or in waiting, e.g. in airport lounges or railway stations. Interestingly, mobile workers often plan for *dead time*, for example by taking documents to read, or by downloading email to a laptop so it can be dealt with offline (ibid.). Furthermore, the authors report that the activities undertaken during *dead time* were more often mediated by the available technology and available resources and space in the environment than by priority or urgency. Perry et al. (2001) conclude that despite being hard to define due to its fragility, *dead time* seems to be key to mobile work and poses a challenge to the designers of mobile technology to enable the users to make better use of it, considering the diversity of context and the available resources.

Furthermore, technology-mediated interaction is often slow-paced and long-term (Dix et al., 1998), without clear beginnings and endings. Long-term interactive tasks are likely to be fragmented over days or weeks (Dix et al., 1996); for instance, consider a conversation via Short Messaging Service (SMS) that lasts for days or longer. In addition, as a consequence of mobility, mobile interaction often limited to short episodes or bursts, as attention is a sparse resource when mobile (Oulasvirta et al., 2005b). Mobile HCI tasks compete for attention with mundane tasks of orienting and navigating in the world (ibid.).

Temporal patterns of interaction may also be related to the level of engagement with the application (Fischer and Benford, 2009). The authors showed that the temporal properties *response time* (player's time to respond to a game message) and *elapsed time* (time between player messages) in the long-term, slow-paced, SMS-based game Day of the Figurines can be used to predict player engagement (Fischer and Benford, 2009). The research has raised several temporal challenges of episodic engagement: Players reported that they often felt 'flooded' with messages while disengaged, and that outdated messages confused them into taking actions that were no longer relevant when re-engaging.

The discussion of ways in which mobile HCI may affect different notions of time, such as *work time*, *private time*, and *dead time* anticipates both the intrusive, disruptive effects as well as the beneficial effects that interruptions delivered on mobile devices may have (see chapter 3). The acknowledgement that technology-mediated interaction often is fragmented, intermittent and long-term appears to make interruptions even more likely to play a central role in initiating the switches between various interactions on different channels, such as social networks, email, IM and location-based services.

2.2.4 *Mobile Context-Awareness*

Bringing context-awareness to mobile devices holds the promise of providing "task-relevant information and/or services to a user" (Abowd et al., 1999, p. 11) wherever they may be. According to Dey and Häkkilä (2008) the term context-awareness was possibly coined by Schilit et al. (1994), but the concept was probably first described in Weiser's vision of ubiquitous computing as taking into account "the human environment [...] and adapt its behaviour in significant ways" (Weiser, 1991).

Despite being at odds with the *interactional* view of context (see 2.1.3), the *representational* definition of context is considered here for its practical impact in informing the design of context-aware systems (Dey et al., 2001). Arguably, the representational nature of system design requires the representation of events in classes and objects. Dey et al. (2001) define context in the following way:

> "Context is any information that can be used to characterize the situation of an entity. An entity is a person, place or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves, and by extension, the environment the user and applications are embedded in." (Dey et al., 2001, p. 106).

Dey and Häkkilä (2008) concede that context-awareness is especially relevant when the user is mobile, as the user's context changes more rapidly. The authors also name reasons for introducing context-awareness on mobile devices. Context-aware mobile applications may help overcome some of the deficits of small screen sizes and limited input functionalities, provide shortcuts to contextually relevant device functionality, and offer location-based and personalised device services (Dey and Häkkilä, 2008).

For instance, ParcTab was arguably one of the first mobile context-aware applications, and pioneered some exemplary context-aware functionality (Schilit et al., 1994), such as:

- *Proximate selection*: Objects in the nearby proximity are emphasised in the User Interface (UI) for easier selection, e.g. to assist with choosing the nearest printer.
- *Automatic contextual reconfiguration*: Software is automatically reconfigured based on the sensed user context, e.g. when a user enters a meeting room, the shared workspace for that room is automatically displayed on the ParcTab.
- *Contextual information and commands*: Information displayed to the user and commands issued by the system may be context dependent, e.g. ParcTab shows files associated with a particular room when the user accesses the filesystem on the device whilst in that room.
- *Context-triggered actions*: Actions for the system to take based on sensed user context may be simple `if-then` rules.

To date, many mobile context-aware systems have been designed and evaluated, too many to be mentioned individually here. However, they seem to share some unifying characteristics:

RELIANCE ON SENSOR AND SOFTWARE USAGE LOG DATA For context-aware systems, there is a strong reliance on *contextual* data, usually in the form of raw data from software or hardware sensors that capture aspects of the users environment (e.g. noise (Krause et al., 2006), light, position), the user's orientation towards the device (e.g. orientation, distance from the device (De Guzman et al., 2007), glances at the device (Miettinen and Oulasvirta, 2007)), and the history of interaction with aspects of the system the context-aware service is interested in (e.g. interaction frequency per contact, calendar events (Horvitz et al., 2005b)).

MACHINE LEARNING AND STATISTICAL REASONING Raw data from sensors, software logs, and user feedback may often be fed into a machine learning toolkit in order to build predictive user models (i.e. which combination of sensor events best predict a certain relevant quality, e.g. *interest*) (Miettinen and Oulasvirta, 2007), that may be pre-computed and uploaded to a mobile device to do real-time reasoning (Horvitz et al., 2005b), or make use of an intermediate machine that broadcasts the inferred context to mobile clients (Krause et al., 2006).

USER FEEDBACK Many machine learning techniques rely on "labels", which label a specific case of data with a specific attribute, e.g. as 'good' or 'bad'. Labels may be acquired by explicit user activity such as voting or ranking items, or by implicit user activity (gleaned from log files), such as time spent on an item or depth of navigation. Most context-aware systems aspire adapting to individual user preference, which can only be learned from individual user feedback.

USER ADAPTATION Context-aware systems make use of the above techniques to adapt to the user's context, such as the user's attentive state (Streefkerk et al., 2006). Such adaptations may include, but are not limited to, adaptation of the content and style of information presentation (Streefkerk et al., 2006), the timing (see chapter 3), or the channelling of information (e.g. a call may be automatically routed to voicemail (Horvitz et al., 2005b)).

The systems studied in this thesis share characteristics with context-aware systems in that they monitor sensor data (device use in chapter 7, location in chapter 8), and as a result adapt the timing of user notifications in a rule-based manner.

2.2.5 *Coordinating remote communication*

Here, techniques for coordinating remote communication (De Guzman et al., 2007) are looked at in more detail because they are (broadly) related to interruption management. The aim of most of the work presented here is to enhance the communication partners' *awareness* of one another. The concept of technologically mediated awareness has been introduced in Computer-Supported Cooperative Work (CSCW) as "an understanding of the activities of others which provides a context of your own activity" (Dourish and Bellotti, 1992). Various techniques have been proposed to achieve mutual awareness and improve remote communication. To survey the related work in this space, this section adopts the differentiation between *receiver-oriented*, *caller-oriented* and *negotiated* techniques (De Guzman et al., 2007).

CALLER-ORIENTED Caller-oriented techniques aim at providing the recipient's status to the potential caller in an awareness display (De Guzman et al., 2007). Hereby, the caller is held accountable to balance their communication needs against the recipient's supplied contextual cues. Thus, complex interpretation such as human *availability* is left to the "human-in-the-loop"[3].

Examples include an awareness display that shows the remote collaborator's workload (Dabbish and Baker, 2003, also see 3.6.2); CALLS.CALM, which provides the caller with receiver-set information in terms of current role, location, and social setting (Pedersen, 2001); and AWARENEX, which displays the contact's locale and activity (Tang et al., 2001). Bardram and Hansen (2004) present AWARE, an architecture for *Context-mediated social awareness*, and AWAREPHONE an implementation of their architecture for mobile devices for hospital staff to support their remote collaboration, communication and coordination. Users are provided with data on location (within the hospital), activity (from a calendar), and status (user provided), as well as messages from other mobile users that can be prioritised. Instead of automating a system based on the sensed contextual cues, this work also follows the approach of distributing the contextual information to other users in order to provide remote collaborators with a sense of social awareness.

Oulasvirta et al. (2005a) implemented CONTEXTCONTACTS for Series 60 Symbian smartphones that displays to the user of the smartphone the contacts' location, time spent in the current location, user-selected ringtone profile and whether or not the mobile phone has been used recently.

De Guzman et al. (2007) and Avrahami et al. (2007b) both investigate the *actual needs* of awareness information to inform the design of an awareness display. De Guzman et al. (2007) conducted a diary study and asked participants to note the contextual information they wish they had when making a call and the information they wished the caller had when being contacted. The most frequently named type

3 For a more detailed discussion of human-in-the-loop techniques, see section 3.6.2, p. 72.

of information receiver's wished their callers had was on their *task status* (34%), *social availability* (15%), and *physical availability* (21%), comprising 71% in total. On the caller's side these categories also dominated, yet they made up only 58% of the information they wished they had about the receiver. *Location* was a more prevalent request here (15%), whereas receiver's wished their callers had this information only 9% of the time – indicating the relative negligibility of location alone in context-awareness as discussed above (see 2.1.4.1).

Avrahami et al. (2007b) present findings from a simulation laboratory study, where callers decided to either call or leave a message based on the *location*, *presence of other people* and *ringer state* of the receiver who had to simultaneously decide whether they preferred to receive a call or a message. In addition to the unsurprising finding that providing contextual awareness information significantly improves the callers' choices, the authors report that the type of contextual information provided leads to different kinds of improvements: more appropriate interruptions and better avoidance of inappropriate interruptions. Namely, callers avoid interrupting inappropriately when in the office, while knowledge that the receiver was alone helped to interrupt at an appropriate time (Avrahami et al., 2007b).

NEGOTIATED De Guzman et al. (2007) suggest that a *negotiated* technique provides a lightweight mechanism for both the caller and the receiver to jointly schedule conversations. An example system, THE NEGOTIATOR, was introduced by Wiberg and Whittaker (2005), based on the requirements they posited to successfully manage availability: *negotiation*, *dependence on awareness information of others*, *brevity*, and *cognitive effort* to switch to the interruption.

RECEIVER-ORIENTED Non-automated receiver-oriented techniques may be the most prevalent and are included in virtually every modern mobile phone. Examples include the ability to change notification modality or specifics based on caller identity and utilise caller identity to inform the decision whether to accept the call or not (Grandhi, 2009; Avrahami et al., 2007b). However, they introduce an overhead of management work on the receiver and do not consider the actual context of the receiver. Bellotti and Edwards (2001) point to experiments that have shown that people are very poor at remembering to update their availability states explicitly.

An example of a system that automates the process of estimating the user's availability is BAYESPHONE, which utilises user models to predict the cost of interruption versus the cost of deferring mobile phone calls (Horvitz et al., 2003). Sensing of the receiver's current context and automated adaptation to the receiver's inferred availability state may rank among the primary real-world problems that motivates research into context-aware computation.

Automated techniques that infer and adapt to the receiver's contextual availability are akin to systems and techniques that manage interruptions and are discussed in detail in the following chapter 3.

2.3 SUMMARY

Section 2.1 introduced a phenomenological foundation for this thesis, which informs the subjective, and situated perspective. On that basis, this thesis develops an understanding of mobile interaction as it is mediated by our involvement in the world. In particular, section 2.1 showed how mobile technology may mediate our perceptual relations with the world, and how these relations are subject to frequent switching. These defining characteristics of mobile HCI illustrate the fundamental ways in which technology may mediate perception. In this view, interruption – the central phenomenon under investigation in this thesis – is an essential constituent of the human experience, potentially guiding our attention and orientation and influencing with our orientation and involvement in the world. This in turn particularly informs the methodological underpinnings of studying interruptions in mobile HCI, detailed in chapter 5.

The consideration of the prevalent paradigms in mobile HCI presented in section 2.2 highlighted some of the social and psychological consequences of mobile technology. Limitations of attention become especially apparent when we are mobile, and some of the social consequences of *access anytime*, *anywhere* have been considered. It was demonstrated from the literature that as a result, mobile HCI becomes both spatially and temporally fragmented. One consequence of fragmented interaction is that the interruptive potential is increased. This observation anticipates the central ways in which interruptions are not just part of mobile HCI, but part of our everyday lives, to be expanded upon in the next chapter.

Having established the general background of this thesis in terms of its philosophical underpinnings, mobile HCI, context-aware computing and its application to coordinate remote communication, the following chapter turns its attention to concepts, techniques, and studies of interruptions and interruption management.

3 INTERRUPTIONS AND INTERRUPTION MANAGEMENT

Following from the previous chapter, it can be concluded that interruptions are central to the human experience. They may instigate a re-orientation in involvement in the world, initiate communication and information exchange and have diverse effects on human behaviour. Interruptions may be detrimental and disrupt an individual's current activity with potentially life-threatening effects, for example in complex multitasking environments such as aviation (Latorella, 1998; McFarlane, 1997) or driving (Iqbal et al., 2010). On the other hand, interruptions may be beneficial for their recipient, in that they initiate the reception of important news, events, personal information and communication in face-to-face settings as well as through ICT during work, private life and whilst being mobile. Interruptions may also initiate engagement with mobile applications, for example as part of an interactive, cultural experience.

In sum, it seems as if interruptions are an inevitable constituent of the human experience, and that they may have detrimental as well as beneficial effects on the current experience. Indeed, there may be a complex tension between the perceived disruptiveness of an interruption and its potential benefits (Hudson et al., 2002). What are the contextual circumstances that characterise the positive or negative experience or effects of an interruption, how do we deal with them and how may technology support interruption management? By discussing related work from different disciplines according to the *pragmatic attitude* (see 3.1), this chapter provides different viewpoints on the interruption process and explores human and technological techniques of interruption management.

After central terminology is formally introduced, related work that studied interruptions in everyday life is presented. Furthermore, a temporal model of the interruption process is devised, which sets the ground for the introduction of related work that looks more specifically at opportune moments for interruption. Then, the range of possible responses to interruptions is introduced and discussed. Systems and techniques to support interruption management are reported and their underlying approaches are contrasted. The discussed techniques, approaches, and motivations to study interruptions and build interruption management systems inform the methodical approach of this thesis and the development of a model of the contextual factors involved in the interruption process detailed in chapter 4.

3.1 DEFINITIONS AND ATTITUDE

An *interruption* has been defined in cognitive psychology as:

> "An externally generated randomly[1] occurring, discrete event that breaks continuity of cognitive focus on a primary task" (Coragio, 1990, p. 19).

The author further defines a *primary task* as:

> "A well defined activity that constitutes the current 'job' for the knowledge worker" (Coragio, 1990, p. 20).

However, for the purpose of this thesis the primary task is defined more broadly as the indivdiual's current activity, or a collection of cognitive processes, where cognitive processes may be any "higher

1 Note that, 'random' here means that its occurence is non-deterministic from the user's perspective, cf. Dix (1990).

mental processes, such as perception, memory, language, problem solving, and abstract thinking"[2]. Even though interruptions may also be caused internally (Miyata and Norman, 1986; McCrickard et al., 2003), research in interruption management usually focuses on effects of external interruptions and strategies to deal with them.

Another definition of interruption introduces the term *interruption task*:

> "Any event or activity that demands attention to be redirected from the primary task toward an interruption task, forcing a task-switch" (Dabbish, 2006, p. 12, paraphrasing McCrickard et al. (2003)).

Where the *interruption task* is defined for the purpose of this thesis as:

> A new activity, or a collection of cognitive processes which are instigated by the occurrence of an interruption.

In this sense, the breakdown event that causes one's consciousness of the computer mouse to switch from ready-to-hand to present-at-hand in Dourish's example (see 2.1.5) is an interruption just like more obvious examples of technology-mediated interruptions, for instance a ringing phone or an incoming SMS.

Furthermore, for the purpose of this thesis, the term *interruption management* is defined both considering human management and technological support of interruptions as:

> The activities by which humans manage interruptions and their effects, including social practices, policies and tools (Andrews, 2004) that support that kind of management.

For the purpose of this thesis, *interruptibility* is mostly understood as it is used in most of the literature, as a measure of a person's general

2 According to the Glossary of Psychological Terms by the American Psychological Association available at `http://www.apa.org`.

state of receptiveness to any form of interruption (Avrahami, 2007, p. 19). Whereas *receptivity* as defined in more detail in section 4.1 may be understood as a measure of receptiveness that more explicitly relates to communication, for which the consideration of sender and the content of the interruption are significant.

ATTITUDE As in the discussion of context as an *interactional achievement* vs. a *representational problem* (see 2.1.3), an ideological tension may become apparent at this point. On the one hand, a phenomenological view of interruptions suggests that the individual's involvement in the world is at its centre, on the other hand a cognitive view of interruptions suggests the centrality of the view of the brain as an information processor.

However, this thesis takes on a pragmatic attitude in that it draws on apparently disparate views and traditions. This pragmatic attitude also informs the mixed method approach to achieve ecological validity in real-world studies (see chapter 5).

As this chapter will show, the proportion of work on interruptions and interruption management primarily drawing on a cognitive view of interruptions suggests that this view has been preferred in previous work. In particular, to inform engineering solutions or *interventions*. This may be due to pragmatic reasons such as the need to represent contextual factors in a system.

Furthermore, interruptions have been investigated in psychological studies of attention long before computers were invented (e.g. Zeigarnik, 1927). Also, HCI has historically drawn upon methods and knowledge in cognitive psychology. This is apparent for example in input-output models of information processing, the cognitive view of the model human processor (Card et al., 1986) and Norman's human action cycle (Norman, 1986), which have strong parallels with the faculties assumed by the cognitivists such as long-term and short-term memory.

Therefore, it appears that a multi-disciplinary thesis on interruptions in HCI benefits from drawing on psychological work as well as having a philosophical foundation.

Despite that this chapter reviews mainly psychologically influenced work, it is developed with the holistic, phenomenological view introduced in chapter 2 in mind. The intention in bringing together these diverse views in a pragmatic attitude is to generate understanding – to let the apparent clefts become spaces in which new possibilities are revealed.

3.2 INTERRUPTIONS IN EVERYDAY LIVE

Regardless of their effects, interruptions are a constituent of everyday life, shaping the ways in which we communicate. Ethnographic studies show that interruptions are a frequent method to initiate informal conversations in the workplace (Whittaker et al., 1994), and mark the beginning of information exchange in the working environments of healthcare professionals (Ash et al., 2004), mobile professionals (O'Conaill and Frohlich, 1995), information workers in investment management (Gonzalez and Mark, 2004; Mark et al., 2005), large corporations (Su and Mark, 2008), and small offices (Rouncefield et al., 1994), amongst others.

In his seminal study of the nature of how managers continuously interacted with their information environment, Sproull not only observes interruptions as characteristic of a manager's workday, but concludes that, because managers often generate their own interruptions, it is difficult to even distinguish between interrupted and interrupting activity, and that it "may be more sensible to think about the manager as a multitask processor" (Sproull, 1984, p. 23).

Ash et al. (2004) report their experiences of deploying patient care information systems and observing the system being used by healthcare professionals. The authors report that the human-computer interface is often not suitable to the highly interruptive context it is used in. Health care professionals were interrupted routinely throughout their day by beepers, telephones, colleagues, and patients. The authors further report that the mismatch between interface and (interruptive) use context have lead to errors in the process of entering and retrieving information, in several cases almost leading to wrong orders for medications. The case illustrates how interruptions are mundane to certain work settings, and how contextually ill-designed information systems may lead to hazardous outcomes.

An early observational study of two mobile professionals showed how they were interrupted over four times per hour, for an average duration of around two minutes (O'Conaill and Frohlich, 1995). The authors state that professionals benefited from the interruption in 64% of the cases. However, the study also suggests that interruptions may indeed be detrimental to the ongoing task: the authors revealed that in over 40% of cases the professionals did not resume the primary task after dealing with the interruption. Yet, it is not clear on what basis the authors decided to flag a primary task as permanently suspended – was it enough if the participants did not resume the task immediately after they completed the interruption task? In contrast, in a more recent and thorough study, Mark et al. (2005) found that 77% of the work suspended due to interruptions was resumed on the same day.

A diary study of task-switching revealed how highly interwoven tasks are common in the workplace of information workers (Czerwinski et al., 2004). Their eleven participants reported an average of 50 task switches during a week, whereby around 40% of the switches were self-initiated, and around 30% were prompted by phone calls, email, diary reminders and other persons (Czerwinski et al., 2004).

Gonzalez and Mark (2004) introduce the concept of *working spheres* to better accommodate the high level of routine discontinuity information workers experience on a day-to-day basis. The authors conducted an observational study at an investment management company in the U.S. over a period of seven months and found that people worked in an average of ten different working spheres and that switches between spheres are initiated equally by both internal (i.e. self-initiated switch) and external interruptions, on average every 11.5 minutes.

Mark et al. (2005) found in a follow-up paper on the same study that working spheres were interrupted externally more often than not (57%). The authors conclude that fragmentation of work is often a way of life for information workers.

Su and Mark (2008) conducted an ethnographic study, where researchers shadowed 19 informants in a large US-based corporation for over 550 hours in total. The authors observed the prevalence of *communication chains* – successive, quick interactions – and described their role as crucial in *aligning* tasks in multitasking environments. When triggered by external interruptions, *communication chains* tend to have more links, more media switches, and more organizational switches. The authors observe that these externally triggered communication chains cause stress, and recommend that future systems that manage interruptions estimate the likelihood that an interruption triggers a communication chain (Su and Mark, 2008).

An ethnographic study of a small office not only confirmed that work in that setting is equally characterised by interruptions, it yielded the interesting finding that the kinds of work triggered by interruptions were mostly the work that workers reported they enjoyed the most (Rouncefield et al., 1994), namely: contact with customers. Conversely, the work that was interrupted was the least enjoyed, the "paperwork". The authors point out the irony that despite being the favourite part of work, interruptions often initiated subsequent paperwork, such as enquiries about bookings.

3.3 temporal models of the interruption process

Despite of different focuses and research aims, the studies reviewed in the previous section all provide evidence that interruptions are part of the everyday environments they studied. But how exactly may the *process* of interruption be described?

This section explores accounts of the interruption process in related work, particularly in terms of its temporal course. This literature informs this thesis' suggestion of a temporal model of the interruption process at the end of this section (see 3.3.5, p. 55). This model is central to this thesis' adaptation of ESM (see 5.5, p. 136) and the experimental design of the studies presented in chapter 6 and chapter 7.

The reviewed literature serves as a basis to understand the temporal phases of the interruption process. This understanding is deployed to devise a temporal model to be applied in this thesis' experimental measuring. In particular, the model introduces the ways in which temporal phases may be experimentally manipulated as independent variables and the ways in which they may be measured as dependent temporal phases of the interruption process.

This then enables the tackling of the main question of this thesis in an experimental manner: What is an opportune moment for an interruption?

Several authors have acknowledged that *interruption is a process*, spanning temporally from interruption notification to perception of the interruption content and potential interruption task performance to potential resumption of the primary task (McFarlane and Latorella, 2002; Altmann and Trafton, 2004; Avrahami et al., 2008). Authors have suggested different temporal models of the interruption process,

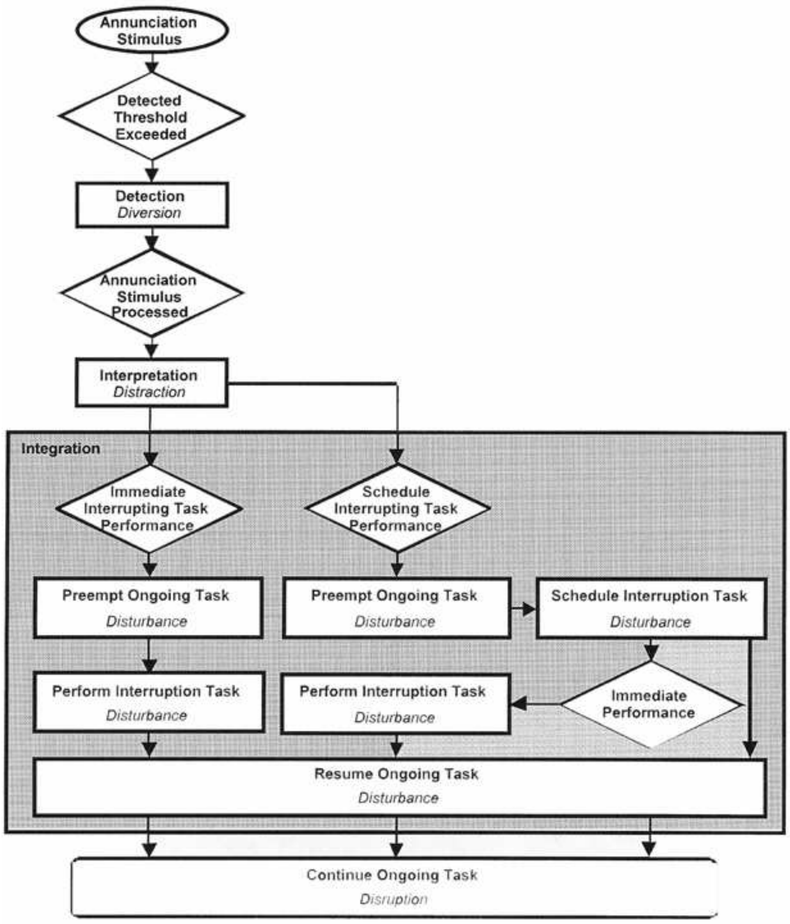


Figure 3.1: Interruption Management Stage Model. Reproduced from McFarlane and Latorella (2002). ©2012 Taylor & Francis, with permission.

depending on the granularity of the studied temporal intervals and research focus, such as the studied medium (e.g. IM in Avrahami et al. (2008)) and the specific research questions. Before presenting the temporal model employed in this thesis, four temporal models that influenced the conception of this thesis' model of the interruption process are presented and discussed.

3.3.1 *Interruption management stage model*

In her Interruption Management Stage Model (IMSM), Latorella characterizes interruption management as information processing stages (Latorella, 1998; McFarlane and Latorella, 2002), with the sequential

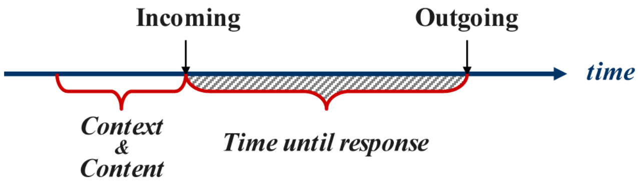


Figure 3.2: Response time as an indicator of responsiveness. Reproduced from Avrahami et al. (2008). ©2012 ACM, with permission.

stages *annunciation stimulus*, *detection*, *interpretation*, *preempt ongoing task*, *perform interruption task*, *resume*, and *continue ongoing task* (see figure 3.1).

IMSM acknowledges cognitive processes such as the requirement that the *annunciation stimulus* must be detected (e.g. heard), and acknowledges the flexibility that the interrupting task may also be intentionally dismissed or scheduled to be completed at a later stage (McFarlane and Latorella, 2002). Furthermore, according to McFarlane and Latorella (2002), the model serves to inform research to identify issues concerning the interrupted and the interrupting task, the "operator", and the environment; and it evokes thinking about dependent measures useful for experimental study of the interruption effects, which will be focussed on in the following subsections.

3.3.2 *Response time*

In a simple model designed to understand responsiveness to incoming instant messages, Avrahami et al. (2008) suggest *response time*, the time between an incoming message and an outgoing message between the same two people as a measure for responsiveness (see figure 3.2).

They showed that responsiveness is affected by the message such as the sender, the content and the presentation of the interruption, and by the time of day and the day of the week, and the desktop activity in the two minutes prior to the interruption (Avrahami et al., 2008).

In this thesis, measuring *response time* provides an important dependent measure of receptivity to interruptions.

3.3.3 *Time course of an interruption*

Altmann and Trafton (2004) suggest a model of the time course of an interruption marked by the discrete events *alert*, *start of secondary task*, *end of secondary task*, and *first action after interruption*. In contrast to the IMSM, the authors explicitly classify the intervals between the discrete events as *primary task performance*, *interruption lag*, *secondary task performance*, *resumption lag*, and *further primary task performance*.

In their experiment, Altmann and Trafton (2004) showed that the *resumption lag* is affected by the *interruption lag*, that they manipulated to vary in length. Thereby, Altmann and Trafton (2004) provide a good example of how time measures can be used both as independent and as dependent variables to show behavioural effects of interruption timing. Their experiment demonstrates the usefulness of temporal dimensions in experiments on interruptions.

The authors elaborate from the viewpoint of cognitive psychology that the interruption lag may be crucial in facilitating the resumption after the interruption, as during that phase goals to be accomplished at resumption may be prospectively encoded. Retrieval cues may play an important role in this process as their encoding may assist the resumption of the task. An example cue for a desktop setting is the visual availability of the primary task window after interruption notification and before responding to the interruption. The experiment of Altmann and Trafton (2004) showed that cue availability during the *interruption lag* affects the time taken to resume the task (*resumption lag*), supporting their theory of prospective memory encoding.

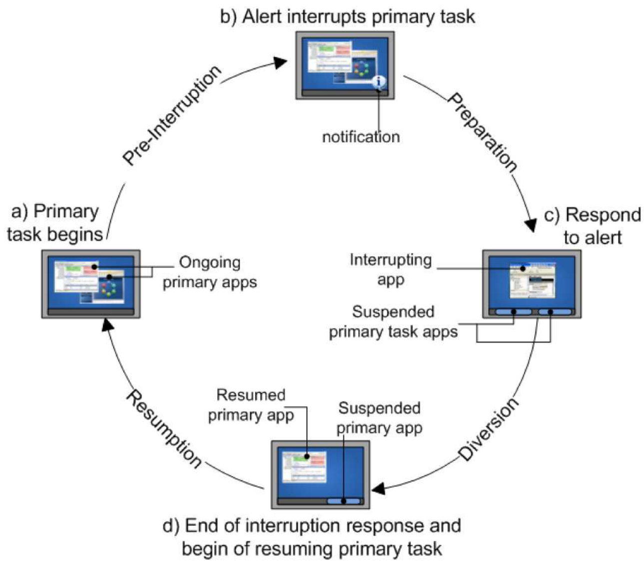


Figure 3.3: Interruption lifecycle model. Reproduced from Iqbal and Horvitz (2007). ©2012 ACM, with permission.

3.3.4 *Interruption lifecycle*

Iqbal and Horvitz (2007) present and investigate in a field study the phases of an interruption lifecycle of users engaged in desktop work (see figure 3.3) that they characterise by the phases:

A) *Pre-interruption*: the user is engaged with their primary task applications.

B) *Preparation*: after the user is notified of an incoming email or IM, they mentally prepare to respond.

C) *Diversion*: the user suspends the primary task and deals with the interruption.

D) *Resumption*: the user returns to resume the suspended primary application.

They distinguish between alert-driven interruptions within 15 seconds of the notification and self-initiated interruptions and find that if users switch to the interruption immediately users perform activities to stabilise the primary task state in the preparation phase – a form of 'state saving', as it were – such as completing a sentence or copy-pasting significantly more often and quicker than in the pre-interruption phase. In the two-week period of the study users switched to 40% of email and 71% of IM alerts within 15 seconds.

Furthermore, Iqbal and Horvitz (2007) report that the visibility of the primary task window assisted the resumption of the primary task after the interruption was dealt with, measured in terms of a significantly shorter resumption lag when more than 75% of the primary task window was visible compared to windows that were less than 25% visible.

The findings of Iqbal and Horvitz (2007) that activities are carried out in a preparation phase to assist the anticipated resumption, and the facilitated resumption when cues were available supports the prospective memory encoding theory presented by Altmann and Trafton (2004). Users utilize the short time lag between alert and switch to prepare the anticipated return to the task after the interruption is responded to.

Whether we regard interruption as a process or as a cycle shall not matter for the further argument of this thesis, as both have their merits. When considering the details of the interruption it may be more appropriate to view interruption as a linear process, whereas if we want to consider how the interruption is embedded in the user's larger activity context the cyclic model may be more appropriate.

Viewing interruption in terms of a temporal process shows how temporal stages of the interruption are interrelated and how they may be measured and utilised to elicit detailed understandings of the effects of contextual aspects of the interruption, such as the presence of a cue (Altmann and Trafton, 2004; Iqbal and Horvitz, 2007), or the sender, content, and presentation of the interruption (Avrahami et al., 2008). In this thesis, the metrics to discern significant impact of the independent variables are measurable differences in the temporal phases of the interruption process as developed in the following subsection.

3.3.5 *A temporal model of the interruption process*

Whereas the time course model (see 3.3.3) and the lifecycle model (see 3.3.4) provide detailed models for experimentation, their implicit requirement to control the primary task limits their applicability to laboratory settings. For example, to measure the interruption lag or the resumption lag the participant's primary task or activity that is interrupted needs to be controlled. However, this thesis' concern to study interruptions in real-world settings implies that temporal dimensions of the interruption need to be studied without knowledge of the primary task, as the user's current involvement in the world cannot be experimentally controlled.

Response time (see 3.3.2) provides a temporal metric that can be measured without knowledge of the user's ongoing activity. For example, by logging the time between delivering an interruption and receiving a response. However, a more detailed model of the interruption process is desired in this thesis. For example, in order to evaluate the time lag between interruption delivery and switching to, or accepting the interruption. Also, depending on the content of the interruption, the interruption task may need more or less time to be completed. These are just two examples of measures that contribute to overall response time.

Consequently, a temporal model of the interruption process was devised that consists of consecutive time periods that can be meaningfully correlated with the experimental manipulation to analyse the data (see figure 3.4). This model allows to study situated effects of contextual aspects of the interruption in sufficient detail without the need to control the participant's primary task. In this thesis, behavi-

Key
<Event> Interruption delivered
<*Phase*> ↓ *Acceptance time*
Interruption accepted
↓ *Decision time*
Task accepted
↓ *Task time*
Task completed

} *Response time*

Figure 3.4: Temporal model of the interruption process developed and applied in this thesis.

oural measures are collected in terms of timestamps logged during interaction with the experiment application.

The model consists of four temporal phases, whereby *response time* represents the sum of the three consecutive phases *acceptance time*, *decision time* and *task time*. Response time describes the temporal process of interruption in its coarsest form. Response time refers to the time lag between the delivery of the interruption and the completion of the interrupting task, where the task may be to respond to a message with a rating (see Experiment 1, chapter 6), or to complete a task (see Experiment 2, chapter 7).

This temporal model also allows the interruption process to be explored on a finer temporal level. In particular, Experiment 2 looks at the interruption process in more detail by studying interrelations of the experimental conditions with *acceptance time* (the time lag to actively accept a notification by pulling it from the inbox), *decision time* (the time lag to decide to accept the task presented when the notification was accepted), and *task time* (the time it takes to complete the task) (see figure 3.4).

In this way, the temporal model allows for conclusions about the situated effects of contextual aspects on a specific phase in the interruption process. For example, analysis according to the model may show that while the content of an interruption influences the subsequent interruption task time, it may also reveal that content does not affect *acceptance time*.

3.4 OPPORTUNE MOMENTS FOR INTERRUPTIONS

This section explores research on the nature of opportune moments for an interruption. In line with the focus on the subjective experience (see 2.1), whether a moment is perceived as opportune for an interruption is understood in terms of the recipient's receptivity to the interruption. On the one hand, the literature reviewed here informs understanding of the cognitive conditions that frame whether the interrupted person is likely to perceive an interruption as timely. On the other hand, the research informs the design of systems that manage interruptions by mitigating detrimental effects that may arise when delivering interruptions at inopportune times. These approaches are typically based on detecting or predicting *opportune moments* for interruption (Iqbal et al., 2005; Iqbal and Bailey, 2007, 2008; Adamczyk and Bailey, 2004; Horvitz et al., 2005a; Bailey et al., 2006). What makes a moment opportune for an interruption and how can it be detected? The question of the constitution of an opportune moment for an interruption is central to this thesis.

In section 3.4.1, work that informs understanding an opportune moment for an interruption in relation to *breakpoints* in the user's primary task is reviewed.

A related body of work that attempts to predict opportune moments more broadly from sensors that capture user activity is reviewed in section 3.4.2.

For this thesis, the research reviewed here is particularly relevant in that it informs the synthesis of the model of contextual factors presented in 4.3. In particular, postulates about the relation between interruptions, opportune moments and the user's cognitive activity are distilled from the review to emphasise this thesis' concern with the embodied processes pertaining to the reception of an interruption. Specific points of correspondence between this section and the postulates in section 4.4.1.2 (p. 94) are identified by [P1-P12].

3.4.1 *Breakpoints and opportune moments*

An influential body of work has associated opportune moments for interruptions with naturally occurring breakpoints in the primary cognitive task. According to Iqbal, "breakpoints reflect transient reduction in cognitive task processing."[3]

Miyata and Norman (1986) relate the user's memory load at different stages of the primary user task to the disruptiveness of interruptions at these stages on the basis of Norman's human action cycle (Norman, 1986) [P1]. Norman (1986, p. 41) approximates the process of performing and evaluating an action to consist of seven stages of user activity:

1. Establishing the goal
2. Forming the intention
3. Specifying the action sequence
4. Executing the action
5. Perceiving the system state
6. Interpreting the state
7. Evaluating the system state with respect to the goals and intentions

3 In a presentation at CHI (Iqbal and Bailey, 2007).

With respect to these stages, Miyata and Norman (1986) posit that interruptions would be least disruptive if they occurred after evaluation and before forming a new goal [P2]. In turn, interruptions would be most disruptive during the cognitively most demanding phases of planning and evaluation [P3]. However, "If the change occurs at the conclusion of the current task or at a natural breaking point, then there is probably no difficulty" (ibid., p. 275).

Since then, a host of laboratory empirical work has largely validated their assumption. For instance, one study has found that the time to attend to an interruption was significantly longer when participants were interrupted between activities within a task, than when interrupted between tasks or before starting or after ending the task (Latorella, 1998). Another study has shown that the time to resume the primary task of programming a VCR after an interruption task was lowest when the interruption occurred right before a new task stage (Monk and Trafton, 2002).

More evidence of the opportuneness of interruptions between tasks is given in a study that looked at the effects of interruptions by IM (Cutrell et al., 2001). Participants attended significantly more slowly to messages delivered during a cognitively more demanding task [P4]. The study concludes that the optimal design solution for a system that manages interruptions is to queue interruptions and deliver them when recipients have completed their current task.

More recently, interruption management has advanced by drawing on models of event perception from neuropsychology, which posit that the brain structures our everyday experience into temporally bounded episodes, where parts and sub-parts are reliably correlated with ecologically relevant features of the action (Zacks et al., 2001). The authors show that patterns of brain activity while watching a video match the pattern in which participants recalled events from the video on both a coarse and a fine level of event hierarchy. Adamczyk and

Bailey (2004) show experimentally, that coarse and fine breakpoints occur between tasks and sub-tasks and that the more opportune moments for interruptions lay at coarse breakpoints [P5].

In an effort to align workload of the primary task with opportune moments for interruptions more precisely, task models were created by using pupillary response as an indicator of workload (Iqbal et al., 2005) and training data of participants' tags of breakpoints in the interaction during the tasks. The study showed that the workload within a task differs by subtask, decreases at subtask boundaries, decreases more at boundaries higher in the task hierarchy than lower in the hierarchy [P7]. Iqbal and Bailey (2005) proceed to compare the appropriateness of interrupting at subtask boundaries with lowest and highest workload and find that interrupting at predicted best moments (i.e. moments with lowest workload) leads to significantly shorter resumption lag, lower annoyance, and higher respect ratings [P6].

In another experiment, participants were interrupted at different subtask boundaries in the task hierarchy (Iqbal and Bailey, 2006). The authors use resumption lag as an indicator of the cost of interruption and find that the cost is higher when the interruption occurs

a) at a deeper level in the task hierarchy,

b) where more information has to be carried over (i.e. remembered) across the boundary to the next subtask,

c) with a more difficult next subtask [P8].

Later research showed that the higher the breakpoint was in the task hierarchy the lower user-reported frustration was (Iqbal and Bailey, 2008) [P9]. This shows that the position of breakpoint plays an important role in whether an interruption is perceived as opportune.

LIMITATIONS AND CONSEQUENCES Interruption systems that base their inference of opportune moments for interruptions on models of the primary task require the designer of mobile technology to create complex task structure descriptions for every potential task the user engages in. Thus, the design of such a system may be daunting. An approach to address this problem relies on building predictive models of the occurrence of breakpoints in tasks (Iqbal and Bailey, 2007). Models were learned from training data of participants' tags of breakpoints in the interactive task execution (ibid.).

However, Iqbal and Bailey (2008) explored how well such models can be used to detect breakpoints in *novel* tasks, i.e. authentic tasks the user engages in that were not used to create the models in the first place. While the models work reasonably well to detect a breakpoint (the authors report 55% accuracy), the system's inference of the type of breakpoint (coarse, medium, fine) differed often and substantially from the users' perception of breakpoint type.

Furthermore, the need to monitor the user's primary task in order to predict breakpoints requires heavy instrumentation with both software (Iqbal et al., 2005; Iqbal and Bailey, 2007, 2008; Horvitz and Apacible, 2003; Horvitz et al., 2005a) and hardware sensors (Iqbal et al., 2005; Avrahami et al., 2007a) in laboratory environments. Few studies have taken on a more ecologically valid approach or looked at the effects of interruption timing in mobile settings, where mobile HCI tasks may often compete with the tasks of orienting and navigating whilst being mobile (Oulasvirta et al., 2005b).

To that end, Ho and Intille report on a study that tested the receptivity to interruption at transitions in physical activity. A body-worn accelerometer sensed transitions such as from sitting to standing/walking. Participant's self-reported receptivity at these breakpoints in physical activity was significantly higher than at random other times (Ho and Intille, 2005) [P12].

This thesis contributes to the exploration of interruptions in ecologically valid studies. With respect to the inspiration particularly drawn from research on breakpoints, the experiment presented in chapter 7 takes on the challenge of investigating potential breakpoints in mobile phone activity as instances of opportune moments for interruptions in a naturalistic setting, i.e. people participate as they go about their everyday lives.

3.4.2 *Predicting interruptibility*

Various authors have conducted empirical studies with the intent to model user behaviour, and ultimately to build context-aware systems that draw on these models to *infer* current user state in (near-) real-time. While modelling is applicable to many domains, in the context of interruptions it has been used to predict responsiveness (Avrahami and Hudson, 2006), interruptibility (Hudson et al., 2003; Fogarty et al., 2004, 2005), unavailability (Begole et al., 2004), the cost of interrupting users (Horvitz and Apacible, 2003; Iqbal and Bailey, 2006), or how busy the user is (Horvitz et al., 2005a).

Authors have instrumented the desktop environment of the interrupted in order to monitor the user's current cognitive task state by proxy of desktop activity (Iqbal and Bailey, 2005, 2006, 2007; Avrahami et al., 2008) or have equipped the office environment with sensors to detect office activity (Fogarty et al., 2004; Avrahami et al., 2007a). This serves multiple purposes: From correlating the activity with measures of responsiveness (Avrahami et al., 2008), to predicting unavailability to messages in real time (Horvitz et al., 2005a).

Avrahami et al. (2007a) report that interruptions are less likely to be perceived as disruptive when standing as opposed to sitting in offices [P11]. Avrahami et al. (2008) found that when participants used a development tool or a word processor when an instant message was delivered lead to significantly slower responsiveness [P10]. The authors also discussed how activities linked to greater work-fragmentation such as more window-tile switches, more keyboard activity and more mouse activity were related to quicker responsiveness. This resonates with the assumption of Miyata and Norman (1986), who stated that when in *interrupt-driven* processing one is more sensitive to external interruptions than when in *task-driven* processing, where one is in a state of decreased sensitivity to events external to the current activity.

Sensor-based statistical models of desktop and office activity have been shown to predict self-reported interruptibility (Fogarty et al., 2004). The authors built models from proximity and noise sensors in offices and desktop events such as mouse movements, keystrokes, and currently used application, and use an experience-sampling approach to label sampling moments with participants' self-reported interruptibility ratings. The authors showed how machine learning techniques can be used to predict non-interruptibility with high accuracy from these models. Interestingly, the features that the machine learning classifier chose as most significant varies between subgroups (managers, researchers and interns). In a follow-up study Fogarty et al. (2005) looked at task engagement in a code developing environment during which negotiated interruptions happen. In this study, the authors built predictive models purely by clustering response time as indicators of interruptibility, that predict interruptibility with an accuracy of around 72%.

Horvitz et al. (2005b) reviewed the construction of probabilistic models that can be used to infer the *expected cost of interruption* of a meeting and the likelihood that users will attend meetings noted on their electronic calendar. These models are applied to inform a cost-benefit analysis of call-handling policies on a smartphone in real time, where the membership of callers in different groups is also taken into account.

LIMITATIONS AND CONSEQUENCES Predicting interruptibility from statistical models is computationally expensive, relies on heavy instrumentation of the user's surrounding with sensors, and requires the designer to build the model for the desired set of sensors and laboriously adapt it to the specific task environment of the user, e.g. by using GOMS (Iqbal and Bailey, 2006). Even though model accuracy between 70% (Fogarty et al., 2005) and 90% (Iqbal and Bailey, 2006), has been reported, the cost of false positives (i.e. wrongly predicting availability) and especially false negatives (i.e. wrongly predicting unavailability) may be high, for example if the system falsely defers an urgent message. Even though the system may make correct inferences about user state 90% of the time, that one time it gets it wrong may lead to the user rejecting the system entirely. Thus, either the system needs to be made aware of the content and sender of the interruption and the recipient's relationship to them, or it needs to be limited to mediating non-timely messages. As Bellotti and Edwards (2001) point out, context-aware systems need to be intelligible in their inferences and accountable for their adaptations.

The discussed limitations of the work in subsection 3.4.1 and this subsection anticipate a critical discussion of the limitations and risks of interruption management systems in chapter 9.

The review of related work in this section has two main consequences for this thesis. Firstly, it anticipates cognitive activity as an important contextual factor of the interruption process that is central to the conceptual model presented in 4.3 and detailed in 4.4. Secondly, the empirical focus on laboratory settings in the extant work reveals an unanswered need to investigate interruptions in everyday life.

3.5 RESPONSES TO INTERRUPTIONS

Regardless of whether or not an interruption comes at an opportune or inopportune moment, and regardless of the cognitive processes and contextual factors (see section 4.4) involved in the decision process, Clark (1996) argues that there are four fundamental ways to respond to an interruption once detected (e.g. heard, seen, felt):

1. *Take-up with full compliance:* Handle the interruption immediately.
2. *Take-up with alteration*: Acknowledge the interruption and prospectively agree to handle it later, e.g. by making a mental note, or by adding an item to a TODO list.
3. *Decline:* Explicitly refuse to handle the interruption, e.g. by refusing a call.
4. *Withdraw:* Implicitly refuse to handle the interruption by ignoring it.

However, there may be some ambiguity in that some of the possible reactions may not be mutually exclusive. Consider the act of receiving a spam email and then deleting it straight away. Would this represent an act of *take-up with full compliance*, because deleting is arguably a way of handling; or does it represent an act of *declination*, as it may be interpreted as refusing to handle the interruption?

In a similar way, McFarlane and Latorella (2002) discuss five different "interruption management behaviors" (ibid., p. 17) in relation to the IMSM introduced previously in section 3.3.1.

1. *Oblivious dismissal:* The notification remains undetected.
2. *Unintentional dismissal*: The significance of the interruption is not interpreted and the interruption task is not performed.

3. *Intentional dismissal:* The significance of the interruption is interpreted but the interruption task is decided not to be performed.

4. *Preemptive integration:* The interruption task is immediately handled and performed to completion before resuming the interrupted task.

5. *Intentional integration:* The interruption task and the primary task are considered as a set and its integration is determined and performed rationally.

In addition to the fact that undetection of the notification is included as an "interruption management behavior" whereas Clark (1996) simply considers detection as a requirement, Latorella includes human error to distinguish between unintentional and intentional dismissal. Whereas Clark's other points are accounted for in Latorella's enumeration, *withdrawal* by ignoring the notification is not explicitly represented in Latorella's account. But then, does not ignoring already require a kind of thought process that could then render it effectively as declination? What if I glance at the phone and see who is currently calling me and then decide to ignore it? It then appears that attempts to classify fundamental ways of reacting to interruptions raise more questions than they answer and need to be considered in relation to the contextual factors at play in situ and on a case-by-case basis during the interruption process. Contextual factors consist of all entities involved in the interruption process, including social entities such as the sender and people potentially co-located with the recipient, which are detailed in section 4.3.

3.6 systems and techniques to support interruption management

With the strong tradition of design and engineering in HCI, it is not surprising that in addition to the studies presented thus far a host of systems has been implemented prototypically that aim at supporting interruption management by sensing and responding to certain contextual factors, such as message content (Avrahami and Hudson, 2004), the user's eye contact (Altosaar et al., 2006), or breakpoints in the user's interaction (Iqbal and Bailey, 2008). Table 3.1 summarises the desktop systems that support interruption management and offers a brief description of the system's main contribution.

System	Description	Author
QnA	An add-on for an instant messaging client that automatically increases the salience of incoming messages that represent questions or answers to previously asked questions by performing relatively simple string matching on the incoming message content.	Avrahami and Hudson (2004)
AuraOrb	A bespoke ambient device for incoming email on the desktop that notifies the user of incoming messages with a light and on sensing the user's eye contact presents the subject heading of the message on a ticker tape display.	Altosaar et al. (2006)
Oasis	A system that uses a defer-to-breakpoint policy to mediate interruptions such as incoming email notifications. Breakpoints are detected in the user's ongoing interaction, and a scheduler delivers queued notifications at breakpoints (for more details see 3.4.2).	Iqbal and Bailey (2008)

Table 3.1: Desktop systems that support interruption management.

In the context of this thesis, the interest is on systems that have been built to support the user's interruption management whilst being mobile. While most modern mobile phones offer ways in which the device can be configured to adapt the ringtone according to caller ID, or to switch between different audio- and vibro-tactile alert "patterns", the focus here is on systems that *infer* context based on sensor input. Research on systems that require explicit user input such as the configuration of a device or setting an availability status, has shown that users often forget to complete this kind of overhead work (Bellotti and Edwards, 2001). User inconvenience such as this arguably motivated the development of context-aware computing in the first place, through the idea of *calm computing*, where the technology moves autmatically between engaging the center and the periphery of our attention (Weiser and Brown, 1996).

Context-aware interruption management systems in the mobile domain may take into account one's likelihood of attending a meeting (Horvitz et al., 2005b), mental workload (Chen and Vertegaal, 2004) or message priority (Sawhney and Schmandt, 2000). For example, in order to automatically defer a phone call or to adapt the presentation of a message to the inferred availability state. Table 3.2 summarises the mobile systems that support interruption management and offers a brief description of the system's main contribution.

This summary of systems suggests that different approaches and techniques to supporting interruption management are inherent in the systems' underlying design rationale. In the following, the underlying principles that inform the design of interruption management systems are classified into approaches that have an underlying orientation towards cognitive psychology and approaches that rely on a *human-in-the-loop*. Finally, this section will conclude by discussing McFarlane's influential four primary methods to coordinate interruptions.

Table 3.2: Mobile systems that support interruption management.

System	Description	Author
MATCH MAKER	A context-aware, pervasive computing environment that consists of several applications that support students' activities around the campus. MATCHMAKER helps students find experts and makes use of the expert's temporal context to determine their availability.	Anhalt et al. (2001)
PAUI	PAUI stands for *Prototype Attentive User Interface*, a wearable system that estimates mental workload from heart rate (ECG) and brain signals (EEG) and actuates one of four states of auditive and vibrotactile notification presentation for incoming mobile calls, email and IMs as well as the IM state, e.g. it switches between *silent*, *vibrate*, and *ring* for incoming calls.	Chen and Vertegaal (2004)
FINGER RING	A system where the called party and all co-located conversation partners may vote in order to reject an incoming call as they are notified by the call through vibrating finger rings. At the time of voting, the users do not know whether it is their own or somebody else's mobile phone that requests to ring, and vetoing the call may be done anonymously by touching the instrumented finger ring.	Marti and Schmandt (2005)
OWNTIME	A system for a headworn display that allows potential visitors to the office to announce their intention to visit and declare a subject to be discussed in order to minimise the alleged disruptiveness of drop-in meetings. The wearer can then agree to the informal face-to-face meeting or ignore the request to signal unavailability.	Rodenstein et al. (1999)

continued on next page

Table 3.2 – continued from previous page

System	Description	Author
Nomadic Radio	Uses different auditory cues to signal the wearer of the mobile and audio only interface of new email, voice mail, news, and personal calendar events. The auditory presentation is scalable on seven increasing levels of notification: *silence, ambient cues, auditory cues, message summary, preview, full body,* and *foreground rendering*, and is based on contextual factors *message priority, usage level,* and *likelihood of conversation*. *Message priority* is computed through content-based email filtering, *usage level* is determined by the time of the user's last interaction with the system, and *likelihood of conversation* is estimated from active microphone input in real time.	Sawhney and Schmandt (2000)
Eye Contact	Uses speech analysis and an additional head-mounted camera to detect whether the user is in a face-to-face conversation. The use of sensors enables the system to detect engagement in nonverbal face-to-face situations, for example when the user is listening for prolonged periods of time. The sensed availability state of the potential recipient of mobile calls is then transferred to potential callers' mobile phones. The idea is that callers can make informed decisions if they want to interrupt the other person.	Vertegaal et al. (2002)
Bayes phone	Utilises precomputed user models on a mobile devices to predict the cost of interruption versus the cost of deferring mobile phone calls. Interruptibility is computed from electronic calendars, likelihood of meeting attendance and caller identity (i.e. they have to have been previously assigned to certain groups).	Horvitz et al. (2005b)

3.6.1 *Approaches inspired by cognitive psychology*

PAUI and OASIS (see table 3.2) represent examples of systems that were designed based on theories in cognitive psychology about the relation between attention, mental workload and resulting state of availability for an interruption. OASIS makes use of the user's instrumented desktop environment to detect naturally occurring breakpoints in the ongoing interaction with the Graphical User Interface (GUI). These breakpoints are seen as transient reductions in mental workload and therefore as an opportune moment to deliver queued interruptions such as system messages or IM (see section 3.4). PAUI aims to estimate mental workload from physiological sensors that are worn on the body, and adapts the presentation of the incoming message's signal to the estimated availability.

In the form of a conceptual framework, Oulasvirta (2005) present the *skilled memory approach* to inform the design of non-disruptive user interfaces. The approach is based on considering memory load of the primary task according to the authors' own *Resource Competition Framework* (RCF) (Oulasvirta et al., 2005b) and encoding of the primary task state prior to the interruption task. RCF presents an analytic tool for the designer of mobile systems to consider typical allocation of cognitive resources such as attention to typical tasks in mobility and to tasks in mobile HCI (see 2.2.1, p. 28).

Systems designed based on cognitive theories may be appealing in theory, yet this thesis is interested in uncovering the constitution of opportune moments for interruptions in real-world studies by exploring the situated effects on users. Nevertheless, for the purpose of designing the adaptations whose effects are studied, interruption management techniques inspired by cognitive psychology play a central role in this thesis (see 4.5).

Interruption techniques based on psychological theories seem to heavily inform the design of automated systems, for example the type that predicts opportune moments for interruptions by machine reasoning on ongoing interaction based on task models (see section 3.4.2 for more detail), while another approach seems to make the reliance on such elaborate models less necessary – systems that leave the psychology of the interruption process to what is referred to as the *human-in-the-loop*.

3.6.2 *Automated vs. Human-in-the-loop*

A fundamental difference in approaches to the design of interruption management lies in the distinction between systems that automate the decision process[4], for example to determine whether a potential recipient is available for an interruption, and those that make explicit use of a human – either the recipient or the sender, or both – to make that decision.

Among the mobile systems in table 3.2, the systems Finger Ring, OwnTime and EyeContact present instances of systems with a *human-in-the-loop*, each with a different role and process of integration. In Finger Ring, the co-located group jointly decides on availability for a call to one of the members of the group; in OwnTime, the recipient decides whether they are available for an informal meeting; and in EyeContact, the system abstracts the availability state from sensors and makes this information available to potential interrupters to leave the final decision whether they want to call to them.

In contrast, automated systems such as Bayesphone and Nomadic Radio employ advanced machine learning techniques to extract and

4 In the literature on context-awareness and machine learning, this is often referred to as *inference*.

reason about information from message content, the user's auditory environment, the user's electronic calendar and the caller's identity.

The value of *human-in-the-loop* systems has been demonstrated empirically. Avrahami et al. (2007a) found that if potential callers have more detailed awareness information about the receiver's state, callers avoid interrupting inappropriately when receivers are in the office, and knowledge that the receiver was alone helped to interrupt at an appropriate time.

Dabbish and Kraut (2003, 2004) conducted an experiment where an awareness display showing the recipient's workload was provided to the sender of an interruption and found that it was beneficial in estimating a recipient's availability to an interruption. Dabbish and Kraut (2003) showed that the provision of an awareness display of the remote collaborator's workload affected how the asker timed their question to the remote collaborator. The sender's timing of the question apparently led to a more appropriate perceived timeliness of the question on the receiver's side, as demonstrated by better performance on their task of playing a simple computer game. This mechanism was at the expense of the asker, who needed more time for their own task by timing their questions sensitively.

Even though the simplicity of the experiment's tasks induce a low ecological validity, the experiment shows the value of a *human-in-the-loop* when it comes to estimating the timeliness of an interruption for the receiver. Dabbish and Kraut's work reveals some of the intricate processes by which the sender negotiates their information need with the recipient's availability. A human element then appears to have the potential to add sensitivity barely conceivable for an automated technique. However, a verdict on the best fit of either an automated or an approach relying on a human-in-the-loop should be made on the basis of an analysis of the system's requirements.

Moreover, it seems that the fundamental differences are aligned with the arguments for (Dey et al., 2001) and against (Bellotti and Edwards, 2001) context-aware computing. Whereas the reliance on a *human-in-the-loop* creates an overhead for the user of the system – part of the reason why context-aware systems have been conceived in the first place – automated systems have an inbuilt risk of drawing the wrong conclusions, for example to defer an important or urgent message to a later moment.

3.6.3 *McFarlane's 4 primary methods to coordinate interruptions*

As part of his taxonomy of human interruption (McFarlane, 1997), McFarlane introduces four primary methods to coordinate interruptions in HCI (McFarlane, 1999), which he compares empirically in his seminal paper (McFarlane, 2002). This work has influenced widespread adoption of the terminology and differentiation of the four interruption strategies:

IMMEDIATE: The interruption is presented immediately and disrupts the recipient's primary task.

NEGOTIATED: The interruption is announced by notification, the recipient retains control when to accept the interruption.

SCHEDULED: Interruptions are delivered according to a prearranged schedule.

MEDIATED: A mediating agent is notified of the interruption request, the agent then decides when and how to interrupt.

In a laboratory experiment (McFarlane, 1999, 2002), participants were interrupted with a simple matching task during their primary task of playing a computer game. Interruptions were presented according to all four strategies, with the *mediated* condition being based on a simple assessment of the participant's current workload in the game based on the number of objects on screen. McFarlane (1999, 2002) concludes that the *negotiated* and *mediated* interruption strategy led to superior performance in the accuracy and efficiency on the primary and interruption task, and to better overall user perception, while being perceived as the least interruptive. The *negotiated* strategy was inferior in the interruption task was completed less often and reduced promptness of task completion, whereas the mediated strategy was inferior in user perception of the predictability of the interruption.

To sum up, the basic finding of McFarlane (1999, 2002) was that the *negotiated* and *mediated* strategies were best for supporting all kinds of human performance, except where immediate completion of the interruption task was critical, in which case the *immediate* strategy was more appropriate. Overall, this shows the importance of supporting some level of human control over the coordination process.

McFarlane's work inspired a considerable amount of further work. Iqbal and Bailey (2008) compared *immediate* notification delivery to their *defer-to-breakpoint* approach, which in McFarlane's typology represents a *mediated* interruption strategy. Interruptions were deferred to predicted breakpoints in the interrupted's current tasks. While participants were coding (task 1), self-reported levels of frustration at breakpoints did not differ from the assumed worse policy of presenting the interruption immediately. However, for the task of editing a diagram (task 2), the defer-to-breakpoint policy did have the desired effect of lowering frustration compared to the immediate strategy. This study indicates that the effectiveness of the coordination strategy is influenced by the interrupted's current task.

Robertson et al. (2004) compared an *immediate* with a *negotiated* interruption strategy in a study of end-user debugging and provide further support with their findings that the *negotiated* strategy is superior. Participants that received notifications according to the *negotiated* strategy were significantly more productive at debugging and were more effective at assessing the success of their debugging.

However, Witt and Drugge (2006) found in their study of interrupting users on their Head-Mounted Display (HMD) while they played with a "hot wire" apparatus, that the *immediate* strategy is more beneficial than the *negotiated* strategy, indicating that different modalities may require different interruption strategies. Nevertheless, their results may be explained by the effect that users could simply ignore the interruption for a short while before completing the interruption task (a matching task similar to the one used by McFarlane (2002)). Effectively, this may turn the *immediate* into a more *negotiated* interruption. This suggests that the experiment design may not have compared an *immediate* to a *negotiated* interruption strategy, but two interruptions (in the *negotiated* condition: the notification and the task) to a single interruption (in the *immediate* condition: the task). As the participants were wearing a semi-transparent HMD the user had the time to mentally prepare the primary task resumption in prospective memory as described above (Altmann and Trafton, 2004), making the interruptions *negotiated* in effect.

The success of the *negotiated* strategy may be explained in a much broader scope by Horvitz's account of the principles of mixed initiative UI (Horvitz, 1999), which gives specific guidelines on how to deal with uncertainty – such as the uncertainty of the appropriateness of an interruption. Horvitz suggests that UI designs should account for uncertainty by considering the costs and benefits of the timing of services for the user, employing dialogue to resolve key uncertainties, and minimise the cost of poor guesses about action and timing. At its heart Horvitz (1999) recommends a UI strategy that relies on *negotiation* between system and user.

McFarlane's definition of the four primary methods has not just proven useful to inspire other authors' work and terminology. This thesis' empirical work may be characterised in terms of the ways in which it employs McFarlane's methods. The study in chapter 6 tests whether scheduling one's own interruptions ahead of time is superior to randomly timed interruptions. The systems studied in chapter 7 and chapter 8 mediate the timing of the interruptions they deliver based on user context. The way these two systems present notifications may be described by McFarlane's negotiated method: the user is alerted but then they have to explicitly switch to an inbox to view the message.

3.7 SUMMARY

This chapter has defined key concepts that inform this thesis (see 3.1) and highlighted the mundane, yet central character that interruptions play in everyday life (see 3.2). The chapter has developed a temporal model of an interruption based on a review of temporal models in the literature (see 3.3), enabling the ESM-adaptation (see 5.5) that underpins in particular the choices and definition of hypotheses and variables in the empirical work presented in chapter 6, 7 and 8.

Related work has been reviewed that seeks to understand opportune moments for interruptions (see 3.4), both in terms of psychological theories of breakpoints in cognitive task processing (section 3.4.1), and more opportunistically with the goal to predict interruptibility from sensor data (see 3.4.2). The question of what might be an opportune moment for an interruption and how we might predict their occurrences is key to this thesis. In support of the focus on the subjective experience as motivated in chapter 2, the literature on cognitive activity anticipates the importance of the *recipient* at the core of the model of contextual factors (see 4.3) synthesised in the next chapter.

The classification of systems to support interruption management by their underlying approaches and techniques (section 3.6) prepares for the discussion of the kind of adaptations interruption management systems make (see 4.5), which provides the focus on timing as the key manipulation of the empirical work.

This concludes Part I of this thesis. While chapter 2 framed the thesis' endeavour with a broader, philosophically critical and holistic view that underpins methodological choices brought forward in chapter 5, this chapter has considered the "nuts and bolts" that explain much of the "mechanics" of the empirical work, so to speak, which feed into the approach detailed in the next chapter, and the experimental approach to studying interruptions in situ outlined in chapter 5.

Part II

APPROACH AND METHODOLOGY

4

APPROACH TO STUDYING RECEPTIVITY TO INTERRUPTIONS

While the previous chapters chart the contextual intricacies of mobile HCI and interruption management in particular, this chapter marks the beginning of this thesis' main contribution. First, *receptivity* to interruptions is introduced and conceptualised as a psychological construct whose exploration is at the core of the empirical studies in chapter 6-8. This conceptualisation is then grounded in a model of the contextual factors that determine a person's receptivity to an interruption. From this conceptual model the methodology of this thesis' studies will be developed in chapter 5.

To inform the development of the model, this chapter considers transmission models of communication and McFarlane's taxonomy of human interruption (1997). The contextual model of receptivity is overviewed before its constituant parts – the contextual factors – are synthesised from the literature in detail in section 4.4. From an examination of the types of adaptations that interruption management systems make in practice, *timing* is identified as the core contextual adaptation to be probed in this thesis. On that basis, the central research questions around timing are then developed for the studies in chapter 6-8.

4.1 RECEPTIVITY

The concept of *receptivity* is the core psychological construct in this thesis. Receptivity is defined to capture the subjective, situated nature of the experience of an interruption. Receptivity has been used as a dependent measure in studies of interruptions (Ho and Intille, 2005), and has been described as one's "willingness to be interrupted" (Begole et al., 2004). Here, this notion is extended by emphasising that receptivity places the recipient's actual experience of the interruption into the focus. Considering a person's receptivity to an interruption then explicitly dismisses the idea that interruption can be studied in its procedural form in an objective fashion – rather, utilising the concept of receptivity underscores both the subjective viewpoint from which to understand interruptions and the situated nature of the interruption process. Hence, receptivity frames interruption as a subjective and situated phenomenon.

Furthermore, receptivity is treated as an *experienced* quality by which the appropriateness of an interruption is assessed. One of the main objectives of this thesis is to understand what constitutes opportune moments for an interruption. This is gauged by means of receptivity. Thus, a moment is defined as opportune for a particular interruption if the participant is receptive to that interruption.

Moreover, this thesis' contribution is not just achieved from an observant stance; strategies for interruption management are developed based on grounded hypotheses about opportune moments for interruptions, and implemented and tested in naturalistic environments. These interventions are partly assessed by measuring participants' receptivity to interruptions in situ. The technical goal is to minimise the negative impact of interruptions by maximising the proportion of interruptions delivered when the user is receptive to them.

In summary, receptivity in this thesis is understood and utilised to:

- conceptually capture and study the subjective, situated nature of the experience of an interruption,

- pinpoint opportune moments for interruptions as indicated by the individual's experience of receptivity, and
- evaluate interruption management strategies by their impact on receptivity.

Studying receptivity to interruptions presupposes the question: What makes someone receptive to an interruption? Following the phenomenological view, receptivity frames interruption as a subjective experience of the involved individual. It was argued that involvement in the world is characterised by a moment-by-moment negotiation of context, which makes it difficult to study the "influence" of context in a systematic way (see 2.1). Therefore, careful dissection of context into *contextual factors* at play in the interruption process is required to enable the systematic study of receptivity. The question thus raised is: What are the factors that influence a person's receptivity to an interruption?

In order to develop the contextual factors, the interruption process is delineated as an instance of a communicative process, drawing in particular on transmission models of communication. The contextual factors then emerge by considering the entities involved in the communicative process.

4.2 models of human communication

One might argue that an interruption is fundamentally caused by the process of transmitting information. Thus, it is posited that interruption is at its heart a communicative process, or more precisely, that an interruption is an instance of a *communicative act*. According to Riley, a *communicative act* is a basic unit of meaning which "can be realised by a wide range of behaviours and conveyed along a number of different channels" (1976, p. 3).

Source⟶Transmitter⟶Channel⟶Receiver⟶Destination

Figure 4.1: Shannon and Weaver's model of communication (1948). Adapted from (Shannon, 2001).

Hence, transmission models of communication are concerned with the process of conveying meaning, which in turn implicates interruptions in exhibiting this function.

4.2.1 *Transmission models*

Shannon and Weaver's (1948) intriguingly simple model of communication (see figure 4.1) shows the central elements involved in a communicative process[1]. However, it was conceived with the intention to develop a mathematical theory of communication and fails to be readily applied to human communication, as it simplifies communication as a relatively unproblematic process of transmitting a message from a sender to a receiver[2].

In contrast, the question brought forward by Lasswell (1948) (see figure 4.2) more flexibly describes the entities involved in an act of communication. Answering this question necessitates thinking about the specific entities involved in the communicative act under consideration. Thus, it introduces a need to consider the *context* of the communicative act.

Whereas Lasswell's intention was to summarise what may be called the five pillars of communication science (the study or analysis of control, content, media, audience and effect), its contextual applicability lends itself well for the purpose of identifying the factors involved in the interruption process.[3] To introduce a focus on the subjective

1 The dysfunctional factor *noise* is omitted for clarity.

2 The model's *source* and *destination* have typically been humanised as *sender* and *recipient* (Chandler, 1994).

3 Both Shannon and Weaver's as well as Lasswell's models have been criticised for over-simplifying the notion of communication as a simple process of transmitting information. It is precisely for its simple essence and the pragmatic attitude that the model is useful to sensitise to the factors involved in the interruption process. A more

Who
Says What
In Which Channel
To Whom
With What Effect?

Figure 4.2: Lasswell's formula of communication. Adapted from (Lasswell, 1948, p. 37).

perspective of being interrupted as developed in chapter 2, Lasswell's question may be asked from a first person point of view, substituting the *To Whom?* by *To Me*. For the purpose of understanding receptivity to interruptions this shifts the focus to *Me*, the *recipient* of the interruption. For each interruption, the individual's receptivity depends on the situated interplay of the factors introduced by Lasswell's formula. Hence, for the purpose of this thesis, the factors are adapted in the following way:

who The recipient's relation to the sender of the interruption. For example, do I (the recipient) trust the sender, what is their authority, status, affiliation etc.?

says what The recipient's relation to the content of the interruption. E.g. is it relevant, urgent, interesting etc. to me?

in which channel The recipient's relation to the medium of the interruption. E.g. am I more receptive to a call or an SMS?

to whom The recipient's current involvement. E.g. what am I doing?

with what effect The anticipated/actual effects of the interruption on the recipient. E.g. does the interruption trigger an activity?

The adaption of Lasswell's formula provides the basis for a model of the contextual factors that influence receptivity to an interruption.

detailed discussion of communication models is out of the scope of this thesis, see for example Chandler (1994).

Descriptive dimension in McFarlane (1997, p. 73)	Entities in Lasswell (1948)	This thesis' adaptation to interruptions
Source of interruption	Who	Sender
Individual characteristic of person receiving interruption	To Whom	Recipient
Method of coordination	NA	Timing
Meaning of interruption	Says What	Content
Method of expression	NA	Presentation
Channel of conveyance	In What Channel	Medium
Human activity changed by the interruption	With What Effect	Effects on the recipient
Effect of interruption	see above	see above

Table 4.1: Comparison of McFarlane's taxonomy of human interruption and Lasswell's formula of communication.

Before developing the contextual model further, it is compared with McFarlane's taxonomy of human interruption.

4.2.2 *McFarlane's taxonomy of human interruption*

McFarlane (1997) aims to develop a unifying definition of interruption to allow researchers to generalise and use existing results in the literature from different fields. Furthermore, he then introduces a taxonomy to augment this unifying definition to make it practically applicable. This supports the analysis and description of the interruption phenomena along important and useful dimensions, and facilitates the integration of related work into one's own work and vice versa. Based on an extensive overview of research predominantly in cognitive psychology and linguistics, McFarlane synthesises his analysis into a taxonomy that he refers to as a "tool for describing instances of human interruption" (McFarlane, 1997, p. 72), and its instances include and expand on the instances identified by adopting Lasswell's formula. Table 4.1 presents a comparison of the factors present in the two models, and this thesis' adaptations.

McFarlane's taxonomy strengthens the motivation to develop a taxonomic model of the factors involved in the interruption process and it adds two factors to the model (*method of coordination* and *method of expression*) that are not represented in a model solely based on Lasswell. The dimension *method of coordination* includes the *immediate, negotiated, scheduled* and *mediated* techniques discussed in 3.6.3 (p. 74). As coordination is commonly achieved through the adaptation of interruption *timing*, this dimension is adopted as the factor *timing* in this thesis' model of receptivity. Similarly, the dimension *method of expression*, which refers to the way in which the interruption is brought forward, is adopted into the model as *presentation of the interruption*.

However, the distinction between the last two factors (human activity changed by the interruption and effect of interruption) is not clear. Is not a change in human activity already an effect of the interruption? Moreover, the major shortcoming of the proposed taxonomy may be the absence of the contextual factors pertaining to the immediate environment of the recipient, and the potentially ongoing social involvement with other people present at the time of the interruption. It follows as a consequence of the phenomenological perspective of the situated individual that we also need to consider environmental factors in order to understand the interruption process.

4.3 A contextual model of receptivity

This section introduces the conceptual model of this thesis – a taxonomy of the contextual factors underlying the interruption process. First, the model is introduced to give an overview of its components. The model's grounding in a classification of empirical work is then detailed (see 4.4). For the sake of clarity, the model is introduced before the contextual factors that comprise it are presented, although the model actually was developed from a synthesis of findings from interruption studies in the literature.

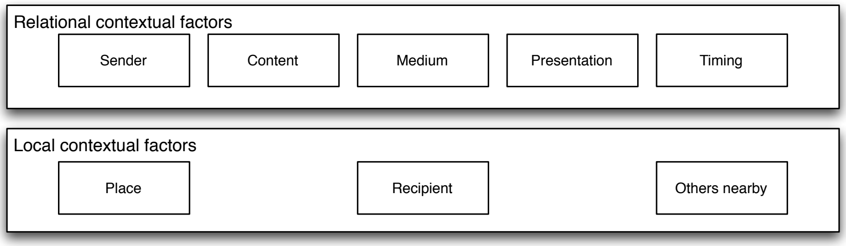


Figure 4.3: Simple depiction of the model of the contextual factors of an interruption process.

The taxonomy of the contextual factors at play during the interruption process presents an analytical tool that may inform related research by providing an overview of the entities involved in the interruption process, and it allows researchers to position their own contribution in relation to a broader taxonomy.

The contextual factors involved in the interruption process necessarily underly receptivity – the core subjective psychological construct that is studied in this thesis. To develop an understanding of receptivity then means to study the influence of the underlying contextual factors. Furthermore, the contextual model is not limited to describing the underlying factors of receptivity, it may be read as a more general model of the contextual factors of an interruption process – and thus by extension any communicative act.

A simple version of the model is shown in figure 4.3. The figure will be extended at various points later on in this thesis to illustrate its relation to the empirical work. Depicted in more detail in figure 4.4, the model broadly distinguishes two types of contextual factors, *relational* and *local* contextual factors. Local factors account for the local context of the recipient of an interruption, while relational factors relate the recipient to aspects of the interruption itself.

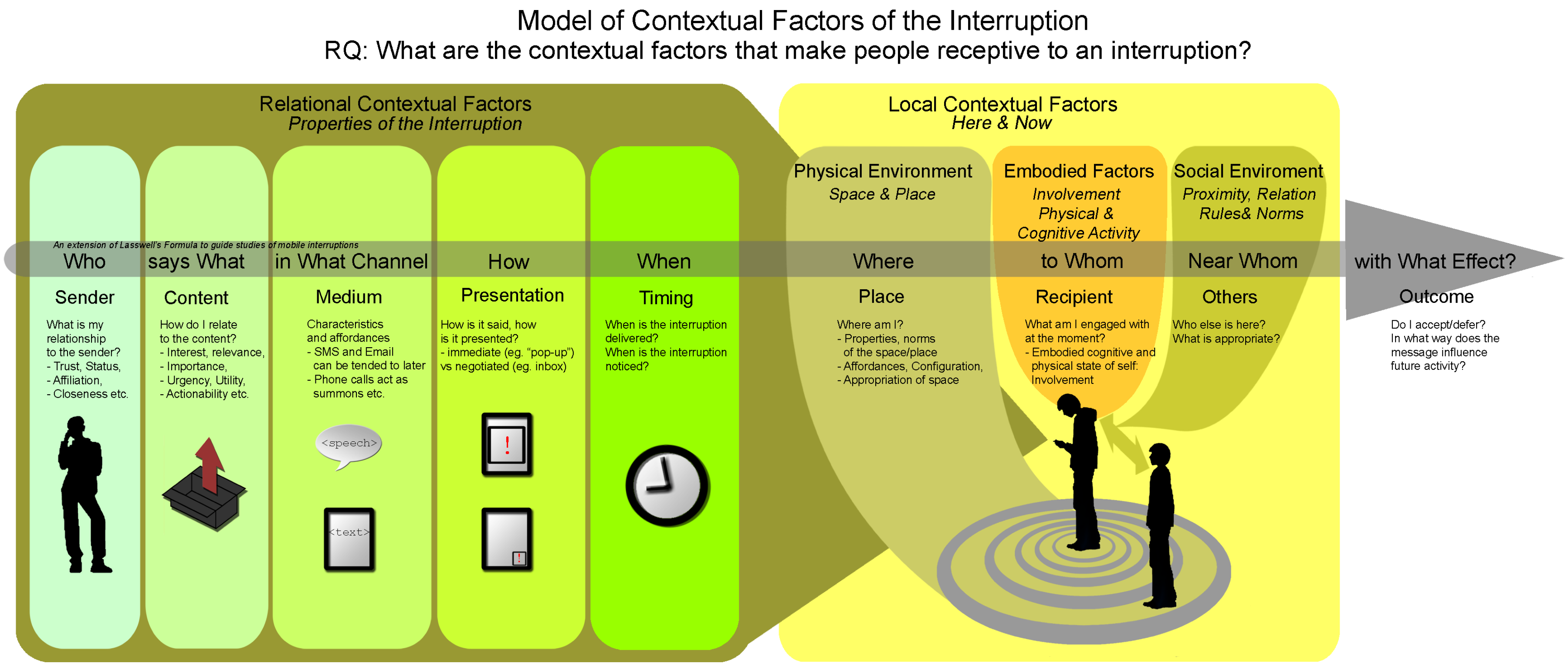


Figure 4.4: Model of the contextual factors of an interruption process.

For example, the environment of the recipient at the time of the interruption is part of the local context, whereas the content of the interruption is part of the relational factors. The model assumes all factors to be interrelated and to interact to define the individual's context, where context is more than the sum of its parts.

In more detail, a recipient's local context is constituted by *environmental*, *social* and *embodied factors*. Embodied factors are key as they account for the recipient's current involvement in the world, including their physical and cognitive activity. It is important to note that it is only through the subjective lens acquired as a result of embodiment that the other factors become meaningful, as they are interpreted in light of the individual's unique viewpoint. Thus, the configuration and appropriation of the space and the presence of other people becomes meaningful in the way they subjectively appear to the recipient, and through the actual social and environmental interactions that take place.

The following example illustrates how the model would be applied in practice: When I am on a bus, local factors that mediate receptivity to an incoming call are a) the environment, or place of the interruption including behavioural norms, rules, etc. b) social proximity, e.g. I may be next to someone or alone on the bus, and c) involvement, e.g. I may be in a conversation with my neighbour, I may be involved in reading a book, or listening to music etc.

The relational factors arise mainly from the consideration of the interruption process as a transmission of information according to the transmission models of communication, and include *sender*, *content*, *medium*, *timing* and *presentation* of the interruption. They are called relational factors following Grandhi (2007) because they receive their value not per se, but in *relation* to and through the interpretation of the recipient. It is the recipient that relates to these entities in a certain, meaning-making way.

Going back to the example of being on a bus, relational factors that mediate receptivity to an incoming call may be a) my relationship to the caller reminded by caller ID, b) the anticipated content of the conversation, c) the fact that it's a call and not a text message (medium), d) whether the call was heard or unnoticed because the phone was set to silent (presentation), and e) the timing of the call.

In the following section, the model will be justified by classifying related interruption studies and synthesising their contributions to the model of contextual factors.

4.4 CONTEXTUAL FACTORS OF INTERRUPTIONS

The holistic view of contextual involvement in the world as described in chapter 2 makes it difficult to study the interplay of context and interruption management in a systematic way. However, following the pragmatic attitude (see 3.1, p. 43), context as achieved by involved interaction in the world is broken down into contextual "factors" in order to increase our understanding and to inform the experimental design in chapter 6, and 7 and the field study in chapter 8. We have seen that context is not just found in the world, but created through involvement and that there are a whole host of factors involved in this process. This section orders and classifies related work on interruption management in terms of its (sometimes implicit) contextual factors that together constitute context as a coherent whole. In this way, the contextual factors that constitute the model of receptivity are synthesised, which are key to the empirical research presented in this thesis.

What are the contextual factors at play during an interruption? Interruptions are accompanied by the situated interplay of a whole host of psychological and environmental factors. For example, Ho and Intille (2005) identify at least 11 factors from related work that influence a person's interruptibility such as "activity of the user", "utility of the message", "emotional state of the user", "modality and frequency of interruption", "task efficiency rate", "previous and future activities", "social engagement of the user", and "history and likelihood of response".

Instead of simply presenting a list of factors, here, the contextual factors are classified by characteristics grounded in a view of the actively involved subject. An initial structure is achieved by distinguishing between *local contextual factors* and contextual factors that have a *relational perspective* (Grandhi, 2007). While local contextual factors pertain to the immediate local context of the person independently of the interruption, relational contextual factors take into account properties of the interruption itself.

4.4.1 *Local contextual factors*

Local contextual factors are local to the recipient of an interruption, and include the social and the physical environment, and embodied factors. Local factors account for the ways in which the individual is actively involved in the immediate (i.e. *perceivable*) environment, how one orients towards or interacts with the structure of the environment and the social encounters within it. Embodied factors include the cognitive and physical state of the interrupted person as influenced by the history of the combined experience in the world. According to Merleau-Ponty, it is through embodiment that the world is encountered and subjectively interpreted.

However, this is not to say that the appearance of the environment or the social encounters can be neglected. The example of the kitchen that affords different kinds of interaction for a chef than for a mechanic (see 2.1.1) illustrates how the environment reveals itself differently to each of us. However, its affordances would not be perceived without the physical structure of the environment.

To demonstrate how local factors could inform an interruption management system, consider the activity of driving a car. Driving requires that eyes be kept on the road and hands on the steering wheel. However, the environment (i.e. the car) affords all sorts of activity, including talking on the mobile phone. An implication for interruption management for this scenario would be that the requirement to keep eyes free and hands on the steering wheel is not violated. In practice this could mean that a text message might be read out by the system.

4.4.1.1 *Environmental factors*

The environment of the recipient is likely to influence the receptivity to an interruption. Environmental factors include the appearance, affordances, configuration and appropriation of the space surrounding the actively involved recipient. Here, *environment* is understood to refer to both the inhabited, social space and the structural, physical space surrounding the recipient. Harrison and Dourish (1996) describe places as spaces "invested with social meaning" (ibid., p. 71). In practice, the social meaning of a place is often embodied by a set of rules or norms of appropriate behaviour for that place. For example, consider how one is likely to be receptive to an interruption on the bus, but not very likely to be receptive in a church. These places acquire such different social practices and norms of appropriate behaviour through the active involvement of social actors. The meaning of a place becomes an interactional achievement (Dourish, 2004).

Organisations are a good example of how places give rise to behavioural norms. Hudson et al. (2002) found in their study that managers' attitudes towards interruptions vary with activity. The authors observed a general temporal pattern of availability to interruptions throughout the day. They state that managers preferred the early hours at work in the morning, the hours right after lunch and in the late afternoon to be interruption-free. Managers reported being more happy to be interrupted during the other hours. Another interesting finding is how a physical artefact may be interpreted as a cue of (un-) availability unintentionally. The authors report that a closed office door often led to managers not being interrupted, despite their stating that they would have been available to an interruption at the time (ibid.).

Others have more directly looked at the mediating role of social proximity. For instance, Avrahami et al. (2007a) have looked at the impact of the presence of others in the recipient's office and report a significant effect on self-reported interruptibility. In this thesis' contextual model of an interruption, social factors capture the recipient's current orientation towards other people present in the immediate environment. Social factors include not just the presence of other people close to the recipient, but also their role, relationship and engagement with the recipient. Consider again our earlier example of being on a bus. Taking physical proximity to other people as a sole indicator of receptivity breaks down when one considers such situations.

Tolmie et al. (2008) report on the social practices around managing interruptions from a mobile SMS-based game. The researchers accompanied two people in a family playing a long-term slow-paced game for a period of four weeks. The study revealed how complex and interwoven interruption management was with the intricacies of the current situation, and how the players managed and maintained their accountabilities towards their social surroundings. The authors identify the role of *social accountability* in influencing whether people engage with technology as a result of an interruption, and suggest that technology should foreground the *grounds of disturbances* and let people continue to manage their own interruptions rather than design technologies that manage interruptions for them (ibid.).

LOCATION A simplistic instance of an environmental factor that is often used is user location. For example, Ashbrook and Starner (2002) build predictive user movement models by monitoring their GPS position over a period of time, and suggest that their models are applicable to applications that manage interruptions based on user location. The authors concede that it may be difficult to sense availability purely based on location, but they suggest that unavailability may be sensed based on location and time, e.g. when entering a lecture theatre at a time when the user has a lecture (Ashbrook and Starner, 2002). In fact, commercial solutions that make use of GPS location have already been developed and are available. For example, LOCALE[4] is an application for the Android platform that allows a user to customise the phone to assume ringtone settings based on specific locations.

However, although location may be an initial indicator of the user's context, it cannot be used effectively to estimate the user's current orientation in space, not to mention the user's current involvement with their environment. Consider our previous example of being interrupted whilst being on a bus – an interruption management system would have no means to estimate the recipient's receptivity on the basis of their (moving) location alone.

It then becomes important to realise that local factors acquire their meaning for the individual only through the process of subjective involvement of that individual in the world, which is captured by the *embodied factors* in the model as outlined in the following.

4.4.1.2 *Embodied factors*

Embodied factors are local to the recipient's body, and include cognitive and physical activity. Embodied factors are key, as they account for the recipient's involvement in the world.

4 Available at `http://www.twofortyfouram.com`.

Although the postulates presented below are derived exclusively from work taking on the view of cognitive psychology, choosing the term *embodied* factors to summarise the factors local to recipient's body resonates with the phenomenological perspective; in particular, with involvement being the core concept. For instance, the proximity of other people alone is not decisive as illustrated by the example of being on the bus. Even taking into account a descriptor of a relationship to other people in the vicinity such as "colleague, friend, or relative" most likely will not sufficiently explicate if one is receptive to an interruption. It is rather the moment-by-moment orientation towards people that matters, for example if one is engaged in a conversation. Even the *degree* of involvement in the activity, which may for example be influenced by the current subject matter at hand, may give rise to fluctuations in receptivity moment-by-moment.

The nature of the recipient's current involvement in the world includes their current mental state, emotions and affect, beliefs, desires, intentions, their collected experiences (see 2.1.5), and their cognitive and physical activity. As reviewed in section 3.4, significant effects on interruption handling in the cognitive space have for instance been found for *mental workload* (Adamczyk and Bailey, 2004), and *attentional focus* (Horvitz and Apacible, 2003). On the physical side, significant effects were found for *transitions between physical activity* (Ho and Intille, 2005), and *body position* (Avrahami et al., 2007a). The postulates detailed in the following were distilled in order to show the veracity of the contextual factors.

postulates from related work In the psychologically influenced literature the recipient's activity is studied in terms of the *primary task*. This literature has been reviewed extensively in 3.4. The following postulates were derived from the literature review, which provide empirical evidence of the importance of cognitive [P1-P10] and physical activity [P11-P12] on the effects of interruptions.

P1: The disruptiveness of an interruption depends on the memory load at different stages of the primary task (Miyata and Norman, 1986).

P2: An interruption is least disruptive after evaluation and before forming a new goal (Miyata and Norman, 1986; Monk and Trafton, 2002).

P3: An interruption is most disruptive during planning and evaluation (Miyata and Norman, 1986; Latorella, 1998).

P4: Interruptions are more disruptive during a cognitively more demanding task (Cutrell et al., 2001).

P5: More opportune moments coincide with coarse breakpoints between tasks than with fine breakpoints between sub-tasks (Adamczyk and Bailey, 2004).

P6: Less workload of the primary task indicates more opportune moments for interruptions (Iqbal et al., 2005; Iqbal and Bailey, 2005).

P7: Workload within a task differs by subtask, decreases at subtask breakpoints, decreases more at breakpoints higher in the task hierarchy than lower in the hierarchy (Iqbal et al., 2005).

P8: The cost of interruption is influenced by the position in the task hierarchy, the carry-over effect, and the difficulty of the next subtask (Iqbal and Bailey, 2006). The cost may be mitigated by external cues (Altmann and Trafton, 2004) or by 'saving state' in the preparatory phase between signal and switching (Iqbal and Horvitz, 2007).

P9: Interruptions higher in the task hierarchy lead to lower frustration (Iqbal and Bailey, 2008).

P10: The nature of the primary task influences the responsiveness to (Fogarty et al., 2004; Avrahami et al., 2008) and frustration (Iqbal and Bailey, 2008) caused by an interruption.

P11: Interruptions are less disruptive when standing as opposed to sitting in offices (Avrahami et al., 2007a).

P12: Interruptions at transitions in physical activity lead to higher receptivity than interruptions at random times (Ho and Intille, 2005).

The postulates P1 to P10 may be summarised to make assertions about the effects of interruptions in relation to what may be depicted as a tempo-hierachical cognitive task model (see figure 4.5). On the top level of the task hierarchy, literature has shown that the kind of the primary task (e.g. kind of application used) influences the responsiveness to an interruption and the frustration it may cause [P10].

On the next level of the task hierarchy, effects of interruptions have been asserted to differ by the generic phases of tasks (e.g. planning, execution, evaluation) [P2, P3]. Overall, the literature seems to agree that interruptions higher in the task hierarchy are more opportune [P5], less costly [P7], and less frustrating [P9].

Traversing the model along the temporal dimension, breakpoints (BP) between the tasks and sub-tasks have been asserted to provide more opportune moments for interruptions [P5, P2, P7]. In more detail, the disruptiveness of the interruption depends on the workload (sometimes called memory load (Miyata and Norman, 1986), cognitive demand (Cutrell et al., 2001) or difficulty (Iqbal and Bailey, 2006)) that the primary task imposes [P1, P4, P6, P7], which varies by subtask and decreases at breakpoints [P7].

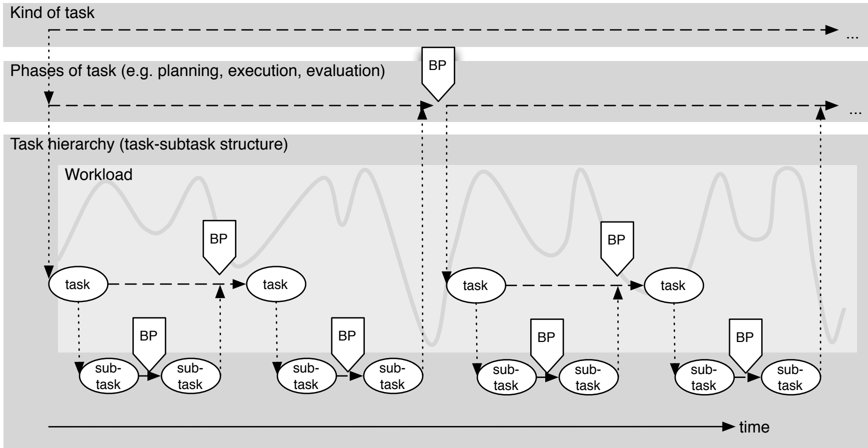


Figure 4.5: Tempo-hierarchical cognitive task model. Cognitive task model illustrating schematic variation of workload relative to hierarchical level and temporal distance from breakpoints (BP).

Last but not least, the carry-over effect influences the disruptiveness of an interruption [P7], i.e. the information needed to be remembered across task boundaries to solve the next task.

While P1-P11 were derived from laboratory studies, only P12 was derived from an ecologically valid study in a real-world setting. Focussing on mobile settings, the consideration of routine physical activity such as walking and wayfinding may affect receptivity to mobile interruptions in a significant way. Oulasvirta et al. (2005b) found that attention while mobile is fragmented and mobile interaction is limited to bursts of 4-8 seconds. This emphasises the attentional cost of mobile interruptions even during routine physical activity.

One of this thesis' goals is to understand the effects of mobile interruptions through real-world studies, both in order to understand receptivity to interruptions in everyday life, and to inform the design of interruption management systems from ecologically valid studies.

4.4.2 *Relational contextual factors*

Relational factors pertain to aspects of the interruption itself. For example, although my local context may suggest that I am not receptive, I may be receptive to a message if the *content* and the relationship to the *sender* justify the interruption; consider the message: "*It's a girl!*".

4.4.2.1 *Sender of the interruption*

In deciding whether we are receptive and responsive to an interruption, the identity of the sender arguably plays a significant role. Studies have shown that the sender's relative status, affiliation, importance and their closeness, interdependence and reciprocity to the recipient influences the way in which two people interact (Dabbish, 2006). Hence, these relationship features are likely to influence the receptivity to an interruption by a particular person.

Further support is given by Dabbish and Baker (2003), who reported that administrative assistants routinely classify the interrupter's importance in relation to their supervisors whose interruptions they coordinated. Another study found the identity of the caller to be the main factor in call handling decisions (Grandhi, 2009). Avrahami et al. (2008) showed that the identity of the sender of an instant message significantly influences the responsiveness of the recipient. As McFarlane (1997) puts it "people will graciously allow themselves to be interrupted by someone whom they hold in high esteem" (ibid., p. 40).

4.4.2.2 *Content of the interruption*

Another important relational factor is the *content* of the interruption. Studies that examine the role of the interruption content have looked at different underlying dimensions according to which we relate to the content, such as relevance, urgency, or complexity of the interruption task.

Some studies have manipulated the relevance of the interruption content to the current task (Czerwinski et al., 2000; Kalyanaramen et al., 2005; Xia and Sudharshan, 2002); and some have found significant effects for relevance (Czerwinski et al., 2000; Kalyanaramen et al., 2005). A study of interruption in the home also found that the urgency of a message is a stronger predictor of the acceptance of an interruption than the people's current engagement in activities (Vastenburg et al., 2004). Studies of acceptance of mobile advertising in China and Taiwan found informativeness, entertainment and credibility to be the most important factors of the advertisement message that influence the consumers' acceptance; and irritation to have a negative effect (Tsang et al., 2004; Shen and Chen, 2008). The study reported in chapter 6 seeks to uncover the relevance of several of these underlying factors of content to receptivity.

Avrahami et al. (2008) show that responsiveness in IM is significantly influenced by the fact if the message contains a question (faster response) or a URL (slower response). In a way this finding is not surprising, as a question represents – in terms of conversation analysis – an invitation to 'take the floor' (Schegloff, 1968).

Iqbal and Bailey (2008) found in a study that interrupting with content of general interest caused more frustration to users than interrupting with content that was relevant to the user's primary task, and that the type of content also interacted with preference in timing: whereas people preferred general interest content to be delivered at coarse breakpoints, they preferred the relevant content to be delivered at medium or fine breakpoints.

The content of interruption has also been shown to affect how interruptions are handled in face-to-face scenarios. In studies of how administrative assistants of managers mediate interruptions, the importance and urgency of the matter at hand were reported among the main factors in handling the interruptions (Dabbish and Baker, 2003; Szóstek and Markopoulos, 2006).

The complexity of the interruption task in terms of information processing and memory demands has been reported to affect the disruptiveness of the interruption (Gillie and Broadbent, 1989). The authors found that interruptions with similar content to the primary task were disruptive even if they were very short. This finding could not be replicated by Mark et al. (2008), who did not find an effect of the similarity of the interruption content to the primary task of responding to dummy emails on the overall task completion time or performance. In contrast, Czerwinski et al. (2000) found in a study that perceived disruption decreases if the incoming message is highly relevant to the interrupted task.

The study reported in chapter 7 also manipulates the content of the message in terms of the interruption task it prompts.

4.4.2.3 *Presentation of the interruption*

The way an interruption is presented may also affect receptivity to an interruption.

McFarlane (2002) introduces four methods of interruption in HCI (as discussed in 3.6.3) of which two relate to the presentation: *immediate*, in which the interruption is delivered to the screen directly; and *negotiated*, in which the user is notified of the interruption and then switches to it explicitly to attend to it. The systems employed in the studies in chapter 7 and chapter 8 implement negotiated presentation. Robertson et al. (2004) have compared the effects of using an *immediate* style interruption strategy and using a *negotiated* style to raise user attention to erroneous input during a spreadsheet debugging task (e.g. values out of range). The authors report that their participants in the *negotiated* condition were not only more productive (i.e. they fixed more bugs), but post-hoc questionnaires showed that they also comprehended the debugging environment better (Robertson et al., 2004). This indicates that the style and placement of an interruption presentation may promote or inhibit learning and orientating in GUI environments.

An experiment by Altmann and Trafton (2004) investigated the effects of visual notifications at different time intervals just before the interruption task took focus automatically on screen. The authors found that the *resumption lag*, the time it took the participants to resume the interrupted task after an interrupting task, was affected by the availability of the visual alert and the time interval for which the alert was shown (Altmann and Trafton, 2004).

In a study that logged IM conversation for more than 6,600 hours in total, Avrahami et al. (2008) found that if the messaging window already existed or already was in focus had a significant effect on the responsiveness to an incoming message.

Gluck et al. (2007) investigate matching the obtrusiveness of a notification of an incoming interruption to the utility of the interruption content. They found that matching the attentional draw of a notification to the utility of the content (e.g. less important content mapped to less obtrusive notification) is perceived to be beneficial and less annoying.

4.4.2.4 *Medium of the interruption*

The channel of the interruption may also play a significant role. In human-human interaction, the interruption is conveyed face-to-face, whereas in technologically mediated communication, the interruption may be conveyed through differing channels that present the notification or interruption in different ways and additionally may constrain the handling of the interruption in different ways.

To state a curious example, Arroyo and Selker (2003) compared different modalities of interruption notification: *light* and *heat*. The authors found, unsurprisingly, it took participants longer to respond to notifications presented by *heat*, and they found higher error rates and lower performance when resuming the primary task for the *heat* modality. In a less conclusive earlier study, Arroyo et al. (2002) have also presented the interruption by *odour, vibration* and *sound*, and found that people's response to *odour* was slowest.

More relevant to the everyday experience of interruptions, synchronous media (e.g. telephone) and asynchronous media (e.g. SMS) have different temporal consequences in how an interruption may be handled. A ringing phone, for example is what McFarlane would call an *immediate* interruption. Schegloff (1968) has reasoned about the immediate character of a ringing phone which may act as a summons. In contrast, asynchronous channels such as email, instant messengers or SMS do not demand attention in such an immediate way. In McFarlane's typology (see 3.6.3) these channels facilitate a *negotiated* interruption strategy, whereby the interruption is announced immediately but may be attended to at a later time of the recipient's choosing.

The difference between interruptions in asynchronous and synchronous media demonstrate the impact that the medium of the interruption may have on receptivity to the interruption. To illustrate by the example of receiving a phone call while on the bus – I may not be receptive to a phone call as it may be considered socially inappropriate to make a phone call in the presence of strangers, while I might have well been receptive to a text message from the same person.

4.4.2.5 *Timing of the interruption*

The timing of the interruption is a particular focus of this thesis, for which a justification is developed in the next section. *Timing* has been a recurrent theme throughout the presentation of related work on understanding opportune moments for interruptions (see 3.4, p. 57), and is at the core of three of the four primary methods to coordinate interruptions in HCI (see 3.6.3, p. 74) as they manipulate aspects of the timing of interruptions (*immediate, scheduled, mediated*) (McFarlane and Latorella, 2002).

Extensive evidence has been compiled especially in the laboratory that the timing of an interruption has significant behavioural and psychological (i.e. self-reported) effects (see 3.4).

To recapitulate, the relational character of the timing of an interruption can be illustrated by the fact that it is typically operationalised *in relation* to the interrupted's activity – or primary task – in order to gauge the effects of interruptions in relation to their position in the primary task (Latorella, 1998; Monk and Trafton, 2002; Cutrell et al., 2001; Adamczyk and Bailey, 2004, see also 4.4.1.2). For instance, studies have shown: the time to attend to an interruption was significantly longer when interrupted between activities within a task, than when interrupted between tasks or before starting or after ending the task (Latorella, 1998); the time to resume the primary task after an interruption task was lowest when the interruption occurred just before a new task stage (Monk and Trafton, 2002); messages delivered during a cognitively more demanding task were attended to significantly more slowly (Cutrell et al., 2001); and coarse and fine breakpoints occur between tasks and sub-tasks, where more opportune moments for interruptions lay at coarse breakpoints (Adamczyk and Bailey, 2004).

However, few studies have looked at the effects of interruption timing in real-world mobile settings. To that end, Ho and Intille tested the receptivity to interruption timed in relation to transitions in physical activity and found that self-reported receptivity at these transitions was significantly higher than at random other times (Ho and Intille, 2005). This thesis also takes on the challenge of investigating timing without controlling the primary task in the laboratory. The effects of adapting interruption timing to phone activity (chapter 7) and to the participants' current place (i.e. timing by proxy of location, chapter 8) are studied in everyday mobile settings.

4.5 focus of this thesis

With the breadth of the conceptual model thus developed, it is now time to narrow the scope to consider the aspects of the model that provide the focus of this thesis. Whereas the interruption process is mediated by the interaction of *all* the contextual factors that were identified previously, a detailed experimental study of the impact of all the factors is considered to be outside the scope of this thesis.

The required focus is achieved in a systematic way by filtering out the factors with the most practical applicability in interruption management systems. Here, practical applicability is understood to be provided by the question of what *adaptations* a system could sensibly make as a result of reasoning about sensed contextual factors to support interruption management. In line with the subjective, situated approach to understanding receptivity, and in order to be worthwhile, these adaptations must be interventions that maximise the overall *receptivity* of the recipient to the system-mediated interruptions as compared to a system that does not mediate the interruptions.

Herein, it is helpful to consider the contextual factors in terms of the *input* data their sensing provides to an interruption management system and the *output adaptations* the system makes to the factors in order to coordinate the interruptions in a context-sensitive way. For example, a system may take the *sender* of the interruption as an *input* parameter, and as a result of reasoning on the appropriateness of an interruption of that person may adapt the *timing* of the interruption delivery as its *output adaptation*.

To elaborate the point further, it was previously suggested that technology is not just seen as the cause of interruptions, it may also be conceived of as a means to coordinate interruptions explicitly. Different systems, approaches and techniques have been introduced (see 3.6) that support interruption management. For the sake of understanding the relation between contextual factors and context-aware engineering of solutions, the account in section 3.6 focussed on the ways in which contextual data served as an *input* to the system. This contextual data may be classified as part of the family of local or relational contextual factors that describe aspects of the experience of the interruption. Example factors that served as system input include message content (Avrahami and Hudson, 2004), eye contact (Altosaar et al., 2006; Vertegaal et al., 2002), or caller identity (Horvitz et al., 2005b).

Before considering which output adaptations are most practical (and practiced) in interruption management systems, the systems' application domain should be defined. The author of this thesis posits that an interruption management technique can be applied in the design of at least two different application domains:

1. The design of systems that manage interruptions from the recipient's existing communication network, where the sender may either be a person in some form of social relationship with the recipient or other information aggregators such as web services or systems that deliver information proactively (Lei et al., 2007).

2. The design of systems that initiate interaction proactively to engage the recipient in a user experience, where interruptions are prompts for action, such as game messages via SMS (Flintham et al., 2007) or messages via a custom system that senses the user's current location (Rowland et al., 2009).

It is argued that all possible adaptations an interruption management system can make in order to coordinate interruptions can be described in terms of system-directed alterations of at least one of the

Contextual factor	Related work	Adaptation
how (presentation)	QnA (Avrahami and Hudson, 2004)	Increases salience of the incoming message
	AuraOrb (Altosaar et al., 2006)	Presents subject heading of message
	Paui (Chen and Vertegaal, 2004)	Actuates one of four auditive/vibrotactile notification presentations
	NomadicRadio (Sawhney and Schmandt, 2000)	Adapts auditory notification signal to one of seven states
when (timing)	Oasis (Iqbal and Bailey, 2008)	Defers notifications to next detected breakpoint
	Bayesphone (Horvitz et al., 2005b)	Defers mobile phone calls to voice mail
where (place)	Matchmaker (Anhalt et al., 2001)	Helps students find nearby experts

Table 4.2: Adaptations of contextual factors in related systems.

contextual factors in the proposed model. For a system that manages interpersonal communication (see 1., above) it does not seem practical to adapt the sender or the content of the interruption. For the case of systems that initiate interaction (see 2., above), the content of the interruption could also be adapted to become more relevant or interesting to the user's system-inferred current context. In both cases, it appears practical for interruption management systems to adapt the *when* (i.e. the timing), the *where* (i.e. delivery according to the recipient's place), and the *how* (i.e. the presentation) of the interruption delivery in order to increase receptivity to the interruption most effectively. This argument is supported by the fact that this classification of adaptations cover all the kinds of adaptations that the automated systems presented in 3.6 make (see table 4.2).

4.5.1 *Focus on timing*

In order to prioritise further and compare the practicality of the three identified candidate factors *presentation*, *place* and *timing*, it may be asked which of these adaptations of interruptions would result in more opportune interruption delivery?

In general, the system is unlikely to be able to influence the user's location *before* delivering a notification. So, in practice the system can only wait until the user moves to a particular place where the notification may be delivered. Thus, in practice the system actually adapts the timing of the interruption – timing becomes a proxy for place.

This argument leaves the adaptation of the timing and the presentation of interruptions to be considered further. It appears that the adaptation of the presentation of the interruption may have the following shortcomings:

A) An interruption management system is likely to play the role of a middleware, thus the final presentation on the user's client device may be subject to the user's configuration. For example, mobile phones already allow the allocation of specific ringtone settings on the basis of incoming caller ID.

B) Even though the presentation may be adapted, the interruption is still delivered and may still be perceived as disruptive. Also, it may be better to make the adaptation earlier in the process to minimise the waste of resources such as processing cycles, bandwidth, and user attention.

C) Modern smart phones often already offer a thought out combined inbox for notifications and messages from various sources. Thereby, they implement a specific kind of presentation (in this case, negotiated). Additional solutions in this space may be seen by the user as further fragmentation of the device's functionality and seen as incoherent with the learned user experience.

Overall, it is not surprising that the *timing* of an interruption has received much attention in the quest for the definition of an opportune moment for an interruption in related work (see 3.4). As stated earlier (see 4.4.2.5), adaptation of timing appears to be central in three out of McFarlane's four primary methods to coordinate interruptions (e.g. *immediate, scheduled, mediated*, see 3.6.3); and the goal to deliver the right thing at the right time (Ho and Intille, 2005) may almost seem like the holy grail of context-aware computing.

4.5.2 *Relation of the model of contextual factors to the empirical work*

While the model of contextual factors of the interruption developed in this chapter seeks to comprehensively frame the factors involved in the interruption process, the empirical work described in Part III of this thesis explores the relation of some of the factors in a more systematic way. Figure 4.6 depicts the synthesised model of the contextual factors in relation to the experimental design of Experiment 1 (see chapter 6) and Experiment 2 (see chapter 7). The relational factors (see 4.4.2) provide the pool of factors that may be manipulated or controlled by experimental design as independent variables (IVs). The experiments test the effects of different manipulations of the factors *content* and *timing* on the dependent variable (DV) *receptivity*, which has been identified as the key psychological construct (see 4.1). Receptivity is depicted as a construct operationalisation of the local factor *recipient*, while latent interaction effects with the recipient's social (others nearby) and physical environment (place) are assumed. Observations of receptivity in the experiments are collected both

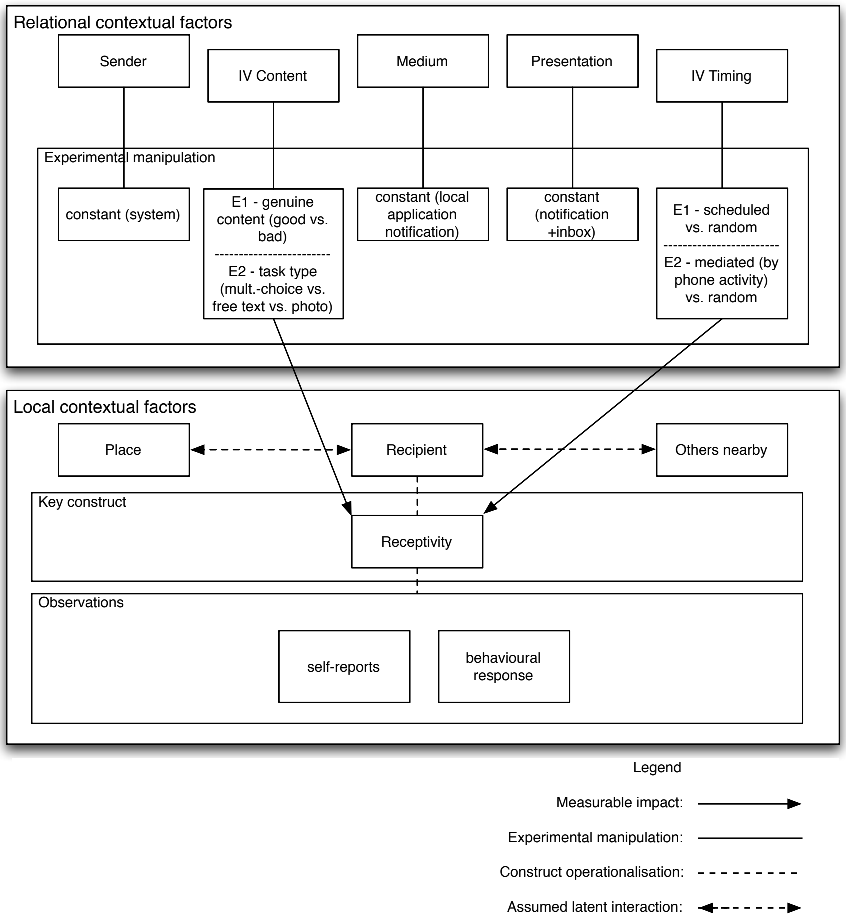


Figure 4.6: Relation of the contextual model of interruptions to experimental design of Experiment 1 (E1) and Experiment 2 (E2).

in terms of participants' self-reports and in terms of behavioural responses.

4.5.3 *Research questions*

At the beginning of this thesis, the principal research question to understand receptivity to interruptions was posed:

$\text{RQ}_{\text{principal}}$: What is an opportune moment for an interruption?

As a result of the focus on timing, the effects of different strategies to adapt the timing of interruptions is investigated in order to understand the ways in which opportune moments surface systematically from

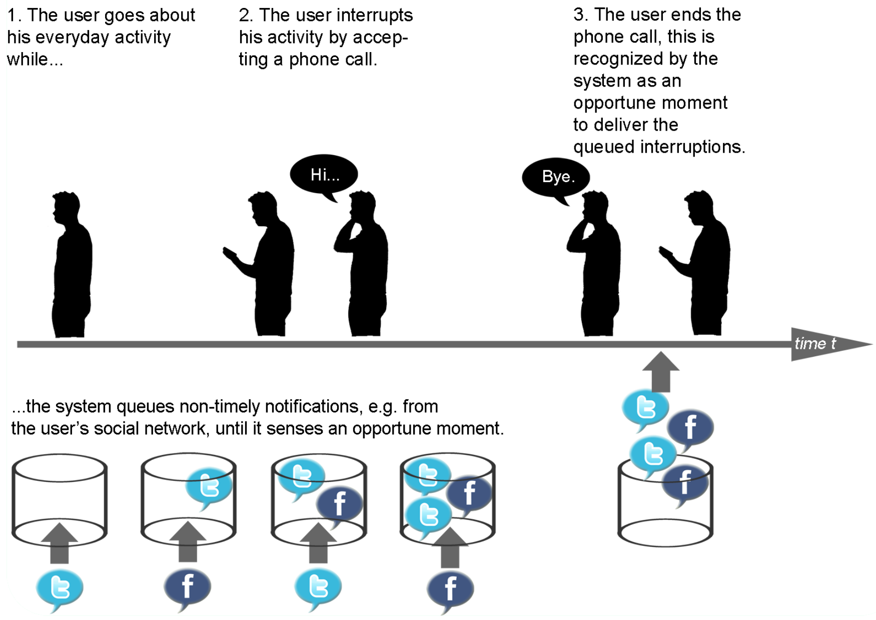


Figure 4.7: Conceptual implementation of the phone activity-mediated coordination method as realised in Experiment 2 (see chapter 7).

the user's current context. The question is studied alternatively by posing three sub-questions $\text{RQ}1_{-3}$ in the two experiments and the field study that comprise the empirical contribution of this thesis. Each study looks at the effects of different temporal methods to coordinate interruptions. The temporal methods employed in the studies roughly correspond to McFarlane's *scheduled* and *mediated* methods (see 3.6.3).

While the first study compares receptivity to randomly timed and scheduled messages, the second experiment explores the appropriateness of interruption timing mediated by the user's activity on the phone; and the final field study investigates interruption timing by location and activity within a theme park.

In Experiment 1, the rationale was to test if the method of coordinating interruptions simply by letting participants schedule them at preferred times would lead to higher receptivity than delivering them at random other times. The research question was:

RQ1: Does scheduling ones interruptions lead to interruptions at opportune moments?

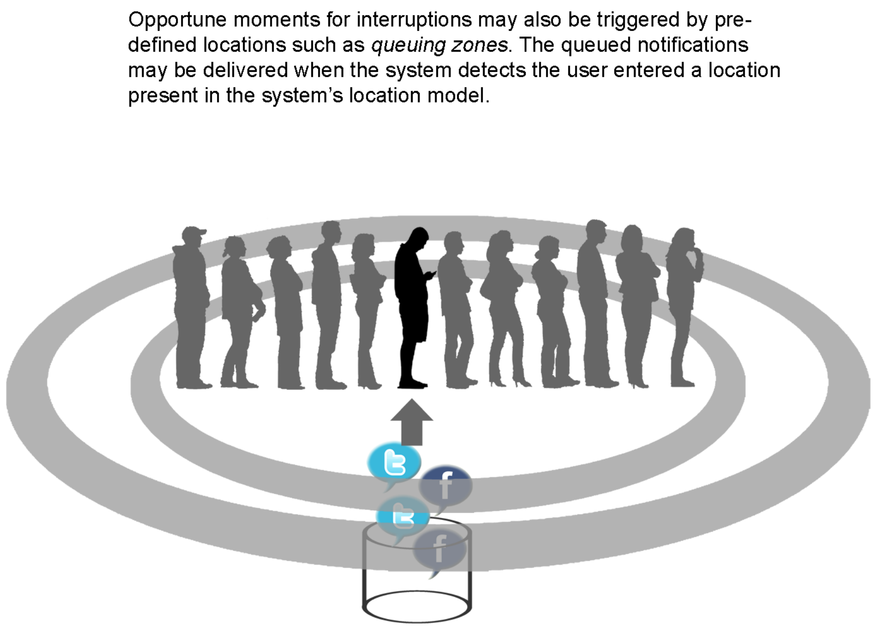


Figure 4.8: Conceptual implementation of the place-mediated coordination method as realised in the real-world application AUTOMICS (see chapter 8).

Experiment 2 and the subsequent study of a real-world application explored the more sophisticated *mediated* coordination method. In Experiment 2, the interruptions were mediated (i.e. triggered) by certain phone activity of the participant (see figure 4.7). In the real-world application, the interruptions were triggered when participants were at certain places in the theme park (see figure 4.8).

To elaborate, Experiment 2 explored the opportuneness of delivering interruptions when the participant is already interacting with their mobile phone, which is sensed by an application which triggers the interruption. The question explored in detail in this study was:

RQ2: Does the end of an episode of mobile interaction represent an opportune moment for an interruption?

Furthermore, the concept's potential was also tested as part of a real-world application. Notification-driven interaction was implemented and studied as part of a mobile service that allows users to create souvenir photo stories of their day in the park. In the system, the interruption delivery was mediated by the user's current location and the interruption tasks were tailored to the user's location. The main research question was:

RQ3: Does the recipient's current place reliably indicate an opportune moment for an interruption?

4.6 SUMMARY AND OUTLOOK

This chapter has developed the conceptual model that underlies the approach of this thesis. Receptivity has been identified as the key psychological construct that frames the experience of interruptions in terms of the phenomenological perspective (see 4.1). The model of contextual factors (see 4.3) has been inspired by models of human communication (see 4.2) and synthesised from a review of the literature (see 4.4). This also revealed a lack of studies that investigate interruptions outside the laboratory, which further motivates the methodology of this thesis to study interruptions in real-world studies.

An analysis of the mechanics of context-aware interruption management systems has added focus to the conceptual framework (see 4.5) and clarified its relation to the design of the experiments of this thesis (see fig. 4.6). The principal research question has been broken down into three main questions that drive the studies in chapter 6, 7 and 8, which in turn revolve around the adaptation of timing as their key manipulation to investigate the constitution of opportune moments for interruptions (see 4.5.3).

The next chapter presents the methodology used to investigate mobile interruptions in real-world studies.

5 METHODOLOGY OF STUDYING MOBILE INTERRUPTIONS IN SITU

This chapter takes an in-depth look at the methodology that underlies the empirical approach in the presented studies.

After discussing the implications the choice of mobile settings has for their empirical investigation, the mixed method approach is introduced and justified. Then, its constituting parts are introduced. The emphasis is on in situ methods. The Experience-Sampling Method (ESM) is central to all three studies, while ethnographic observations are an important part of the field study. Furthermore, interviews, questionnaires and diaries are introduced as supplementing the in situ methods.

In particular, ESM is introduced as the central method of choice, and its conceptual relatedness to phenomenology is developed. The choice of ESM is in line with the pragmatic attitude of this thesis – ESM may be seen as a *systematic phenomenology* (Hektner et al., 2007), but it is very much applicable in the established scientific approach of experimental hypothesis testing. Its merits, limitations and applications to HCI and studies of interruptions in particular will be discussed. Following a methodological critique of ESM in related work this thesis contributes an adaptation of ESM to study behavioural response in situ. Before part II is summed up, the use of the methods in the three studies are briefly summarised and contrasted.

5.1 SETTINGS

As the settings of HCI become increasingly mobile and temporally fragmented (Oulasvirta et al., 2005b) and its constituent parts become increasingly distributed across devices and locations (Crabtree et al., 2006) new challenges for the study and evaluation of these settings arise.

The subjective and situated perspective on experience adopted in this thesis makes it necessary to study interaction in real-world settings. The reasons for choosing to do research in real-world settings over laboratory settings are manifold; for example to account for the situated nature of how action is negotiated and produced as an interactional achievement (Suchman, 1987), and to understand how these settings may be exploited to develop context-aware technologies that dynamically adapt to their users, places and activities (Crabtree et al., 2006). Also, real-world studies provide greater ecological validity (Scollon et al., 2009), for the real-world context of the studied activities and phenomena may be considered.

In turn, this type of research "in the wild" raises a set of challenges.

Challenges are imposed by the practical difficulties of merging the typically fragmented, distributed and mobile data sources and perspectives of participants and researchers (Crabtree et al., 2006), for example when the user experience is fragmented over a longer period of time and is interwoven with the participant's everyday life (Benford and Giannachi, 2008).

The two experiments in this thesis are set amidst the participants' everyday lives. Participants are locally dispersed and do not interact with one another through the technology whilst they take part in the experiment for a period of two weeks. The focus is on the individual experience of the interruption.

In the field study, on the other hand, participants interact with one another through and around the mobile system to create souvenirs while they are observed during their day in a theme park. The focus is especially on the social practices around system usage and implications of the proactive initiation of interaction by location-based notifications.

As a result of the challenges mobile settings impose on their study, and as a result of the focus on the subjective experience of interruptions whilst participants go about their everyday lives, two methodological research aims for the empirical studies in this thesis emerge:

A) Support the study of participants irrespective of their geographical location.

B) Allow the observation of systematic changes in behaviour and/or attitudes over time.

In order to satisfy these requirements, a mixed method approach is adopted to support the studies in real-world settings in this thesis, as detailed in the following.

5.2 MIXED METHOD APPROACH

Following the pragmatic stance that this thesis adopts, the empirical approach to study interruptions in real-world settings is characterised by employing *mixed methods.* In particular, the experiments' quantitative findings from the analysis of experience-sampling data are contextualised with qualitative insights from semi-structured interviews. Observations in the field study are accompanied by interviews and analysis of application system logs that capture interaction. The technique of using multiple sources to provide corroborating evidence is also known as *data triangulation* and serves to increase confidence in their validity and help to deal with concerns about the quality of data provided by a single source (Lazar et al., 2010).

Before developing the thrust of the methodology by looking at the employed in situ methods, some additional methods and their uses in this thesis are briefly delineated.

5.2.1 *Interviews*

Semi-structured interviews were conducted with each participant after Experiment 1 and 2 had been completed, and with the visiting groups in the field study. The interviews served the important function of contextualising the quantitative data with qualitative accounts and rich descriptions of situated issues with the experience of interruptions, whereby emergence of unanticipated issues was fostered by asking open-ended questions.

Whereas some questions were posed to specifically contextualise the experimental treatment, qualitative interview data was also coded, and then collated into concepts and categories according to the data-led processes of grounded theory (Glaser and Strauss, 1967).

5.2.2 *Questionnaires*

In addition to the questions delivered by means of the ESM, experiment 2 also made use of the standardised questionnaire NASA Task Load Index (TLX)[1] to gauge a broad picture of the interruption task type's overall perceived workload after the experiment. Iqbal and Bailey (2005) used NASA-TLX to assess subjective workload of their interruption tasks. NASA TLX is a standardised instrument to assess individual's perceived workload of tasks. Participants rate their per-

1 Available at `http://humansystems.arc.nasa.gov/groups/TLX`.

ception of a task's workload by weighing the contribution of the six factors *temporal demand, mental demand, physical demand, effort, frustration* and *performance*. For the experiment presented in chapter 7, this served the purpose of a manipulation check that the levels of the IV task type were genuinely perceived as different as indicated by significant differences in workload.

5.2.3 *Diaries*

Experiment 1 utilised an end-of-day diary to get a more detailed picture on the participants' relation to the content of the interruption. Participants were reminded by email at the end of each experiment day to fill out an online questionnaire which showed them the messages they had received throughout the day and asked them to rate several aspects of the content (see chapter 6).

End-of-day diaries in this thesis complement in situ sampling with a method to add more detail to the sample by asking the participant to remember the sampling situation and to answer additional questions about it. Due to their repetitive nature, experience-sampling questionnaires are limited to short questions, while end-of-day diaries serve to overcome this limitation.

5.2.4 *In situ methods*

The method of observing participants in the field study is informed by Ethnomethodology (EM), whereas the Experience-Sampling Method (ESM) is used as the primary method to gauge receptivity to interruptions in the experiments.

Traditionally, EM-informed ethnographies focus on the observable, *overt* action and social interaction by using techniques such as the observation of interaction in situ (Crabtree et al., 2000). ESM on the other hand, is concerned with the experience that is *covert* to the eye of the observer, as it is subjectively perceived (Hektner et al., 2007). Thus, the use of ESM responds to the shortcoming induced by solely relying on behavioural observations, whereby "no information is gained on how people are actually experiencing [...]. The cognitive and affective dimensions of experience are lost." (Hektner et al., 2007, p. 6). Hence, while differing in their focus, ESM and observations in the field can be seen as complimentary for the purposes of this thesis.

Like ethnography, ESM is a method that allows for longitudinal studies, as the participants are repeatedly prompted to assess their experience over a desired timeframe. They differ in that the cost of an ESM-based study is on the side of the technical implementation, whereas the cost of the ethnography is on the side of the execution, especially in distributed settings where potentially many ethnographers need to be in the field for a prolonged period of time.

Where ethnography focusses on the thick description of the observable qualities of experience, ESM makes the idiographic and subjective experience measurable and available for analysis with established statistical methods. For the purposes of unpacking the experience of interruptions in situ, a mixed methodology that employs both qualitative as well as quantitative methods is most beneficial to achieve a broad and multi-faceted understanding.

While hypotheses testing allows for the identification of relationships between variables, ethnographies allow for the study of the context and interactions of individuals in groups (Lazar et al., 2010). Furthermore, as distinct as they may seem at first sight, EM-informed ethnographies as well as ESM have common roots in the ideas of phenomenology. In the following, their conceptual overlap in their emphasis on the situated experience and their complementary nature is considered in more detail.

5.3 ethnomethodological perspective

The ethnomethodological perspective is grounded in phenomenological thought in that it takes as its topic the methods people employ to overcome the problem of intersubjectivity – that is, in the face of the subjective experience of our everyday activity, how do we make sense of one another and achieve some commonality of understanding? This critical contribution by Schutz (1932) to phenomenology arguably paved the way for phenomenological thought to be applied to the sciences concerned with the social world. Today, EM and the problem of intersubjectivity is a central concern to sociology. Within the field of HCI, the perspective has gained momentum in the shift towards understanding groups of people and their interactions with and around systems, which lead to ethnomethodologic accounts of socio-technical systems (e.g. Crabtree et al., 2006; Tolmie et al., 2008).

These kinds of ethnographies may be said to "attend to the interactional work of a setting (to what people do), and work-practice to the methodical ways in which people accountably organize whatever it is that they do." (Crabtree et al., 2009, p. 881). It is through the notion of accountability of actions through which intersubjective understanding is achieved, hence it is the accountability of actions which is at the forefront of the ethnographer's interest. This then explains that observation is the method of choice for the ethnographer. As Garfinkel (1967), the founder of ethnomethodology put it, activities are made "account-able" by the fact that they are "observable-and-reportable, i.e. available to members as situated practices of looking-and-telling" (ibid., p. 1).

For the the purpose of this thesis, the social embedding of activities was argued to be a critical contextual factor to understand the interruption process. Observation of the social situation surrounding the interaction with and around the mobile technology was a key method in the field study. The technology-in-use was observed and filmed for later analysis – video is widely recognised as an important resource for ethnographies around technology use (Crabtree et al., 2006). The main question this thesis asks from observing the usage setting is: How does the system's location-based and event-based messaging interact with the social situation and how the system is used? The approach is detailed in chapter 8.

5.4 EXPERIENCE-SAMPLING METHOD

By introducing the Experience-Sampling Method (ESM), Csikszentmihalyi et al. (1977) invented an approach to assess the subjective quality of the individual's current experience without the need to observe participants. At the heart of the method is the prompting of participants with a signalling device to fill out questionnaires about their current experience over longer periods of time. Whereas Csikszentmihalyi et al. (1977) used a beeper as their signalling device, the ubiquity of the mobile phone makes it an ideal vehicle for ESM studies today.

ESM has been designed to accommodate for the shortcomings of post-hoc evaluation techniques such as diaries and questionnaires that rely on the retrospective assessment of the experience, which may lead to reporting bias (Stone and Shiffman, 2002) and distorted assessments of the actual experience due to how individuals reconstruct episodic memories (Hektner et al., 2007), e.g. through social or cultural connotations that are applied post-hoc (Csikszentmihalyi and LeFevre, 1989).

ESM has been introduced to study the quality of experience in everyday life, but has subsequently been applied to such diverse fields as the experience of work, the examination of cross-cultural differences (Hektner et al., 2007), and educational and clinical research, for example on stress related diseases (Yoshiuchi et al., 2008), often under its pseudonym Ecological Momentary Assessment (EMA). With a strong tradition in psychology and diverse application fields, it has more recently been adopted by HCI researchers to study mobile and ubiquitous applications in the wild (e.g. Ho and Intille, 2005; Consolvo and Walker, 2003).

5.4.1 *Idiography and nomotheticity*

A challenge for methods that take an idiographic perspective – a focus on the subjective experience – is that it may be difficult to reliably compare the findings for one person with the findings for another. On the other hand, the application of nomothetic quantitative techniques such as comparing different treatments to groups of people to derive rules that apply for populations may be problematic in light of this thesis' primacy of the subjective nature of involvement in the world.

ESM reconciles these apparent opposite approaches to knowledge. Hurlburt and Akhter (2006) illustrate the principle in which the ideographic and the nomothetic approaches co-exist in ESM: By conducting an idiographic study with several participants, salient characteristics might emerge across the collection of participants. Thus, ESM is a method with which nomothetic insight can be achieved through a series of idiographic descriptions of subjective experiences. Practically speaking, ESM allows both for the investigation of within-person processes, and a between-person comparison of aggregated momentary data that reflects trait conceptions more broadly (Scollon et al., 2009).

The founders of ESM have described it as a *systematic phenomenology* as it makes idiographic data available for statistical reasoning (Hektner et al., 2007).

> "The measurements of experience made possible by ESM can be called a systematic phenomenology in that it departs from pure

phenomenology by combining a focus on lived experience with an attempt to use the tools of empirical investigation – including available technologies, research designs, and statistical analyses" (Hektner et al., 2007, p. 6).

Hektner et al. (2007) concede that Husserl would probably strongly disagree with their "nomothetic contamination of his ideographic approach" (ibid., p. 6), but their pragmatic view that "standing on the shoulder of giants one might see views that the giants themselves could not perceive" (ibid., p. 6) is very much in line with the pragmatic approach present in this thesis.

5.4.2 *Types of ESM*

Scollon et al. (2009) identify three types of ESM, according to the differing underlying principles of the timing of prompts:

SIGNAL-CONTINGENT SAMPLING: The signal is randomly timed, hence it is also known as random sampling.

INTERVAL-CONTINGENT SAMPLING: The signal occurs after a pre-determined amount of time.

EVENT-CONTINGENT SAMPLING: The signal occurs of after a certain event happens.

Mixed forms are also possible. For example, Experiment 2 makes use of all three types of sampling. In one condition, participants are signalled at random times three times a day (signal-contingent); however the signals are also constrained to occur at least 30 minutes apart from each other (interval-contingent). In the other condition, notifications are delivered after participants interacted with the device (event-contingent). In fact, the responsiveness to signal- and event-contingent sampling is compared to provide behavioural evidence for hypotheses testing.

Experiment 1 compares receptivity to interval-contingent sampling at participant selected times with random sampling. The mobile technology deployed in the field study uses location-based messages to initiate interaction, and system event-contingent sampling to notify users of newly available images.

In addition to the methods by which signalling occurs, ESM may also be categorised by the type of data that is collected by it. It has been argued that ESM is traditionally more concerned with collecting self-reports of subjective experience, and that Ecological Momentary Assessment (EMA) includes more broadly other types of data such as the recording of events in the environment and physiological data (Stone and Shiffman, 2002). For the purpose of this thesis, the distinction between ESM and EMA as a characterisation by the type of data recorded is not useful. Instead, it is treated as a feature of its flexibility that ESM can be used to record *multi-modal* data (Ebner-Priemer and Trull, 2009), such as psychological, physiological and behavioural data.

Multi-modality and flexibility in the design of the method guiding the timing of the prompts are two important benefits of ESM. Further benefits and limitations are outlined in the following.

5.4.3 *Benefits and Limitations of ESM*

Ebner-Priemer and Trull (2009) identify the following major benefits of ESM:

a) Real-time assessments increase accuracy and minimize retrospective bias;

b) repeated assessments can reveal dynamic processes;

c) multi-modal assessments can integrate psychological, physiological, and behavioural data;

d) setting- or context-specific relationships can be identified;

e) interactive feedback and monitoring is possible in real time; and

f) assessments in real-life situations enhance generalisability.

It seems that the benefits of ESM mainly relate to the situatedness of research it allows (a, d, f), the sound way to assess experience on the individual level (b, e), and the multi-modality of the data that can be gathered by it (c). Hence, ESM is an optimal method for the purposes of studying the subjective experience of interruptions in situ.

However, ESM is not without its limitations, which need to be addressed to develop a full picture of the method. Scollon et al. (2009); Hektner et al. (2007) and Stone and Shiffman (2002) all report issues to be addressed when conducting ESM studies that relate to the participants, the situation, the treatment (i.e. the sampling), the data analysis and ethical concerns. In the following, the issues are considered and the ways of dealing with them in the empirical work presented in this thesis are discussed.

5.4.3.1 *Participant issues*

Hektner et al. (2007) and Scollon et al. (2009) mention self-selection bias as a potential problem for ESM studies. Who is willing to take part in a study that lasts weeks? Who completes them and who provides the most data? Who is asked to participate? This may mean that certain types of individuals are over- or under-represented in the data. Individual motivation and demographic factors such as education and exposure to technology may also confound compliance and then appropriateness of experience-sampling, e.g. it may be situationally inappropriate for "truck drivers" to participate (Scollon et al., 2009).

However, this is less of an issue for the within-subject design of the two experiments in this thesis, which mitigates individual differences, as the unit of analysis is the individual. Fluctuations in within-subject measurements are compared to each other only per individual; the aggregated reactions of people are not compared to one another. Furthermore, statistical analysis techniques such as Linear Mixed Models (LMM)[2] allow for the inclusion of *participant* as a random factor into the model, as applied in the analysis of Experiment 2 in section 7.4.

In addition to statistical analysis as a source for dealing with participant issues, monetary incentives have been shown to significantly improve compliance (Scollon et al., 2009). However, caution must be exercised in deciding on the amount of compensation, as too much may attract participants who are not intrinsically motivated (Stone et al., 1991). For the experiments in this thesis, the amount of reimbursement the participants were paid was estimated based on the recommendation of Hulley et al. (1988), who suggest to pay the current hourly wage for unskilled labour. For the groups participating in the field study, the prospect of a free visit to a theme park and the reimbursement of subsistence and travel costs to the park was seen as enough of an incentive.

5.4.3.2 *Compliance*

A further issue that may be both related to the individual participant or even to the treatment is compliance. Compliance in ESM is an important issue because the success of a study (Stone and Shiffman, 2002) and its ecological strength (Scollon et al., 2009) may depend on the degree of compliance.

2 They are called *mixed* because they model both the fixed effects, i.e. the effect of the treatment/IV and the random effects, i.e. the incalculable effect of the individual participant.

Compliance in ESM studies can be expressed as response rates, and response rates are reported for all the studies in this thesis. In addition, statistical tests are employed to reveal if compliance or non-compliance are systematically related to the treatment. For example, in the case of Experiment 2, noncompliance is significantly higher for the baseline condition than for the treatment condition – indicating a significant effect of the treatment on compliance. This case illustrates how the measure of non-/compliance may be explicitly included as a dependent variable and not just as a confounding factor.

Another issue related to compliance is the time lag between signal and response (Scollon et al., 2009), or "not completing momentary reports according to the protocol but [...] completing reports at a later time" (Stone and Shiffman, 2002, p. 240). Again, for the empirical studies in this thesis this is not seen as a problem, but is explicitly included as the dependent behavioural measure *response time*.

However, for the analysis of self-reports, which often aim at the moment of the delivery of the interruption, a significantly delayed response diminishes the validity and increases the recall bias. It is unclear whether that response can still be seen as related to the past moment of delivery. Scollon et al. (2009) cite several studies whose authors seem to suggest a somewhat arbitrary response window of 30 minutes to accept responses as valid and include them in their analysis.

To follow the trend in the literature, in this thesis responses delayed for more than 30 minutes are excluded from the analysis of self-reports. As the distribution of response time typically follows an inverse power function – with the dominant proportion of responses within seconds or minutes of the interruption – the number of responses lost by cutting off the long tail at 30 min should be small.

These issues of compliance are discussed in more detail in the analysis of the respective studies, for example as response rate bias and response time bias in section 6.6.1 (p. 170).

5.4.3.3 *Habituation and experiment fatigue*

A further challenge of the validity of ESM studies is imposed by the long-term nature and the repeated prompting that occurs throughout the day. Participants in the experiments routinely reported that they got more annoyed with the study towards the end of the time period of two weeks. The danger herein is that the quality of responses may diminish beyond validity. To become aware if habituation affected the data, statistical tests can be used to assess the effects of the study's temporal course on the responses. This is achieved by looking for significant differences in response time and response rate as the studies progressed.

Moreover, the ideal solution to experiment fatigue would be to design it in such a way that it does not occur in the first place. This is then very much a question of balancing the number of interruptions over the day and limiting them to a tolerable number. The details of this design question are discussed as part of the individual experiments, as this balance needs to consider a number of factors with the experimental conditions and statistical reliability on the one hand and the burden on participants and threats to compliance and habituation on the other.

5.4.3.4 *Reactivity*

Reactivity is the psychologists version of Heisenberg's Uncertainty Principle – the potential for the studied phenomenon to "change as a result of measurement or reporting" (Scollon et al., 2009, p. 19).

For example, Picard and Liu (2007) devised an experience-sampling study that prompted the participants to self-report stress levels. Their participants agreed to placing bluetooth beacons at home and in the office, agreed to wearing a heart rate monitor, an accelerometer and a pedometer and agreed to carrying around a Personal Digital Assistant (PDA) for a period of two weeks, and were interrupted to report their stress levels and current activity. The authors do acknowledge the problem of reactivity (without calling it that) faced in experience-sampling studies: the challenge to build a device that is highly interruptive, and in their case, supports monitoring of stress level without at the same time being a *cause* for increased stress itself. In a way, the measuring device becomes the treatment itself.

It is especially for the investigation of health-related issues that ESM-based studies impose the danger of leading people to pay unusual attention to their internal states and behaviour and thus may serve as an amplifier of the condition's symptoms (Scollon et al., 2009). However, studies that explicitly looked for a relation between ESM studies and reactivity on the whole could not conclusively establish its presence (ibid.).

For the studies presented here, it is argued in section 5.4.4.4 that reactivity is not a problem, but a *feature* of the experimental design. That is, the signal that prompts participants to react is equivalent with the treatment, i.e. the interruption.

5.4.3.5 *Data issues*

Once all the data is collected by means of the ESM, the researcher faces not only problems common to all statistical analysis such as the identification of erroneous data, outliers, and the appropriate analyses for the variables' scale level etc., but specifically problems that arise from the within-subjects repeated measurements design.

Does one choose an aggregate or a disaggregate analysis technique? That is, is the participant the unit of analysis and therefore do the repeated measurements have to be aggregated (e.g. by averaging) to the individual level, or is the sample the unit of analysis and therefore is a disaggregate procedure used?

Whereas aggregation might be the more straightforward and manageable technique that also increases reliability and gives higher correlations, a concern is that response sets might get amplified through aggregation (Scollon et al., 2009). For this reason, aggregate procedures such as repeated-measures Analysis of Variance (ANOVA) require equal amounts of repeated measurements per participant and condition. Disaggregate procedures such as LMM do not have this requirement (Garson, 2010). In addition LMM allows the inclusion of both fixed and random effects, and allows the consideration of both within- and between-subjects effects. However, the choice of the appropriate procedure depends on characteristics of the data set, such as the shape of the distribution and the scale level of the variables, which will be discussed in more detail on a per experiment basis.

5.4.3.6 *Ethical issues*

For studies involving human participants both Palo Alto Research Center (PARC), where the first experiment was supervised, and the Mixed Reality Laboratory of the University of Nottingham, where the second experiment and the field study were supervised have guidelines for appropriate ethical conduct. In all cases, the studies' aims, methods and procedures were approved by an ethical committee, especially guaranteeing anonymisation of data and prior informed consent of the participants.

5.4.4 *Objectives and Applications of ESM in HCI*

ESM has been applied in HCI to problems relating to different phases in the iterative design process, from requirements engineering (Consolvo and Walker, 2003) to evaluation (Iachello et al., 2006; Lee et al., 2010), to the labelling of data in user modelling (Kapoor and Horvitz, 2008). ESM tools and ESM in studies of interruptions are also reviewed in this section.

5.4.4.1 *ESM in user modelling*

Kapoor and Horvitz (2008) suggest the Experience-Sampling Method to assist in the building of predictive user models. The experience-sampling method can provide a situated technique to gather *labels*[3] of certain aspects of a continuous data stream with a relatively low overhead on the user. They compare different methods for guiding experience sampling including:

RANDOM PROBE: Participants are prompted at random times.

UNCERTAINTY PROBE: Participants are prompted based on a predictive model to gather labels for situations associated with the highest uncertainty.

DECISION-THEORETIC PROBE: In addition to detecting uncertainty of the model, the expected value of the information and the expected cost of the interruption the probe would cause is taken into account.

DECISION THEORETIC DYNAMIC PROBE: In addition to the above item, this methodology adds advanced dynamic generation of models should an existing model not fit the situation.

Data for the purpose of user modelling is often referred to as "ground truth" (e.g. Klasnja et al., 2008) because of its representation of behaviour, for example device usage or real world activity through sensor data. In order to make meaning of this data it is necessary to label it, e.g. by ascribing feelings, values and thoughts, or even

3 Machine learning techniques use *labels* on data to train classifiers or other techniques such as *supervised learning* to make predictions or classifications.

just a description of what succession of data points in an accelerometer log represents a human step.

Even though this thesis does not strictly employ any machine learning techniques, the behavioural data describing the interruption process can be seen as ground truth, and the self-reports, for example of receptivity levels can be understood as labels. This opens up possibilities for the research presented here to be enriched with machine learning techniques in the future.

5.4.4.2 *ESM in mobile experiences*

ESM has been applied in the design and evaluation of mobile applications (Consolvo and Walker, 2003; Iachello et al., 2006). For example, Consolvo and Walker (2003) show how ESM can be used early in the design process as part of a requirements analysis for a ubiquitous computing application by studying users' information needs. ESM is used to inform the requirements for the design of a "Personal Server" for example by asking the users on a PDA about information they needed but did not have available or wanted to retain from their environment for future use (Consolvo and Walker, 2003).

Iachello et al. (2006) show exemplarily how ESM can be used to evaluate a prototyped mobile technology. The authors use an experience sampling approach to gauge users' opinions on the impact on privacy their mobile system has. The system called PERSONAL AUDIO LOOP continuously records audio from the environment in order to assist the user's recall of recent auditory information. The findings point out how important the current situation of a mobile technology is for privacy concerns and at the same time underscore the importance of an in situ method such as ESM to enable that kind of situated research.

Klasnja et al. (2008) use mobile phone-based experience sampling in conjunction with a continuous mobile sensing platform (MSP) to establish the optimal trade-off between participants' accuracy of recall of routine activities (by the example of sitting and walking as inferred from the MSP's accelerometer data) and annoyance of prompting multiple times per day. Interesting in the context of this thesis is the finding that 5-8 regular prompts per day is the optimal prompting frequency to achieve fairly accurate recall and keep annoyance relatively low.

ter Hofte (2007) used ESM on mobile phones to investigate the context of his participants (whether they were in conversation; their location, e.g. home, office, transit; and with how many people), and which of this context information participants are willing to share with potential callers.

The popularity of ESM in HCI has led to the design of generic toolkits to support the otherwise costly implementation of ESM-based studies, which often enable further research. For example, the investigation by Consolvo and Walker (2003) employs the ESM tool iESP (Barrett and Barrett, 2001), the study presented by Klasnja et al. (2008) makes use of the freely available toolkit MyExperience (Froehlich et al., 2007), and ter Hofte (2007) research is based on the tool Xensor (ter Hofte et al., 2006). The main ESM tools have been reviewed for potential adoption in this thesis.

5.4.4.3 *Experience-Sampling Tools*

Due to the extensive interest in the HCI and broader research community, several tools to support experience-sampling have been developed. Some of these tools have been reviewed in order to decide whether they were appropriate to support the experience-sampling endeavour for this thesis[4]. An overview of the tools is given in table 5.1[5].

4 For a more detailed review see Fischer (2009).

5 At the time of writing, a more comprehensive list can be obtained at http://amsterdamlivinglab.novay.nl/learn/tools.

Tool	Description	Platform	Public availability	Authors
caes	Context-Aware Experience Sampling tool which allows questionnaire triggering by contextual clues captured with the mobile device.	PDA (PocketPC)	no	Intille et al. (2002)
ESP/iESP	The Experience Sampling Program consists of a native mobile application to trigger and run the ESM questionnaires, and a browser-based desktop application to create the logic for the timing of the prompts, the content and structure of the questionnaires and to deploy the studies to the PDAs.	PDA (Palm Pilot); Windows or Linux desktop	yes	Barrett and Barrett (2001)
MyExperience	Native mobile application, which triggers prompts to answer questionnaires, can also be configured to collect sensor data and user activity.	Windows Mobile 5.0	yes	Froehlich et al. (2007)
Momento	Client-server experience-sampling architecture which consists of a desktop platform to configure the experiment logic on a server and a mobile client to deliver the prompts and send data via SMS/MMS or HTTP to the server where it can be monitored.	client: flexible, Windows Mobile, J2ME	yes	Carter et al. (2007)
Reconexp	Native mobile ESM application and combined web-based diary to reduce the data loss of ESM-based studies by a supporting diary method known as the Day Reconstruction Method.	Windows Mobile & website	no	Khan et al. (2008)
Xensor	Native mobile application for sensor logging, device usage and experience-sampling; data is periodically uploaded to a server.	Windows Mobile	no	ter Hofte (2007)

Table 5.1: Overview of experience-sampling tools.

Primarily for practical reasons it was decided to implement the experience-sampling from scratch, or in the case of the first experiment, to extend an existing sampling system (see chapter 6). The main reasons for this were the requirements for phone activity-contingent sampling and logic of timing, the desired interruption task structure, full control of time data logging in experiment 1, and the tight integration with a mobile system for user-generated souvenir creation employed in the field study.

5.4.4.4 *ESM in studies of interruptions*

ESM has a special role in studies of interruptions, not just because interruptions are the phenomenon investigated in this thesis, but because the prompts that are an essential constituent of ESM *are* a particular kind of interruption[6]. As pointed out above in section 5.4.3.4, the fact that the signal is an interruption might be a potential limitation to the experimental validity of studies that look at stress levels or other health-related issues. However, in case of the study of interruptions themselves, this particularity is actually advantageous in that the ESM signal serves the dual purpose of being the treatment (i.e. interruption) and the prompt to self-report on its effects at the same time.

In the literature, ESM has been applied to an examination of attitudes towards availability and interruption in offices (Hudson et al., 2002), and in the home (Nagel et al., 2004); it has been applied to label simulated sensor logs from offices with estimates of interruptibility to enable machine learning techniques to predict interruptibility from simulated (Hudson et al., 2003) and real sensors (Fogarty et al., 2004); and it has been used to investigate the differences in self-estimated interruptibility and others' estimation of that interruptibility (Avrahami et al., 2007a).

6 In this view, interruptions are seen as neutral – they may be seen as disruptive, welcome or even go unnoticed, but this is decided in the eye of the beholder, i.e. the recipient.

Away from stationary settings such as offices and homes, ESM has been used to solicit receptivity scores in relation to breakpoints in physical activity (Ho and Intille, 2005), and the method has been used in a study of the factors influencing call handling decisions (Grandhi, 2009).

Even though all of these studies use ESM, the research focus and methodical details are quite distinct from the ones in this thesis. While the goal of this thesis is to more broadly understand the interruption process both in terms of self-reports and behavioural measures of the actual response to the interruption as it may delineate receptivity in situ, these studies treat their ESM approach solely as a method to collect self-reports, and miss the opportunity to measure actual behavioural responses to the prompting signal that represents an interruption.

This important discovery is related to the research approach present here and its acknowledgement of the relational factors of an interruption, which leads to practical research design implications, as discussed in the next section.

5.5 ADAPTATION OF ESM IN THIS THESIS

Based on a methodological critique of ESM in related work, this section develops an adaptation of ESM as a methodological contribution of this thesis. At the core of the adaptation is the temporal model of the process of the interruption developed in section 3.3.5 (p. 55).

5.5.1 *Methodological critique of related work*

All of the reviewed studies that use ESM to investigate interruptions (Hudson et al., 2002, 2003; Nagel et al., 2004; Fogarty et al., 2004;

Avrahami et al., 2007a; Ho and Intille, 2005; Grandhi, 2009) fail to acknowledge the dual nature of the method's prompting signal, which represents both an interruption and a prompt to answer an experience-sampling questionnaire at the same time. Thus, the opportunity is missed to record details of how and when the recipient responds to the prompt, and thereby collecting behavioural data of the interruption process in situ. The collection of multiple data sources to triangulate the data process is central to the empirical studies presented in this thesis. It is practised for various reasons, e.g. to:

A) increase (ecological) validity,

B) accommodate multiple measures of experience,

C) test hypotheses, and

D) achieve a multi-dimensional view of the interruption process.

Furthermore, another shortcoming from solely relying on gathering self-reports in the reviewed work is the failure to deliver the sampling prompts according to an experimental design. That is, while specific aspects of the interruption in the studies presented here are designed to represent an experimental *treatment* (e.g. timing and content in Experiment 1 and 2) in order to test the experimental hypotheses, the ESM studies in the related work fail to do so, and as a consequence remain exploratory in nature. While exploratory work has its place and reason, it is often limited to identifying correlation between variables – the identification of causal relations remains the realm of experiments (Lazar et al., 2010).

Another challenge with ESM is the generic type of question the researcher is often limited to ask in the ESM questionnaire as a consequence of trying to make the question non-specific enough to be answerable in any possible context. For example, ter Hofte (2007) asked their participants if they were willing to accept a call by *someone* and what type of context information they are willing to share with that *someone*. The problem with this type of hypothetical and non-specific question is that it disregards the relational factors of an interruption (see Grandhi, 2007, and 4.4.2). Intuitively, to that type of question one might respond: "*it depends*". It depends on who is calling, the topic of the conversation, and the resulting shared context that arises, which may make it reasonable to share contextual information such as my current location with the caller.

In a similar vein, Hudson et al. (2002) ask generically how much time the interrupted would have for an interruption, or ask their participants to rate their interruptibility on a scale (Hudson et al., 2003). Ho and Intille (2005) ask their participants the generic question of how receptive they would be to a "phone call" or a "reminder". Studies also used fictional messages (e.g. "the fire alarm in the shed has detected smoke")(Vastenburg et al., 2004). Again, what is lost here is the chance to consider the influence of genuine relational factors of an interruption. As pointed out in section 4.4.2, factors such as the identity of the sender or the *genuine* content of the interruption play an important role in influencing how receptive we are to the interruption.

However, it is acknowledged that the investigation of relational factors does not come without its costs and limitations. For example, it may diminish the comparability of messages and thereby endanger the validity of the treatment. Hence, where in Experiment 1 interruption content is manipulated in a way to study the effects of *genuine* content, in Experiment 2 content takes the form of task requests as the focus is more on the interactional effect of timing and task type.

In summary, this critique emphasises that the phenomenon under scrutiny may have a profound impact on the methodological choices and the design of empirical study. In the following, the adaptation of ESM underlying the experimental approach of the empirical work (Part III) in this thesis is detailed.

5.5.2 *ESM-adaptation to capture behavioural response in situ*

As a consequence of the methodological critique above, this thesis adapts ESM to shape the experimental design with the following insights:

A) *The dual nature of the prompting signal* – The notification represents both the method's prompting signal and the treatment (i.e. interruption) of the experiments (chapter 6-8).

B) *Signal delivery according to an experimental design* – The notifications are manipulated according to the experimental design to represent conditions of the independent variable (e.g. interruption content or timing)(chapter 6, 7).

C) *Notifications may contain genuine content* – The notifications may be manipulated to contain *genuine*, rather than generic content so as to test the influence of genuine *relational contextual factors* (chapter 6).

The important consequence of insight A) is that in addition to collecting repeated self-reports on their current experience, the recording of behavioural data in terms of the temporal dimensions of the model developed in section 3.3.5 (p. 55) is made possible. The temporal model of the interruption process was devised (see 3.3.5) based on a review of several temporal models in the literature (see 3.3). Furthermore, the wealth of research was reviewed (see 3.4) that operationalises this kind of behavioural data as dependent variables to measure and analyse temporal aspects of the interruption (see table 5.2 for a summary) or manipulates it as independent variables in their experimental design.

The temporal model enables the adaptation of ESM to analyse behavioural data in a meaningful way. To illustrate, as the temporal dimensions are captured for each interaction on the individual level

Operationalisation	Prior work
Time to attend to an interruption	Latorella (1998); Cutrell et al. (2001)
Time on the primary task (completion time)	Adamczyk and Bailey (2004); Cutrell et al. (2000)
Time on the interruption	Adamczyk and Bailey (2004)
Time to resume the primary task	Monk and Trafton (2002); Adamczyk and Bailey (2004)
Pupillary response	Iqbal et al. (2005)
Forgetting the primary task goal	Cutrell et al. (2001)
Self-reported receptivity rating	Ho and Intille (2005)
Self-reported interruptibility rating	Avrahami et al. (2007a)
Self-reported emotional state	Adamczyk and Bailey (2004)

Table 5.2: Dependent measures to assess interruption timing in related experimental work.

they can be analysed for systematic differences and associations over time and by the conditions of the experimental manipulation. For example, this allows to compare differences in response time to notifications delivered at a random moment versus a hypothesised opportune moment as an indicator of receptivity. Quicker responsiveness is associated with higher levels of receptivity, as the active process of responding is arguably indicative of receptivity to the message that is responded upon.

The potential discovery of significant differences in the temporal dimensions is faciliated by signal delivery according to experimental design (insight B) above). In the studies, observations may be made in relation to the manipulations of the independent variables, for example in relation to the content of the interruption, where content may be genuine information (Experiment 1, chapter 6), or an interactive task such as a multiple-choice, a free-text or a photo task (Experiment 2, chapter 7), or a task to annotate a photo or create a photo story (field study, chapter 8).

Hence, the temporal model allows for conclusions about the effects of the independent variables on certain parts in the process of the interruption. For example, while the *timing* of an interruption may influence *acceptance time*, analysis according to the model may show that it does not influence the *task time*.

In addition, behavioural data allows for more robust statistical tests, and helps to triangulate the data from multiple perspectives. The collection and compilation of a multi-layered data model is central to the approach of this thesis.

Behavioural responses also compliment and contrast self-report data in that they are not subject to self-report bias, according to which participants answer in what they expect to be a socially desirable way (Donaldson and Grant-Vallone, 2002). According to Scollon et al. (2009), the collection of such multi-modal data may be the best remedy to self-report problems. This approach has been reported in the literature on interruptions, for example Iqbal and Bailey (2008) measure reaction time and resumption time and let users self-report frustration after the study.

In the adaptation of ESM in this thesis, while the participants' behavioural response is logged in terms of timestamps during their interaction with the experiment applications, participants also provide self-reports by completing the interruption tasks the application has prompted them with. In summary, the ESM-adaptation extends existing approaches by

- exploiting the notification's dual nature to be both prompting signal and experimental treatment, and
- delivering the signals according to an experimenal design (independent variable manipulation), and
- collecting both participants' self-reports and temporal metrics that represent their behavioural response.

Having explored the approach and the methodology that underlie the empirical work of this thesis, the three studies are finally briefly contrasted and summarised in the following.

5.6 studies in this thesis

While the research questions that link the three studies were stated in 4.5.3, here, the methods of the studies are briefly contrasted and summarised.

The two experiments reported in chapter 6 and 7 are conceptually similar in that they test aspects of the interruption process in a hypothesis-driven fashion with controlled and counterbalanced independent variables in a repeated-measurements, within subject design. Also, the experiments take place in a real-world setting, as each participant goes about their everyday life as they are interrupted on their mobile phone throughout the day for around two weeks. Both behavioural as well as self-reported data is gathered in situ for later data analysis, while post-hoc interviews serve to contextualise and enrich the data qualitatively.

The adapted ESM (see 5.5) is employed to repeatedly prompt participants on a mobile phone to record their current receptivity to an interruption. For the case of Experiment 1, participants are prompted and interruption content is delivered by SMS to their own mobile phones, and responses are sent back by participants in the same way. Participants were asked to rate their receptivity to the message's genuine content – general interest information such as news – in terms of the timing and their *gut reaction* to the message (see chapter 6).

In Experiment 2 (chapter 7) and in the field study (chapter 8), mobile phones with a custom application were lent to participants. In Experiment 2, the custom application prompted participants to complete one of three short interactive tasks with the application: to take a photo, to answer a multiple-choice question about the moment of the interruption, or to state their current activity. The custom application allowed for the detailed recording of the four temporal dimensions outlined in the temporal model (see 3.3.5, p. 55).

In the field study, notifications suggested the completion of interactive tasks in order for participants to create their own photo-story souvenirs of their day in the park (Durrant et al., 2011). Rather than seeing this work as disparate from the previous two studies, it can be seen as a natural progression and an evaluation of the developed interruption technique in a prototypical real-world application. In addition to recording the interaction in system logs in the familiar way, in contrast to the experiments the interaction with and around the real-world application was directly observed in the field.

It was argued (see 4.5, p. 105) that an interruption management system may also be used to initiate interactions with the system proactively – the system used in the field study implements this approach. Hence, whereas the experiments may be seen as studying effects of mediating interruption timing abstracted away from an actual application, the field study tests the applicability of mediating interruption timing in a real-world application.

To pick up the research questions posed in 4.5.3, the overarching theme of the studies is the exploration of the ways in which opportune moments surface from different strategies of adapting interruption timing. Experiments 1 and 2 explicitly compare *scheduled* and *mediated* strategies (McFarlane, 2002) with a random baseline, using temporal dimensions as dependent measures (see figure 4.6, p. 110). In particular, Experiment 2 triggers notifications by the participants' phone activity. The real-world application also implements the *mediated* method in that it notifies participants to prompt interaction at opportune places.

5.7 summary

This concludes Part II of this thesis. In chapter 4, the approach to studying receptivity to interruptions was developed, with the psychological construct of receptivity and a conceptual model of the contextual factors of the interruption at its centre. The argument of practical applicability served to define the focus of this thesis on investigating the effects of manipulating interruption timing as the key adaptation of interruption management systems. The definition of this focus then gave rise to the principal research question "What is an opportune moment for an interruption?" and three variations of it that will be at the heart of the empirical studies.

Whereas chapter 4 detailed the *what* of this thesis, this chapter concerned itself with the *how*. Specifically, how does this thesis study the experience of being interrupted in mobile real-world settings (see 5.1)? This chapter outlined the methodology underlying the studies in this thesis, with a focus on the knitting together of mixed methods to understand the subjective, situated experience of being interrupted. At the heart of the mixed-method approach (see 5.2), *in situ* methods that promise high ecological validity are employed in the empirical work, with ethnographic observations (see 5.3) on the qualitative side, and ESM on the quantitative side (see 5.4). With an emphasis on experimentation "in the wild" this chapter outlined this thesis' contribution to studying interruption in situ: an ESM-adaptation to study both self-reported as well as behavioural response, with a temporal model of the interruption process in its centre (see 5.5). Finally, the application of the methodology to the empirical studies was contrasted and summarised (see 5.6).

The three chapters of the following Part III together comprise the account of the empirical work in this thesis.

Part III

EMPIRICAL WORK

6

EFFECTS OF CONTENT AND TIME OF DELIVERY ON RECEPTIVITY

This chapter presents a study completed during an internship at PARC to investigate effects of the content of interruptions (SMS messages) and of the time of day of interruption delivery on mobile phones. A naturalistic experiment using experience-sampling showed that content had a significant effect on receptivity to interruptions whereas participants' self-scheduled times of delivery did not significantly increase receptivity compared to a random baseline. Testing of underlying variables relating to perceived quality of content established that the factors *interest*, *entertainment*, *relevance* and *actionability* influence people's receptivity significantly.

6.1 INTRODUCTION

The absence of the consideration of relational contextual factors (see 4.4.2, p. 99) such as *sender* and *content* of interruptions in most of the interruption research has been noted previously (see 5.5.1). In addition, it has been argued, this absence is reflected in the experimental methods of related work: instead of sending genuine content most of the studies that investigate interruptibility interrupt their participants only with a request to confirm their interruptibility either on a scale (Hudson et al., 2003); or they ask how much time the interrupted would have for an interruption (Hudson et al., 2002); or they use generic types such as "how receptive are you to a 'phone call' or a 'reminder'?" (Ho and Intille, 2005). Whereas the sender is more difficult to control or manipulate for an experimental study, the potential of considering *genuine content* inspired the design of this study.

Although content has been acknowledged in related work as an important characteristic of the interruption (see 4.4.2.2), the reviewed work operationalises content somewhat narrowly as the relevance of the interruption content to the current task. However, this thesis argues that one can be receptive to an interruption even though the content is not relevant to the current task. For example, it was argued (see 4.4.2), if the content and the relationship to the sender justify the interruption; consider the message: "*It's a boy!*".

Typically, reviewed work operationalised both timing (see 4.4.2.5) and content (see 4.4.2.2) in relation to the primary task. In turn, their experiments inherit a requirement to control or monitor the primary task; thus calling for a controlled laboratory study, which makes it difficult to transfer the methods to a naturalistic mobile setting. However, this thesis chooses to study interruption in the real world, which means that control of the participant's primary task is sacrificed for ecological validity.

Thus, timing also needs to be studied without having to control the participant's primary task. Therefore, this study explores the impact of a simple operationalisation of timing. Namely, by letting people schedule the times for their own interruptions in advance. In the broader context of this thesis' focus on the adaptation of timing (see 4.5, p. 105), the experiment tests if McFarlane's *scheduled* coordination method is superior to a random baseline condition (McFarlane, 2002, see 3.6.3).

In the literature, no study was found that interrupted participants with real messages containing genuine content in mobile settings. Therefore, in this study participants are interrupted with real messages with genuine content on their mobile phones in order to examine how manipulation of interruption content and time of day of delivery affects their receptivity in situ.

6.2 RESEARCH QUESTIONS AND HYPOTHESES

Receptivity is studied in a naturalistic mobile setting where participants do not have to carry additional equipment. The experiment simply operationalises timing (Independent Variable (IV)1) as *time of day of delivery* (henceforth called *time of delivery* or *timing*). In more detail, this is tested by letting participants schedule the time of delivery in advance. In terms of the primary methods of coordinating interruptions (McFarlane, 2002, see section 3.6.3, p. 74) the experiment tests how the *scheduled* method compares to a random baseline condition. As a control measure, the participants' self-reports of *timeliness* at participant-selected *good times* for interruptions are compared with random *other times*, where times are specific clock-times, e.g. 11:15am.

In relation to the research questions introduced in 4.5.3, this experiment attempts to answer RQ1: Does scheduling one's interruptions lead to interruptions at opportune moments?

Furthermore, to evaluate the impact of a relational factor on receptivity this experiment explores the impact genuine content has on the receptivity, thus the research question that combines *timing* (IV1) and *content* (IV2) is:

RQ: How do the interruption content and the time of delivery impact receptivity to pushed information on mobile devices?

In more detail, what properties of the content make people more receptive to information pushed to their personal mobiles? In line with the relational contextual factors (see 4.4.2), rather than properties of the content per se, these are investigated as relations people have to the content; namely, how people relate to *what* is said, to *how* it is said, and to the *effect* of the message. The factors chosen to be studied as probable underlying factors of receptivity are interest in the content of a message (*what*) and *actionability* of a message (*effect*) in addition to the factors *relevance* and *entertainment* (*how*) identified by previous research as outlined above (Czerwinski et al., 2000; Shen and Chen, 2008; Tsang et al., 2004; Vastenburg et al., 2004; Xia and Sudharshan, 2002; Kalyanaramen et al., 2005, see 4.4.2.2). *Actionability* is understood to be the likelihood of influencing the recipient's future actions. For example, a message apprasing a recent film or album release may prompt the recipient to look up show times for the local cinema, or go to their local record shop to listen to the album.

The identified four factors are hypothesised to be strong indicators of receptivity. The study therefore builds on the following two major hypotheses:

h1: Interest in content is a predictor of receptivity: People are more receptive to "good content" than to "bad content", based on their individual interest ratings.

h2: Interest in content, perceived entertainment value, relevance and actionability of a message's content influence people's receptivity.

6.3 experiment design

In a 2x2 within-subjects design, *type of content* (good and bad content) and *time of delivery* (good and other time) were manipulated. Both the ESM (see 5.4, p. 121) and a diary method were employed over a period of 10 working days. Six text messages were delivered each day to the participant's mobile phone. Participants were instructed to rate their receptivity to the content by replying with a rating as soon as possible after they noticed the message.

As the levels of the variables "good" and "bad" are subjective, it is important to stress that the messages were different for each participant. The messages had in common only that their respective content categories were rated highly (for "good" content) or lowly (for "bad" content) by the participants in the setup interview.

The daily survey was set up to unpack what factors of the content make people more receptive, i.e. it looked at relational qualities of the content. The initial hypothesis H2 stated that *interest*, *perceived entertainment*, *relevance* and *actionability* all influence the receptivity. Are the receptivity ratings correlated with how interesting and entertaining they find the content, and with how relevant and actionable a message is perceived to be?

The fact that people's daily information consumption routines differ and the fact that repeated measures were collected requires a within-subjects analysis approach of the data.

Figure 6.1 gives an overview of the experimental design in relation to the model of the contextual factors of the interruption (see figure 4.4, p. 88). Two relational factors, *timing* (IV1) and *content* (IV2) are experimentally manipulated in a 2x2 factorial design. In situ, the experiment measures the impact of that manipulation on the operationalisation of receptivity employed here; namely, by participants' self-reports of *gut reaction* and *timeliness* (explained below), and by measuring response time. Furthermore, participants complete a daily *diary* for analysis of four specific factors assumed to qualify the receptivity to content specifically; namely, participants' *interest* in the content and their perception of the content's *entertainment*, *relevance* and *actionability*. Other relational factors are held constant by experimental design, as other local contextual factors are not in the focus of this experiment, while still assumed to interact to influence receptivity (see figure 6.1).

The next two sections outline how pre-study interviews were used to compile a list of content categories to provide the content for the

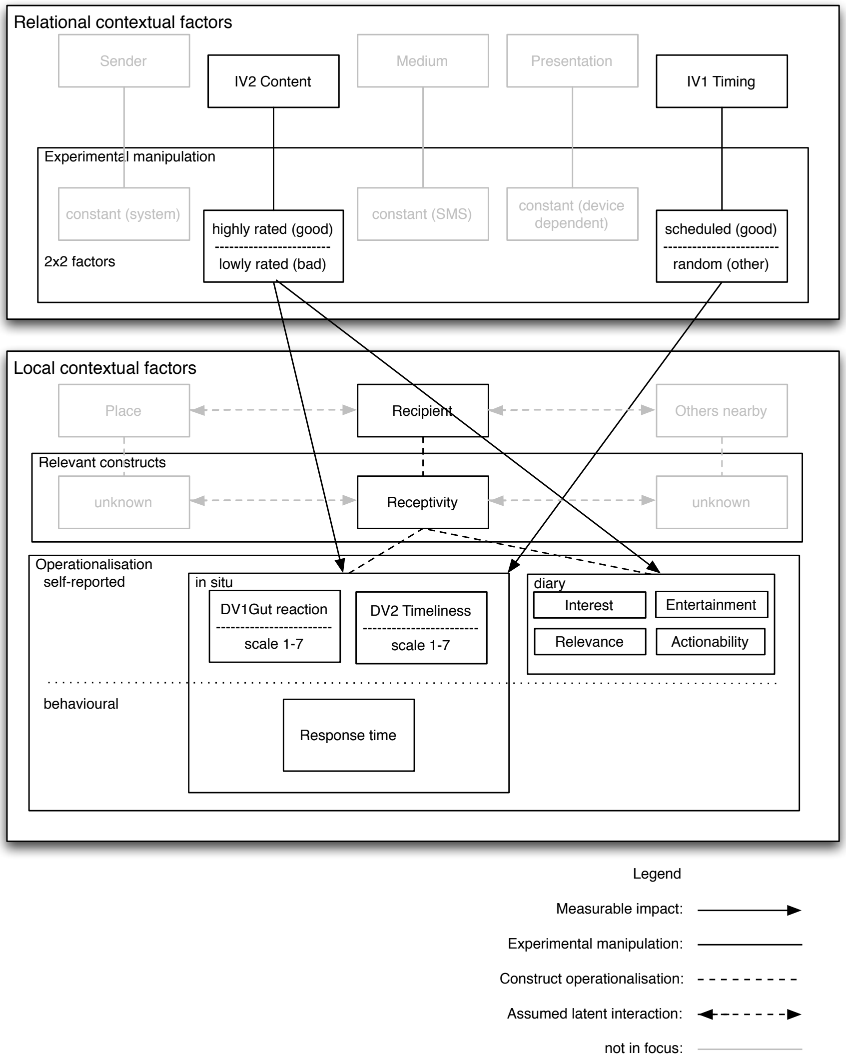


Figure 6.1: Overview of the experimental design of Experiment 1 in relation to the model of contextual factors of the interruption.

experiment (see 6.3.1), and how participants rated their interest in these categories in post-experiment questionnaires to inform a per-participant manipulation of *content* (see 6.3.2). The rationale of the operationalisation of receptivity is explained subsequently to that (see 6.3.3).

6.3.1 *Compiling content categories from pre-study interviews*

The goal for the content categories presented to the participants was for all of the participants to be interested in some and not interested in other categories. To arrive at the content categories interviews were conducted to study the everyday information consumption behaviour of the sample.

In a set-up study, informal interviews with 10 co-workers elicited their information consumption habits. They were asked to itemise what categories of information they consumed frequently, if it was on a regular or on an irregular basis and if they consumed the information out of general interest or out of a situational need. As information categories consumed out of general interest, they named news, sports, hobby/leisure-related and work-related information ("to stay on top of what's happening"[1]). Information tied to the co-workers' situation was named more often and the category terms seemed more specific, e.g. information on traffic, weather, movies, products and places to go. Even though the categories were diverse, what they shared was that they had an immediate impact on the interviewee's situation and may have influenced subsequent actions. For example, weather reports were followed to inform what to pack for a future travel; movie reviews were read to see if there was anything worth watching and eventually leisure plans for the weekend may have been influenced by this; product reviews and prices were compared to inform a buying decision.

On a broad level, the set-up study identified two modes of information consumption. The first is general interest or serendipity in information consumption, having no immediate impact on activities. The second is situationally directed information use that explicitly

1 All quotes by participants unless stated otherwise

informs future actions. A more detailed explication of information needs has been presented by Sohn et al. (2008).

Since it would be an ambitious technical project in its own right to predict and cater to situational information needs, limited systems development resources restricted the study to general interest information categories which were easier to serve to volunteers for the study.

6.3.2 *Generating good and bad content for participants*

Individual diversity in the ways people relate to information forbids sending the same type of information to every participant. In order to compare the relational impact of the factor *content*, a constant across participants' relation to the content was needed. A straightforward way to introduce a constant was to ask participants how interested they were in the particular content category, and thereby inform what type of information they would get as *good* and as *bad content* throughout the study.

Based on the findings from the set-up study interviews balanced with availability of Really Simple Syndication (RSS) feeds on the web, a fairly extensive list of 28 general content categories was developed. The list was offered to the main study participants to allow them to express both interest and disinterest in different categories (see content categories in figure 6.2).

11 co-workers were recruited as participants (3 female and 8 male) for the main study. None of them were involved in or knew details about the study. As part of the setup process, participants were asked to rate their interest in the 28 content categories[2] provided on a scale from 1 (not at all interested) to 7 (very interested) (see figure 6.2). For each participant, 3 categories they rated highly (6-7) were chosen to

2 The questionnaire is supplemented in appendix A, p. 290.

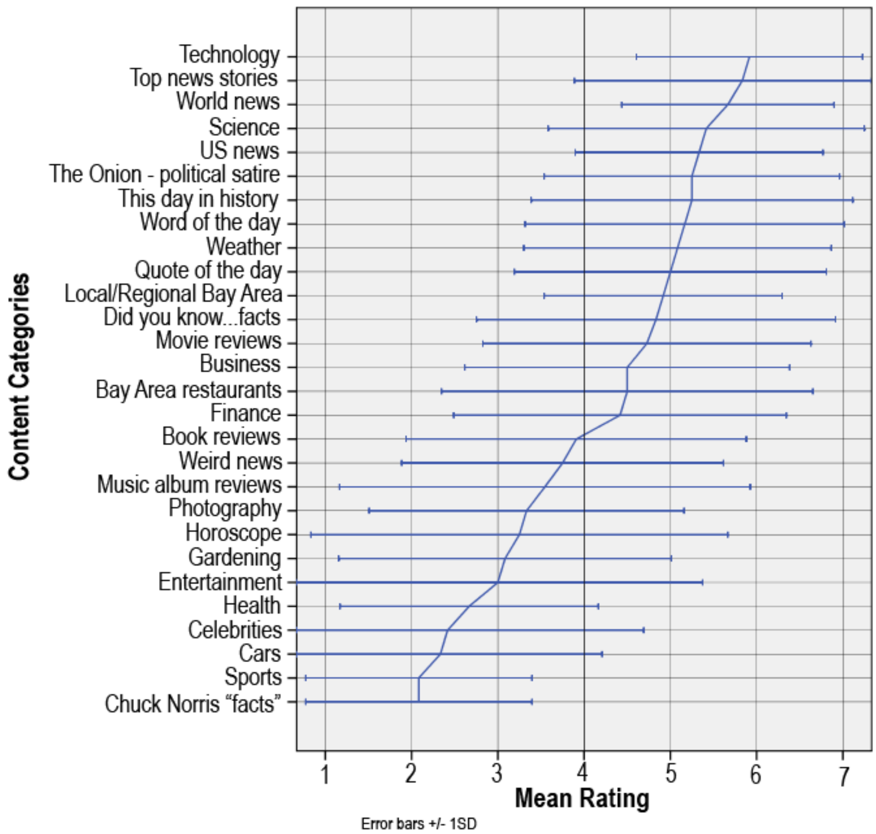


Figure 6.2: The sample's content preferences (mean +/- SD) as collected in the setup procedure.

provide the *good content*. Another 3 categories they rated lowly (1-2) were chosen to provide the *bad content*. Figure 6.2 shows the mean interest ratings of the 11 participants, with the content categories ordered according to descending interest rating.

In addition, they provided a general time window of 10hrs during the day where they would be willing to receive messages on their mobile phones. Within this window, they were specifically asked to state three times where they thought they would be receptive to messages. These 3 times would serve as *good times*. In addition, three *other times* were chosen, whereby an even distribution of messages throughout the specified 10hr window was attempted.

6.3.3 *Framing receptivity: gut reaction and perceived timeliness*

In the study, participants were asked to reply to the experiment's text messages as soon as possible with a rating of their receptivity (via a text message reply). The challenge was to instruct the participants in a way that the ratings reflected their perception of receptivity. Participants were asked to rate their *gut reaction* (Dependent Variable (DV) 1) to the message on a scale from 1 (low) to 7 (high), because the concept it refers to is both subjective and situated in the particular time it occurs and is tied to the circumstances of the situation. Participants were explicitly instructed to include such guidelines in the assessment of their gut reaction as the initial feeling towards the content of the message at that particular time; how the interruption of the incoming SMS felt; the particular circumstances of the environment such as appropriateness of location or who else is present.

In order to test whether participants' self-scheduled interruptions were actually perceived as more timely than interruptions at other times, they were also asked to provide a rating of the *perceived timeliness* (DV2) of the time of delivery on the same scale of 1 (low) to 7 (high). This dependent measure also serves as a manipulation control whether the participants' chosen good times actually represented good times.

6.3.4 *Procedure*

An SMS server was set up to deliver the text messages according to a daily schedule for each individual participant (see appendix A, p. 290 for a system diagram). When a time to send a message was detected in the daily schedule, a task was performed that looked up the RSS feed that provided the content according to the individual preferences for each participant. For each participant, three content categories provided the good content and three categories provided the bad content. They were randomly rotated in their assignment to the three good times they said they would be receptive and three other times, so that participants would not always get the same type of message at the same time of day. This rationale was informed by the goal to get a sufficient number of ratings for each of the possible 2x2 factor combinations for a within-subjects analysis. With six messages per day sent for 10 days and 4 different combinations of the factors, each participant would get an average of 15 messages per combination.

6.3.4.1 *Role*

Under supervision at PARC, the main responsibilities of the author in this experiment ranged from experimental design, including questionnaire and interview design, to recruiting and setting up participants, managing orchestration of the experiment, supporting participants, and interviewing them after the study. The system component that looked up message content from RSS feeds (see appendix A) was also written by the author.

6.3.5 *Daily survey*

So as not to overload participants by asking them to send a whole host of ratings by text, a web-based survey was set up to reveal the role of further underlying factors of content. At the end of each day, an online survey presented each of the six messages as a reminder. For each message, the participants were asked to provide a rating of their interest and entertainment value of the content on a scale of 1 (low) to 7 (high). Furthermore, they were asked to pick categories that best described the relevance and actionability of each message.

	Good time	Other time	Total
Good content	98 (23.4%)	107 (25.6%)	205 (49%)
Bad content	108 (25.8%)	105 (25.2%)	213 (51%)
Total	206 (49.2%)	212 (50.8%)	418 (100%)

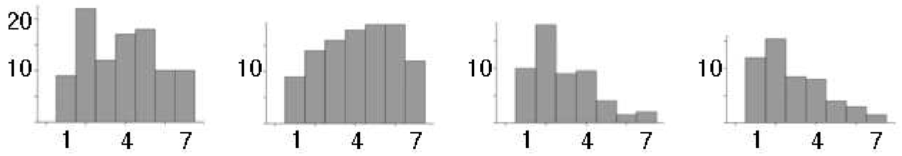
Table 6.1: Distribution of the independent variables across messages that received a response.

6.3.6 *Exit interviews*

Finally, participants were interviewed to learn about their experience during the study. Semi-closed or open ended questions were asked to get an overall picture of the experience, looking at change over time, and specific moments, e.g. particularly good or bad timing of messages. Also, it was asked whether and how the messages ever impacted the recipients' activity or were relevant to it, and their general receptivity to information pushed to their mobile phones were investigated among possible alternatives, e.g. "information pull". The full exit interview questions are listed in appendix A.

6.4 Analysis and results

The 11 participants received a total of 641 text messages (56 pp on average) over a period of 10 working days. The intended 60 messages per person were not quite achieved due to technical difficulties with message delivery. 418 (38 per person on average) responses in total were received (i.e. to 65.2% of the sent messages). Table 6.1 shows the distribution of the independent variables (IVs) content and time of delivery among the messages that received a response.



(a) good content, good time (b) good content, other time (c) bad content, good time (d) bad content, other time

Figure 6.3: Frequencies of *gut reaction* ratings by conditions of the independent variables.

6.4.1 *Gut reaction*

The mean gut reaction to all of the 418 messages sent was slightly below neutral with a mean rating of 3.42 on a 7-point Likert scale, the median was 3, the SD was 1.8.

In order to analyse whether the effect of the manipulation of *timing* (IV1) and *content* (IV2) yielded significantly different *gut reaction* (DV1) scores per participant, within-subjects (repeated measures) ANOVA was selected. To perform a within-subjects ANOVA the gut reaction scores were aggregated on the individual level by averaging per condition (see table 6.2 and figure 6.3 a-d); a common approach for dealing with unequal amounts of repeated measures per individual (Ho and Intille, 2005; Stone and Shiffman, 2002). The dataset that was derived in this manner comprised one averaged score of gut reaction for each of the four conditions per participant. Inspection of Q-Q plots showed that deviation from normality was tolerable and skewness (.253) showed that the distribution was still approximately symmetric, according to a widely used rule of thumb (Bulmer, 1979).

The within-subjects ANOVA showed that content had a significant effect on gut reaction, with F (1, 10) = 30.95, $p < .001$. Partial eta squared = .76, which is a large effect.

Pairwise comparison showed that the mean gut reaction for *good content* (3.96) was significantly higher than for *bad content* (2.86) ($p <$ .05, see fig. 6.4). The factor *time of delivery* (*good* vs. *other time*) did not

IV conditions	Mean	SD	N
Good content, good time	3.85	1.84	98
Good content, other time	4.21	1.82	107
Bad content, good time	2.85	1.55	108
Bad content, other time	2.84	1.62	105

Table 6.2: *Gut reaction* descriptives between subjects by conditions of the independent variables.

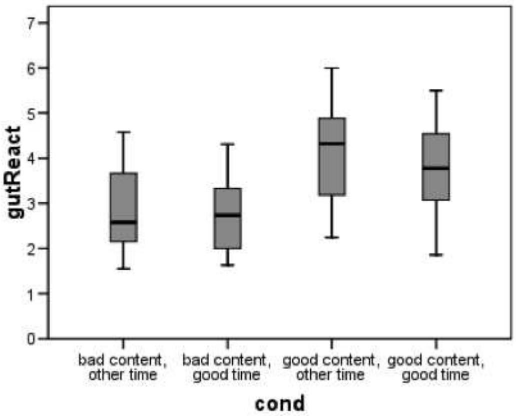


Figure 6.4: Boxplots of *gut reaction* by conditions of the factors *content* and *time of delivery*.

lead to significant differences in *gut reaction* (F (1, 10) = 2.8, *p* = .125); and there was no significant interaction of content and time of delivery (F (1, 10) = 1.93, *p* = .195). Table 6.2 and figure 6.4 show the means and SDs of *gut reaction* across the conditions.

The finding that content has a significant effect on people's gut reaction supports the hypothesis that people are more receptive to content they have expressed higher interest in (H1). Furthermore, the findings suggest that we can predict receptivity by studying individual information consumption behaviour, as the manipulation of the factor content was informed by the previously collected interest ratings in the content categories.

6.4.2 *Perceived timeliness*

The control measure *perceived timeliness* revealed that the perceived timeliness for *good times* was not significantly different than for *other times*. A within-subjects ANOVA done in a similar manner as for *gut reaction* above showed that neither *content* (F (1, 10) = .92, p = .36) nor *time of delivery* (F (1, 10) = 2.82, p = .12) had a significant effect on *perceived timeliness*. The interaction was also not significant, with F (1, 10) = .65, p = .44.

The fact that the perceived timeliness does not differ between self-scheduled times and randomly selected other times may mean that time of delivery in an asynchronous medium such as SMS is less important as messages can be tended to later. It is equally plausible that people are simply bad at predicting the times when they are receptive and therefore, that the levels "good" vs. "other" are not actually different from each other.

6.4.3 *Discussion with qualitative findings*

To review the statistical findings above, it was found that when pushing content to people's cell phones, their gut reaction to content they had previously expressed interest in was significantly higher than to bad content. The fact that participants' gut reaction to good content was significantly higher than to bad content shows that content matters.

Furthermore, the fact that significant effects were found for content and not for time of delivery indicates that interest in content is a more stable predictor of the gut reaction to a message than scheduling the time of delivery. As one participant so poignantly put it:

> Content really trumps the timing.

This is widely echoed throughout the interviews. When asked directly, 8 participants said that content was more important to the gut reaction than time of delivery; one said that both were important and two said that timing was more important – however one of these last two immediately qualified his statement in an interesting way:

> Timing first, content second, I mean, well, it's interesting...I never got a message while I was holding my phone, it was always while I was doing something else. So the first problem is, that I'm doing something else, that's why timing is the first part, and then I look at the message, and then *if the content is not important to me... it was not a good moment to interrupt me* with that. But I'm likely to be a little bit more forgiving if the content is interesting to me.

On the one hand, the participant suggests that the timing is more important than content; on the other hand, he mitigates his assertion by what he says in the highlighted area about his receptivity. This statement points to a notable quality: the assessment of the timing of an interruption is related to the content of the interruption and presumably other factors, supporting the notion that contextual factors in the model interact to determine the individual's context (see 4.3). Even if the timing is bad, to use the participant's words, we are likely to be more forgiving if the content adheres to certain qualities; in this case, interest. The following statement from another participant emphasises how timeliness of an interruption is assessed through its content.

> Sometimes it was OK to be interrupted, but not with that particular kind of snippet...depending on what the interruption is, it is a good time or it's not a good time.

Through statements such as these, it can be seen that the impact of an interruption is assessed ad hoc via its content. This echoes Dourish's insight that contextuality cannot be determined a priori; "It is an emergent feature of the interaction, determined in the moment and in the doing" (Dourish, 2004, p. 23). Research on interruption management has put a lot of effort into predicting the user's interruptibility by sensing their local context (see 3.4.2). The implication of the finding is that interruption management has to take into account the relational factor content to assess receptivity to the interruption.

Since all participants provided supposed good times to interrupt, the nonexistent statistical effect of time of delivery on the participant's perceived timeliness suggests that people are bad at predicting the times at which they will be receptive to messages and shows that scheduling time of delivery of one's messages does not lead to being more receptive – not that timing doesn't matter. In the exit interview, when asked about the times he had said he would be receptive, one participant replied:

> [...] Those are basically times where I thought I wouldn't be too engaged in things – it turns out that most of the times I actually was engaged with things.

The participants picked these times without complaining about the difficulty of this task and with some care, which can be seen by the fact that, at the end of the study, only one participant didn't remember the times he said he would be receptive to messages. Only 5 participants said in the interviews that the majority of messages were delivered at a bad time (as opposed to 9 that disliked the content of the majority of the messages). This further supports the finding that in the context of this study participants developed stronger views on the content than on the timing:

> When the content was bad, I felt annoyed, when the timing was bad, I just let it go.

This "letting it go" in this participant's statement points to a key quality of interruption specific to the SMS medium. This is that, perhaps unlike some modes of communication, receipt of a message on one's mobile phone can be ignored. It can be equally well tended to later – at a convenient time. The *medium of the interruption* (e.g. *asynchronous* versus *synchronous*) has been acknowledged as one of the contextual factors that interact to determine receptivity (see 4.4.2.4). Asynchronous modes of delivery such as email, SMS or IM do not demand attention in such an immediate way. They even afford *plausible deniability* (Nardi et al., 2000), in that it is socially acceptable not to respond immediately.

Presumably, the cost of ignoring an incoming SMS was less as its content can be tended to at a convenient later moment whereas the content of a missed phone call may be lost indefinitely. Hence, the timing of interruptions in an asynchronous channel may be evaluated more liberally than in a synchronous channel. The asynchronous nature of SMS effectively enables the individual to tend to the interruption at a convenient moment – even if it is just seconds after the interruption was noticed, it effectively becomes a self-selected attention shift rather than a forced attention draw.

Furthermore, that interruptions can be ignored when focussing on the task at hand has been acknowledged as a property of being engaged in *task-driven processing* (Miyata and Norman, 1986). And with SMS content delivery, the majority of the participants appeared to be able to stay in that mode with little disruption from an incoming item as was echoed in numerous quotes:

> Timing...I mean, if it's a little bit off, it's OK, because I can always push it off 'til later. The thing buzzing or beeping once doesn't bother me too much.

> When I wasn't available I could just not look at it. And that's fine.

> There were plenty of times, where the timing was not good, but I just ignored it, it's not like I really disliked it. The thing came in...the timing was not good, so I didn't look at it.

> Timing wise... you don't have to look at it immediately, you can go back to it later. Phone allows that, so timing is probably less important than interest. I like the flexibility of that, text messages come in, they are lightweight, you don't have to respond. If you do have to respond, the person will probably phone you.

The quality of SMS communication this person describes as "flexible" and "lightweight" in that "you don't have to respond" alludes to the existence of plausible deniability (Nardi et al., 2000) in SMS communication.

6.5 further analysis of factor content

Now, results from the web-based daily survey are presented and discussed. In order to reveal possible underlying factors of receptivity, participants were asked to rate their interest in the content, how entertaining they thought the content was, and the relevance and actionability of the content for each message.

Participants completed the daily survey for only 268 messages they responded to (out of 418 in total). A reason may have been that the web-based survey could only be accessed for the current and the previous day. Despite the fact that a reminder was sent out every evening surveys were often forgotten, and each omission meant that 6 messages at once were without survey responses.

6.5.1 *Content factors*

In the following, factors anticipated to contribute to the perception of the content are analysed in more detail.

6.5.1.1 *Interest*

The median *interest* rating was 3 (M=3.24) on a 7-point Likert scale. A within-subjects ANOVA showed that participants were, unsurprisingly, significantly more interested in *good content* than in *bad content*, with F (1, 10) = 21.09, p = .001. Interest ratings for *good content* (M=3.97) were significantly higher than for *bad content* (M=2.53). This finding supports the hypothesis that interest is a predictor of receptivity (H1). Also, this strengthens the validity of the factor *content*: content in the *good content* category really was significantly more interesting than content in the *bad content* category. Again, *time of delivery* did not have a significant effect (F (1, 10) = .083, p = .779), nor did the interaction of *content* and *time of delivery*, with F = .115, p = .742.

6.5.1.2 *Entertainment*

The median for *entertainment* was 2 (M=2.41) on a 7-point Likert scale. Even though *good content* (M=2.74) was rated slightly more entertaining than *bad content* (M=2.09), a within-subjects ANOVA showed that the differences were not significant, with F (1, 10) = 3.39, p = .095. *Time of delivery* was not a significant factor, nor was the two factors' interaction.

6.5.1.3 *Relevance*

The majority of messages rated (66.4%) was not relevant to the participants, 29.5% of the messages were rated as being generally relevant regardless of the time read, and 4.1% of the messages were rated as being relevant at the time read or soon after.

6.5.1.4 *Actionability*

Actionability referred to the likelihood that the receiver of the message would subsequently take action or alter his or her actions in some way informed by the content of the message. Actionability is an ordinal variable. In the daily survey, the participants were asked: To what extent did/will the content of the message change/influence your actions?

Participants rated an overwhelming majority of 75% of the messages as unlikely to influence their future actions, and just 19.4% to be somewhat likely to influence future actions. Only 4.5% of the messages were rated as being very likely to influence future actions, while only 3 messages (1.1%) were rated to have already influenced their actions.

6.5.2 *Underlying factors of receptivity*

It was already shown that individually perceived good content fosters significantly higher receptivity than does bad content. But what makes good content good? To shed light on this, the interrelation of the factors measured in the daily survey and the participant's gut reaction was studied. How much do people's interest in the content, its perceived entertainment value, the relevance and actionability of the content influence people's receptivity?

A word of caution is probably appropriate here. The reader should keep in mind that the content factor ratings were collected in a different context (web-based, at the end of the day) than the gut reaction ratings that form the main part of the study (mobile, throughout the day). Furthermore, the chosen factors may not conclusively explain the latent factors that underlie receptivity.

Factor	Levels	Mean gut reaction	SD	N	$\chi^2(2)$	*Exact p*	Kendall's W
Interest	Low (1-2)	$2.51^{2,3}$	1.09 (1.2)	11 (138)	18.7	<.001	.851
	Medium (3-5)	$3.87^{1,3}$	0.82 (1.1)	11 (113)			
	High (6-7)	$6.04^{1,2}$	0.57 (0.78)	11 (48)			
Entertainment	Low (1-2)	$2.97^{2,3}$	1.09 (1.6)	11 (190)	16.2	<.001	.738
	Medium (3-5)	$3.93^{1,3}$	1.44 (1.4)	11 (87)			
	High (6-7)	$6.28^{1,2}$	0.27 (0.86)	11 (22)			
Relevance	Not relevant	$2.30^{2,3}$	1.05 (1.6)	11 (199)	17.6	<.001	.802
	Currently rel.	4.17^{1}	0.34 (.81)	11 (12)			
	Generally rel.	4.83^{1}	1.54 (1.7)	11 (88)			
Actionability	Unlikely	$2.87^{2,3}$	1.02 (1.6)	11 (228)	19.6	<.001	.89
	Somewhat likely	$4.44^{1,3}$	1.69 (1.8)	11 (56)			
	Very likely	$6.04^{1,2}$	0.64 (1.5)	11 (12)			

[1] The mean difference to the first level is significant at the .05 level.

[2] The mean difference to the second level is significant at the .05 level.

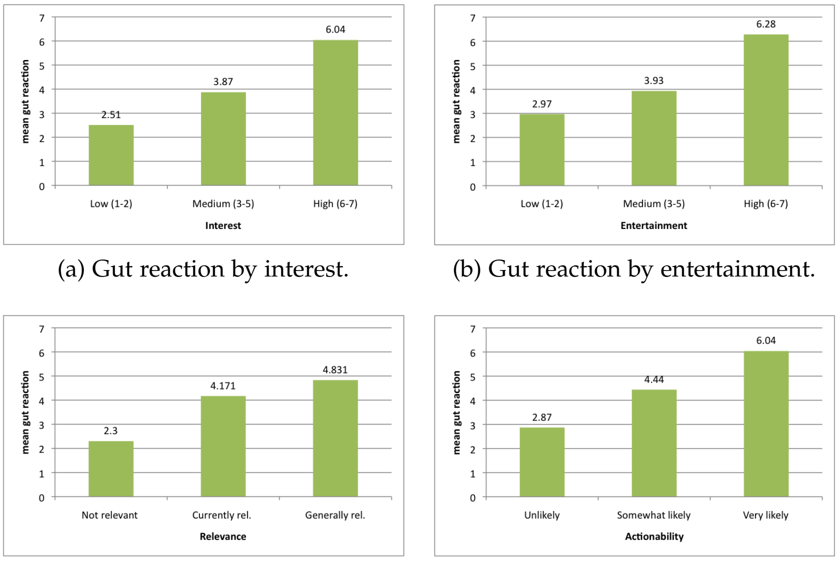
[3] The mean difference to the third level is significant at the .05 level.

Table 6.3: Mean gut reaction scores and within-subject effects by factors. Values before averaging shown in parentheses.

6.5.2.1 *Unpacking the factors*

To unpack the interrelations, the gut reaction for each participant was averaged per each level of the content factors. In order to be able to compare the effect size of each factor, *interest* and *entertainment* was collapsed from 7 into 3 levels. The binning introduced unequal variance between the levels (see the differences in the SD-column in table 6.3), so nonparametric Friedman tests for ordinal repeated measures (*k-related* samples) suitable for non-normally distributed populations were conducted for the four factors. To analyse the differences among the factor categories, pair-wise comparisons were conducted by means of a Wilcoxon test for 2 related samples. Kendall's W is reported as a measure of effect size.

The Friedman tests showed that differences in *gut reaction* across different levels of the four factors were all significant at the .01 level. Furthermore, pair-wise comparison indicates that the mean *gut reaction* increased along with rising factor levels (see figure 6.5 a-d). For the



(a) Gut reaction by interest.

(b) Gut reaction by entertainment.

(c) Gut reaction by relevance.

(d) Gut reaction by actionability.

Figure 6.5: Mean gut reaction by *interest*, *entertainment*, *relevance* and *actionability* of content

ordinal variables *interest*, *entertainment* and *actionability*, the pair wise comparison shows that the mean *gut reaction* differs significantly for each level at the .05 level. Table 6.3 depicts the mean *gut reaction* per level of factor, including pair-wise comparisons and effect size (Kendall's W).

6.5.2.2 *Discussion*

The analysis supports the hypothesis (H2) that *interest*, *entertainment*, *relevance* and *actionability* of the content are all underlying factors of receptivity. This finding implies that qualities of the content influence receptivity in a substantial way and need consideration when studying interruptibility.

Interest was found to be a strong factor to influence receptivity; as hypothesised, it suggests itself as a predictor of receptivity. This is also supported by the many statements in the interviews indicating how "interesting" content lead to a high rating of gut reaction (see 6.4.3, above).

The fact that content did not have a significant effect on the participants' rating of entertainment (see 6.5.1.2) showed that the content used was not well distributed on the scale of entertainment, i.e. just not that entertaining. However, it proved to have a significant effect on the participant's gut reaction, indicating that the entertainment factor of content does play a significant role in the receptivity to the content, as suggested by earlier work (Tsang et al., 2004; Shen and Chen, 2008).

Only 79 out of 268 messages (29.5%) were rated as relevant (of which only 55 were classified as good content). Note that the average *gut reaction* for messages that people considered being *generally relevant* (4.83) is higher than their average gut reaction to *good content* (4.02). Therefore, the provision of content from categories people said they would be interested in does not necessarily imply that the gut reaction is always good, but if you manage to say something relevant, chances are higher that people will be more receptive, i.e. have a better gut reaction. In the interviews, some views about receiving irrelevant messages emerged:

> I found it annoying to get messages that were irrelevant.

> ...content that was irrelevant to me I have no interest in getting pushed to me.

This is complementary to the finding that relevant interruptions are less disruptive (Cutrell et al., 2000).

Regarding the actionability of content, the fact that the majority of messages were classified as unlikely to influence future actions indicates a low signal-to-noise ratio of the content. So, while the data tentatively suggests that achieving high actionability may be a good predictor of receptivity (it had the largest effect), the chances of delivering a message with likelihood to influence the recipient's actions are low, at least with general content categories such as the ones used in this study.

Anecdotes from the interviews exemplify situations where the participants' activities actually were influenced by a message. Messages that influenced the person's activities often coincided with high gut reaction ratings of the messages. As a result of a message, one participant shared the content with his friends and thus used it as a "conversation starter". Two mentioned the weather report in this context: "I think the weather thing let me form my plans differently a couple of times" and "Influenced in a way that it confirmed that I could do something, the weather ones". Most commonly, five participants reported to have looked up the content of the news story, a movie or a music review online as a result of the message.

6.6 DISCUSSION

6.6.1 *Limitations*

The experience-sampling method is challenging in that the researcher has to deal with an incomplete description of the experience. Almost never can a response rate of 100% be achieved.

6.6.1.1 *Response rate bias*

The problem of incompleteness is intensified by the fact that timing is investigated. A critique of the design may be that the factor *time of delivery* is biased, as it may well be the case that non-responses are simply a result of bad times of delivery. Thus, the distribution of the independent variables among the 223 messages that were not responded upon was looked at. An almost equal amount of non-responded messages was sent at good times (112) and at *other times* (111). A chi-square goodness-of-fit test showed that the distribution of the frequencies was not biased towards *good* or *other times*, with $\chi^2(1) = .004$, *exact* $p = 1.0$. This shows that non-responded messages (i.e.,

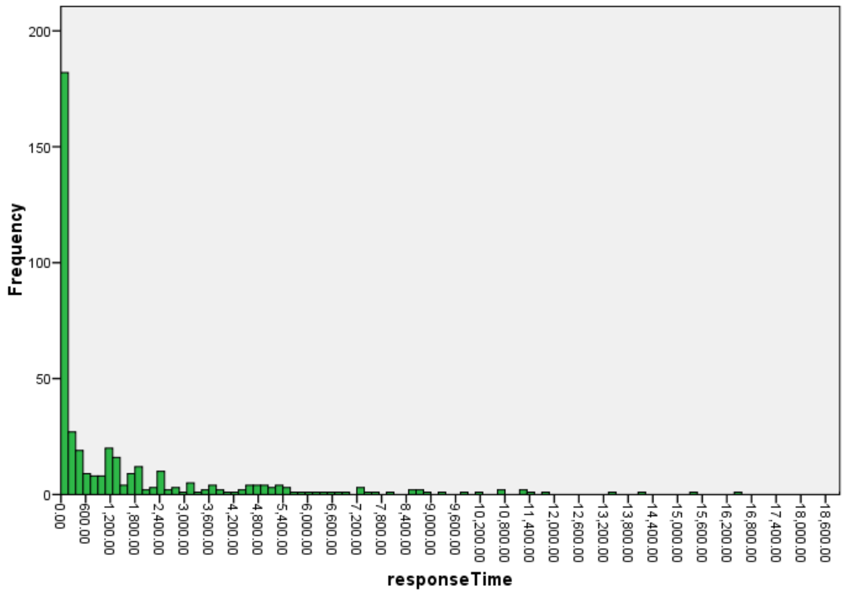


Figure 6.6: The distribution of response time in seconds (interval = 180s).

not replied to) were not systematically biased toward bad times of delivery.

Likewise, the distribution of non-responded messages between *good content* (117) and *bad content* (106) is not significant; A chi-square goodness-of-fit test showed no significant bias towards *good* or *bad content*, with $\chi^2(1) = .543$, *exact* $p = .503$.

6.6.1.2 *Response time bias*

The ESM is particularly designed to be an ad hoc method by which the quality of an experience is assessed in the moment. The time participants took to respond to the prompts was analysed to elicit if the reaction of the participants actually happened in situ.

The distribution of *response time* (see fig. 6.6) was encouraging in terms of timeliness of response: 50% of the messages were responded to within 373 seconds (6.21 mins), 76.1% were answered within 40 minutes after reception, and still 90% of the messages got responded to within 2 hrs. A within-subjects ANOVA did not yield any significant differences in average log-normalised response time for *perceived* *timeliness* (F (6, 12) = .31, p = .92) and *gut reaction* (F (6, 6) = 2.56, p = .16).

Different practices with data exceeding a certain response time are common; a lag of 30 minutes has been suggested as a boundary for valid data (see 5.4.3.2, p. 126). Filtering out all data points that exceed a response time of 30 minutes, the results do not change in a significant way. A within-subjects analysis on the basis of the 298 messages that were responded to in less than 30 minutes does not change the fact that *content* is a significant factor for *gut reaction* (F (1, 10) = 21.02, p = .001) and *time of delivery* is not (F = .01, p = .915).

6.6.2 *Push vs. pull*

Even though a comparison of modes of delivery as in "push vs. pull" was not the focus of this study, questions about the general receptivity to pushed information in the interviews inevitably led to the discussion of alternatives.

Also, the purpose of this study was not to evaluate an "information push system", as content participants were not interested in was deliberately delivered – obviously a property strongly discouraged in the design of a "real" system. Therefore, participants' negative reactions towards getting content pushed to their mobile phones did not come as a surprise.

Two participants developed strong aversion against getting content pushed to them; one even said that she "...grew to really dislike hearing my phone tell me a text came in". The other one said that he realized within three messages he disliked it: "the phone is only for keeping in touch with my own, close personal network". This supports the notion that the delivery of content to mobiles might potentially be perceived as very intrusive, breaching the owner's privacy (Lei et al., 2007). Interestingly, the other person with strong negative feelings also mentions her close social network when describing her disappointment:

> I was just not excited about receiving the texts. I get excited from texts from friends and family, but not these.

It seems that the expectation that a device is for the sole purpose of communication with a close personal circle influences the acceptance of it as a provider of other services. Only three participants mentioned that they prefer to get content pushed, as opposed to pull it themselves on demand. Six out of the 11 participants prefer to pull content, and frequently use explanations such as that they favour being in control of what kind of content they consume at what time.

However, they often qualified under which conditions information push would be acceptable. Information push seems to have its merits too, one participant said:

> I like it pushed, because it frees myself from having to remember to go pull for it.

Example content that participants see fit to be pushed to them includes reminders, event-driven news, information with high priority or urgency, and alerts of unpredictable events such as a rain. A sense of control seems important here as well, examples given are either on the side of explicit user feedback, such as mechanisms for the user to subscribe to certain content or signal unavailability, or on the implicit system side, such as algorithms that learn the user's content preferences. Asked if he prefers information push or pull one participant said:

> Pull it, absolutely. Although, if it says, here is something new for you, pick it up when you're ready, that would be something different than getting a text message with an alert.

This participant's statement points out that information push is acceptable under the condition that its delivery is not interruptive and that the content can be accessed on demand. This is certainly not a new paradigm of information delivery and has for instance been realized in subscribe and consume-on-demand RSS technology. However, it is interesting that many of the participants describe this kind of system as the ideal solution after taking part in the study.

All in all, the experience with an information push system from the study did not put the participants off the idea of information push in general, quite to our surprise. Instead, it seemed to lead them to conclude that an ideal system would be a hybrid between push and pull, a system that proactively provides smart and individually pre-filtered content unobtrusively and lets users consume content on demand at their own pace.

6.6.3 *Receptivity vs. interruptibility*

The concepts of receptivity and interruptibility have often been treated synonymously in related work. However, this thesis has already argued for a subtle but important difference (see 4.1). Interruptibility is the more prevalent concept that seems to be mostly used to describe a person's general state of receptiveness to any form of interruption (Avrahami, 2007), often also with the intention to predict an opportune moment when to interrupt someone (Avrahami et al., 2007a; Ho and Intille, 2005; Hudson et al., 2003). Hence, it is a concept that focuses on helping the sender of a message by reasoning about the receiver's local contextual factors, such as current activity.

It was argued that receptivity adds an important quality to the study of the interruption process: it places the anticipation of the receiver's experience of the interruption into the focus of attention, i.e. it is inherently a concept related to the relational contextual factors (see

4.4.2). It caters for the subjective viewpoint from which interruptions are studied in this thesis, while respecting the situated nature of the interruption process. In terms of communication models, it places the *recipient* into the focus of attention. Choosing the terminology of studying *receptivity* rather than *interruptibility* then may at least serve the purpose of avoiding the simplistic notion that *interruptibility* can be simply measured from an objective stance. *Receptivity* accounts for the moment-by-moment negotiation of context as an interactional achievement (Dourish, 2004) and the inherent difficulty to predict opportune moments for interruptions.

Note that the goals of the two approaches remain the same: to ease some of the burden of inevitable interruptions. The concept of receptivity is preferred, as it is more user-centred in that it includes and anticipates the subjective and inherently uncertain and hard to predict experience of and reaction to the interruption.

6.7 conclusions

The results from the study show that the content of a message plays an important role in influencing people's receptivity. The participants' gut reaction to good content was significantly better than to bad content. Also, the results show that interest, entertainment, relevance and actionability of the content all have a significant effect on the gut reaction, giving some evidence that they are important informational qualities that influence people's receptivity.

6.7.1 *Content trumps time of delivery*

The main finding from the study is that in the context of the asynchronous delivery method SMS, interruption content trumps self-scheduling their times of delivery. This finding is arrived at mainly by contextualising the results from the statistical analysis with the participant's qualitative responses from the interviews. Furthermore, people's comments from the exit interviews suggest that content interacts with other contextual factors to influence people's gut reaction. It appears that the content of an interruption may also influence people's perceived timeliness. This finding may inform any endeavour of understanding interruption management and building context-aware systems. Interruptibility is not a fixed property of time. While scheduling one's interruptions was shown to be an insufficient method to coordinate interruptions, the findings support the veracity of the model of contextual factors. Receptivity to an interruption is influenced by local and relational contextual factors including, but not limited to, the recipient's relation to the content of the interruption.

Findings from further analysis of content imply that attention other properties of the content of an interruption should be taken into account when trying to assess someone's interruptibility in addition to other local and relational context factors. For example, we have to consider the following kinds of questions: Is the content interesting to the interrupted? Is it relevant or actionable to the interrupted? And is it urgent or does it have a high priority for the interrupted?

6.7.2 *Dealing with timing*

However, by no means does the finding that content trumps time of delivery imply that the timing of an interruption can be neglected. The fact that time of delivery was not a significant factor shows that scheduling is an insufficient coordination method to predict times one is receptive, not that timing in general doesn't matter. This is supported by the fact that the control measure *perceived timeliness* showed that participants' own interruption schedules did not validly represent good times for people, they were just other times. Experiment 2 addresses the problem of finding ways to manipulate the timing of interruptions in mobile settings in relation to the recipient's activity, in order to accommodate for the situated uncertainty of timing (see next chapter).

Another reason why in this study the time of delivery did not have an impact on the gut reaction may be that most of the participants seemed to be well-trained in ignoring the interruption caused by the incoming message. People are accustomed to dealing with interruptions in their everyday lives; they have developed methods to cope with them (Tolmie et al., 2008). An incoming message is not necessarily perceived as an interruption; the disruptiveness of an interruption is determined not only by the timing of a signal, the pressure to respond seemed to be more interruptive. Also, the medium of delivery plays an important role. Whereas the disruptiveness of a synchronous medium such as an incoming phone call is undoubted, SMS is an asynchronous medium, where messages can be tended to later with little or no decrement in convenience.

Nevertheless, the qualitative responses suggest that the timing of an interruption does influence people's receptivity. Good timing is more variable and complex to predict as it is mediated by the local context of the recipients, such as their situated engagement in ongoing activities and their social surroundings. To develop an understanding of this immediate context, it becomes necesary to observe the recipient in their social and physical environment as undertaken in the field study (see chapter 8). Conversely, properties of the content as discussed in this chapter may be more stable and straightforward predictors of a person's receptivity.

6.8 summary and outlook

In the broader context of this thesis, this experiment highlights the importance of the relational factor *content*. Further analysis also showed that content factors (interest, entertainment, relevance and actionability) underlie receptivity and their consideration qualify and increase the likelihood of higher receptivity.

On the other hand, the experiment also reveals some of the complexity involved with prediciting opportune moments for interruptions. Simply scheduling one's own interruptions then, does not guarantee that one will be more receptive to these interruptions. More hypothetically, considering a coordination method that simply *schedules* interruptions from our personal communication network would mean that urgent, time-critical or vital information may reach us too late. All in all, this seems to render scheduling interruptions as an impractical coordination method. The following chapter will explore an implementation of a more sophisticated technique of McFarlane (2002); namely, the *mediated* coordination method, which is implemented so as to predict opportune moments for interruptions by sensing participants' phone activity.

Methodologically, this experiment shows some of the potential of conducting studies of interruptions in real-world settings. A lesson learned from this study is that simply measuring one behavioural temporal metric (response time) in situ does not allow for understanding the interruption process in the desired detail – there are simply too many unknowns: When did the recipient actually look at the message? How long did it take them to respond? Follow-up discussions with supervisors and colleagues identified that more fine-grained temporal metrics such as acceptance time would enable more detailed analysis in future studies. The design of the following Experiment 2 responds to this issue by employing all of the temporal metrics developed in the temporal model (see 3.3.5).

7

INVESTIGATING EPISODES OF MOBILE PHONE ACTIVITY AS INDICATORS OF OPPORTUNE MOMENTS FOR INTERRUPTIONS

This chapter presents a study that investigates whether opportune moments to deliver notifications surface at the endings of episodes of mobile interaction (making voice calls or receiving SMS) based on the assumption that the endings collocate with naturally occurring breakpoint in the user's primary task. Findings from a naturalistic experiment show that interruptions (notifications) are attended to and dealt with significantly more quickly after a user has finished an episode of mobile interaction compared to a random baseline condition, supporting the potential utility of this notification strategy. It is also discovered that the workload and situational appropriateness of the interruption task significantly affect subsequent delay and completion rate of the tasks. In situ self-reports and interviews reveal complexities in the subjective experience of the interruption which suggest that a more nuanced classification of the particular call or SMS and its relationship to the primary task(s) would be desirable.

7.1 INTRODUCTION

It has been argued that three out of the four primary methods to coordinate interruptions (McFarlane, 2002) were related to the timing of the interruption (see 4.4.2.5, p. 103), and the previous Experiment 1 has shown that the method of *scheduled* interruptions is insufficient to deliver more interruptions at opportune moments, i.e. when the user is receptive to them. The study presented in this chapter explores in how far the user's phone activity gives rise to opportune moments for interruptions that may be leveraged by a *mediated* coordination method (McFarlane, 2002, see 3.6.3, p. 74), i.e. to contextually *mediate* the timing of interruption by user activity. Hence, whereas the previous Experiment 1 employed the *scheduled* coordination method of interruptions, this experiment explores the implications of interruptions being *mediated* by the user's activity.

As reviewed in section 3.4 (p. 57) an influential body of work that has largely verified that opportune moments for interruptions are located at naturally occurring breakpoints in the primary cognitive task. Their findings are typically based on laboratory studies, where the detection of breakpoints in the primary task has been shown to approximate opportune moments (Iqbal et al., 2005; Iqbal and Bailey, 2007; Adamczyk and Bailey, 2004). Previous work has shown that the episodic nature of the human everyday experience (Zacks et al., 2001) provides opportune moments for interruptions (Adamczyk and Bailey, 2004), and that transitions between physical activities are indicative of such breakpoints in mobile experience (Ho and Intille, 2005).

However, it is in the relative chaos of everyday activity where such opportune moments must be routinely identified if this concept is to be applied in practical systems. The temporal and spatial mobility of mobile device users, the large range of possible egocentric mobile device positions (e.g. in hand, in pocket, in bag, on desk etc.), and a desire to avoid invasive (e.g. body-worn) sensors make it extremely difficult to control or observe their primary task. This leaves an unanswered need to identify opportune moments routinely.

Inspired by the work on breakpoints, this chapter explores the hypothesis that episodes of mobile phone activity indicate opportune moments for interruptions as the attention shifts to the mobile interaction episode at the beginning and away from it at the end in an ecologically valid setting. An opportunistic notification delivery mechanism using a *defer-to-breakpoint* interruption management strategy (Miyata and Norman, 1986; McFarlane, 2002; Iqbal and Bailey, 2007, see 3.6.3, p. 74) would then defer interruptions until the end of an episode of mobile interaction, which might provide an opportune moment before the user attention shifts away from the device.

7.2 RESEARCH QUESTION AND HYPOTHESES

To test the effects of interrupting after representative episodes of mobile interaction, the primary tasks of calling and reading SMS were chosen to test our hypotheses, because they are arguably among the most common examples of episodes of mobile interaction. This approach arguably provides a more ecologically valid alternative to using bodily worn sensors (e.g. Ho and Intille, 2005). The principal research question (RQ2 in the overall thesis) of this study is:

RQ: Does the end of an episode of mobile interaction represent an opportune moment for an interruption?

To answer the research question, a naturalistic experiment was designed that relies on an application on a mobile phone to infer opportunities for interruptions from phone activity. Testable hypotheses were formulated (see below) for a mix of behavioural (H1, H2, H4) and self-reported (H3, H5) dependent measures inspired by related work (see table 5.2, p. 140). Whereas the first set of hypotheses focuses on the impact of the *timing* strategy (IV1), H4-H5 are aimed at testing the influence of the *task type* (IV2) of the interruption, as explicated below. These hypotheses may support the assumption, which inspired this experiment: That cognitive breakpoints are located at the endings of episodes of interaction.

7.2.1 *Timing*

In order to evaluate the temporal coordination method that mediates interruptions contextually by the participant's phone activity, the following hypotheses are tested to gauge the behavioural impact (H1, H2) and perceived (i.e. self-reported) impact (H3) of the employed coordination method.

H1: People will be quicker to accept the notification of an interruption at the end of an episode of mobile interaction than at random other times.

H2: People are significantly more responsive to interruptions at the end of an episode of mobile interaction than at random other times.

H3: People will perceive completing the task at random times as a higher burden than after episodes of mobile interaction, and people will rate the appropriateness of the timing of a notification after an episode of mobile interaction higher than at random other times.

7.2.2 *Task type*

Whereas H1-H3 are aimed at testing the impact of the *timing* strategy (IV1), H4-H5 are aimed at testing the influence of the *task type* (IV2) of the interruption. Whereas the nature of the primary task has had much attention in the literature (see 4.4.1.2, p. 95), the effects of the nature of the secondary task – or the cognitive task or otherwise activity initiated by the interruption – on the primary task have been neglected. The benefits of considering the temporal *process* of the interruption (e.g. McFarlane and Latorella, 2002) has been highlighted (see 3.3, p. 49).

In the context of this study, the advantage of such models is that the interruption task itself shifts into focus. The complexity of the interruption task has been reported to affect the disruptiveness (Gillie and Broadbent, 1989, see 4.4.2.2, p. 99), and the observation of two mobile professionals showed that in over 40% of their interruptions they engaged in a new activity as a result (O'Conaill and Frohlich, 1995, see 3.2, p. 46).

Clearly, an interruption may not only affect the original primary task, but it may become the starting point for a new primary task, effectively becoming a task switch. In addition, the attention demanding nature of mobility, where mobile HCI tasks may often compete with tasks such as orienting and navigating, may lead to fragmentation of mobile HCI into second-long bursts (Oulasvirta et al., 2005b, see 4.4.1.2, p. 94), indicating that length and attention resource demands of the interrupting task may play a significant role for mobile settings. Therefore, in addition to interruption timing, the effect the type of interruption task has on the perceived workload of the interruption task and the resulting perceived burden to complete the task is investigated.

As a result of the dynamically changing contextual settings and changing attentional demand whilst being mobile, this study assumes that attentional and cognitive demand of the interruption task as indicated by perceived workload, and its social and situational appropriateness influence the perceived disruptiveness of an interruption and the completion rate of the task. This leads to the following hypotheses, to test behavioural (H4) and perceived (H5) impact of *task type*.

H4: Interruption tasks with a higher perceived workload, and/or situational inappropriateness are delayed longer before being started and have a lower completion rate.

H5: Interruption tasks with a higher perceived workload, and/or situational inappropriateness are perceived as more burdensome to complete and less appropriate when mobile.

7.3 experiment design

In a 3x3 within-subjects design, *task type* (multiple-choice, free-text and photo) and *timing* (random, opportune: after SMS, after call) was manipulated. The experience-sampling method (see 5.4) was employed over a period of two weeks and post-hoc interviews and the NASA TLX questionnaire (see 5.2.2, p. 117) were used to assess perceived workload of the tasks.

Figure 7.1 gives an overview of the experimental design (in the style of figure 6.1, p. 151) in relation to the model of the contextual factors of the interruption. Two relational factors, timing (IV1) and content (IV2) are experimentally manipulated in a 3x3 factorial design. In relation to the model, *task type* is an instance of content, for the content of the interruptions are user tasks.

The experiment measures the impact of that manipulation on the operationalisation of receptivity employed here; namely, by participants' self-reports of *appropriateness* of timing (DV5) and the *task burden* (DV6) (explained below), and by measuring *acceptance time* (DV1), *response time* (DV2), *decision time* (DV3) and *task time* (DV4). Other relational factors are held constant by experimental design, as other local contextual factors are not in the experimental focus but may be enlightened through interview statements (see figure 7.1).

After briefly expanding on the methods used in this study, the experimental app(aratus) is introduced, the task design is explicated and the way the standardised NASA-TLX questionnaire is used to identify differences in the perceived workload of the tasks as a manipulation control is outlined (see 7.3.2).

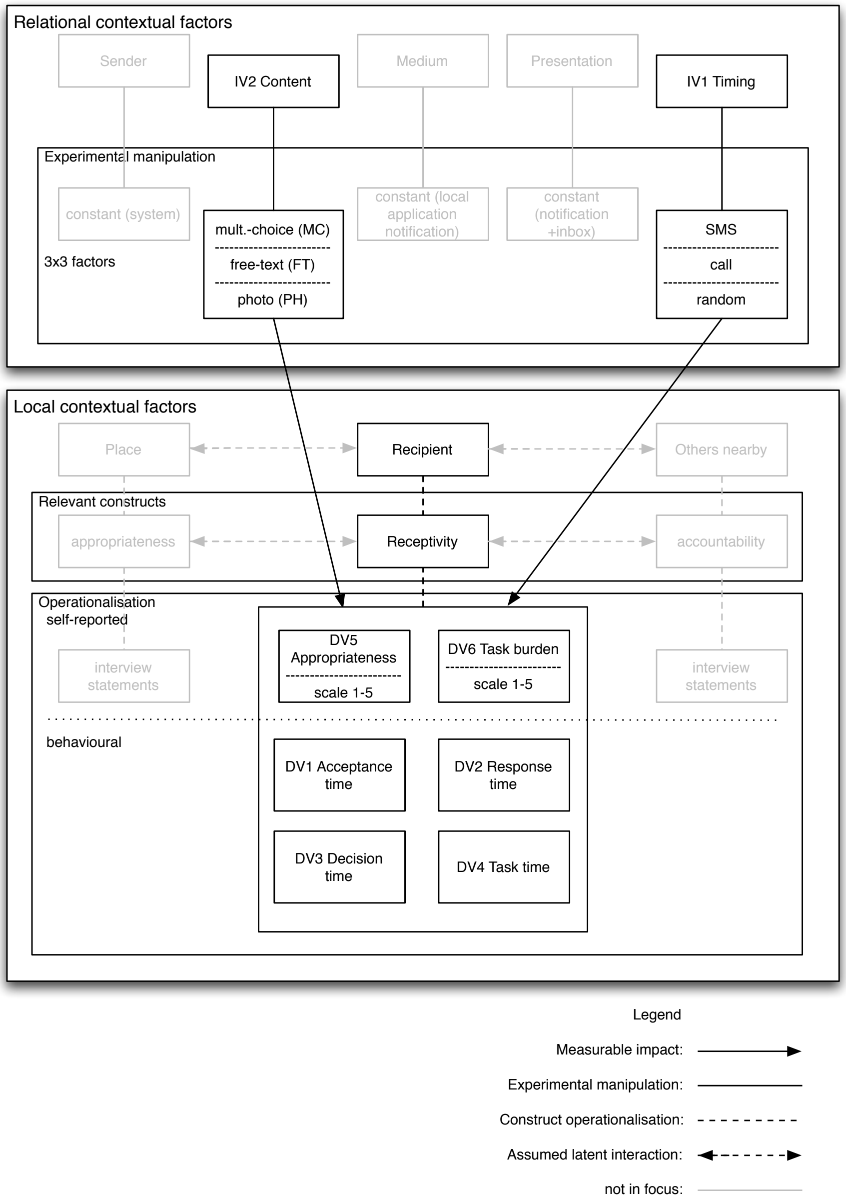


Figure 7.1: Overview of the experimental design of Experiment 2 in relation to the model of contextual factors of the interruption.

7.3.1 *Methods*

In addition to collecting self-reported ratings of the appropriateness of the timing of the interruption and the burden to complete the task by ESM, behavioural data describing device usage was also collected as ground truth for parametric data analysis. To reflect and study different phases of the interruption process in more detail, several

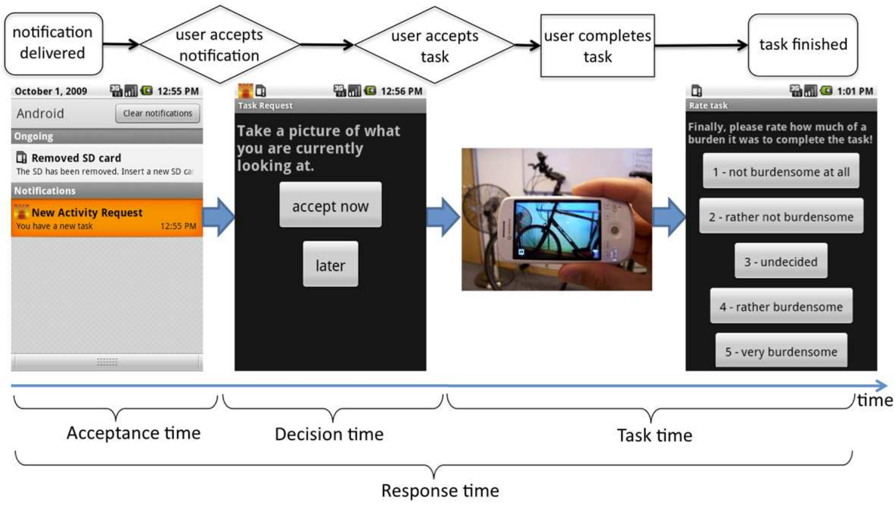


Figure 7.2: Temporal metrics (bottom) to analyse user behaviour (top decision flow) in phases of the interruption process.

temporal metrics were computed from timestamps of the interruption process (see figure 7.2) according to the temporal model introduced in 3.3.5 (p. 55), and again similar to dependent measures found in related work (see table 5.2, p. 140).

- First, *acceptance time* is the time between notification delivery and the participant's acceptance of the notification. So as not to convolute acceptance time by task type, a generic notification "new activity request" had to be clicked to accept the notification after pulling down a task bar equivalent to checking the SMS inbox.
- Then, *decision time* is the time between the task type being displayed to the participant and them accepting the task.
- Then, the *task time* is the time the user spent on the task and the rating of the burden of the task, which concluded every task.
- Finally, the *response time* is the sum of the three times above, the interruption process from notification delivery to completed response.

After the study period, an assessment of the perceived workload of the task types by means of the NASA TLX questionnaire was conducted. What is the perceived contribution of each workload factor to the overall workload by task and for which factors do the tasks differ? This also serves the purpose of a manipulation control for the intended task design: did the manipulation of the independent variable *task type* succeed?

To contextualise the quantitative findings, the study was concluded with semi-structured interviews around themes such as appropriateness and disruptiveness of task and timing, anecdotal experience of interruptions in context, and social implications of the interruptions[1].

7.3.2 *Procedure, App(aratus), and Manipulation Control*

After instructing participants about the procedure and gaining their informed consent, participants were given a mobile phone running the experiment application and they were asked to use it for two weeks with their own SIM card as their everyday phone. Participants were told to attend to the experiment notifications as they would normally attend to their personal messages and they were discouraged from attending to the notifications when it might be unsafe, e.g. when driving a car or walking in an unsafe area.

7.3.2.1 *MActivityMonitor*

The app *M(obile)ActivityMonitor* was designed for Android 1.5 and sent random and user activity-triggered notifications by monitoring broadcast events such as when the user made a phone call or received an SMS (see appendix B, p. 296 for a detailed system diagram). To monitor experiment progress remotely and to minimise the risk of

1 The interview questions are reproduced in appendix B, p. 296.

data loss, collected data was transmitted to a server when the phone connected to Wi-Fi, to minimise participants' costs.

The app would send around six SMS-style notifications to the participant's phone between 9am and 9pm. Three messages were sent at a (pseudo-) random time with at least one hour between each message. Additionally, notifications were sent after the user had completed or attempted to make a phone call, and after they had read a previously unread text message from their inbox. An algorithm attempts to balance the distribution of notifications over the day so that the participants could not predict notification delivery. It determines if the historic pattern of opportune moments of the participant's previous days shows enough opportunities over the course of the day to defer to a later moment. If the participant did not respond to the notification within 30 minutes it timed out (disappeared).

7.3.2.2 *Task design*

When participants clicked on the generic notification "new activity request" (see figure 7.2), they were prompted to complete one of three tasks:

1. A multiple-choice task: "How good was the timing of the interruption of this task when you first noticed it?"
2. A free text task: "What are you doing at the moment?"
3. A photo task: "Take a picture of what you are looking at." (see figure 7.2).

The tasks were designed to impose varying attentional and cognitive demand. In keeping with the requirements for relatively short episodes of interaction on mobile devices (Oulasvirta, 2005) and repeated prompting in an ESM study (Csikszentmihalyi et al., 1977), none of them should take longer than one minute to complete. Task order and balance was counterbalanced in order to avoid learning effects and predictability of task type. This was achieved by randomly assigning even numbered tasks, while odd numbered tasks were assigned to balance task distribution.

In addition, tasks each had varying characteristics. The Multiple-Choice (MC) task was designed to be the quickest to complete so it would absorb attentional resources for the shortest time, but it did require some cognitive resources to reason about the appropriateness of the timing of the interruption. The Free-Text (FT) task was designed to absorb the most attentional resources, as it required typing on the phone's virtual keyboard. It was probably most demanding cognitively too, as it required the participant to reflect on what they were doing at the moment and to compose it into a short statement. The Photo (PH) task added an extra quality. Instead of interacting solely with the device screen, participants interacted with the environment through the device by being forced to select a motif/subject and take a photo. In addition, this task was expected to be most confounded by the social context of the participant's current setting.

7.3.2.3 *Task design manipulation control: NASA TLX*

The task design was assessed by NASA TLX. Participants rated the procedure's six factors of workload (mental demand, physical demand, temporal demand, performance, effort and frustration) after the study for each of the three task types. A repeated measures ANOVA showed that the mean aggregated workload differed significantly by task, with $F(2, 40) = 13.19$; $p < .01$. Hence, the task type had a significant effect on perceived workload. The perceived workload for each task was then compared by means of a pairwise comparisons by the Bonferroni procedure. It showed that the mean workload of the FT task (57.8) was significantly higher than the mean workload of the MC task (31.0; $p < .01$) and significantly higher than the mean workload of the PH task

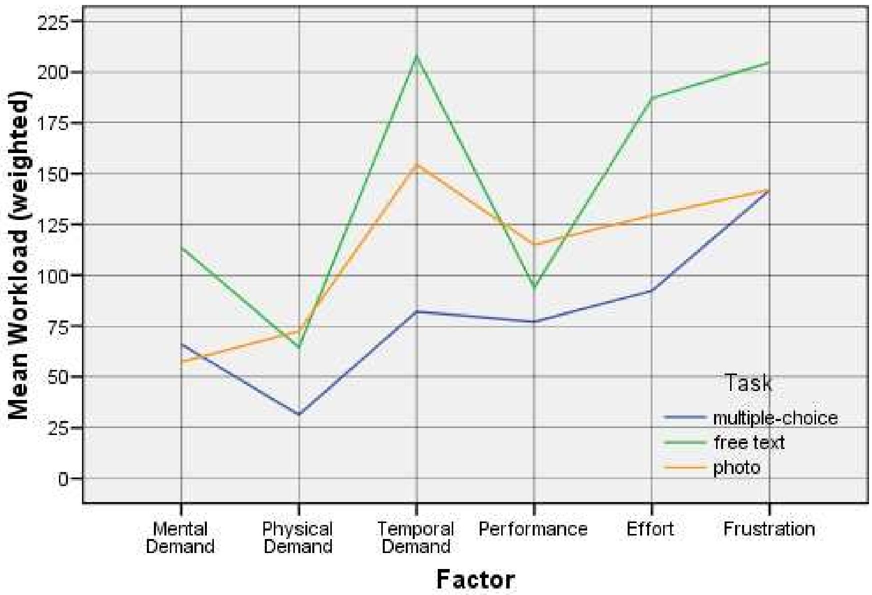


Figure 7.3: Workload contribution of factors by task type.

(40.3; $p < .01$). MC task and PH tasks did not differ significantly. This supports the intended manipulation of the independent variable task type.

In order to compare the amount of each of the six factor's contribution (e.g. temporal demand etc.) to the perceived workload of the tasks, a further analysis of the contribution of the individual workload factors to overall workload (see figure 7.3) was conducted. It showed a significant effect of the individual factor, with $F (5, 95) = 6.38$; $p < .01$. Task type also contributed significantly, with $F (2, 38) = 10.57$; $p < .01$. The interaction of the factors was not significant.

Pairwise comparison showed that the mean rating of temporal demand for the MC task (79.0) was significantly lower ($p < .01$) than for the FT task (205.8) and the photo task (149.3; $p < .01$). Also, effort of the FT task (185.3) was rated significantly higher ($p < .01$) than effort for the MC task (88.3) and significantly higher ($p < .05$) than the mean rating of effort for the PH task (125.0).

7.3.2.4 *Role*

Carried out under supervision at the Mixed Reality Lab at the University of Nottingham, the author had full responsiblity and in agreement with supervision performed all activities necessary to conduct this experiment. This ranged from experiment design, orchestration of the experiment, participant recruitment, set up and support during the experiment, debrief and exit interviews. Also, the technical system to support the experiment was designed and implemented by the author, both the Android application, and the server-side data storage.

7.3.3 *Participants*

20 participants, (10 male, 10 female) were recruited through email lists and subsequent snowballing. The participants were between 21 and 48 years old (M=30, median = 27.5). 10 participants were postgraduate students, five were employed at the university, and three were employed in sales, one in health and one in the environmental sector (see appendix B for details on the demographics). Participation was reimbursed with £20.

7.4 ANALYSIS AND RESULTS

Each of the 20 participants took part in the experiment for two weeks. In total, they received 2002 notifications and completed 1380 of the tasks (i.e. a response rate of 68.9%). Table 7.1 shows the distribution of the independent variables (IVs) across the notifications and responses.

To check if the distribution of responses and non-responses were biased by the timing strategy (random vs. opportune) by which they were sent, a chi-square analysis on the resulting contingency table was conducted (see table 7.2). The analysis showed that the association

IVs			Task type						Total	
	Levels		*Multiple-choice*		*Free-text*		*Photo*			
			s	r	s	r	s	r	s	r
Timing	*Random*		359	262 (73%)	370	260 (70.3%)	271	201 (54.2%)	1100	723 (65.7%)
	Opportune	*SMS*	154	132 (85.7%)	134	101 (74.5%)	140	85 (60.7%)	428	318 (74.3%)
		Call	154	126 (81.8%)	168	124 (73.8%)	152	89 (58.6%)	474	339 (71.5%)
Total			667	520 (78%)	672	485 (72.2%)	563	375 (66.7%)	2002	1380 (68.9%)

Table 7.1: Distribution of sent (s) and responded upon (r) notifications and response rates across levels of the independent variables (IVs).

between the distribution of responses and the timing strategy was significant, with χ^2 (1) = 11.70, *exact p* = .001. A ϕ value of .076 indicates a weak association.

Inspecting the crossabulation in table 7.2 further, it should be noted that the actual frequencies and the frequencies expected by the statistical test[2] follow a pattern that supports hypothesis H2: notifications sent at random times are less likely to be responded upon (and more likely to not be responded upon) than expected. Furthermore, notifications sent at (hypothesised) opportune times are more likely to be responded upon (and less likely to not be responded upon) than expected.

To summarise, participants were more likely *not* to respond to a notification if it was sent at a random time than at an opportune time, and they were more likely to respond if it was sent at a (hypothesised) opportune time.

Furthermore, in order to test if participants were more likely to complete tasks of a certain type, a chi-square goodness-of-fit test was conducted on the distribution of frequencies of the task type prompted by the notifications. The null hypothesis that the three tasks are equally

2 In the employed chi-square test for association, the expected count for each cell is calculated using estimates of probability computed from the marginal totals and the total frequency *N* (cf. Kinnear and Gray, 2008, p. 395). Thus, the *estimation* accounts for the differences in *observations* in rows and columns.

Timing		Responses	
		Responded	Not responded
Random	actual count	723	377
	expected count[2]	758	342
Opportune	actual count	657	245
	expected count	622	280
Total		1380	622

Table 7.2: Crosstabulation of responses and non-responses by timing.

likely to be completed proved to be significant and thus has to be rejected, with χ^2 (2) = 24.89; *exact p* < .001. Participants were 5.8% more likely to complete an MC task over a FT task, and 11.3% more likely to complete an MC task over a PH task, and still 5.5% more likely to complete a FT task over the PH task (see table 7.1). This supports the part of hypothesis H4 that tasks with a higher workload and/or situational inappropriateness receive a lower completion rate.

7.4.1 *Behavioural data*

The four primary behavioural dependent variables *acceptance time* (DV1), *response time* (DV2), *decision time* (DV3), and *task time* (DV4) were computed from timestamps recorded each time when a participant went through the process of responding to a notification (see figure 7.2).

As argued in 5.4.3.5 (p. 129), repeated-measures ANOVA would be the familiar choice of data analysis technique, but it has a major drawback: it requires subjects to have equal numbers of repeated measurements (Garson, 2010). In a study where measurements are collected based on individual behaviour the analyst would be forced to shrink the dataset down to the lowest common denominator dictated by the subject with the fewest repeated measures, or to exclude those subjects entirely. In any case, this would affect a loss of richness of the data and analytic power and may even lead to false conclusions.

Linear Mixed Models (LMM) was adopted as an alternative approach, which has been applied to HCI research before (Wang et al., 2009). LMM is a disaggregate procedure which does not require equal amounts of measurements per subject and condition, and the variances do not need to be uniformly distributed (Garson, 2010). LMM have the advantages that variance in the data is not lost by averaging as in an aggregate procedure such as repeated-measures ANOVA, and that they account for the individual participant as a random effect, i.e. *participant* can be included as part of the model to reveal if individual differences have any significant effects on the result.

Dependent measures were log-normalised to meet the assumption of normality, and then a linear mixed model was created for each of the measures. For *acceptance time*, the only fixed effect was *timing*, as *task type* could not have had an effect because it was unknown to the participant. For the other three cases, the independent variables *task type* and *timing* were modelled as fixed effects; *participant* was always included as a random effect. Note that in linear mixed models, Satterthwaite's approximation of degrees of freedom may yield non-integer denominator degrees of freedom (Wang et al., 2009).

In addition to results from LMM pairwise comparisons from the Bonferroni procedure for significant effects are reported, where log-normalised values are used to derive significance levels but provide median values per level in seconds for the sake of readability and sense-making.

7.4.1.1 *Acceptance time*

Timing had a significant main effect on *acceptance time*, with $F(2, 1526.9) = 104.59$, $p < .001$. In addition, *participant* was a significant random effect by the Wald test, with Wald $Z = 2.51$, $p < .05$. According to Garson (2010), the percentage of variance in the dependent measure can be computed by dividing the variance of the random effect (i.e. *participant*) by the total variance. This computation showed that the percentage of variance in acceptance time explained by between-subjects effects was 5.9% in the employed default variance component model.

Pairwise comparison showed that the mean *acceptance time* was significantly higher (at the .01 level) when the notification was delivered at a random time (median (med.) = 36 seconds) than when the notification was delivered after the participant had read an SMS (med. = 19s) or had made a phone call (med. = 10s). *Acceptance time* for the *opportune* conditions (SMS vs. call) also differed at the .01 level. *Task type* could not (and did not) have an effect on *acceptance time* as the task type is unknown to the user at the time of accepting the notification.

The result that *acceptance time* is significantly higher for random than opportune times support hypothesis H1 that people attend to notifications on their mobile phones significantly quicker when they have just completed an episode of interaction.

7.4.1.2 *Response time*

Figure 7.4 shows the distribution of response time in seconds – 50% of the notifications (that did prompt a response) were responded upon within 45.3s, 75% were responded upon within 269s (4.5min), and 90% were responded upon within 845s (14min).

Both the manipulation of *timing* and *task type* had a significant main effect on *response time*. *Timing* had a significant effect on *response time*, with $F(2, 1374.92) = 73.71$, $p < .001$. The effect of *task type* on *response time* was significant, with $F(2, 1360.76) = 13.03$, $p < .001$. Again, *participant* was a significant random effect by the Wald test, with $Z = 2.44$, $p < .05$. The variance in *response time* attributable to participant was 6.2%. Pairwise comparison showed that the response

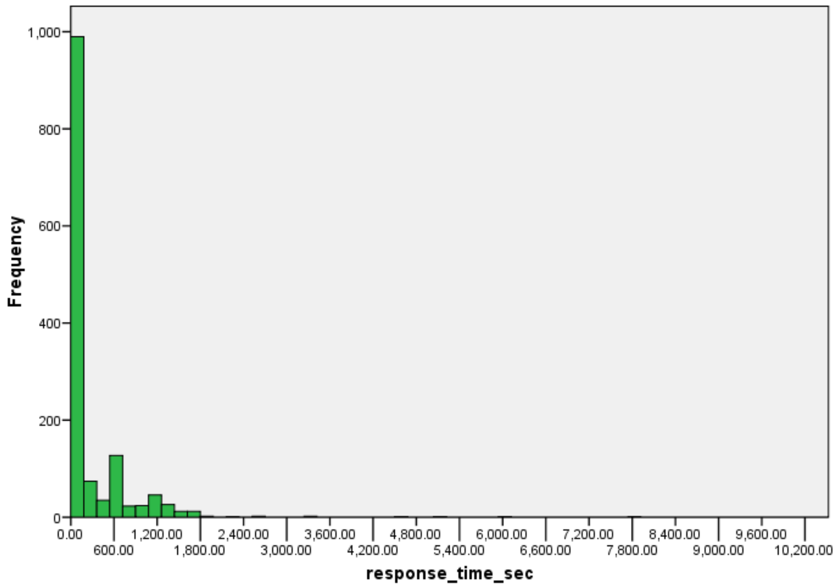


Figure 7.4: The distribution of response time in seconds (interval = 180s).

time for random timing of notifications (med. = 66s) was significantly higher (at the .01 level) than after reading an SMS (med. = 38s) or after making a call (med. = 29s). Response time did not differ significantly for notifications after reading an SMS or making a call.

Furthermore, *response time* was significantly lower (at the .01 level) for the multiple-choice (MC) task type (med. = 29s) than for the free-text (FT) task (med. = 53s) or the photo (PH) task (med. = 48s). As response time is a composite temporal metric (see figure 7.2), this may be explained by the significantly shorter *decision* and *task time* for the MC task, as discussed below.

Results regarding response time support hypothesis H2. People's response time to notifications sent after completing an episode of mobile interaction is significantly lower than to notifications send at random times.

7.4.1.3 *Decision time*

Task type had a significant effect on *decision time*, with $F(2, 1450.51) = 16.38$, $p < .001$. *Timing* did not have a significant effect on *decision time*.

A Wald test showed that *participant* was a significant random effect, with Z = 2.63, $p < .01$. Individual differences explained 15.3% of the variance in *decision time*. Pairwise comparison showed that the mean decision time for the MC task (4 seconds) was significantly lower (at the .01 level) than for the PH task (36 seconds). The difference to the mean for the FT task (17 seconds) was not significant for either task.

The fact that the decision time was significantly lower for the MC task than for the other tasks completes the support of our hypothesis H4 that tasks with a higher workload and/or social inappropriateness are delayed longer before being started and have a significantly lower completion rate.

7.4.1.4 *Task time*

Task type had a significant effect on *task time*, with F (2, 1353.67) = 875.43, $p < .001$. *Timing* did not have a significant effect on *task time*. Again, individual differences contributed by *participant* showed to be a significant random effect, with Wald Z = 2.76, $p < .01$. The variance in *task time* attributable to *participant* was quite high (20%).

The task time spent on the MC task (med. = 3.4s) was significantly lower (at the .01 level) than the time spent on the FT task (med. = 17.6s) or the PH task (med. = 12.5s). Likewise, the difference between the time spent on the FT and the PH task was significant at the .01 level.

The results regarding task time further stress achievement of the goal of task design: tasks with distinct characteristics. With respect to *task time*, the tasks differed significantly.

7.4.2 *Self-reported data*

Ratings of the participants' perception of *appropriateness of the timing* (DV5) of the notification were collected by means of the MC task and ratings of the perceived *burden of completing the task* (DV6) at the end of every task. Both dependent measures were Likert scales with 5 ranks (timing: from 'very good' to 'not good at all'; burden: from 'not burdensome at all' to 'very burdensome' (see figure 1)). To analyse the data, the median rating was obtained per participant per category, and nonparametric Friedman tests for ordinal repeated measures were conducted.

7.4.2.1 *Appropriateness of timing*

Participants' self-reported *appropriateness of timing* did not differ significantly by timing (χ^2 (2) = 5.65, *exact* p = 0.068). The median rating for the random notifications was 'undecided', whereas for the conditions *SMS* and *call* it was between 'rather not good' and 'undecided'.

7.4.2.2 *Burden of response*

Participants' self-reported *burden of completing the task* did not differ by *task type* (χ^2 (2) = 4.51, exact p = 0.1) or by *timing* (χ^2 (2) = 0.46, *exact* p = 0.8). The median reported burden to complete the MC and the PH task was 'rather not burdensome', whereas for the FT task it was 'undecided'. The median burden for randomly timed notifications was 'undecided', whereas for after an SMS it was 'rather not burdensome' and for call it was between 'rather not burdensome' and 'undecided'.

7.4.2.3 *Summary*

The results of the nonparametric tests on the self-reported perception of the *burden of completing the task* and the perception of the *appropriateness of timing* do not support our hypotheses H3 and H5. The timing of the notifications did not make a difference in how much of a burden participants saw in completing the task, or in how appropriate they rated the timing of the notification (H3). Also, *task type* did not influence the perception of the burden of completing the tasks in a significant way (H5).

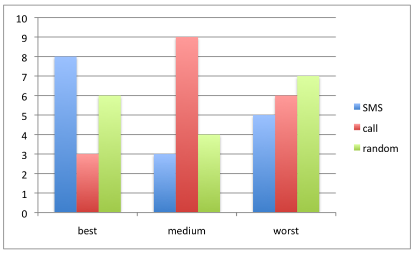
7.4.3 *Interview data*

After the experiment 18 participants were interviewed in a semi-structured fashion. Interview responses were also coded for statistical analysis that are reported here.

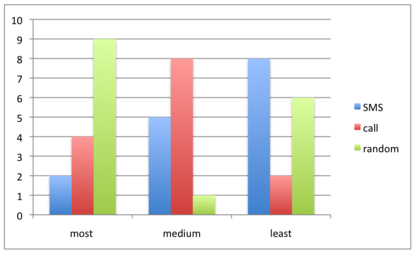
None of the participants felt that they could predict the timing of a notification in advance, but eight participants noticed the notifications were triggered by their phone activity and five of them correctly identified phone calls and SMS as triggering the notifications. In keeping with the results from the ESM, a Friedman test of the rankings of the appropriateness (best, medium, worst) and the disruptiveness (most, medium, least) of the three types of timing (*random, SMS, call*) during the interview (see figure 7.5) failed to produce significant results.

Despite the statistical insignificance, the random condition was still ranked as the most disruptive condition 9 out of 16 times (6 times as least disruptive, see fig. 7.5 a), and the least appropriate 7 out of 17 times (6 times as most appropriate, see fig. 7.5 b). The *SMS* condition trumped the other ones in terms of appropriateness (most: 8, least: 5 out of 16 times) and least disruptiveness (least: 8, most: 2 out of 15 times).

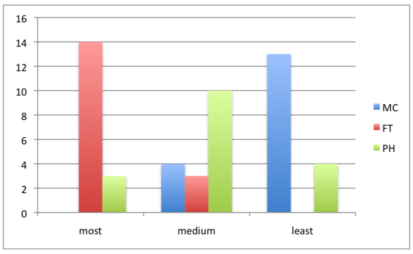
In contrast to the *in situ* ratings of the burden to complete a task, but in accordance with intended task design and the findings on task workload and *task time*, participants reported in the interviews that they perceived the tasks as quite distinct from one another. When asked to rank the burden of the tasks in the interview again, the free text task was ranked as the most burdensome task in 14 out of 17 cases (82%, see fig. 7.5 c), the photo task was ranked in the middle



(a) Appropriateness of timing.



(b) Disruptiveness of timing.



(c) Burden of task.

Figure 7.5: Interview-reported ranking of the appropriateness (a), and disruptiveness (b) of timing strategy, and perceived burden of task type (c).

with 10 mentions of medium burden (59%) and the multiple choice task was in 13 cases considered the least burdensome task (76%). A Friedman test showed that burden was ranked significantly different for the tasks, with χ^2 (2) = 24.8, exact p < .001, Kendall's W = 0.73. Pairwise comparisons by Wilcoxon's test also showed the three tasks were all ranked significantly different from each other on a .01 level.

7.4.4 *Summary and hypotheses*

7.4.4.1 *Effects of timing*

- Hypothesis H1 (quicker acceptance) is supported by the finding that acceptance time is significantly higher for *random* than *opportune* times.
- Hypothesis H2 (quicker completion) is supported by the finding that *timing* had a significant effect on the *response time*.

- In relation to hypothesis H3 (perceived *appropriateness of timing*) no significant effect was found in the self-reported or interview data.

7.4.4.2 *Effects of task type*

- Hypothesis H4 (task delay and non-completion) is supported by the finding that *task type* significantly affected the likelihood of completing the task, and the finding that *decision time* was significantly lower for the MC task than for the other tasks.
- In relation to hypothesis H5 (perceived *burden*) no significant effect was found in the self-reported data. However, interview data suggets that tasks with a higher workload were perceived as more burdensome.

To summarise the results, our hypotheses related to the participants' behaviour were supported by the analyses, whereas the hypotheses related to the participants' self-reports were not supported by the analyses. In the following, this disparity is unpacked by discussing the findings from the interviews at the end of the study.

7.5 DISCUSSION

Now, the findings are discussed by contextualising the quantitative results with more in-depth qualitative descriptions from the interviews. Participants' quotes are used to illustrate and issues are related back to broader concerns on interruption management.

7.5.1 *Contextual sensitivity to the timeliness of interruptions*

In the interviews there was substantial disagreement between the participants regarding the timing strategy of the notifications. This is reflected in the lack of significant support for hypothesis H3. Here, some of the factors that participants reported as relevant are laid out.

7.5.1.1 *Egocentric device position*

The activities of making a call and reading an SMS were both characterized by holding the phone in hand. Participants mentioned this as being beneficial for dealing with the interruption task.

> If I've already got it in my hand, in that position there's more chance of me responding to it. If it gave me time to put the phone down, then chance is less of me responding immediately, because I already put the phone down and went to a different task.
>
> When you're already using the phone, it's probably not as burdensome (...), less of an imposition. You were using it rather than it's gone off in your pocket and you were perhaps doing something else and you had to check to see what it was.

With regards to egocentric mobile device positions it seems that to time the mobile interruption appropriately means to time it so that the device is still *in hand*, but no longer *in use* for a task such as calling or texting. It was argued that notifications have the potential to interfere with our orientation towards the world, as they for example may 'get in the way' of interacting through a device (cf. 2.1.5, p. 23). Borrowing Heidegger's insight that obstacles can cause *unreadiness-to-hand* (Heidegger, 1962) one might conjecture that delaying notifications to when the mobile phone is no longer being used while still being in hand may help to avoid obstinate interruptions.

7.5.1.2 *Locus of control*

However, as suggested by the insignificant differences in participants' *in situ* self-reports, phone activity related timing was not always preferred to random timing.

> Random was less annoying because they didn't interfere with my actual phone usage. Because I'm a bit "controlly" I felt less hostile towards these because I had actively chosen to look at my phone to see if there were any notifications.

Preference of the random condition raises issues with balancing control of awareness and interruption (Horvitz et al., 2005a). Also, people who preferred the random condition often also said they generally let messages accumulate over time to "batch process" them at a convenient time.

7.5.1.3 *Task context*

In addition, it may not have been the random timing alone that people preferred *per se*, but they may have found it less bad compared to situations where the phone activity-related notifications actually interrupted their phone activity.

> While I'm reading a text it is quite annoying, it's like a little child poking you while you're doing something else, but pretty good after I sent one. Random ones...changed over time. (...) After a phone call was quite interruptive because sometimes you want to call someone else, or you didn't reach the person and need to call back. Then came the message. Was quite interrupting.

Apparently, the notifications triggered by phone activity were more prone to interrupt phone activity that consisted of multiple sub-tasks; such as making several phone calls or exchanging several SMS in succession.

> When I was sending multiple text messages...Just the tasks in general were more annoying then. When I was having almost like an IM style text conversation with my mother or someone who expects a reply.

> Timing was often quite good if it was just after I sent a message. Because then I just dealt with one thing, I just cleared one thing off my agenda, and then a notification would pop up, I'd still be sort of set in my processing mode of dealing with stuff on my phone.

This is in line with the argument inspired by Heidegger brought forward earlier, that the device should no longer be *ready-to-hand*, i.e. used for an ongoing task, when the notification is delivered.

7.5.1.4 *Task coherence*

Further anecdotes from the interviews suggest that similarity in the activities of texting and replying to a task request may have made the SMS-triggered task notification more appropriate than a call-triggered notification.

> Best is after a text because chances are you still have your keyboard out. With the phone you're just holding it to your ear, then you put it away.

The mode of interaction when composing an SMS and when responding to the notifications has similar physical requirements – having eyes and hands free. On the contrary, people often reported using their mobile phones for calling when this requirement could not be met, for example whilst driving or walking, or generally physically moving between activities, e.g. on their way to the car to confirm a meeting.

7.5.2 *Behavioural versus self-reported evaluation of timeliness*

Whereas the quantitative analyses of the behavioural data support the assumption that opportune moments for interruptions are located at the end of episodes of mobile interaction, the analysis of the self-reports fails to give further support. It appears that the benefit of the presented interruption delivery method may be on the side of the *sender*: The interruption gets tended to and acted upon significantly quicker, which suggest that endings of episodes of mobile interaction are indeed opportune moments, but to *deliver* rather than to receive an interruption. However, it was argued that the understanding of receptivity to interruptions is gained from a subjective viewpoint that places the recipient's experience of the interruption in the focus (see 4.1).

This participant's reasoning about the experience of the interruption delivery method illustrates the difficulty in predicting whether the ending of an episode of mobile interaction actually represents a breakpoint that is an opportune moment for an interruption purely based on sensing phone activity:

> It was quite good when I got a text message that wouldn't require a response. It was a big difference there between if you wanted to carry on with another text message or wanted to make another call or if it was the end of a conversation. If it was at the end it was a quite good time and no problem at all, but if it was in the middle of a conversation or the middle of texting, if you're doing 2 to 3 texts, that didn't seem good. I suppose the end of a conversation, yeah, that's possibly good.

As the participant illustrates, the ongoing information exchange exploited by the application may have already informed the intention of carrying out a new activity, which is a good example of how an interrupting task can become the onset of a new primary task (O'Conaill and Frohlich, 1995). Hence, even though endings of interactional episodes were assumed to collocate with cognitive breakpoints, the interviews show that:

A) The breakpoint may have been missed and the user is already in a state of processing a subsequent task, or

B) The breakpoint is at a sub-task and may still be considered more disruptive than a randomly timed notification.

7.5.3 *Experience of the interruption tasks*

Now, the behavioural findings are related back to the participant's perceptions of the interrupting task and its burden, and the range of reported factors are briefly summarised.

It was found that interruption tasks with a higher workload are delayed longer. The NASA-TLX analysis (see figure 7.3) showed significant differences in *temporal demand* and *effort* for the three task types, which accords with some of the interview comments. The FT task was reported as taking the most time, and requiring more cognitive resources than the other tasks (paralleling the tasks' assessment of workload and actual *task time*). 12 people said they deferred the FT task when asked if the task type influenced whether they accepted the task right away or deferred it. The MC task was generally preferred to the other two tasks for requring the "least effort", and being "quick". However it also required reflecting on the moment of interruption, implying a degree of mental demand:

> It wasn't just the time the task took. It was a little more effort to sit and think about the MC task, whereas a photo task you didn't have to think, you could just take a photo of anything. Whereas for the MC, I had to put more thought into it.

However, the other workload factors show less differentiation and in some case contrasting ordering (e.g. *performance*, for which the FT mean workload is less than that for PH), or may be confounded by social, affective and environmental factors, which were also reported to affect the appropriateness of completing a task. For example, the free-text task was reported to be inappropriate while driving or walking, or difficult to do in sunlight because of diminished visibility of the screen.

Highlighting an affective aspect to the tasks, the photo task was described as being "fun", "interesting", "enjoyable", "engaging with the environment" and an "easy" task, which may well mitigate the perceived burden of completing the task. In contrast, the MC task did not allow creative completion, which may have made it less enjoyable.

As anticipated, it was also found that the photo task introduces an element of social accountability, which affects the timing and completion rate of that task. The PH task differed from the other two tasks in that it not only involves interaction with the device but with the environment through the camera; a fact that may have rendered the task socially inappropriate in some situations, as this participant points out:

> Probably the photo task I would defer to a later time. Depending on who I was with. So if I sat in a meeting and it goes off... to take a photograph of the person that I'm looking at, it's not very socially acceptable, is it?

In addition, as opposed to real world interruptions, the study relied on fabricated content of the interruptions (the tasks). However, studies presented previously showed factors such as the content (see 4.4.2.2 or Experiment 1 in chapter 6) and sender (see 4.4.2.1) of the interruption play a significant role in how receptive a person will be to the interruption.

In contrast to the analysis of behavioural data, self reports of the burden of response did not differ significantly by task type. The interviews show that lack of significance may be due in part to the multi-dimensional character of *task burden*, including multiple workload factors, environmental factors, social accountability, and affect, which precludes the effective use of a single measure.

Interestingly, the findings regarding the assessment of the appropriateness of interruption timing is somewhat similar to the findings from a diary study that investigated which contextual information receivers wished their callers had at the time of calling them (De Guzman et al., 2007, see 2.2.5). The most frequently named type of information they wished their caller had was their *task status* (34%), *social availability* (15%), and *physical availability* (21%), pointing towards some level of generalisability of the qualitative findings presented here. Furthermore, De Guzman et al. (2007) look at the distribution of ratings where their participants said receiving a call was very disruptive, and find that this correlates with the most frequently named contextual information they wished their callers had. Namely, *physical availability* was mentioned at almost 30% of the times where they said the call was very disruptive; *social availability* was mentioned over 20%, and *task status* was mentioned in just under 15% of the very disruptive calls. Notably, at further 18% of the very disruptive calls, receivers wished their callers had information relating to their orientation towards the device, such as *distance to phone* and *frequency of recent interaction* (De Guzman et al., 2007), again providing some support of the effectiveness of the methods of contextual inference used here.

7.5.4 *Practical considerations*

Finally, some practical issues for applying the presented interruption strategy to application domains as outlined in 4.5 (p. 105) are anticipated.

A challenge for systems that defer potential interruptions to an anticipated opportune moment is that systems may hold back an interruption indefinitely. To deal with this, the concept of *bounded deferral* was introduced (Horvitz et al., 2005a), according to which notification of an interruption is held back until a maximum deferral time has been reached.

A further problem is posed by the fact that the content of the interruption may be urgent or time critical to the recipient. Therefore, it is assumed that most people do not want a mediating service that interferes with their first-order communication, such as phone calls, SMS and email. Consequently, either system design must incorporate the difficult problem of robust semantic content analysis, or its application must be limited to non-time critical messages.

Alternatively, the presented interruption strategy could be applied to mobile applications that aggregate and deliver information from the user's second-order communication networks, such as social networks' *activity streams*, or other information sources the user has subscribed to, such as RSS feeds. The mechanism could also be used by services that deliver a dedicated user-experience or prompts for interaction, such as location-based services or games.

In summary, the presented strategy mediates interruptions by deferring them until an episode of interaction provides an opportune moment and messages are made available in an inbox-and-notification style. In McFarlane's typology (see 3.6.3, p. 74), the employed strategy represents a mix of the mediated and negotiated interruption strategies.

7.6 CONCLUSIONS

Using a naturalistic study to test novel but simple interruption coordination based on sensing mobile phone activity it was found that mobile users tend to accept and reply to notifications significantly more quickly after they finish an episode of mobile interaction than at random other times. This suggests that the presented strategy may be effective for applications that aggregate and deliver content proactively, or for systems that manage interruptions from the user's second-order communication network. In comparison to Experiment 1, findings certainly show that contextual mediation of interruptions is superior to letting participants schedule their own interruptions.

However, *in situ* self-reports did not show the subjective experience of activity-triggered timing to be superior to the random condition. The qualitative analysis exposes some of the situated complexity of interruption handling that can influence whether the phone activity-triggered notification is considered timely. In particular, three major task/activity cases are revealed that influence perceived timeliness, i.e. whether at the moment of interruption the user:

BEST CASE: Just finished a task – physically but especially cognitively – and is therefore available to an interruption;

INTERMEDIATE CASE: Has only finished a sub-task within a larger activity; or

WORST CASE: Has already instituted a new task or has entered the *planning* phase (Miyata and Norman, 1986) for that task, which is therefore being interrupted.

On the one hand, the best and the intermediate case support the assumption that cognitive breakpoints may be located at the endings of episodes of mobile interaction, due to parallel findings that breakpoints higher in the task hierarchy may be more opportune than between sub-tasks, supporting postulates P5, and indirectly P7 and P8 (see 4.4.1.2, p. 94). On the other hand, the worst case qualifies the assumption by uncovering that breakpoints and endings of mobile episodes do not always collocate, which means that opportune moments may have been missed or not reached yet, supporting P3; while the best case supports P2 and P12.

Distinguishing these cases in a mobile context is a question for future work, which may also be inspired by the consideration whether the device was still *in use*, *in hand* (Dourish, 2001), or neither. Whereas if the device was still in use, e.g. to mediate a phone call, a notification would likely be perceived as disruptive, having the device still in hand may provide a more opportune moment than when it has been put away.

With regards to the interruption task, the study showed that its character has a significant effect on the time to decide whether to accept the task and the overall completion rate. Furthermore, while the appropriateness of completing an interruption task depends on the factors that comprise workload of the interruption task (esp. *temporal demand*, *effort*, *frustration*), the true "workload" of a task is determined through involvement the moment (see 2.1.5) and is influenced by its situated social accountability (e.g. taking photos in a meeting), and cognitive and attentional demands (e.g. typing while walking). Overall, whereas social, cognitive and attentional demands may contribute to the burden of dealing with an interruption task, affective factors such as a sense of fun may mitigate the sense of burden.

7.7 summary and outlook

The findings, in particular the discovery of task/activity cases that influence perceived timeliness furthers the understanding of the complexity of receptivity to interruptions in this thesis. However, whereas the employed strategy to mediate interruption by phone activity shows some promising potential, the uncovered caveats suggest that the sensing of phone activity alone is not sufficient to apply this strategy in interruption management to routinely and reliably predict opportune moments. To raise confidence in the prediction strategy, more contextual factors need to be taken into account to estimate receptivity, such as the recipent's current place and their social surrounding. The potential of mediating interruptions by the recipents' location is explored in the following chapter 8.

Furthermore, while the previous Experiment 1 (chapter 6) and this experiment explore temporal coordination of interruptions as *experiments*, what are the potential benefits and shortcomings of applying temporal coordination of interruptions to a real-world system? It was argued that interruption management is valuable for two kinds of real-world systems (see 4.5). AUTOMICS, the system reported in the next chapter implements temporal coordination of its notifications to users. Hence, the field study of AUTOMICS provides insight into the applicability of developed temporal interruption management in a real-world application.

Methodologically, while the employed temporal metrics allowed for a detailed analysis of the impact of task and the temporal coordination method, evidence of the ways in which local contextual factors such as the social and the physical environment mediate receptivity to the interruption remains anecdotal, as participants were locally dispersed during this study. The following field study addresses this shortcoming by focussing on the observation of groups, and their interaction with and around the application, and especially the effects of the application's notification-driven interaction on the groups.

8

STUDYING NOTIFICATION-DRIVEN INTERACTION IN THE FIELD

This chapter presents a field study that explored the socio-technical effects of notification-driven interaction in Automics (AUTOmated and assisted photo-stories and COMICS). Automics is a prototype mobile application which enables theme park visitors to create photo-story souvenirs of their day in the park. Findings were generated from a mixed methods approach of observations in the field, interviews and an analysis of system logs. The findings show that participants appreciated location-based notifications as reminders of the tasks necessary to create a photo-story. Also, notifications worked well to foreground the experience during relative 'downtime' of the theme park experience while queuing. However, the large quantity of notifications that informed the visitors of each others' activity meant they soon habituated to and eventually ignored them. Moreover, the applications' ignorance of the visitor's current system activity led to redundant notifications. The analysis of system logs offers formal support for the qualitative observations in terms of the participants' acceptance rate, and acceptance time by notification type and over the course of the day. Furthermore, the designer's notion of the notifications' utility (to coordinate the pacing and ordering of user actions while at the same time assisting the user in negotiating an optimal balance between the experience of being in the park and the experience of the souvenir system) is evaluated and contrasted with the findings from the field study. From this, implications for the system's next potential iteration and the broader use of notifications in interactive systems are drawn.

8.1 INTRODUCTION

While the former two chapters present experimental work, this chapter differs in that it explores research questions by means of a field study of a real-world application. The field study especially allowed the observation of the social and the environmental context of the interaction with and around the technology. The environment and the social situation represent local contextual factors in the contextual framework that could not have been studied in the previously presented experiments yet, as the participants were not observed as they went about their everyday life whilst participating in the study.

In chapter 4 it was argued that interruption management can not only be applied to applications that manage a user's incoming communication, but also to systems that initiate interaction proactively in order to engage the user in an interactive experience. Notifications and messages are frequently used as design elements in interactive cultural, mobile or pervasive experiences. For example, *Day of the Figurines* is a long-term SMS based game that is played purely by text messaging (Flintham et al., 2007). Particularly, location triggers are used to drive notifications to engage with digital services and media, such as mobile advertising (Bellotti et al., 2008), tour guides (Cheverst et al., 2000; Brown et al., 2005; Rowland et al., 2009), and location-based mobile gaming (Paelke et al., 2007). For example, the *Silitoe Trail* is experienced by cycling through an urban area, while the system plays location-triggered audio clips to convey a historic sense of the surroundings at the time of the famous writer Silitoe (Rowland et al., 2009). The study of this kind of system in use presents an opportunity to explore the effects of the interruptions that the technology causes for its users, which has led for example to a detailed account of how players manage their accountabilities to their social surrounding (Tolmie et al., 2008).

Automics employs location-based messages at anticipated opportune places that invite users to complete interactive tasks such as taking photos or editing or selecting them for the photo-stories they create as a souvenir.

In addition, Automics automatically generates and distributes notifications that inform the members of the visiting group when one of them has shared a newly captured or edited an existing photo. These notifications have been designed to convey a sense of the actions that members contribute to the shared media space. Thus, their function is to raise mutual awareness of the users' activities to provide a context for the individual user's own activity (Dourish and Bellotti, 1992).

The usage and effects of location-based activity prompts and awareness messages in Automics were studied to elicit information about the contextual settings that may render these kinds of interruptions opportune or inopportune. In contrast to the previous experiments, the approach of the field study is exploratory rather than hypothesis-driven to cater for a wider array of research topics that allows for findings to emerge from observations in the field. In the following, Automics is described and on that basis the research questions are developed.

8.1.1 Automics

The work presented here is part of broader research in theme parks and mobile photoware (Durrant et al., 2011), but is distinct in its focus on notification-driven interaction, which was developed for this thesis[1]. From a sensitising ethnographic study in a popular theme park in the UK reported by Durrant et al. (2011), several design requirements for an enhanced souvenir system for the theme park setting were synthesised:

1 For clarification of the contributions of this thesis' author see 8.3.3.1, p. 227.

- Combine professional and personal photos
- Enable playful sharing of photos
- Support personal narratives
- Involve spectators
- Providing physical and digital forms

The requirements informed the design of AUTOMICS, a mobile application that enables groups of visitors to 'co-create' photo-stories from a shared pool of photographs as a souvenir of a visit to a theme park. Each member of the group uses the application to take photographs, annotate them with captions and speech or thought bubbles if they wish, while all of this user-generated media is shared across the users' devices. Additionally, the system also incorporates photographs into the shared media pool that are taken by photo systems that are installed on the rides, which are normally part of a commercial service from which visitors may purchase their "on-ride" photo at the exit of the ride. To create their personal photo-story a user selects three images from the mixed media pool to make up the final automic (i.e., photo-story).

8.1.1.1 *Task design in Automics*

There are three tasks the user has to accomplish in order to create an automic:

1. Take photos[2]
2. Caption/annotate photos [3]
3. Combine photos into a photo-story

2 For a detailed interaction flow diagram see figure C.3, p. 305.
3 For a detailed interaction flow diagram see figure C.4, p. 306.

These tasks can be accessed from the applications "home screen" at any time. However, successful use of the application to create an Automic requires a certain order of actions. In order to create one Automic, the user needs to have taken and prepared (i.e. annotated and/or cropped) at least two photographs, and must also have been on one of the rides selected for the experience in order to include the on-ride photo.

Only when these conditions are met can the user combine the three photos into an automic. Hence, despite the seemingly simple tasks, the application imposes some particular requirements on task order and pacing in relation to the visitors' activities within the park. To assist the user in achieving the necessary requirements, location-based messages encourage the user to complete certain tasks at particular places in the park, as detailed in the following section.

8.1.1.2 *Location-based messages*

The location-based messages invite the user to interact with Automics. The locations were chosen with a notion of the activities that are likely to take place at that location as identified in the sensitising study, for example queuing, walking and resting. In addition, the task types were mapped to the locations based on an estimation of the attentional demands associated with those activities. This mapping is achieved by weighing carefully the attentional demands of a task with the expected attentional demands in the environment as inspired by the *Resource Competition Framework* (Oulasvirta et al., 2005b, see p. 28).

The mapping follows the following rationale:

- It was assumed that visitors were receptive to taking photos whilst walking between rides and just having been on rides. For this task, the user only needs to briefly focus their attention on the device's screen (ca. 4 seconds (s)). Arguably, the view of the world through the viewfinder on screen let users re-focus their attention to orienting and navigating in their environment.

- It was assumed that users would be most comfortable with the annotation/caption whilst queuing, as this task requires prolonged touch-interaction with the device (at least 45s). Many people were observed interacting with their mobile devices in the sensitising study, which suggested that the attention required whilst queuing was likely to be substantially lower than whilst walking.

- It was assumed that visitors were receptive to completing the photo-story task in "break zones" (cafés, picnic areas etc.), as not only required it prolonged interaction, but resting was more likely after the initial requirements had been met (having been on a ride and having taken at least two photos).

This mapping was achieved in practice through the use of a bespoke pre-authored location model that serves GPS locations as "hooks" in the physical space to coordinate the tasks of Automics and the theme park experience. The location model mapped discrete GPS coordinates to suggest engagement with the application at places that were identified as photo opportunities, end of ride zones, break zones and queuing zones[4]. The application then triggered notifications for the user based on their location in the park, e.g. it suggested taking a photo when close to a "photo opportunity" or suggested annotating a photo on entering a "queuing zone"[5]. Some of those messages uniquely relate the tasks to the environment by referring to objects in the proximity, i.e. the ride they are queuing for or the object they may take a photo of.

4 See appendix C, p. 304 for a map of the GPS triggers.
5 A detailed system diagram is provided in appendix C, p. 302.

8.1.1.3 *Design objective*

Overall, the utility of the notifications is to coordinate and assist the users with the pacing and ordering of the tasks required to achieve the goal of creating a souvenir photo-story, while at the same time facilitating an optimally interleaved experience of both the theme park and the AUTOMICS application.

8.2 RESEARCH QUESTIONS

Because of its exploratory nature, instead of experimental hypotheses a set of research questions is developed that explores a subset of factors of the model of contextual factors of interruptions (see table 4.4, p. 88). The exploration of opportune moments for interruptions continues the theme of investigating the timing of interruptions in interruption management techniques. Observing the groups in the field enables the study of local contextual factors of the interruption, including the social and the environmental situation. Although it is not the focus of this thesis to evaluate technology, it is inevitable that a study of the application in use also represents an opportunity to evaluate the extent to which the experience of the technology is actually coherent with the designer's objectives. Therefore, in addition to questions relating to the model, a question relating to technology evaluation is posed.

8.2.1 *Questions related to the model*

In terms of the contextual model the factors that were investigated in the field study are the *timing*, the *content*, the *presentation*, the *social* and the *environmental* situation of the interruption.

timing As reviewed in section 4.5 (p. 105), Interruption management systems typically make one of three types of adaptations: they may adapt the *when*, the *where* and the *how* of the interruption (see table 4.2, p. 107). According to this taxonomy, by implementing location-based messages Automics adapts the *where* of the interruption. However, as argued in section 4.5.1 (p. 108) this may also be understood as adapting the *timing* of an interruption by proxy of the user's current place. In the broader context of this thesis, this field study aims to answer the research question (see 4.5.3, p. 110) RQ3: Does the recipient's current place reliably indicate an opportune moment for an interruption?

Note that, with the example of Automics, the locations that trigger location-based messages are deliberately chosen with a notion of the activities that take place at that location, such as walking, queuing and resting. Particularly, the queue zones in the park have been equipped with for example arcade style games along the queues, and are often designed to extend the thrill of the ride to the spectator in the queue, for example by channeling queues beneath the roller coaster's rail. This structuring of the space creates opportunities to become a place (Harrison and Dourish, 1996, see 2.1.4.1 p. 20), spaces invested with cultural expectations, a sense of behavioural appropriateness and so on. This (somewhat predictable) shared understanding allows the development of heuristics about the appropriateness of interruptions and the interactive tasks they prompt based on the likelihood of activities and the attentional demands associated with those activities.

Specifically, the activity of queuing and the endpoints of rides as potentially opportune moments for notifications (i.e. interruptions) are investigated. The sensitising ethnographic study showed that many people seemed to interact with their mobile phones to text and check messages while queuing to combat boredom or to make use of this "downtime" (Durrant et al., 2011). In another study, similar waiting time (e.g. on train platforms) has been coined *dead time* and has seen mobile information workers make use of that time to keep on top of their work by resuming information work or communication tasks (Perry et al., 2001).

The endpoint of rides may be conceptually likened to breakpoints (see 3.4.1, p. 58), as discussed in chapter 7. These moments right after the participants have been on a ride may provide an opportunity to interrupt and engage them before they embark on the next episode of their theme park experience.

This raises the following research question:

RQa To what extent are location-based notifications experienced as contextually opportune by the visitors?

CONTENT The data from the field study will show effects of the content of the notification on receptivity and responsiveness. By suggesting to the user to complete certain tasks in certain places, AUTOMICS also adapts the *what*, i.e. the content of the notification, according to another design heuristic: tasks that require longer periods of interaction with the device's screen are suggested by notifications whilst the user is likely to require less attentional resources to orientate in the environment, for example when queuing and resting.

This raises another question:

RQb To what extent was the completion of the suggested tasks in reality compatible with the attentional demands in the environment?

In addition to these task requests, which represent *actionable* notifications (i.e. they prompt subsequent action), AUTOMICS also notifies its users when another member of the group has shared a new image. It will be interesting to study the effects of both pre-authored location-based task requests on the one hand and notifications that are generated to raise mutual awareness on the other hand. The question that follows from this is:

RQ_c What observations can be made regarding the receptivity to either type of notification?

presentation In addition to content, another relational factor (see 4.4.2) that is investigated is the presentation of the interruption. The audio alert for the location-based notification is different from the alert for the "new image" notification. Also, the vibrating pattern of the notifications delivered in queue zones is more prominent and that notification is repeated every 30 seconds.

In terms of the contextual framework, different alert types represent a variation of the presentation of the interruption (see 4.4.2.3, p. 101). The question that arises is:

RQ_d What effects do these different alert types have on receptivity and their contextual appropriateness?

social and physical enviroment In keeping to the theme of this thesis, rates and temporal aspects of receptivity and responsiveness are measured. However, the emphasis of the study is on qualitative field observations and interviews. In contrast to the experiments presented in chapter 6, and chapter 7, the character of this work is more exploratory in nature. A qualitative focus allows especially for the observation of local contextual factors such as the social and physical environment (see 4.4.1.1, p. 92), factors that could not yet be studied in the earlier experiments because of their distributed setting.

The situated character of the field study provides an opportunity to observe application usage and dealing with the notifications. More specifically, ways in which the interruptions and the social accountabilities are managed in situ may be observed along with the impact of physical factors such as noise, light, and weather conditions.

The research question thus raised is:

RQ_e To what extent does the social and the physical environment mediate the visitors' receptivity to the location-based and awareness messages?

8.2.2 *Question related to evaluation*

Automics adapts both the content and the timing of its interruptions in order to meet the stated design objective (see 8.1.1.3). The interruption technique is evaluated as stated in section 4.1, by discerning its impact on the user's receptivity. The research question that guides the evaluation is accordingly stated as:

RQ_f To what extent does the employed interruption management technique help in meeting the system's design objectives and what is its impact (benefits and shortcomings) on the user experience?

8.3 study design

We evaluated Automics in a naturalistic field study. Three groups of voluntary participants visited a popular theme park in the UK for a day and used the Automics application extensively throughout their day in the park. In addition to the research questions outlined in the previous section, the field study was conducted in order to investigate research into attitudes, potentials and business models for the 'co-creation' of souvenirs. This discussion is outside of the scope of this thesis, but is presented by Durrant et al. (2011).

8.3.1 *Methods*

As described in section 5.2, mixed methods were employed to gather data in situ about the user experience with a focus on the research questions as outlined in the previous section. The data gathered by the various methods then can be triangulated (see 5.2), for example in order to indicate quantitative support for a qualitative observation or statement made by a participant.

8.3.1.1 *Ethnographic Observation*

Two researchers accompanied the group of visiting participants for large parts of their day in the park and recorded their observations by means of photos, video and notes. This was partly in order to document participants usage of the AUTOMICS prototype and partly in order to analyse the data in more detail. The research topics presented earlier were used to create an observation guideline for the field (see table C.1 in appendix C). To enable the discovery of emergent issues the technique of grounded theory (see 5.2) was later applied to the data.

8.3.1.2 *Usage logging*

To continue the technique of analysing temporal aspects of the interruption process, usage of AUTOMICS was logged. Thus, not only the output of the application (photos and Automics souvenirs) could be analysed, but also the process of its creation. Timestamps of every episode of interaction with the applications were recorded for example in order to look at the receptivity and responsiveness to the different kinds of notifications and the ratio of user-initiated compared to system-initiated interaction.

8.3.1.3 *Interviews*

At the end of their day in the park with AUTOMICS the group of participants was interviewed about their experience. The interview was semi-structured and focussed not only on the research questions posed in this thesis, but also on research questions of interest in the broader research project on the experience and potential of 'co-creating' one's own souvenirs.

8.3.2 *Participants*

Participants were recruited from participants of the previous sensitising ethnographic study and via a snowballed advertisement. The first group to test our prototype were Jim (11) and Beth (8) and their grandparents Alf (58) and Helen (59). It was decided in consent with her grandparents that Beth was too young to be given her own device for the day and she occasionally shared the phone with her brother or her grandparents. This group had already participated in the sensitising ethnographic study about a year earlier.

The second group were made up of colleagues James (28), Mike (30), Martha (26) and Kate (28), who were also friends.

The third group comprised members of a nuclear family group, Gary (51) and his three children Jenny (20), Paul (19) and Daisy (14).

Participants were recruited with the goal to cover a wide age spread and different social constellations. Four participants were between 10 and 20 years old, four were between 20 and 30 years old, and three were in their fifties. The mean age was 29, the median 27 years old.

Participants were reimbursed with complementary tickets and their travel expenses were covered.

8.3.3 *Procedure*

The groups were met by the research team at the park entrance on separate days and provided with complementary tickets. Before freely exploring the park and the attractions, they were taken to a room to be briefed and each person was given a Nexus One smartphone running AUTOMICS to use. The smartphone functionality and the application was explained to them, and an example photo-story was shown. They were then invited to use AUTOMICS freely in the park, while they were also told that the device would occasionally notify them to prompt a certain activity depending on their location.

The author accompanied them for the first hour to assist with any technical queries and to observe. Another researcher accompanied them intermittently during the day to observe. Towards the end of the day participants were invited back to the room to discuss their experiences. Each individual was shown all of their automics on a tablet computer and asked to choose one to be framed and taken away as their souvenir.

A final interview was then conducted at a large table with all of the media (including all iterations of annotated photos) present as paper prints. At this point each group was asked to carry out a simple triaging task in which they had to agree the best 12 images, then the best 6, and finally the best 1, that represented their collective day. They were also asked to lay these out in a suitable format and describe why that had been chosen.

The entire procedure lasted for about 5 to 6 hours.

8.3.3.1 *Role*

AUTOMICS was developed and evaluated as part of the project Day in the Park by the Horizon Research Institute co-located at the University of Nottingham. Through discussions with the research team it was decided that substantial overlap between the project's goals and the theme of this thesis justified a collaboration. Throughout the process it was ensured that both the author of this thesis and the project team would benefit from each other's contributions while maintaining a separation of concerns that would warrant a clear individual contribution for this thesis.

The author's role was the design and implementation of the mobile application and the development of research and interview questions and an observation guideline (see table C.1 in appendix A) to illuminate the research questions. During the actual field study, the author took on the role as an observer and documenter and supported the participants if they had questions or problems with the usage of the application.

8.4 RESULTS

In this section the participants' experience of AUTOMICS is quantified in terms of the number of photos, annotations and photo stories produced, and the number and types of notifications participants received. In addition, the acceptance time and acceptance rate is analysed in order to shed light on the question of receptivity and appropriateness of the interruptions. A comparison of self-initiated and system-initiated interactions concludes the section before contextualising and detailing the experience qualitatively in the next section.

8.4.1 *User activity*

Table 8.1 gives an overview of the user activity of the three groups. Group One consisted of three active users and they jointly uploaded 72 images (both raw and annotated and/or cropped) and they created

Group	Raw photos	Annotated and/or cropped	Total shared	Automics created
1	54 (75%)	18 (25%)	72	6
2	98 (65%)	53 (35%)	151	8
3	110 (75%)	36 (25%)	146	12

Table 8.1: Distribution of user activity in Automics by groups.

6 photo stories. The four members of Group Two uploaded 151 images and they created 8 Automics, while the four members of Group Three captured and shared 146 images and they created 12 Automics.

8.4.2 *Notifications*

Table 8.2 shows the number and type of notification each group received during their day. The table shows that the groups received the vast majority of notifications as a result of sharing images. Respectively, 78.8%, 89.3% and 85.9% of the notifications the groups received notified them that a member of the group had shared a new image. The location-based messages that prompted the groups to take a photo ("end of ride", "photo opportunity"), annotate and/or crop ("queue zone") or create a photo-story ("lunch/break") comprised only 21.2%, 10.7% and 14.1% of the notifications the groups received respectively.

8.4.2.1 *Acceptance rate*

Table 8.2 furthermore includes the number and percentage of notifications that were accepted by the users. "Acceptance" here is used to describe the explicit user activity of opening a notification from the device's notification inbox. Throughout this thesis, acceptance rate has been applied as a behavioural measure of the receptivity to interruptions.

Even though the overall acceptance rate is rather low (36.1%) compared to the acceptance rate reported in chapter 6 and 7, subtract-

Group	End of ride		Lunch/break		Photo opportunity		Queue zone		New images		Total	
	(t)	(a)	(t)	(a)	(t)	(a)	(t)	(a)	(t)	(a)	(t)	(a)
1	3	2 (66.7%)	4	3 (75%)	14	13 (93%)	7	6 (85.7%)	104	22 (21%)	132	46 (37.1%)
2	5	2 (40%)	10	5 (50%)	15	10 (66.7%)	8	8 (100%)	317	104 (32.8%)	355	129 (36.3%)
3	7	4 (57.1%)	7	3 (42.9%)	26	14 (53.8%)	15	9 (60%)	334	111 (33.2%)	389	141 (36.2%)
Total	15	8 (53.3%)	21	11 (52.4%)	55	37 (67.3%)	30	23 (76.7%)	755	237 (31.4%)	876	316 (36.1%)

Table 8.2: Distribution of total (t) and accepted (a) notifications in Automics by groups and notification type.

ing the non-actionable "new image" notifications changes the rate substantially. Acceptance rate of location-based messages (all other notification types) was 63.6% in total (groups One, Two, and Three accepted 85.7%, 60.5% and 54.5% respectively). Hence, the acceptance rate to the actionable notifications is similar to the rate observed in the experiments.

ACCEPTANCE RATE BY NOTIFICATION TYPE For the purpose of the evaluation of the location-based notification rationale, the analysis presented here provide empirical evidence of the actual opportuneness of the notifications.

A chi-square analysis of the contingency table between notification type and acceptance (see table 8.3) confirmed that there is an association between the type of the notification and the likelihood that the notification was accepted, with χ^2 (4) = 56.2, exact $p < .001$. A ϕ value of .253 indicates a moderate strength of association[6].

Inspecting the table more closely, the actual frequencies and the frequencies expected by the statistical test[7] indicate that the "queue

6 Moreover, three independent tests for each group confirms the significance of the association for each group. For group One, χ^2 (4) = 41.5, *exact p* < .001; for group Two χ^2 (4) = 22.5, *exact p* < .001; and for group Three χ^2 (4) = 9.9, *exact p* < .05.

7 As stated in chapter 7, in the employed chi-square test for association, the expected count for each cell is calculated using estimates of probability computed from the marginal totals and the total frequency *N* (cf. Kinnear and Gray, 2008, p. 395). Thus, the *estimation* accounts for the differences in *observations* in rows and columns.

Notification type		Accepted? No	Accepted? Yes
Queue zone	actual count	7	23
	expected count[7]	19	11
Photo opportunity	actual count	18	37
	expected count	35	20
End of ride	actual count	7	8
	expected count	10	5
Lunch/break	actual count	10	11
	expected count	13	8
New images	actual count	518	237
	expected count	483	272
Total		560	316

Table 8.3: Crosstabulation of accepted and non-accepted notifications by type.

zone" and "photo opportunity" are notably more often accepted than expected, while the "new image" notifications are notably less often accepted than expected.

Thus, further separate binominal analyses for each of the notification types were conducted to find out whether the users were significantly more likely to accept a notification of a certain type.

The binominal analysis in table 8.4 shows that the notifications of the types "photo opportunity" and "queue zone" were significantly more likely to get accepted at the .05 level. Unsurprisingly, because of their non-actionable content and their overwhelming amount, the "new images" notifications were significantly more likely not to get accepted. Comparisons for "end of ride" and "lunch/break" notifications were not significant.

ACCEPTANCE RATE OVER THE COURSE OF THE DAY The acceptance rate decreased significantly over the course of the day. The analysis was conducted by splitting notifications that were accepted chronologically per person into three parts, and then grouping all the notifications according to which third of acceptance they belonged to

Notification type	Total number	Acceptance rate	Sig. difference in acceptance rate?
Queue zone	30	77%	Yes, $p < .01$ (2-tailed)
Photo opportunity	55	67%	Yes, $p < .01$ (2-tailed)
End of ride	15	53%	No
Lunch/break	21	52%	No
New images	755	31%	Yes, $p < .01$ (2-tailed)

Table 8.4: Binominal analyses of the acceptance rate of notifications by type.

(1^{st}: 108, 2^{nd}: 98, 3^{rd}: 110 104)[8]. Thus, if there is no habituation effect (H0) the number of notifications not accepted in the categories should be equally distributed. This was not the case. Participants did not accept roughly more than 1.5 times the number of notifications (164) whilst accepting the first third of all notifications. During the second third, the number of notifications they did not accept decreased to 123 which might be indicative of a small learning effect. However, during the final third of notifications they accepted the non-acceptances soared to 273, which might be indicative of fatigue or habituation.

A chi-square analysis of the crosstabulation confirmed the association between the day's progress and the likelihood that a notification was not accepted, with χ^2 (2) = 17.1, exact $p < .01$. A ϕ value of .14 indicates a weak association.

In summary, despite an initial rise in acceptance rate towards the middle of the day which might be indicative of a learning effect, as the day progressed and the participants continued to receive notifications, the ratio of accepted notifications declined increasingly.

8.4.2.2 *Acceptance time*

To continue the theme of inspecting the behavioural response to the notifications a closer look is taken at *acceptance time*, the time the user took to accept a notification. Differences in acceptance time by notification type may shed light on the contextual appropriateness of the

8 Thirds differ in numbers because they are computed per person and subject to rounding; one participant had to be excluded as they only accepted two notifications.

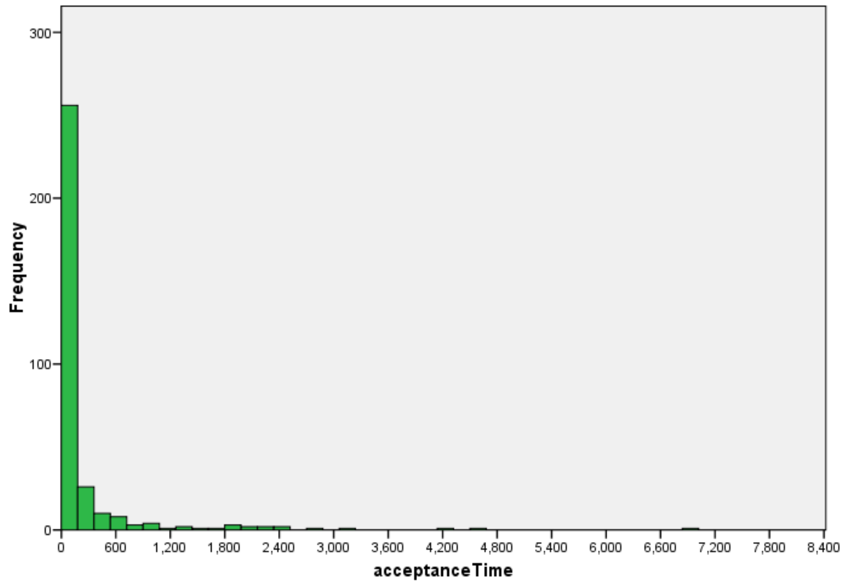


Figure 8.1: The distribution of acceptanc time in seconds (interval = 180s).

different types of notifications. Furthermore, an analysis of the differences in acceptance time as day progresses anticipates a discussion of the participants' increasing habituation to the notifications as the day progresses.

Figure 8.1 shows the distribution of acceptance time. 50% of the notifications (that were accepted) were accepted within 33s, while 75% were accepted within 145s and 90% were still accepted within 559s (9.3min).

In statistical terms, the behavioural data logged from the participants' interaction with AUTOMICS represents within-subjects repeated measurements. For reasons outlined in section 7.4.1 (p. 194), Linear Mixed Models (LMM) were used to analyse the repeated measurements.

ACCEPTANCE TIME BY NOTIFICATION TYPE *Acceptance time* was log-normalised to meet the assumption of normality. Then, *notification type* was included in the model as a fixed effect, and participant was included as a random effect. In addition, pairwise comparison were computed by means of the Bonferroni procedure and are reported

Notification type	Acceptance time in seconds Mean	Median	Std. Error	Log-norm. acceptance time Mean	Std. Error
End of ride	622.06	206.5	358.35	2.22	.26
Lunch/break	1294.78	1345.5	327.51	2.81	.23
Photo opportunity	282.14	105.5	258.53	1.99	.13
Queue zone	1053.64	105.0	253.16	2.33	.16
New images	142.36	21.0	239.89	1.49	.08

Table 8.5: Acceptance time by notification type.

where effects are significant, where log-normalised values are used in the LMM, but mean values are reported in seconds for the sake of sense-making and readability.

Notification type had a significant main effect on *acceptance time*, with F (4, 88.8) = 12.29, $p < .001$, i.e. at least one of the types of notifications was significantly related to *acceptance time*. The estimates of the fixed effects table showed that the "new image" notification was significantly more related to *acceptance time* than the other types of notifications. *Participant* did not turn out to be a significant random effect by the Wald test, with Z = 1.37, $p = .17$.

Pairwise comparison of the mean log-normalised acceptance time for the different types of notifications (see table 8.5) in ordering from lowest to highest showed that

- the *acceptance time* of the "new image" notification was significantly lower than that of the "lunch/break", the "queue zone" (both $p < .001$) and the "photo opportunity" notification ($p < .05$).
- the *acceptance time* of the "photo opportunity" notification was significantly lower than of the "lunch/break" notification ($p <$ .05).

Notable in this pairwise comparison is that the non-actionable "new image" notification had the lowest average acceptance time, while the queue zone notifications had a high acceptance time even though they were predicted to be especially opportune because they were delivered during downtime when the participants were predicted to be receptive to engage with AUTOMICS.

ACCEPTANCE TIME OVER THE COURSE OF THE DAY *Acceptance time* increased notably, however due to high standard errors (see table 8.5) not significantly over the course of the day. To conduct the analysis the accepted notifications were split into thirds per user in order of their occurrence (as explained on p. 231), then a LMM was built with *thirds* as a fixed effect. The LMM did not yield a statistically significant main effect of *thirds* on acceptance time, with $F(2, 315) = 1.6$, $p = .2$. However, acceptance time increased over the course of the day. Whereas the participants accepted the first third of notifications within 159s on average, this increased to 191s for the second third and even to 407s on average for the final third of notifications that they accepted.

8.4.3 *User- vs. system-initiated interaction*

AUTOMICS allows both for the interaction with it to be initiated by the user at convenient times and by the system through location-based prompts that are invitations to interact with it to take photos, annotate them or create photo stories from the shared media pool that has accumulated through the users' previous activity. These different starting points provide for alternative inroads into the user experience that have previously been discussed under the rhetoric of "push vs. pull" (Franklin and Zdonik, 1998) or from a more neutral stance as mixed initiative user interfaces (Horvitz, 1999).

To contribute to the debate of "push vs. pull" the amount of user-initiated activity needs to be quantified and compared against the amount of system-initiated activity, i.e. the interaction initiated by the user at self-selected moments and the interaction prompted by the location-based and "new image" notifications.

Group	user-initiated	system-initiated
with "new image" notification		
1	168 (56%)	132 (44%)
2	255 (41.8%)	355 (58.2%)
3	343 (46.9%)	389 (53.1%)
Total	766 (46.7%)	876 (53.3%)
without "new image" notifications		
1	168 (85.7%)	28 (14.3%)
2	255 (87%)	38 (13%)
3	343 (86.2%)	55 (13.8%)
Total	766 (86.4%)	121 (13.6%)

Table 8.6: Distribution of user- vs. system-initiated interaction.

Table 8.6 summarises the distribution of user- and system-initiated interaction. Under consideration of the "new image" notification the amount of times interaction was initiated or attempted to be initiated by the system is almost equal to the amount of times it was initiated by the user. Leaving aside the non-actionable "new image" notifications that were generated through user activity (see lower part of the table), the dominating ratio of user-initiated activity compared to (attempted or actual) system-initiated interaction becomes apparent. Almost 90% of the time the user initiated the interaction with AUTOMICS. This cleft is further widened by the fact that only 77 of the 121 (63.6%) location-based notifications were accepted.

It was of course the amount of user-initiated activity of taking photos that caused the generation of the vast amount of "new image" notifications. We shall see in the following discussion that the relationship between user activity and generation of notifications was seen as a factor that diminished the user experience of AUTOMICS.

8.5 QUALITATIVE FINDINGS

In the following, the observations during the groups' visits and findings from the final interviews with participants are stated thematically as derived by an analysis led by the research questions as outlined above. Support from the quantitative analysis is given where appropriate to strengthen the findings.

8.5.1 *The experience of location-based messages*

Location-based messages were designed to assist the user in the pacing and ordering of tasks to successfully use AUTOMICS while maintaining a sensitivity to the user's attentional demands in their surroundings (see 8.1.1.3). To what extent was this design objective achieved in practice (RQ_f)? The following sections are structured by the observations for the different types of location-based notifications.

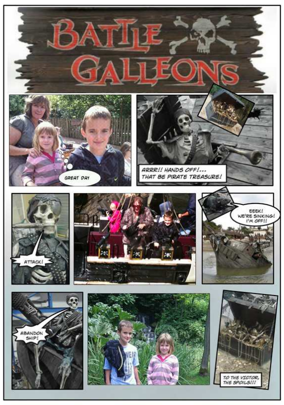
8.5.1.1 *Photo opportunities*

Notifications of photo opportunities instigated mixed responses in people. Generally, the idea of being reminded to take photos on a day out seemed to be welcome by participants, as Kate pointed out:

> I think it was a good idea, because whenever I leave the house for a day out I have really good intentions of taking loads of pictures that day, but most of the time I completely forget [...] if I just got the odd notification I would think actually, that's good because it's reminding me to take a picture. So I think, it was good.

Viewing these kinds of notifications as negative seemed to be associated with instances of people proactively choosing to take and share photos in advance of being prompted, as demonstrated by the quant-

(a) capturing near a photo opportunity



(b) the final automic includes the photo

Figure 8.2: A user captures a photo in the theme park which is later incorporated into an Automic.

ities of images shared (see table 8.1, p. 229). In effect, notifications suggesting to take a photo often appeared to be redundant.

> **Kate:** I think at that point we were all so busy taking photos anyway, we didn't need a reminder.
>
> **Helen:** Well, actually because we were taking a lot of photos anyways, so I don't think we needed the 'this might be a good time' prompt.

Unfortunately, this kind of redundancy was a recurrent theme in accounts of experiencing the AUTOMICS notifications.

Overall, the acceptance rate of notifications of photo opportunities was high (67%), and the median acceptance time was among the lowest (105.5s) of the location-based notifications. It may be fair to suggest that the photo task prompts helped more than hindered users in creating a multi-faceted souvenir that depicts the day beyond ride experiences. The instance in figure 8.2 captures this well.

8.5.1.2 *Queue zone notifications*

Notifications in queue zones prompted participants to caption images with thought- or speech-bubbles and/or crop them to fit the template of an automic. For the successful usage of AUTOMICS this was an essential task, as only edited images could be included in the final automics because of the constraints of the template. Furthermore, fewer opportune moments for this task were anticipated as the task required more focussed and longer interaction with the application than for example the photo task. Queuing was proposed as providing opportune moments for the user to complete the annotation task, but during a busy day a group may only queue for two big rides. Thus, so as not to miss the precious opportunities, the notifications were implemented to vibrate for a longer period of time than the other notification (6 seconds), and to be repeated every 30 seconds were they not attended to.

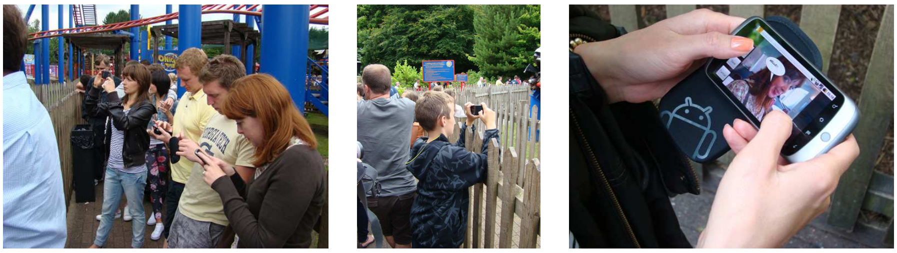
As a result, queue zone notifications had the highest acceptance rate (77%) of all types of notifications, and its median acceptance time (105s) is the lowest of the location-based notifications.

Again, responses about the usefulness of these kinds of notifications were mixed. First, its reminding character was appreciated by James, which supports the notification's utility of coordinating the task requirements:

> I think when we were queuing for Oblivion it was the first time that happened, I thought, 'oh yeah' that could be quite useful to remind me to do that sort of task.

Gary acknowledged implicitly the fit of the task-context mapping, and supports the general usefulness of utilising queuing time to complete the annotation task:

> Probably when you're in the queue, I suppose you haven't got that much to do anyway, apart from talking to other people, [*jokingly*] which is of course, fairly dreadful, ahh..ahem. So, you're more focussed, it's also an opportunity to start and annotate things, so that was quite good actually, because it made the queuing time go quicker.



(a) playing with AUTOMICS (b) capturing (c) annotating while queuing

Figure 8.3: Groups use AUTOMICS while queuing.

Our observations were that AUTOMICS was used almost overwhelmingly while the groups were queuing (see figure 8.3), suggesting that queuing did provide opportune moments to use the application indeed. However, it was this overwhelming usage that often made the notifications redundant, as users were already using the application to annotate or take photos.

> **James:** So I think that, because when we were in the queue we weren't doing anything else but messing with the application we were generating tons of notifications whereas when we were walking around we probably wouldn't. So anything that you received while you were walking around is probably more likely to get your attention.

This may also explain the high mean acceptance time: users may have ignored the notifications that came in while they were already using the application for another task.

8.5.1.3 *End of ride notifications*

Notifications were triggered in the exit zones of rides to prompt the users to take a photo of themselves after they have finished a ride. The quantitative analysis showed a low acceptance rate (53%) and that users were no more likely to accept this type of notification than not to accept it. If they did get accepted, they often were not looked at until some time later (the median acceptance time was 206.5s).

The notifications were originally designed to assist the user with capturing images to fit the sequential structure of the automics, consisting of a "before", "during" and "after" ride image. James mentioned that a scarcity of "after" ride images resulted from the inconvenience of taking photos when coming off a ride:

> The problem is with the threefold, before, during, after. Because most things happened before, while we were queuing, whereas after, when you're just getting off and you're all dizzy and stuff – taking photos then.

So, participants reported feeling temporally unwell, which suggests that just after finishing a ride it is rather inopportune to interact with the device.

In addition to physical inconvenience, dialogue and social engagement seemed to liven up after having finished a ride as well. Our observation was that the time was characterised by lively chat about the ride experience, and then shifted towards a discussion of where to go next. Also, sometimes the group had split up and while a subgroup went on the ride, the other group waited. Then, the moments after they have finished a ride were characterised by coordinating the rejoining of the subgroups, and typically a question along the lines of "how was it?" was followed by an account of the experience.

All in all, the notification to take a photo after just coming off a ride did seem inopportune, and contrary to the initial idea that this moment may be opportune for an interruption as it may be likened to a breakpoint in that it might represent a transition in the participants' theme park experience.

8.5.1.4 *Lunch/break notifications*

These kinds of notifications were triggered if the user spent at least 10 minutes in one of the many designated break areas of the theme park. The 10 minute-rule would keep the notifications from triggering when the group was just passing through the area. The notification invited the user to create an automic by selecting the "before", "during" and "after" images they wanted to to include. Logically, this is the last of the three tasks in AUTOMICS, i.e. the user needed to have already captured and edited photos as a prerequisite to this task. In order to assist the user with the ordering of the tasks as put forward in the design objectives, the notification should not be triggered before the conditions were met. We could assume that the participants' day would follow a certain temporal structure, where their first break was unlikely to occur before they have been on a ride, and therefore were likely to have met all the conditions to create one automic. This was the case for all of the groups.

However, with an acceptance rate of 52% these kinds of notifications were the least likely of all the location-based notifications to be accepted. Also, with a median acceptance time of 1345.5s they did not only have the highest acceptance time, but they were the only type of notification, for which the distribution did not have a long tail as indicated by the mean acceptance time being lower than the median (see table 8.5).

Nevertheless, when noticed at the right time the notification did fulfil its purpose of reminding people to compile a photo-story from the shared photo material as Helen pointed out:

> I did think that the prompt that says 'now might be a good time to go and have a look at the panels and see if you choose the pictures..' Yeah. Because it's reminding you that there is an end product. Which focuses your mind then, because I slipped away into just, 'oh, look at this, Beth looks so cute on this', but actually, it's about what's at the end of this. So it focuses you on this.

In the field, we often had to remind our participants towards the end of their day to create some photo-stories so they could be printed out for them in time. It may be due to a mixture of inopportune timing of the notifications and the fact that the other two tasks of capturing and annotating images dominated the experience that these notifications went undetected or were deliberately ignored.

8.5.2 *The experience of user-generated messages*

The other class of notifications were generated by the system as a result of one of the users sharing an image with the rest of the group, i.e. after they have captured or annotated an image it was automatically shared with the group's other members. Hence, this kind of notification had a different function; it did not prompt action directly, rather, it made the user aware of what the other users of the system were doing.

As the quantitative analysis showed, this kind of "new image" notification comprised the overwhelming majority of notifications that the users got (755, 86.2%). The analysis also showed that unsurprisingly, the acceptance rate of this kind of notification was the lowest of all kinds of notifications (31%). This may be down to the fact that a new notification of the same kind always overwrote old ones, which consequently means for more frequent notifications that they are more likely to get overwritten. Furthermore, the notification was not as directly actionable in nature as the location-based messages which prompted the user to do one of the tasks in AUTOMICS. This is how Jenny explained it:

> Just about missed it every single time, actually. Because it vibrates every single time you get a picture or something. Because every one is taking pictures it vibrates all the time. So you're obviously not pulling it out every single time going 'oh, look more pictures' because then you're going to miss the day. Obviously, we were today, because it's always, a new toy

> and things like that, but, if this was to be something you had all the time you're going to miss things.

On the other hand, if these notifications were accepted, they were accepted within a significantly shorter time than the other types of notifications, with a median acceptance time of 21s. Observations of the group's usage made apparent why this was the case. When one participant started to interact with AUTOMICS, which in turn generated these "new image" notifications this seemed to have a rippling effect on the other members of the group and they also pulled their device out. This effect may have correlated with the notifications rather than having been caused by it, meaning that the group's decision to start using the application was probably also influenced by other factors, such as (self-selected) photo opportunities or downtime whilst queuing. Consequently, participants often had their device in hand when the notification came in, which explains the low median acceptance time.

This then also had the negative side-effect that the notifications were seen as redundant, as the group was together and everyone could see what the others were doing, as also mirrored by this excerpt from the interview of group Three:

> **Daisy:** We were all taking pictures so you knew that it was us that were taking pictures.
>
> **Jenny:** Yeah, I already knew what was going on. If I would've pulled it out all the time...it would've just seemed to an extent quite anti-social as well.
>
> **Gary:** Largely, I missed those [prompts]. [...] like Jenny said, after a while I just started ignoring the buzzing because I just assumed it was another photo coming through. [...] Probably, the idea of prompting is OK, but the notifications for photos, because you kind of know they are coming anyway. Assuming that it's working and you can see people taking photos.

> So you can assume that these photos are coming through. So you didn't need the buzzing for that.

The fact that the "new image" notification had a different audio alert did not make them sufficiently distinguishable from the other kinds of notifications. This seemed to go unnoticed in the noisy and distracting environment of the theme park. As an observer, I hardly ever heard a phone going off, despite being close to the participants. As a result, the number of "new image" notifications then seemed to also have a detrimental impact on the noticing of the other notifications, as Jenny observed:

> The amount of time that we had a 'photo opportunity' and it was for Oblivion or Sonic and we were miles away from it. So, it might have just been slow, but it also might have been the fact that the phone was vibrating so much, so many notifications.

It became apparent that the number of "new image" notifications were a key problem that was echoed across groups to the point where participants became increasingly frustrated and even started to ignoring them. Jenny speculated about ways in which the problem could be addressed:

> I suppose an option could be to notify if there were sort of 5 or 10 new photos would be really helpful and I prefer that to it being constantly going. Also would have been good for if we'd split up, we could've just taken a picture of something, and of course it would have gone to everyone else so you could've gone 'oh, ok, they're by so and so ride' so it's quite handy like that actually.

The idea that the notifications would have greater value as part of remote communication in helping the coordination of the sub-group's activities evokes the notion that a system's context-awareness would benefit from considering the current social constellation of the group.

8.5.3 *Habituation*

Probably as a result of the number of notifications participants received one might conjecture that they habituated to the device's beeping and vibrating and paid less attention to it as the day progressed. This is apparent in the decreasing acceptance rate and increasing acceptance time as the day progressed as outlined in the quantitative analysis. Also, this is supported by the participants' statements in the group interview, as this excerpt from group One shows:

> **James:** I think at that point we had pretty much habituated to the phone making noises and vibrating at us, because it was doing it so much. I mean you would look at them when you were taking the phone out, but you didn't take it out to look at them.
>
> **Mike:** It got the context right, when it was asking the question, but there was just too much stuff going on. At that point, like James said, you just habituated to it. Just ended up ignoring it.
>
> **Kate:** I think it might have been useful if it wasn't vibrating every time someone had uploaded picture. We would've noticed it a lot more, I think. After a while you sort of got used to it and didn't bother taking it out.

So, as these quotes suggests the amount of "new image" notifications led to ignoring all kinds of notifications as they were not distinguishable by their presentation, or it led to checking them "in batches" at convenient times. This may have meant that the location-based notifications were viewed at a later time where they appeared to be incongruent with the current environment.

8.6 DISCUSSION

This section considers the implications of the presented findings in relation to the original design objective (see 8.1.1.3) of the system, reflecting upon some of the prominent issues that they raise and relating these issues back to wider HCI concerns with interruption management systems. Hereby, the research questions RQ_{a-f} (see 8.2) provide the lens through which the findings are discussed.

8.6.1 *Coordinating tasks in* AUTOMICS

Participants voiced their appreciation that the location-based notifications helped them to remember to take photos of photo opportunities and annotate them while queuing. The idea to leverage downtime while queuing to "push" authoring tasks to the users seemed to be welcome by users and suggests that the mapping of the tasks to phases of queuing worked well to foreground the AUTOMICS experience during relative 'downtime' of the theme park experience. We arrive at this conclusion by triangulating our observations and participants' statements with the quantitative findings that suggest higher acceptance rate and lower acceptance time for these kinds of notifications. With regards to the broader discussion of opportune moments for interruptions the notion of actively sensing downtime to prompt engagement with a system may hold interesting opportunities for future system design. *Dead time* had been shown to be used for productivity tasks (Perry et al., 2001), and future systems may utilise the approach applied here to suggest to the user to take on short productivity tasks while sensing that they have 'downtime', e.g. while they are idly waiting at a bus stop or on a train platform.

Participants were less receptive to the notifications delivered when participants came off a ride or the ones delivered in break areas. The moments when participants came off a ride were characterised by more lively social engagement as the thrill experiences were exchanged and the next destination was negotiated; at times with a subgroup that had not gone on the ride. This was accompanied for some by a mild nausea from having been physically shaken, which did not favour interaction with an application on a small device. Notifications of this kind were significantly more likely not to be accepted, and if they were accepted, they were only after longer delays. The same is true for notifications delivered in break areas. The observation was that participants may have forgotten that the task to compile their shared media into a photo-story was the vital final task to create an automic as the experience was dominated by capturing photos and annotating and/or cropping them. Participants often had to be reminded to author their automics by the researchers towards the end of the day.

With respect to the design objective (see 8.1.1.3) of coordinating the pacing and ordering of tasks required to achieve the goal of creating photo-story souvenirs while at the same time optimally interleaving the experience of being in the park and the experience of the system, one may conclude that the notifications of photo opportunities and whilst queuing contributed to the achievement of the objectives, while the notifications at the end of rides and in break areas left something to be desired.

8.6.2 *Redundancy*

The intended mapping of the annotation task to queuing zones was so intuitive that it did not require a reminder for the participants as they had often already started to use AUTOMICS as soon as they started queuing. A similar redundancy of notifications may be seen for photo opportunities. In many cases, participants were capturing images so avidly that they hardly needed to be reminded of photo opportunities.

Participants' avid use of the application at what were anticipated to be opportune places supports the rationale of the task-location mapping. However, the observation that this can lead to redundant notifications entails an important lesson. What can be learnt from this then is that a consideration of the participants' current and past activity should also be used to determine whether a notification is appropriate or not. If the participant is currently using the application or has been using it to a sufficient extent to be likely to achieve the goal of creating a souvenir, then notifications to coordinate the tasks to achieve the goal are no longer required.

This finding may be related to the contextual model (see 4.3), despite that at first sight the model appears to focus on a single interruption in time. Redundancy and the following observations on habituation may be explained by the recipient's embodied involvement at the centre of the model. Involvement is also characterised by past experience and memories (see 4.4.1.2, p. 94). In more technical terms, the current experience is shaped by the retentions of the past and protensions of the future (see 2.1.4.2, p. 22).

8.6.3 *Habituation*

The same avid "voluntary" usage of the application that led to redundant notifications also led to an over-abundance of messages notifying a user that another member of the group had shared an image. This happened automatically when a user had taken a photo or had edited one and saved it; all of the media was shared across all of the devices in the group. Although media sharing was highlighted by many users as the most beneficial element of the application, the notifications may have been of greater value in the case that the group had split into subgroups that were physically distant. Effectively, this could have supported the visitor's mutual awareness of their whereabouts in the park, and may have become a useful tool to coordinate their rejoining. Being implemented without this sensitivity, notifications of new shared media were not only redundant, as users could observe each other creating the media, but led to habituation to the beeping and buzzing of the device that led to frustration and eventually to participants ignoring all notifications for some of the time.

Again, this latter argument evokes a sense that the notification mechanism would have profited from a real-time sensing of the group's current status, e.g. if all members of the group were present, there would be no need to notify everybody of newly shared images. The observation that a group's engagement in conversation, for example after having finished a ride, indicated an inopportune moment for a notification also suggests another way in which sensing of the group's status might be beneficial to infer (non-) availability. For example, speech detection has been used by Sawhney and Schmandt (2000) and Vertegaal et al. (2002) to detect if a user is in a conversation and may not be available to incoming messages.

Real-time sensing of the members collocated presence and their conversational status may have informed a more sensitive notification technique that would rely on assessing the *current* situation before notifying. In this way, flooding with notifications and the resulting habituation and eventual ignoring of them may have been avoided.

8.7 CONCLUSIONS

This chapter presented a study of a real-world application that employs an interruption-driven interaction paradigm. The study showed that the number of notifications generated through an unanticipated high usage of the application led to participants' habituating and increasingly ignoring the notifications as the day progressed. The following subsections present the conclusions with regards to the research questions related to the contextual model (RQ_{a-e}) and the evaluation of AUTOMICS' design objectives (RQ_f).

8.7.1 *Conclusions related to the contextual model*

With regards to *timing* (RQ_a), the finding that queuing 'downtime' was generally seen as an opportune time to engage with AUTOMICS (see 8.6.1) may have wider implications for the design of applications that seek to proactively engage their users. The use of AUTOMICS was reported to make "queuing time go quicker", which highlights the potential for future applications to leverage this 'downtime' as signalled to the system by its knowledge of the place it is used in, for example to "push" productivity tasks to the user or to deliver messages.

An essentially appropriate mapping of AUTOMICS' tasks to the attentional demands in the enviroment (RQ_b) however, led to place-based notifications often being redundant, as participants were often already using the application and did not need a reminder (RQ_c, see 8.6.2). "New image" notifications were also observed to be redundant, as participants could observe each other creating the images, which suggests an interactional effect of the social enviroment (RQ_e, see 8.6.3). Whether a group was in a conversation seemed to also mediate receptivity to the notifications, while the environmental distraction (visual and auditory) appeared to make differences in the alert types of the presentation indistinguishable (RQ_d), suggesting the environmental factors of the model exhibit the assumed interactional effects on the recipient (see 4.3, p. 86).

8.7.2 *Conclusions related to the evaluation of* Automics' *design objective*

It was argued that the utility of the notifications was to coordinate and assist the users with the pacing and ordering of the tasks required to achieve the goal of creating a souvenir photo-story, while at the same time facilitating an optimally interleaved experience of both the theme park and the Automics application (see 8.1.1.3). While it may be fair to say that overall the notifications helped the participants to create their souvenir (RQ_f), the discussion revealed problems of redundancy of and habituation to notifications, which may have been avoided had the system considered:

a) the individual's current system activity – to avoid redundant notifications in case they are currently using the system;

b) the group's historic pattern of activity – to glean if the user may already have been on track to create a photo-story in any case;

c) the group's members' presence and conversational status – to avoid redundant notifications of user activity if it can be mutually observed because of collocation and to avoid inappropriate notifications while socially engaged.

Consequently, the implications for design of future iterations or similar experiences that arise when taking on the perspective of an evaluator may be formulated by altering the initial design rationale by extending them by *contextual conditions*:

- Prompt users to assist with coordination of pacing and ordering of tasks to achieve the system goal *only if minimal goal requirement is not achieved and the group is not in a conversation.*

- Prompt to suggest interaction with the application at opportune places *only if they are not using it already and the group is not in a conversation.*

- Notify users of newly uploaded images *only if the group has split up, and then not too frequently.*

8.8 summary

This chapter concludes Part III of this thesis – the empirical work done to arrive at a deeper understanding of what it means to be receptive to interruptions, and the potential benefits and shortcomings of temporal methods to coordinate interruptions.

While Experiment 1 reported in chapter 6 confirmed that the content of an interruption has a significant effect on our receptivity, the experiment also showed that scheduling one's own interruptions in advance is not a sufficient method to achieve better receptivity than delivering interruptions at random times. Experiment 2 took on some of the shortcomings raised by Experiment 1, investigated the temporal details of the interruption process, and provided support for the utility of the method to coordinate interruptions based on the user's phone activity. However, the findings not only show that opportune moments for interruptions at breakpoint did not always collocate with endings of episodes of mobile interaction; the findings also revealed some of the situated complexity that mediates receptivity and responsiveness to the interruptions. The field study was conducted in order to observe some of that social and physical complexity surrounding notification-driven interaction with a real-world application.

The field study has taken the investigation in this thesis into new territory. For the first time in this thesis, a temporal method to coordinate interruptions (by proxy of location) was applied to a real-world application, and studied as participants used the application with a purpose other than to take part in an experiment. This development presented a logical step in line with the pragmatic attitude of this thesis, the practical justification of the focus on *timing* (see 4.5), and the grounding in methodological approaches used to study interaction 'in the wild' (see chapter 5). Inevitably, this development shifts the focus from a scientific investigation of a technological intervention to a practical design exercise. With the study of AUTOMICS, the potential of temporal coordination of notifications in system design comes into focus, rather than the 'purer', experimental study of receptivity to interruptions.

This discovery has implications for the following Part IV of this thesis, where the thesis' contributions are discussed in detail and related to one another, and the overall conclusions are drawn. The study of AUTOMICS in particular is drawn upon to inform the discussion of implications for design.

Part IV

SYNOPSIS

9

DISCUSSION

It is now time to weave the separate discussions of the three different studies together into a broader discussion of the findings brought forward in this thesis and their implications for context-aware computing and the design of interruption management systems in particular.

9.1 OPPORTUNE MOMENTS FOR INTERRUPTIONS

This thesis posed the apparently simple question: What is an opportune moment for an interruption?

The empirical studies in this thesis suggest that the prediction of opportune moments for interruptions is a complex endeavour. The studies reveal diverse ways in which the contextual factors captured by the model (see 4.3, p. 86) interact to mediate the actual receptivity to interruption in situ. To exemplify, the following findings thwart the apparent simplicity of the principal research question.

Although Experiment 2 showed that participants were quicker to accept and respond to interruptions triggered by phone activity, their actual perception of opportune moments surfaced from more complex factorial interaction, for example by a) the contextual task coherence (e.g. fit of the cognitive and attentional demands in the environment with the demands of the interactive task), b) the social appropriateness of completing the task (e.g. taking photos in meetings) and c) affective factors (e.g. sense of fun).

Opportune moments for interruptions have been related to breakpoints in the user's primary task in the literature (see 3.4.1) that relied on controlled laboratory studies.

However, a central methodological choice in this thesis was to sacrifice control of the user's primary task in the laboratory in order to achieve ecological validity in real-world studies. The study of the situated effects of adapting the interruption timing on the user's receptivity were in the focus of this thesis (see 4.5, p. 105) for two primary reasons: 1. The goal was to study the systematic ways in which opportune moments arise, and 2. the adaptation of the interruption timing was argued to be the most practical interruption management strategy.

With regards to the effects of adapting the interruption timing, Experiment 1 compared participants' receptivity to interruptions they *scheduled* themselves (by time of day) to a random baseline and showed that scheduling one's own interruptions in advance does not imply that one is more receptive to the interruptions than when they are delivered at random times. Both the system in Experiment 2 and the mobile souvenir system AUTOMICS implemented a *mediated* coordination method in that the interruptions delivered by the systems are mediated by aspects of the user's context. The system in Experiment 2 delivered its notifications after users had completed an episode of interaction with their mobile phones, and AUTOMICS delivered its notifications as triggered by the user's current location in the theme park. Experiment 2 showed that users were significantly quicker to accept and deal with notifications triggered by their mobile phone activity compared to randomly triggered notifications, whereas their self-reports did not reflect a systematic preference. Observations of AUTOMICS in use showed that notifications delivered in the user's 'downtime', such as when queuing, were preferred and were significantly more likely to be accepted.

This thesis makes the following contributions towards an answer to the question of the constitution of an opportune moment.

The analysis of the participants' behavioural response suggests that contextual mediation of the timing of interruptions (by phone activity in Experiment 2 and opportune places in Automics) are more likely to lead to interruptions at opportune moments than when participants schedule their own interruptions.

In addition, observations and interview statements qualify the ways in which opportune moments arise from the situated interplay of local and relational contextual factors. From Experiment 2, it was concluded that phone activity-triggered notifications resulted in one of three distinct cases that influence whether the users were receptive to the notification. The worst case was when the user had already instituted a new task or had entered the cognitive planning phase (Miyata and Norman, 1986) for that task; the intermediate case was when the user had only finished a sub-task within a larger activity, e.g. had finished one phone call in an activity that consisted of many phone calls; and the best case was when a task had just been completed and therefore a coarse breakpoint within the user's activity context had been reached. The field study showed that users welcomed the idea of leveraging their downtime (e.g. while queuing) to engage with the interactive experience, indicating opportunities for future research to more systematically investigate how downtime may be sensed and what it may be used for, for example to push productivity tasks to the user.

Moreover, the concept of readiness-to-hand of equipment (cf. Heidegger, 1962), as introduced to HCI by Winograd and Flores (1985) and Dourish (2001) has also proven fruitful in thinking about the constitution of an opportune moment for an interruption. In contrast to Winograd and Flores (1985) and Dourish (2001) this is not just achieved by considering how *unreadiness-to-hand* becomes visible in damage or *breakdown*. Rather, it is the understanding of how *obstacles* can 'get in the way' of our concern where the phenomenon of interruption is grounded. For example, if the device was still *ready-to-hand*, e.g. to mediate a phone call, a notification would likely be perceived as disruptive, as an obstacle that 'gets in the way' of our current concern, i.e. the current task or activity that we are enaged in. On the other hand, having the device still in hand may provide a more opportune moment than when the device has been put away. However, whether a device is *ready-to-hand* is not determined by its physical orientation, but by the fluent ways in which the user may in one moment use the device to enable their current concernful dealings, and in the next experience an interruption as an obstacle that brings to the fore the device's *unreadiness-to-hand* for their current concern (cf. 2.1.5, p. 23). The fact that our current technological tools enable a multitude of concernful dealings makes it ever more likely that their notifications create obstacles for our current concern.

Whether an interruption is perceived as opportune is determined by the nature of the recipient's involvement in the here and now (see 4.4.1.2, p. 94), to which context provides "the horizon within which the user makes sense of the world" (Svanaes, 2001, see 2.1.3, p. 18). The difficulty for any endeavour to build context-aware systems is that rather than being entirely represented in the world, context is created as an individual interactional achievement in the moment (Dourish, 2004, see 2.1.3). This follows from the acknowledgement of our embodied intentionality towards the world (Merleau-Ponty, 1962, see 2.1.5), which is fundamental to our way of *being-in-the-world* (Heidegger, 1962, see 2.1.5).

9.2 MODEL OF CONTEXTUAL FACTORS

In the face of adopting this holistic view of context, how is it justifiable to break down context into components? Components such as the contextual factors that make up the model introduced as the conceptual core in this thesis. The challenges of contributing to the field of HCI often run along the clefts that divide the disciplines it proclaims to unite. Engineers often seek to understand problems in terms of system requirements, often having to simplify along the way, having to deal with constraining factors such as limitations in technologies and resources. Social scientists view the implementation of systems as interventions which are worth studying and observing for their goal of understanding people's interactions, but only reluctantly are willing and able to give recommendations or draw implications for design (Dourish, 2006a). In the past, this has famously caused an internal team of ethnographers to "fall out" with their employer (Suchman, 1985), and continues to excite the research community until today (Crabtree et al., 2009).

Hence, the justification for the proposed model of contextual factors (figure 4.4, p. 88) arises from the pragmatic attitude towards the disciplines this thesis draws upon to contribute to knowledge in HCI. The breakdown of context into contextual factors enables the systematic study of controllable "chunks" of context, and facilitates the component-based engineering of systems that either take these contextual factors as input by sensing them, or adapt them as a result of their contextual inference.

The model reconciles the seemingly disparate views of cognitive psychology, context-aware computing and phenomenological philosophy by centring on the recipient and their subjective perception of the interruption process. The nature of their current involvement in the world, which includes their current mental state, emotions and affect, beliefs, desires, their collected experiences and their cognitive and physical activity, guides how receptive they are to an interruption as an interactional achievement within their current local context, i.e. their social and environmental surrounding. At the same time, relational factors that account for the recipient's relation to the sender, their relation to the content, the medium and the presentation of the interruption mediate and influence the individual's receptivity.

The model has been synthesised from an extensive survey of the literature (see chapter 4), but also iteratively refined from the findings from the studies presented in this thesis.

9.2.1 *Empirical contributions to the model*

The empirical work comprising Part III of this thesis supports the veracity of virtually every factor in the model of contextual factors (see table 9.1, p. 262) – if not always as expected. The table shows the contextual factors and the section(s) in which the factor is synthesised from the literature, and the extension to Lasswell's formula. Moreover, the table shows the manipulations (or instantiations) of the relational factors in the design of Experiment 1 and 2 and the system deployed in the field study. Where Experiment 1 and 2 only controlled content and timing – the other factors were held constant by design – in Automics, the presentation and the sender (indirectly) also varied by notification type. Furthermore, the table shows the kind of data gathered to understand the interaction of and with the local contextual factors. The most obvious results are also stated.

In summary, the tables' columns provide the contextual factors and their empirical manifestations in this thesis, whereas the rows provide the coverage of the contextual factors in relation to the three studies.

Row	Contextual factors of the interruption process								
Extension of Lasswell's formula	Who	says What	In What channel	How	When	Where	Who else is here	To Whom	With What effect?
Factor family	Relational factors					Local factors			
Factor group	*Pertaining to the interruption itself*					*Environmental*		*Embodied*	*Emergent*
Factor	Sender	Content	Medium	Presentation	Timing	Space & Place	Social environment	Recipient	Outcome
Related work presented in	4.4.2.1	4.4.2.2	4.4.2.4	4.4.2.3	3.4, 3.6.3	4.4.1.1	4.4.1.1	2.1 and 4.4.1.2	throughout chapter 2, 3 and section 4.4
Coverage of this thesis	Manipulations (i.e. controlled by design)					Data (i.e. observed/measured)			Main findings
Experiment 1 chapter 6	– constant – system	genuine information (good vs. bad)	– constant – SMS	– constant – dependent on user's own phone	random vs. scheduled	some anecdotes	some anecdotes	self-reports & measure of response time & rate	significant differences in self-reports for good vs. bad content
Experiment 2 chapter 7	– constant – system	task type (mult-choice vs. free-text vs. photo)	– constant – notification and application (Android)	– constant – auditory/vibrotactile & inbox (negotiated)	random vs. mediated (triggered by phone activity)	interview statements	interview statements	self-reports & measures of acceptance, decision task and response time & rate	significant differences in behavioural measures for timing strategy
Field study chapter 8	system vs. user (indirectly, auto-generated by user activity)	awareness (new image) vs. task (photo opp. vs. caption vs. photo-story)	– constant – notification and/or application (Android)	differing auditory/vibrotactile for new image vs. caption vs. other tasks & inbox (negotiated)	self-selected vs. mediated (triggered by location)	observations & interview statements	group observations & interview statements	observations, self-reports & measures of acceptance time & rate	observed preference towards foregrounding experience in 'downtime'

Table 9.1: The model of contextual factors of interruptions and its grounding in related work and this thesis' empirical contribution to the model.

9.2.1.1 *Local contextual factors*

At the heart of the contextual model of interruptions is the *recipient* herself – the recipient's current involvement in the world shapes their receptivity to interruptions. Unfortunately for the design of context-aware systems, it is also the hardest factor to sense. As Greenberg states,

> "External things – the artifacts, the physical environment, the people – are relatively simple to capture [...], internal things – individuals' interests in that contextual setting, their history of interaction, their current objectives, and the state of the activity they are pursuing – are extremely difficult to capture" (Greenberg, 2001, p. 262).

As argued throughout this thesis, involvement is a highly dynamic and volatile affair that may change from one moment to the next without any 'sense-able' change in the environment. Experiment 2, for example, has shown how the participant's intention to make another phone call renders the ending of the first phone call inappropriate for an interruption. The structure and shape of cognitive activity, and breakpoints within it to pinpoint opportune moments may be among the hardest things to measure or predict for context-aware systems. Yet, an accurate estimation of the recipient's current involvement as a result of their physical and cognitive activity, their history, emotions, affect, beliefs, desires and intentions would appear likely to be the most powerful and accurate way of predicting an opportune moment for an interruption.

With regards to the environmental factors, the study of AUTOMICS showed that heuristics about the matching of interactive tasks to certain *places* – i.e. occasioned and appropriated spaces (Harrison and Dourish, 1996) such as queuing areas in the theme park – according to estimates of the attentional resources required to navigate the environment might be an effective way to gauge opportune moments for participants to interact with the application.

The *social* situation was reported by participants to be an important factor in mediating whether a moment was opportune to tend to an interruption or to complete a task, for example to take a photo whilst in a meeting was reported as inappropriate (Experiment 2). Observations of AUTOMICS in use suggested that periods of increased social engagement in the group, for example when chatting about the experience of rides, indicated less opportune moments to be interrupted, and co-location of group members meant that notifications about other members' activity were redundant because it could be observed.

9.2.1.2 *Relational contextual factors*

With regards to the relational contextual factors, the field study of AUTOMICS showed how the notifications were perceived to have potential social value in creating mutual awareness of the location of others in the theme park. This became apparent when the system automatically generated and sent notifications that inform others that another member of the group has shared a new image. This points to the importance of the relational factor *sender*.

The *content* of the messages in Experiment 1 had a significant impact on the receptivity self-reported in situ, and further analysis revealed how the factors *entertainment*, *relevance*, *actionability* and *interest* in the genuine content were all positively correlated with higher receptivity of the message. The *task type* sent as the content of the interruption in Experiment 2 influenced the *decision*, *task* and *response time*, and the likelihood of completion.

The *medium* employed to deliver the interruptions was held constant for each of the studies; for the first one, it was SMS, while for the second and third one, it was a native application that triggered a notification and inbox-style interruption.

The *presentation* of the interruption in the studies followed what McFarlane has found to be the superior method for coordinating an interruption (McFarlane, 2002): the *negotiated* method according to which the user is notified but tends to the interruption actively by opening it from their inbox. Furthermore, the usage of different audio-alerts and vibro-tactile patterns in AUTOMICS to indicate the different types of notifications was not detected by participants in the noisy and distractive environment of the theme park.

Strategies to adapt the *timing* of the interruptions were identified as the most practical adaptation for interruption management systems (see 4.5). The investigation of the effects of adapting the timing of interruptions present the core contribution of this thesis.

On the one hand the studies showed that contextual sensing of phone activity (Experiment 2) and opportune places, such as queuing zones (field study), lead to a better approximation of an opportune moment than when participant simply schedule the interruptions (Experiment 1). On the other hand, the observations and the qualitative findings point out that the perception of an opportune moment arises out of the situated interplay of several local and relational contextual factors, as highlighted in the following.

9.2.1.3 *Factorial interaction*

Qualitative analysis often revealed the interactional interplay of several factors jointly contributing to whether an interruption was perceived as opportune. In Experiment 1, one participant reported that they felt that the message was more timely if they were more interested in the content of the message, which points out the interaction of the factors in determining an interruption's appropriateness.

Experiment 2 showed that the coherence of the modality of the interruption task and the modality of the previous interactive episode contributed to the assessment of its timeliness. For example, if they had "the keyboard out" and were already typing it seemed less of an imposition to answer the free text question. Furthermore, the interaction of the cognitive attentional demand and the task type may have rendered an interruption as inopportune. Whereas the photo task and the multiple choice task consumed attention for so little time that the tasks were feasible to do while walking, the free text task consumed attention for too long and thus was often delayed.

The study of AUTOMICS emphasised how the the social and environmental surroundings mediated and gave rise to opportune moments to interact with the mobile souvenir system.

It may be concluded that a situated interplay of dynamically changing contextual factors determines the individual's receptivity and responsiveness to interruptions. Therefore, it appears that an automated prediction of opportune moments for interruptions can never be 100% accurate, and that an optimal prediction can only be achieved by combining the latest software and hardware sensors, the most advanced content analysis, machine learning, and statistical reasoning techniques to capture and analyse all of the contextual factors all the time in an effort that may not be practically feasible. In the following, the implications for the potential of practical systems design are developed further, outlining the limitations and opportunities for systems that employ an interruption management technique.

9.3 IMPLICATIONS FOR DESIGN

The practical implications for the design of interruption management systems are discussed along the lines of the proposed application domains. It was proposed in section 4.5 that both systems that aim to coordinate personal communication and pervasive experiences may benefit from an interruption management technique. Implications for these two application domains will be considered in turn. In more detail, these two application domains include:

1. Systems that manage interruptions from the recipient's existing communication network where the sender may either be a person in some form of social relationship with the recipient or other information aggregators such as web services or systems that deliver information proactively (Lei et al., 2007).

2. Systems that initiate interaction proactively to engage the recipient in a cultural, mobile, playful or pervasive experience, where interruptions are prompts for action, such as game messages via SMS (Flintham et al., 2007) or messages via a custom system that senses the user's current location (Cheverst et al., 2000; Rowland et al., 2009).

9.3.1 *Mediating interruptions in personal communication*

The design of a system that mediates the interruptions from the recipient's existing communication network may be an especially challenging endeavour. For example, Dabbish and Baker (2003) studied the complex process by which administrative assistants mediate the interruptions for their supportand. Assistants carefully weigh against each other the importance of the interruption, the importance of the interrupter, and the supportand's "interruption threshold". In terms of the contextual model, in order to decide how to handle the interruption the assistant interprets the importance of the sender and the content of the interruption and compares that to an estimation of the recipient's current availability.

Thinking about the design of a system that aims to replicate this seemingly simple technique of interruption management then quickly confronts the designer with seemingly insurmountable technical challenges. How does a system determine the importance of the content of the interruption? Even if the most advanced content analysis is applied, is it not likely that the system may get it wrong at times? What if that misinterpretation leads to a deferral of a message about an emergency in the family?

Equally challenging for the system designer will be to make the system infer the importance of the sender of the interruption, or to determine the availability state of the recipient. This is not to discourage system designers and computer scientists to think about these "hard problems". In fact, these challenges often motivated the related research on interruption management presented throughout this thesis, and the many invaluable contributions it makes to knowledge. However, a discussions of the risks and limitations of the things we build may be as important as the discussion of its benefits and advantages.

The concerns about building systems that manage interruptions for us have been voiced from a broader perspective about context-aware systems in general (Bellotti and Edwards, 2001; Bellotti et al., 2002; Greenberg, 2001). Greenberg (2001, p. 265) recommends that

> "[...] context-aware systems should be fairly conservative in the actions they take, should make these actions highly visible, and should leave "risky" actions to user control."

In their critical essay on context-awareness, Bellotti and Edwards (2001) propose that two key features that empower the users themselves to make informed decisions on context must be supported in any such system. *Intelligibility* of these systems warrant that users are presented with information to make sense of the system's processes and results of sensing and inferring, and the resulting actions it proposes to take on the user's behalf. *Accountability* must be enforced particularly by systems that mediate interpersonal relations, so that mutual accountability is guaranteed. It appears that these principles are especially relevant for the design of interruption management systems that mediate personal communication.

Bellotti and Edwards (2001) and Greenberg (2001) suggest several design recommendations how these principles might be achieved in practice, which may for example be adapted to a system that mediates interpersonal communication by deferring interruptions in the following way:

INFORM: Inform the sender that their message delivery is delayed by the system until the recipient is receptive, or a maximum of X minutes.

FEEDBACK AND FEEDFORWARD: Provide information on what the system does (system state) and why (to provide intelligibility to create trust) on demand to the recipient and the sender, e.g. on queuing of messages and message status.

CONTROL: Delegate control to resolve conflicts to the user in case of uncertainty if a message is urgent/important etc. and allow recipient to request messages from queue on demand.

ADJUST: Allow recipient to adjust which contextual data is collected/-processed, and facilitate user feedback and (machine) learning of individual user preferences.

However, systems that mediate interruptions from the recipient's social network interfere with interpersonal communication. Bellotti and Edwards (2001, p. 199) conjecture that "mediation between people is an ambitious and potentially threatening aim for context-aware systems". Even if our system is 99% accurate, the scenario of an incorrectly deferred emergency message hints at the potential implications of a false positive or type I error (in this case, falsely deferring an urgent message). As Greenberg (2001, p. 266) points out, "a single inappropriate system action may be enough to preclude people from using it further".

As a consequence, there are several ways in which a designer of such a system could deal with these inherent risks. It was suggested in the discussion of chapter 7 that the application could be limited to the mediation of interruptions in the user's *second order* communication network, a user-definable list of social and location-based networks and other information aggregators about which the user is confident that they will not transmit any communication of immediate, vital importance.

Another way to deal with the inherent risk imposed by designing a system that manages interruptions from the user's personal network is not to do it at all. The discussion now turns to the application domain suggested at the beginning of this section that leaves the designer with more freedom, and less risks involved in their choices: the design of systems that initiate interaction proactively to engage the recipient in a cultural, mobile, playful or pervasive experience.

9.3.2 *Interweaving experiences through contextual notifications*

It was argued that interruption management strategies could also come to play in the design of interactive pervasive experiences, to inform and deliberately design for the ways in which they for example prompt player engagement by SMS (Flintham et al., 2007), or trigger audio clips for the participant near relevant places in a cultural experience played by cycling through a city (Rowland et al., 2009).

Notifications may have different functions in the design of these interactive experiences. For example, in *Day of the Figurines* (Flintham et al., 2007), game messages confirmed the player's actions (feedback), notified the player of game reactions to their activity (response), conveyed game events and storyline (narrative), informed of other player's actions, and reminded the player once a day of the game status (Flintham, 2011, personal communication).

Here, particular interest is paid to the ways in which notifications foreground the interactive experience whilst the user goes about other activities in their everyday life. Examples of how players achieve a balanced interweaving with their everyday life include playing a location-based game during their commute to work (Bell et al., 2006), or managing their accountabilities to their social surrounding (Tolmie et al., 2008). Accounts like these have recently contributed to the synthesis of the design framework of *interactional trajectories* (Benford and Giannachi, 2008; Benford et al., 2009) as a conceptual tool for designers to think and reason about the continuity of the experience in spite of the seemingly disjoint hybrid structures of the experience. The framework is designed to capture the individual routes through interactive experiences that combine multiple roles and interfaces and that span space and time. It shows how trajectories can be steered, shaped and interwoven with one another (Benford et al., 2009).

Automics is another example of a kind of interactive experience that is interleaved with an ongoing experience, namely that of being in a theme park. Its design uses notifications as a feature to steer and shape the user's activities within the application. As stated in section 8.1.1.3, the design objective for the notifications was to coordinate pacing and ordering of application tasks while facilitating an optimally interleaved experience of the theme park and the Automics application.

Thereby, two kinds of challenges are addressed that routinely arise when designing mobile, pervasive experiences: *temporal* and *pervasive* challenges, as elaborated in the following. First, notifications that encourage the user to complete certain tasks of the application may assist the participant in adhering to a required order, and may help them to pace their interaction with the task – a mismatch in pacing of task and user has been reported as a common cause of breakdowns in interaction (Dix, 1993). Secondly, the objective tackles the *pervasive* challenge of enhancing the experience of visiting the theme park without intruding upon it or distracting unpleasantly from it by making the system aware of opportune moments to become proactive.

In order to interleave the system experience with the user's ongoing real-world experience, the latter needs to be studied and analysed for opportunities for interleaving. To inform the design of AUTOMICS, the "pure" theme park experience was studied in a sensitising ethnographic study (Durrant et al., 2011), and analysed by means of the trajectories framework.

There are different kinds of trajectories. *Canonical trajectories* express the authors or the designers intended trajectory through the experience, while *participant trajectories* express the participant's actual route through the experience (Benford and Giannachi, 2008). In terms of the framework, the objective was to let the participant trajectory of the enhanced theme park experience arise from an optimal negotiation of both the canonical trajectory of the theme park experience and the canonical trajectory of the AUTOMICS experience.

The ethnographic study revealed that physical structures in the park constrained human activities in predictable ways. Queuing zones, ride locations and picnic benches, refreshment facilities, restaurants and cafés all have a definite spatial locations and afford certain activities such as queuing and eating/resting in a certain physical location. As a result, the activities the user engages in is predictable to a certain extent given the user's current location. For the design of AUTOMICS this represents an ideal articulation of *canonical contextual states* (Greenberg, 2001), some of which the user may typically perceive as opportune to engage with the application. This is expressed in the task-to-place mapping as suggested to the users by location-based notifications (see 8.1.1.2).

Hence, location-based messages play the role of coordinating the two canonical trajectories to facilitate the emergence of the participant trajectory by encouraging intermitent engagement with the application at hypothesised opportune places in the park. For experiences situated in other settings, these kinds of *contextual notifications* may be based on other definitions of canonical contextual states as appropriately informed by a sensitising ethnographic investigation.

So, in theory, contextual notifications aim to leverage the natural flow of the ongoing real-world experience to engage participants at opportune moments such as downtime whilst queuing. For the case of Automics, the role of contextual notifications was to shape the participant trajectory by mediating between the potentially competing canonical trajectories of the theme park and the application. Notifications may also be understood to initiate transitions from the theme park trajectory to foreground the Automics trajectory. In this way, contextual notifications become a utility in design to manage trajectories.

On a final note, a lesson learned from the field study is that the current goings-on in and around the use of the application (see 8.7) may be an important source of contextual information that the notification system should consider in deciding whether a notification should be delivered or not. A sequence of user activities gives rise to an individual *historic trajectory* (Benford and Giannachi, 2008) as system events are recorded. Taking into account aspects of this emergent historic trajectory might have informed a more accurate notification system. For example, reminders to take or annotate photos could have been recognised as redundant, had a real-time analysis of recent user activity identified that a sufficient number of images had already been taken and annotated.

The field study showed that prior study and analysis can identify *potential* opportune moments to engage with the experience. However, it also revealed that sensing of the *actual* contextual situation (that may render the moment as inappropriate) should be incorporated into the design. Thus, for example the system could have sensed that the recipient was socially engaged, or inferred that a notification was redundant, when the user was already engaging with the experience.

9.4 SUMMARY

This discussion wove together the main threads of this thesis, in particular findings on opportune moments for interruptions (see 9.1), the empirical contributions to the proposed model of contextual factors (see 9.2), and implications for the design of systems that employ temporal methods to coordinate interruptions (adaptation of timing, see 9.3).

While especially the quantitative results from the studies showed that contextual mediation is more likely to lead to interruptions at opportune moments, the discussion of the qualitative findings highlighted the dynamic ways in which they arise out of the situated interplay of local and relational contextual factors (see 9.1).

The findings regarding the individual contextual factors of the model were then revisited and their interaction was discussed in more detail, which arguably supports the model's veracity (see 9.2).

Furthermore, the lessons learned from adapting the timing of interruptions in the studies led to considering the potential implications of applying interruption management to mediating personal communication and to mobile, interactive and pervasive experiences (see 9.3).

10

CONCLUSIONS

This chapter briefly summarises the approach of this thesis, and then goes on to state the key conclusions in three identified key areas, namely, methodology, understanding receptivity and implications for design. Finally, the thesis concludes with an outlook on future work that this thesis might inspire.

10.1 SUMMARY

This thesis has argued for the ubiquity of interruptions in everyday live and their profound impact on the *conditio humana*. Mobility has been identified as a defining characteristic of many modern societies, and mobile technology as its accompanying blessing and curse, so to speak. A curse, because it causes disruptive interruptions at awkward moments and may entail unintended social and personal consequences; a blessing, because of the sense of connectedness with people and information in the world that it affords. This thesis bridges the study of social, psychological, and organisational implications of interruptions, and the design of solutions and systems to support interruption management that are intended to improve our everyday lives.

A holistic view of the phenomenology of perception and involvement in the world has shaped the viewpoint that the subjective, situated experience of interruptions is paramount to understanding the way in which our receptivity and responsiveness to interruptions arises as a contextual interactional achievement. Receptivity was suggested as a key concept, on the one hand to capture and study the subjective, situated nature of the experience of an interruption. On the other hand, to pinpoint opportune moments for interruptions, and to evaluate timing strategies used to coordinate interruptions in interactive experiences.

Overall, the studies in this thesis inform a multi-faceted understanding of receptivity to interruptions, while the adaptation of timing in the studies prompt discussions of design strategies for context-aware interruption management. Whereas the focus of the thesis is on the systematic study of interruptions in the real world, the discussion chapter also highlighted ways in which *contextual notifications* may be used to design interactive experiences that are interwoven with our everyday lives.

10.2 CONTRIBUTIONS AND KEY CONCLUSIONS

This thesis makes contributions in three key areas:

1. The *methodology* of studying mobile technology use in naturalistic environments;
2. *Understanding receptivity* and how contextual factors mediate the interruption process;
3. *Implications for design* of systems that interweave interactive experiences with participants' ongoing real-world experiences.

These contributions are summarised in the following.

10.2.1 *Methodology*

The methodology employed in this thesis is conceptually founded upon an emphasis on the subjective experience in situ as inspired by phenomenological philosophy and the pragmatic and established methods of experimental design and analysis as practised in cognitive and social psychology.

The real-world settings of the experiments on interruptions (see chapter 6, 7) posed a methodological challenge, in that the participants' primary task (or activity) could not be controlled. It was demonstrated how the Experience-Sampling Method (ESM) was used for the collection of data in situ without the need for laboratory control, and how this data could be triangulated to elicit multiple perspectives with data from ethnographic field observations, interviews, diaries and questionnaires. Although the employed methods are established tools used to generate understanding, their broader impact lies in their cross-referential and simultaneous application, and how ESM was adapted to collect both behavioural as well as self-reported data. At its core, a temporal model of the interruption process was developed to inform the operationalisation of receptivity in experiments with interruptions (see 3.3.5, p. 55).

The practical contribution made to methodology is:

- An adaptation of ESM to capture self-reported and behavioural responses to interruptions in situ (see 5.5).

10.2.2 *Understanding receptivity*

The main conclusion of this thesis is that whether the recipient perceives an interruption as timely surfaces from the multi-faceted way the contextual factors interact in the moment of interruption. In its goal to understand receptivity, this thesis answers the research question "What are the contextual factors that make people receptive to an interruption?" by making the following conceptual contribution:

- A taxonomic model of the contextual factors of interruptions synthesised from related work and supported by the empirical studies (see 4.3, p. 86).

The, findings from the empirical studies support the veracity of the model (see 9.2.1), in that *local* and *relational contextual factors* interact to mediate our receptivity to interruptions moment-by-moment.

Local contextual factors have the *recipient* herself at their centre – her involvement in the moment include her physical and cognitive activity, her beliefs, intentions, desires, emotions and history of experience. It was argued that involvement may be the optimal way to predict someone's receptivity, but at the same time, for a system it would be the most difficult factor of the model to sense or infer.

The studies have also shown the importance of the physical and social environment of the recipient at the moment of the interruption. For example, observations of AUTOMICS in use showed the impact of noise, the structure of the physical environment, and the degree of social engagement on whether or not notifications were noticed or deemed appropriate.

The importance of relational contextual factors, which has been argued throughout the thesis, was largely supported by the empirical studies. Experiment 1 in particular demonstrated the importance of genuine *content* of an interruption. The adaptation of the *timing* of interruptions was investigated as the primary means to answer the principal research question:

$RQ_{principal}$: What is an opportune moment for an interruption?

In that respect, the studies have shown that:

A) Contextual mediation of the timing of interruptions (by phone activity in Experiment 2 and opportune places in AUTOMICS) are more likely to lead to interruptions at opportune moments than when participants schedule the delivery time of their interruptions as indicated by analysis of participants' behavioural response.

B) The perception of an opportune moment arises from the situated interplay of several local and relational contextual factors.

With regards to A), Experiment 2 showed that users were significantly quicker to accept and respond to a notification triggered by their activity compared to a random baseline, partially supporting its RQ2 (Does the end of an episode of mobile interaction represent an opportune moment for an interruption?).

The study of AUTOMICS indicated that notifications delivered in the visitors' 'downtime', e.g. when queuing, were preferred and more likely to get accepted, also partially supporting its RQ3 (Does the recipient's current place reliably indicate an opportune moment for an interruption?).

With regards to B), interviews and self-reports did not support the superiority of the mediated timing strategy, but suggest that the actual appropriateness of interruptions was evaluated in a more complex way, as a result of the situated interplay of contextual factors. The following phenomena emerged as important contextual mediators.

- *Modality coherence* – The coherence between the modality of the interruption task and the modality of the previous interactive episode (Experiment 2).
- *Attentional demand coherence* – The coherence between the attentional demand to navigate the environment and the attentional demand to complete the interactive task (Experiment 2 and AUTOMICS).
- *Position of breakpoint* – Whether the interruption was delivered a) during execution or planning of a task (worst), b) at a breakpoint

between tasks (best), or c) at a breakpoint at a sub-task boundary (intermediate) (Experiment 2).

- *Orientation towards device* – Whether the device was *in use* (worst case), e.g. mediating a phone call or typing a message, or *in hand* but not in use (best case), or has been put away (Experiment 2).
- *Social appropriateness* – Whether it was socially appropriate to complete the task, e.g. taking photos in a meeting (Experiment 2), or tending to the phone while in a conversation (AUTOMICS).
- *Affective factors* – Whether the task was experienced as fun (Experiment 2), or work, or whether participants became annoyed by the amount of notifications (AUTOMICS).

The main conclusion is that a situated interplay of dynamically changing contextual factors determines the individual's receptivity and responsiveness to interruptions. Therefore, it appears that an automated prediction of opportune moments for interruptions is a highly uncertain endeavour that may never be 100% accurate.

For the last time, the relationship between the model of contextual factors and the empirical work in this thesis is depicted (see fig. 10.1). The figure seeks to capture not just the measured statistical effects, but also the moderating interplay of the local contextual factors that determines receptivity in situ as observed (both quantitatively and qualitatively) in the experiments and the field study, and as summarised in the phenomena above. While future work is needed to more fully understand the complexities of contextual interaction underlying interruptions, the examples in this thesis of how the proposed model can be flexibly adapted to all kinds of research questions make the author hopeful that the model may sensitise future research endeavours in this field.

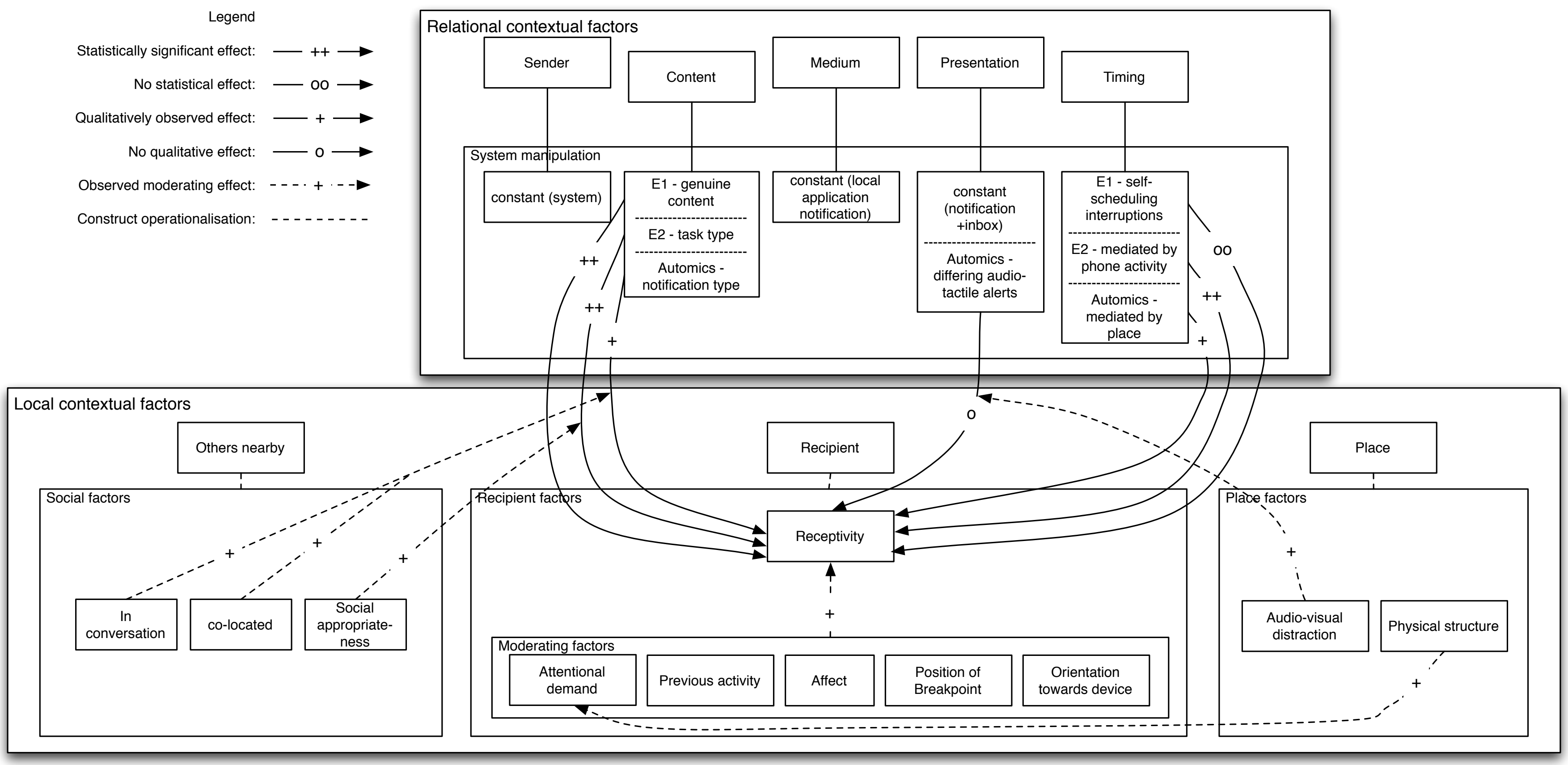


Figure 10.1: Relation of the model of contextual factors and this thesis' findings from empirical work.

10.2.3 *Implications for design*

At the beginning of the thesis, a research question regarding the implications of the studies for the design of interruption management systems was posed. The adaptation strategy that was most successful in the empirical work was to coordinate the interruptions by *mediating the timing* according to the user's current context. However, the empirical studies demonstrated the complexity of the endeavour and the difficulty of getting it right.

The discussion of systems that mediate interruptions in personal communication highlighted the trade-offs inherent in a strategy that *defers* interruptions, for which a number of design recommendation were suggested. As a consequence of the uncertainties and risks (see 9.3.1) imposed by such a system, the application might be limited to mediating interruptions from the user's *second order* communication network – a user-definable set of communication services that will not transmit any timely or vital information.

Opportunities of how mobile device development may be exploited to create systems that manage interruptions in a way that is sensitive and responsive to its settings were presented. Partly, they were demonstrated by the second experiment's sensing and responding to user activity on the phone, and partly by AUTOMICS' use of location-based messaging. Furthermore, a broader strategy of context-sensitive design that abstracts away from the implementation details was highlighted by the discussion of the design of *contextual notifications* in pervasive experiences (see 9.3.2).

The latter discussion illustrated how *contextual notifications* might be employed in order to deal with the temporal and pervasive challenges of mobile, interactive experiences. It was argued that contextual notifications may help to overcome the challenges, and that they specifically facilitate alignment of the participant trajectory by mediating between the canonical trajectories of the interactive experience and the real-world experience.

10.3 CRITICAL REFLECTION

A multi-disciplinary education in Computer Science, Psychology, Communication- and Media Sciences, and Social Sciences has shaped the disciplinary background of the author of this thesis. Therefore, this thesis is characterised by multiple views onto its core questions, experimental methods and approaches to gain knowledge about the problem, and evaluations of computational techniques to engineer solutions. A pragmatic attitude towards the disciplines and their paradigms influenced the tone of the thesis, rather than strictly following the approaches of a single discipline. Many contributions to HCI are characterised by their interdisciplinary view, and it is the author's firm belief that knowledge in HCI arises most effectively out of combining the practices from multiple disciplines that have proven to fit the problem well.

However, combining multiple perspectives is challenging. The thesis does not simply adopt a single interdisciplinary view – rather, it remains multi-disciplinary in that it contrasts and discusses approaches and findings from various perspectives. While some disciplines seek to study the world *as is*, other disciplines seek to engineer solutions that *change* the world. Consequently, these perspectives may not be compatible. So, what this thesis sometimes does is to point out the tensions where disciplines collide. For example, in retaining a view of the recipient's involvement in the world while breaking down the context into contextual factors may seem contradictory. Nevertheless, the justification lies in the pragmatic attitude that is required to experimentally study contextual aspects in detail, or to suggest a framework that simplifies complexity into components to inform the engineering of context-sensitive solutions.

Consequently, this work has some inevitable limitations, as outlined in the following.

10.3.1 *Limitations of this work*

Limitations arise from a number of tensions in this thesis, which are considered in turn.

First, the conceptual framework of the contextual factors of an interruption can only be used to inform a system's operationalisation of the factors by acknowledging the trade-offs this entails. As it was argued in the discussion, it may be extremely difficult to measure the factors that comprise *human context*, or to design a system to reason and deduce human intent from sensor input. In particular, internal mental processes that are so decisive for our involvement in the world and thus receptivity to interruptions are difficult to study directly and are likely to remain unknown factors for future systems that aim to sense human context.

Secondly, the methodological choices to study the individual, subjective experience while participants go about their everyday lives or to observe small groups in the field may bring ecological validity to the results, but at the same time make it difficult to generalise results to a wider population. However, one of the central reflections of this thesis is that not only does the study of experience need to take place *in situ*, but also the design needs to accommodate this situated perspective. This means that sensitising ethnographic studies of the settings and users, and iterative prototyping and testing in situ need to become part of the design process. Thus, arguing for a situated design practice mitigates the need for wider generalisability.

Thirdly, the author encountered the dilemma that has troubled many authors of studies of socio-technical systems: How do you derive design recommendations from empirical findings? This dilemma stems directly from the collision of disciplines – whereas the scientific study is considered complete with an account of the how's and what's of the goings-on, engineers are interested in the *implications* for design the findings may have, often in the form of a speculation about the ideal specification of a system.

In the following section, design guidelines are presented that were derived from the experience of devising and studying a real-world application designed to prompt user engagement at opportune moments (Automics). These are formulated as imperative statements that suggest certain design activities conceived through a subjective reflection on the empirical findings. Hence, the guidelines represent subjective interpretations, in contrast to the aspiration of maximum objectivity present in the scientific method underlying the experiments.

10.3.2 *Reflection on designing for contextual interweaving*

Here, a set of design guidelines to achieve contextual interweaving is presented as an outcome of the author's reflection of the design process of Automics. The design guidelines may be read as design activities to be conducted in temporal sequence:

1. *Conduct sensitising ethnographic studies* of people in the environment that you want to augment with your interactive experience, to uncover the routine and predictable canonical contextual states. This follows the finding and the recommendation (Greenberg, 2001) that we must study the expected context-of-use carefully.

2. *Distil the findings into a typical participant trajectory* (or family of participant trajectories as guided by the trajectories framework by Benford et al. (2009)) to identify recurrent transitions and episodes in the real-world experience that may indicate opportunities to engage with the interactive experience.

3. *Analyse the interactive sequences in your application* to identify their temporal challenges of pacing, ordering, and scheduling and analyse the interaction in terms of its attentional demands and technology-world relations it creates (see section 2.1.2).

4. *Map the interactive tasks to contextually opportune situations*, as indicated by an opportune matching of the attentional demands in your application and the anticipated demands in the environment, as articulated for example by the *Resource Competition Framework* (Oulasvirta et al., 2005b), or the structure of the physical envrionment.

5. *Consider the user's current and the historic interaction* with your application in real-time in order to avoid redundancy and habituation, and to align your application to the user in a personalised way.

6. *Implement contextual reasoning about the actual opportuneness* of the situation, for example to assess the social appropriateness (e.g. collocation, and conversational state), and the physical state (e.g. "potentially dizzy").

7. *Acknowledge that failure is inevitable* (Greenberg, 2001) and context-awareness can never be 100% accurate, and think of ways in which your system can provide intelligibility and accountability (Bellotti and Edwards, 2001) to its users.

These guidelines were synthesised from the experience gained through the process of writing this thesis, and are likely not complete or comprehensive. Most importantly, the discussion of implications (see 9.3) provides us with a sensitivity to the complex endeavour of designing and building useful context-aware systems, and emphasises that its design must be as domain- and context-specific as the reasoning and accuracy is aiming to achieve.

10.4 FUTURE WORK

Following from the thesis' contribution, its impact in those areas may be extended into future work in several possible ways.

With regards to the *methodology*, the mixed method approach and especially its *in situ* aspects may be applied to other research fields that would benefit from real-time, repeated assessment of multi-modal qualities of people's situated experiences. For example, knowledge transfer opportunities exist to apply and develop these techniques to the domain of *healthcare*. There may be an engineering challenge to create a tool that empowers the healthcare professionals and researchers to design and conduct their own studies and interventions according to ESM or EMA, without the need to design and implement the entire system.

The studies have highlighted the challenge of determining the appropriate set of *canonical contextual states* (Greenberg, 2001), and their analyses have revealed the complex ways in which appropriate moments to interrupt or engage with an interactive task surface as an interactional achievement. A more structured approach to reveal such canonical contextual states routinely as part of the design process for context-aware systems may be needed, for example by conducting sensitising ethnographic studies of the application domain and its usage setting, and by several iterations of prototyping to arrive at a sufficiently adequate representation of human context in systems.

The critical discussion of the applicability of interruption management to application domains raises the challenge of tackling the hard problem of mediating interpersonal communication in a sensible, useful and safe way. Possible approaches may involve the entire array of techniques available to design interactive systems, such as mixed initiative approaches to resolve critical decisions by involving the users, artificial intelligence and autonomous agents for more sophisticated context reasoning and inference, speech- and content- analysis, and innovative interface design to make the systems accountable and intelligible.

The modest findings offered, and the challenges illuminated and discussed in this thesis foreshadow the challenges that designers of context-aware and autonomous systems are likely to face on a much larger scale in the future. Consider systems that are embedded and connected in our homes, cars, at work and in the environment and that proactively offer their services, or exchange information and interpret it on our behalf to assist us as we work, consume, travel, relax, and encounter each other.

This thesis is only an initial step in the research and development effort required to take us into a future in which we can trust ubiquitous systems, be comfortable or even enjoy having them around us, and in which they can assist us in intelligible and accountable ways.

Part V

APPENDIX

A

APPENDIX TO EXPERIMENT 1

This appendix includes additional material to Experiment 1 (see chapter 6):

- Figure A.1 depicts the system architecture.
- Figure A.2 shows the pre-experiment questionnaire.
- Figure A.3 shows the exit interview questions.

Additional contents on the accompanying CD include:

- The source code of the ContentFetcher in directory `code/Experiment1`.
- The quantitative data from the experiment in directory `data/Experiment1/quant`.
- Audio recordings and thematic transcripts from the interviews in `data/Experiment1/qual`.
- A list of the RSS sources used in directory `assets/Experiment1`.

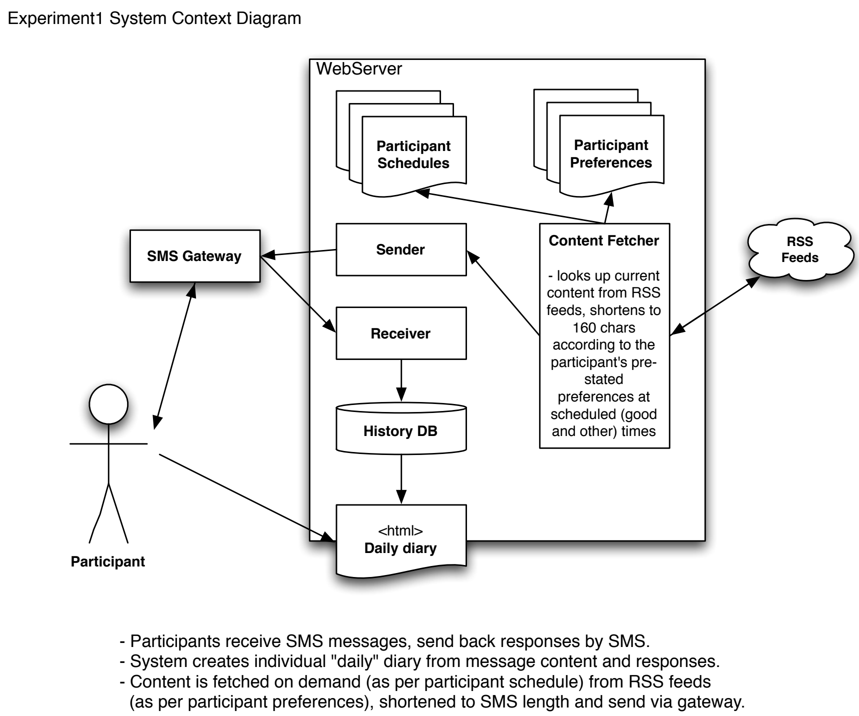


Figure A.1: System context diagram of the architecture of Experiment 1.

Participant choices

Participant # ____________

Name: ______________________

Mobile #: ______________________

Preferred times

Please think about daily times or time windows during which you think you would be receptive to content on your mobile phone.

Time 1 ____________________

Time 2 ____________________

Time 3 ____________________

Time 4 ____________________

Time 5 ____________________

Please specify a 10hr window in which you are willing to receive msgs!

[ ] 7am-5pm
[ ] 8am-6pm
[ ] 9am-7pm
[ ] 10am-8pm
[ ] 11am-9pm
[ ] 12pm-10pm

Content rating

Please rate how interested you are in the following content categories on a scale from 1-7, where 1 means not at all interested and 7 means very interested. Lines are optional to specify for content you are highly interested in.

NEWS	**not at all…**			**medium…**			**very interested**
Top News stories	1	2	3	4	5	6	7
Local/Regional Bay Area	1	2	3	4	5	6	7
World News	1	2	3	4	5	6	7
US News	1	2	3	4	5	6	7
Business	1	2	3	4	5	6	7
Finance	1	2	3	4	5	6	7
Health	1	2	3	4	5	6	7
Weird News	1	2	3	4	5	6	7
Sports	1	2	3	4	5	6	7
Science	1	2	3	4	5	6	7
Technology	1	2	3	4	5	6	7

other..please specify ______________________________

____

HOBBIES							
Gardening	1	2	3	4	5	6	7
Cars	1	2	3	4	5	6	7
Photography	1	2	3	4	5	6	7

other..please specify ______________________________

____

continued overleaf

	not at all...			medium...			very interested
REVIEWS							
Music album reviews	1	2	3	4	5	6	7
Book reviews	1	2	3	4	5	6	7
Movie reviews	1	2	3	4	5	6	7
other..please specify	______						
	______						
GOSSIP							
Celebrities	1	2	3	4	5	6	7
Entertainment	1	2	3	4	5	6	7
other..please specify	______						
	______						
WEATHER	1	2	3	4	5	6	7
your zip code:	______						
TRIVIA							
Horoscope	1	2	3	4	5	6	7
your sign:	______						
Word of the day	1	2	3	4	5	6	7
Quote of the day	1	2	3	4	5	6	7
This day in history	1	2	3	4	5	6	7
Did you know...? Facts	1	2	3	4	5	6	7
SATIRE and HUMOR							
The Onion: political satire	1	2	3	4	5	6	7
Chuck Norris facts	1	2	3	4	5	6	7
RESTAURANTS							
Bay Area rest. News&reviews	1	2	3	4	5	6	7

Thanks for completing the form!

Figure A.2: Questionnaire used to gauge participants' preferred times to receive messages and interest in content categories.

Exit interviews

1. Before I get into the specific questions, do you have any comments or thoughts you'd especially like to share about participating in the study?

Change over time

2. Did you change the way your phone alerted you (ringtone, vibration etc) as a result of receiving 6 msgs a day? In which way?

- y
- n

3. Did you feel differently about the messages or the study as the study progressed over the past 2 weeks? Any differences in how you thought about or reacted to the messages?

- y
- n

Specific moments

4. Do you remember the times you named you would be receptive to msgs when I set you up for the study?

- y
- n

5. Do you remember a moment when you really liked the timing of a msg? And disliked?

- y
- n

6. Do you remember a moment when your gut reaction to the rcvd msg was really good/bad? What was the message?

- y
- n

7. Were there any other particularly memorable moments?

- y
- n

Consequences

8. Did a message ever influence your subsequent activity in any way? How?

- y
- n

9. Was a message ever relevant to your activity? In which way?

- y
- n

General receptivity

10. How is your overall feeling about receiving messages on your phone on a regular basis after this study?

11. What seemed more important to your gut reaction: the timing or the content of a message?

- timing:

- content:

12. If you could have chosen any content for yourself, what are the top content types you would have chosen?

13. Overall, did you get more messages you liked or more messages you disliked?

- liked

- disliked

14. Overall, was the timing of the messages on the good or on the bad side?

- good

- bad

15. Do you think there is much variation in what content you want at various times, or is it predictable and if so what is predictable about it? Explain.

- Variation:

- Predictable:

16. Do you think you might like to have content pushed to you automatically or would you prefer to pull it yourself? Explain.

- What kind of modality? Prompt for: audio, video, text, pictures, etc...

17. Did a msg ever raise your interest that you wanted to see more on the topic?

- y
- n

If, yes: If you could have clicked a "more on.." or "see more.." button at the end of a msg, would that have been appealing to you / would that have been a feature you wanted?

Exit

18. Is there anything we didn't touch upon that you would like to comment on?

Figure A.3: Exit questionnaire for Experiment 1.

B
APPENDIX TO EXPERIMENT 2

This appendix includes additional material to Experiment 2 (see chapter 7):

- Figure B.1 depicts the system architecture.
- Table B.1 shows the demography of the sample.
- Figure B.2 shows the exit interview questions.

Additional contents on the accompanying CD include:

- The source code of the MActivityMonitor Android project is in directory `code/Experiment2`.
- The quantitative data from the experiment in directory `data/Experiment2/quant`.
- Audio recordings and thematic transcripts from the interviews in `data/Experiment2/qual`.

partID	age	occupation	gender
8133	26	sales team manager	m
277	30	research associate	m
1441	29	research fellow	m
8043	28	research associate	m
8134	38	sales manager	f
8334	25	part time PhD student	m
8483	25	product tech support	f
8484	25	consultant ecologist	f
9472	26	ft Phd student	m
9473	33	ft Phd student	f
2317	21	ft Phd student	m
7070	30	ft Phd student	f
626	27	ft Phd student	f
8135	26	ft Phd student	f
8136	26	ft Phd student	m
8137	48	ft Phd student	m
8487	38	genetic counselor, NHS	f
8485	37	ft Phd student	f
8486	26	ft Phd student	m
9474	29	Lecturer	f
Average age:	29.65		
Median	27.5		
		10 postgrad students	
		5 uni staff	
		5 employees	

Table B.1: Demography of the sample of Experiment 2.

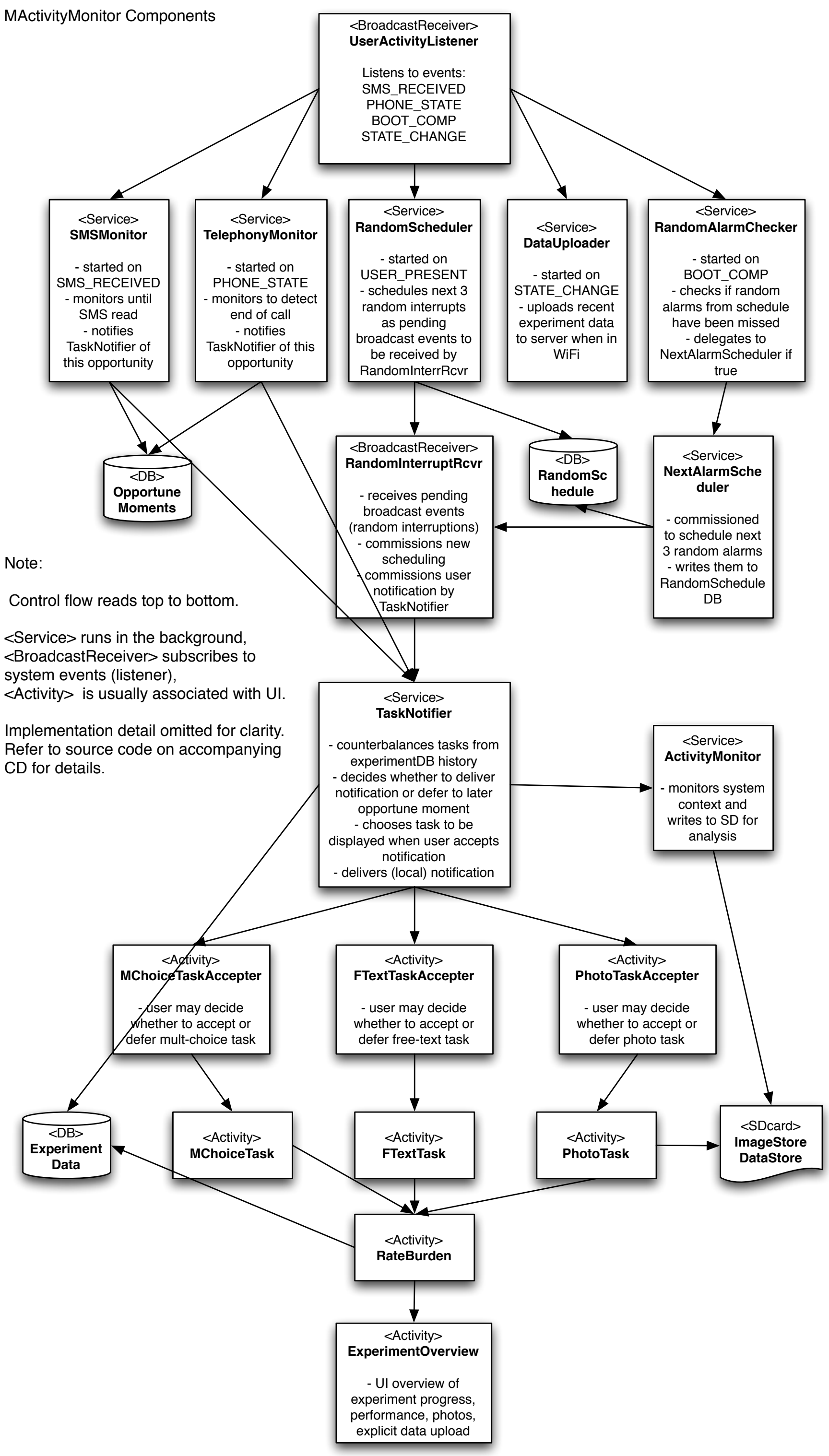


Figure B.1: MActivityMonitor component diagram with control flow.

Participant ID: ____________

Experiment 2 Exit Interview

Before I get into the specific questions, do you have any comments or thoughts about the study you would like to share before we get started?

Tasks

When you looked at the notification and saw which task it was, did that sometimes influence if you accepted it right away or not? If yes, can you give an example? Would you sometimes have accepted another task at that time then, just not that particular one?

Were you generally more likely to accept one of the tasks over one or both of the others or was that dependent on your context/where you were, what you were doing etc.? If context-dependent, in which way? E.g. where: home vs. work vs. travel etc.

About the multiple choice task, do you remember a moment where it was particularly inappropriate to take a photo or where you delayed or ignored the task because of that?

Would you say it was inappropriate:

Always / almost always / often / sometimes /almost never / never

Can you describe the situation, please?

About the photo task, do you remember a moment where it was particularly inappropriate to take a photo or where you delayed or ignored the task because of that?

Would you say it was inappropriate:

Always / almost always / often / sometimes /almost never / never

Can you describe the situation, please?

About the free text task, do you remember a moment where it was particularly inappropriate to type an answer or where you delayed or ignored the task because of that?

Would you say it was inappropriate:

Always / almost always / often / sometimes /almost never / never

Can you describe the situation, please?

If you would have to rank the tasks from most burdensome to least burdensome, how would you rank them?

most: middle: least:

Timing

Do you remember a particular moment where the timing of the interruption was really good? Can you describe where you were and what you were doing at the time? Where was the phone? Do you remember any more moments..?

What was a typical moment where you looked at the notification immediately when you heard it coming in?

What was a typical moment where you would ignore the notification when you heard it coming in?

Did you sometimes look at the phone without hearing a notification coming in to check if there was an idle notification?

The application interrupted at different moments during the day. Did you notice any pattern about the timing of the notifications? Did you feel you could predict when the notification happened?

- Debrief about timing:
  There were usually three random notifications per day and the other ones were after you read a text send by SMS or after you made or tried to make or received a phone call.

What did you think about the timing of a notification when it came after a phone call/after reading an SMS/at random other times?

If you had to rank these general types of timing from best to worst, how would you rank them?

Best: middle: worst:

Why would you say XY is best and ZY is wost?

Can you think of any other times or situation when it would be sensible to receive notifications in general?

Interruptiveness and Receptivity

Do you feel that at any of these moments, you were generally more receptive to a notification? Which one/s?

At what times were the notifications the most disruptive to what you were doing at the time?

If a notification was disruptive, in which way was it disruptive?

How hard or easy was it to pick up what you were doing before the interruption?

If you had to rank the interruptiveness of the different timings SMS/call/random, how would you rank them from least to worst interruptive:

Least: middle: worst:

Can you explain, please?

Motivation/Social

What normally motivated you to respond to the messages? Did you look at the "Overview" screen?

Did you talk about the experiment with others? Did you show them the application or the photos?

Did the notifications sometimes interrupt others? What was the occasion?

Did you feel you sometimes had to excuse your interaction with the experiment, did you get held accountable?

General usage and habituation

What type of mobile have you normally got?

How many notifications do you feel you typically got per day?

How many texts did you typically send/receive per day?

How many phone calls did you typically make during the day?

What do you normally use your phone for?

Do you feel that because of the study you tended to your phone more regularly and paid more attention to it than usual?

Did the way you felt about participating in the study change over the course of the study?

When you think about your preferred way of working, do you in general prefer to do one task at time or do you do several things intermittently?

Would you say you typically check your incoming notifications, such as email and texts right away?

Did bugs or errors make a difference to the experience?

(If own phone:) Did the application have a noticeable impact on your device's performance? E.g. network issues or battery live?

Finally, what did you generally like and dislike about the experiment?

Context (only if not covered yet)

Can you describe some situations where you were notified by the application? Where were you, what were you doing, who else was there, etc.

Do you remember any particular moments where you found the study really annoying you? Can you describe the situation?

Figure B.2: Exit interview questions for Experiment 2.

C APPENDIX TO THE FIELD STUDY

This appendix includes additional material to the field study of AUTOMICS (see chapter 8):

- Figure C.1 depicts the system architecture.
- Figure C.2 shows the location model of the system.
- Figure C.3 depicts the control flow of the photo task.
- Figure C.4 depicts the control flow of the annotation task.
- Table C.1 shows the observation guideline for the field study.

Additional contents on the accompanying CD include:

- The source code of the AUTOMICS Android project, including the location model and the location-based messages is in directory `code/Automics`.
- The quantitative data from the experiment in directory `data/FieldStudy/quant`.
- Video recordings and thematic transcripts from the interviews in `data/FieldStudy/qual`.

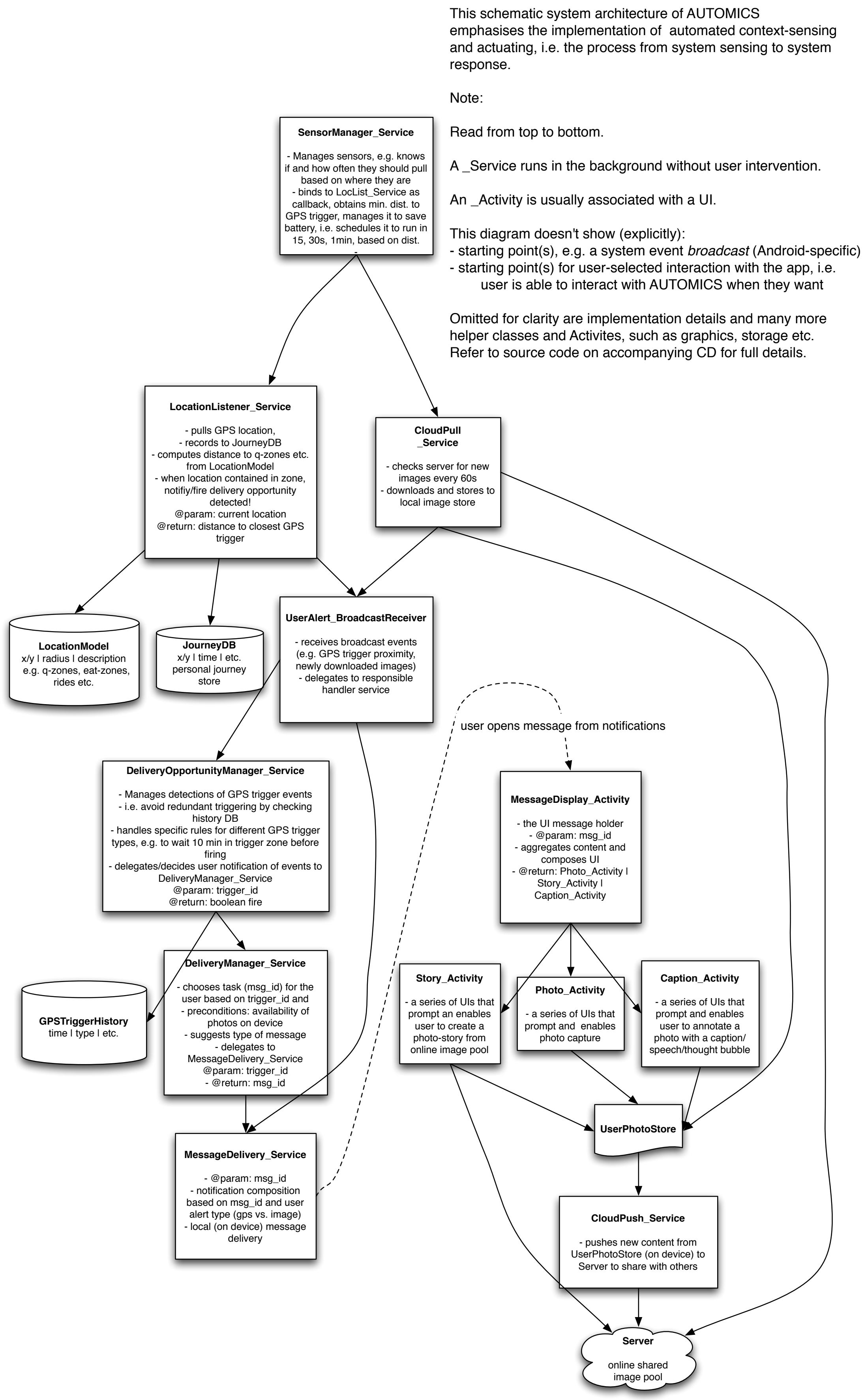


Figure C.1: Context-aware architecture of Automics with control flow.

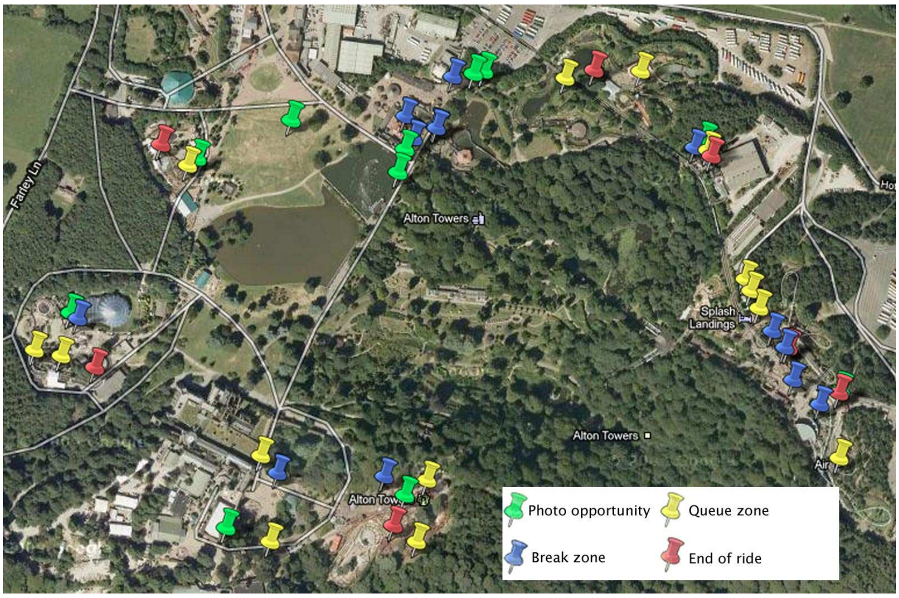


Figure C.2: Location model and GPS triggers of Automics.

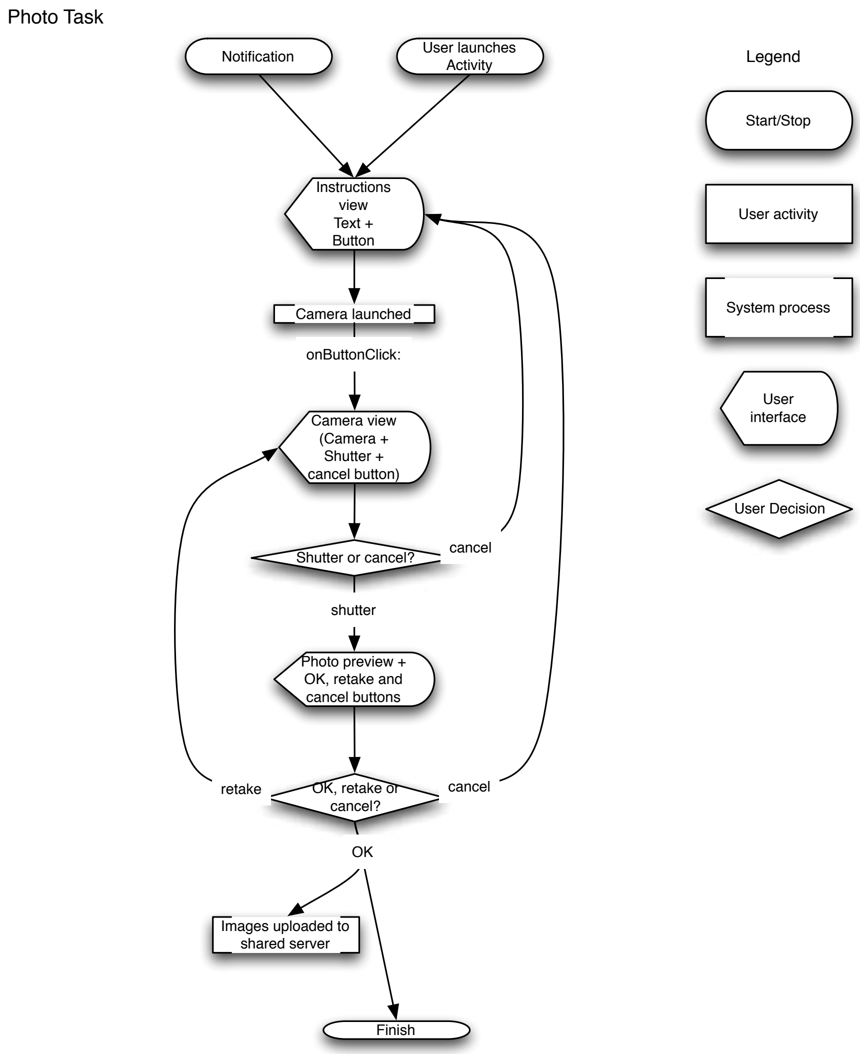


Figure C.3: Task flow of the photo task in AUTOMICS.

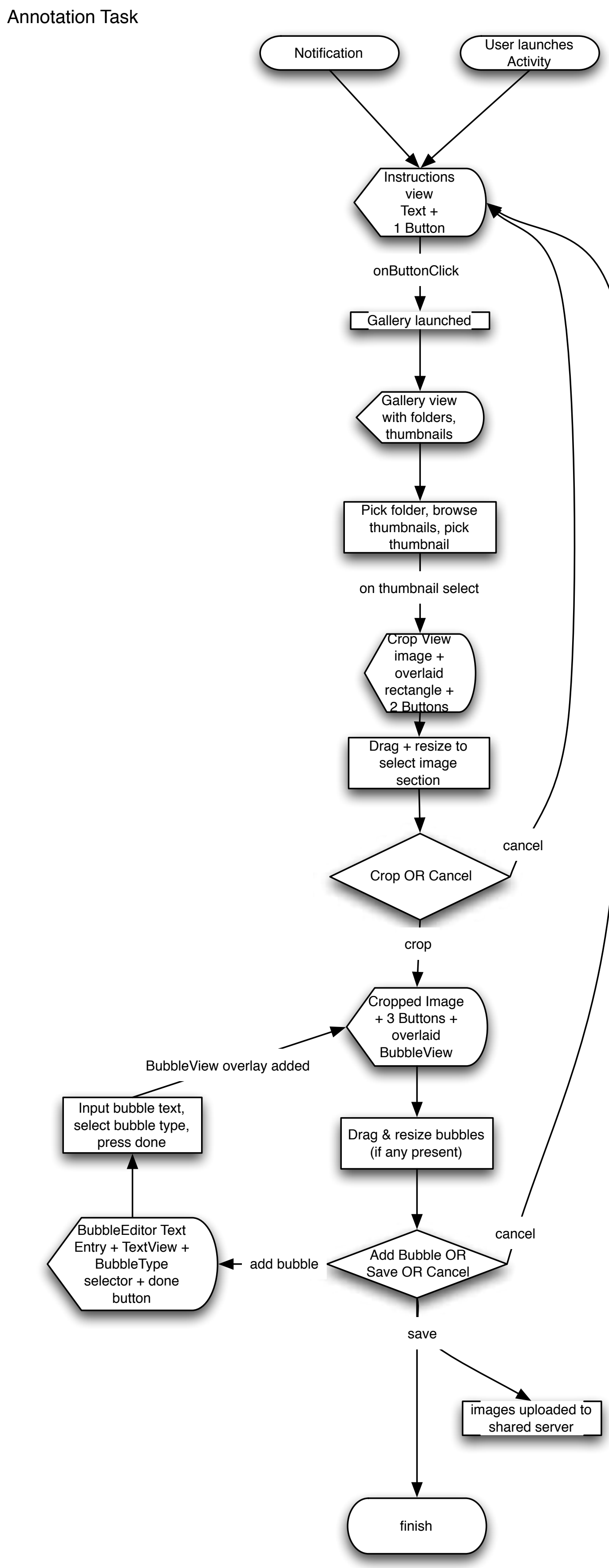


Figure C.4: Task flow of the annotation task in AUTOMICS.

Contextual factor	Implementation	Questions
Content	Location-based messages prompt one of three tasks, refer to the surrounding in the task wording (e.g. they mention the ride in the vicinity, objects as photo opportunities).	Are there any benefits in the experience of the task wording referring to the surrounding?
	When a user shares a photo the other members are notified on their phones.	What are the users' perceived benefits and shortcomings of this type of notification? How does the reception of this type of user-generated notification differ to the actionable location-based messages?
Presentation	Location-based messages and "new image" notification differ by audio alert.	Does this differentiation help the user to distinguish the notification type by its sound?
	Location-based messages triggered in queue zones vibrate longer (6 seconds), repeated every 30 seconds.	Does this feature make the user miss less notifications? What is the tradeoff between annoyance and task completion?
Timing	Location-based messages triggered according to heuristic about opportuneness of task-context mapping.	Task coherence: In how far did the rationale succeed to prompt tasks at opportune places in accordance with the attentional demands imposed by orientating in the environment? Are there any observable instances when task completion fails because of attentional demands to navigate or orientate in the environment?
Social	emergent	What is the observed social impact of the interruptions? What kind of interaction with and around the mobile technology can be observed, how are accountabilities managed?
Environment	emergent	What is the impact of environmental factors such as light, noise, weather conditions, structures and other people around on the observed user experience?
Push vs. pull	emergent	What are the benefits and shortcomings of initiating interaction at user-selected moments compared to system-selected moments?

Table C.1: Observation guideline for AUTOMICS.

BIBLIOGRAPHY

Abowd, G. D., Dey, A. K., Brown, P. J., Davies, N., Smith, M., and Steggles, P. (1999). Towards a better understanding of context and context-awareness. In *HUC '99: Proceedings of the 1st international symposium on Handheld and Ubiquitous Computing*, pages 304–307, London, UK. Springer-Verlag.

Adamczyk, P. D. and Bailey, B. P. (2004). If not now, when?: The effects of interruption at different moments within task execution. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '04, New York, NY, USA. ACM.

Agre, P. and Horswill, I. (1997). Lifeworld analysis. *Journal of Artificial Intelligence Research*, 6:111–145.

Altmann, E. M. and Trafton, J. G. (2004). Task interruption: Resumption lag and the role of cues. In *In Proceedings of the 26th Annual Conference of the Cognitive Science Society*, pages 42–47.

Altosaar, M., Vertegaal, R., Sohn, C., and Cheng, D. (2006). AuraOrb: Social notification appliance. In *CHI '06 extended abstracts on Human factors in computing systems*, CHI '06, pages 381–386, New York, NY, USA. ACM.

Andrews, P. (2004). Vying for your attention: Interruption management. Executive Technology Report 7, IBM Business Consulting Services.

Anhalt, J., Smailagic, A., Siewiorek, D., Gemperle, F., Salber, D., Weber, S., Beck, J., and Jennings, J. (2001). Toward context-aware computing: Experiences and lessons. *Intelligent Systems, IEEE*, 16(3):38–46.

Aoki, P. M. (2005). Making space for stories: ambiguity in the design of personal communication systems. In *In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '05, pages 181–190, New York, NY, USA. ACM Press.

Arroyo, E. and Selker, T. (2003). Self-adaptive multimodal-interruption interfaces. In *IUI '03: Proceedings of the 8th international conference on Intelligent user interfaces*, pages 6–11, New York, NY, USA. ACM.

Arroyo, E., Selker, T., and Stouffs, A. (2002). Interruptions as multimodal outputs: Which are the less disruptive? In *ICMI '02: Proceedings of the 4th IEEE International Conference on Multimodal Interfaces*, page 479, Washington, DC, USA. IEEE Computer Society.

Ash, J. S., Berg, M., and Coiera, E. (2004). Some unintended consequences of information technology in health care: The nature of patient care information system-related errors. *Journal of the American Medical Informatics Association*, 11(2):104–112.

Ashbrook, D. and Starner, T. (2002). Learning significant locations and predicting user movement with GPS. In *Proceedings of the International Symposium on Wearable Computers*, ISWC 2002, pages 101–108.

Avrahami, D. (2007). *Enhancing Technology-Mediated Communication: Tools, Analyses, and Predictive Models*. PhD thesis, Human-Computer Interaction Institute, School of Computer Science, Carnegie Mellon University.

Avrahami, D., Fogarty, J., and Hudson, S. E. (2007a). Biases in human estimation of interruptibility: Effects and implications for practice. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '07, New York, NY, USA. ACM.

Avrahami, D., Fussell, S. R., and Hudson, S. E. (2008). IM waiting: Timing and responsiveness in semi-synchronous communication. In *Proceedings of the ACM conference on Computer supported cooperative work*, CSCW '08, pages 285–294, New York, NY, USA. ACM.

Avrahami, D., Gergle, D., Hudson, S. E., and Kiesler, S. (2007b). Improving the match between callers and receivers: A study on the effect of contextual information on cell phone interruptions. *Behav. Inf. Technol.*, 26(3):247–259.

Avrahami, D. and Hudson, S. E. (2004). QnA: Augmenting an instant messaging client to balance user responsiveness and performance. In *Proceedings of the ACM conference on Computer supported cooperative work*, CSCW '04, pages 515–518, New York, NY, USA. ACM.

Avrahami, D. and Hudson, S. E. (2006). Responsiveness in instant messaging: predictive models supporting inter-personal communication. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, CHI '06, New York, NY, USA. ACM.

Bailey, B., Adamczyk, P., Chang, T., and Chilson, N. (2006). A framework for specifying and monitoring user tasks. *Computers in Human Behavior*, 22(4):709–732. A framework for specifying and monitoring user tasks.

Bardram, J. E. and Hansen, T. R. (2004). The AWARE architecture: Supporting context-mediated social awareness in mobile cooperation. In *Proceedings of the ACM conference on Computer supported cooperative work*, CSCW '04, pages 192–201, New York, NY, USA. ACM.

Barrett, L. F. and Barrett, D. J. (2001). An introduction to computerized experience sampling in psychology. *Social Science Computer Review*, 19(2):175–185.

Begole, J. B., Matsakis, N. E., and Tang, J. C. (2004). Lilsys: Sensing unavailability. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work*, pages 511–514, Chicago, Illinois, USA. ACM.

Bell, M., Chalmers, M., Barkhuus, L., Hall, M., Sherwood, S., Tennent, P., Brown, B., Rowland, D., Benford, S., Capra, M., and Hampshire, A. (2006). Interweaving mobile games with everyday life. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, CHI '06, pages 417–426, New York, NY, USA. ACM.

Bellotti, V., Back, M., Edwards, W. K., Grinter, R. E., Henderson, A., and Lopes, C. (2002). Making sense of sensing systems: Five questions for designers and researchers. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '02, pages 415–422, New York, NY, USA. ACM.

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 *Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, CHI '08, pages 1157–1166, New York, NY, USA. ACM.

Bellotti, V. and Edwards, K. (2001). Intelligibility and accountability: Human considerations in context-aware systems. *Hum.-Comput. Interact.*, 16:193–212.

Benford, S. and Giannachi, G. (2008). Temporal trajectories in shared interactive narratives. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '08, pages 73–82, New York, NY, USA. ACM.

Benford, S., Giannachi, G., Koleva, B., and Rodden, T. (2009). From interaction to trajectories: designing coherent journeys through user experiences. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '09, pages 709–718, New York, NY, USA. ACM.

Brown, B., Chalmers, M., Bell, M., Hall, M., MacColl, I., and Rudman, P. (2005). Sharing the square: collaborative leisure in the city streets. In *Proceedings of the ninth conference on European Conference on Computer Supported Cooperative Work*, pages 427–447, New York, NY, USA. Springer-Verlag New York, Inc.

Bulmer, M. G. (1979). *Principles of Statistics*. Dover Publications. Excellent introductory text on statistics and probability. Provides historical context for a number of distributions.

Card, S., Moran, T., and Newell, A. (1986). *The Model Human Processor: An Engineering Model of Human Performance*, volume 2 of *Handbook of Perception and Human Performance*, pages 1–35. Wiley, New York, NY, USA.

Carter, S., Mankoff, J., and Heer, J. (2007). Momento: support for situated ubicomp experimentation. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '07, pages 125–134, New York, NY, USA. ACM.

Chandler, D. (1994). The transmission model of communication. In *http://www.aber.ac.uk/media/Documents/short/trans.html*. University of Wales, Aberystwyth, UK.

Chen, D. and Vertegaal, R. (2004). Using mental load for managing interruptions in physiologically attentive user interfaces. In *CHI '04 extended abstracts on Human factors in computing systems*, CHI '04, pages 1513–1516, New York, NY, USA. ACM.

Cheverst, K., Davies, N., Mitchell, K., Friday, A., and Efsratiou, C. (2000). Developing a context-aware electronic tourist guide: Some issues and experiences. In *Proceeding of the SIGCHI conference on Human factors in computing systems*, CHI '00, New York, NY, USA. ACM Press.

Clark, H. H. (1996). *Using language*. Cambridge University Press, Cambridge, England.

Consolvo, S. and Walker, M. (2003). Using the experience sampling method to evaluate ubicomp applications. *Pervasive Computing, IEEE*, 2(2):24–31.

Coragio, L. (1990). *Deleterious effects of intermittent interruptions on the task performance of knowledge workers: A laboratory investigation*. PhD thesis, University of Arizon.

Crabtree, A., Benford, S., Greenhalgh, C., Tennent, P., Chalmers, M., and Brown, B. (2006). Supporting ethnographic studies of ubiquitous computing in the wild. In *Proceedings of the 6th conference on Designing Interactive systems*, DIS '06, pages 60–69, New York, NY, USA. ACM.

Crabtree, A., Nichols, D. M., O'Brien, J., Rouncefield, M., and Twidale, M. B. (2000). Ethnomethodologically informed ethnography and information system design. *J. Am. Soc. Inf. Sci.*, 51:666–682.

Crabtree, A., Rodden, T., Tolmie, P., and Button, G. (2009). Ethnography considered harmful. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '09, pages 879–888, New York, NY, USA. ACM.

Csikszentmihalyi, M., Larson, R., and Prescott, S. (1977). The ecology of adolescent activity and experience. *Journal of Youth and Adolescence*, 6(3):281–294.

Csikszentmihalyi, M. and LeFevre, J. (1989). Optimal experience in work and leisure. *Journal of Personality and Social Psychology*, 56(5):815–822.

Cutrell, E., Czerwinski, M., and Horvitz, E. (2001). Notification, disruption, and memory: Effects of messaging interruptions on memory and performance. pages 263–269. IOS Press.

Cutrell, E. B., Czerwinski, M., and Horvitz, E. (2000). Effects of instant messaging interruptions on computing tasks. In *CHI '00 extended abstracts on Human factors in computing systems*, CHI '00, New York, NY, USA. ACM.

Czerwinski, M., Cutrell, E., and Horvitz, E. (2000). Instant messaging: Effects of relevance and time. volume 2, pages 71–76. British Computer Society.

Czerwinski, M., Horvitz, E., and Wilhite, S. (2004). A diary study of task switching and interruptions. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '04, New York, NY, USA. ACM.

Dabbish, L. and Kraut, R. (2003). Coordinating communication: Awareness displays and interruption. In *CHI '03 extended abstracts on Human factors in computing systems*, CHI '03, pages 786–787, New York, NY, USA. ACM.

Dabbish, L. and Kraut, R. E. (2004). Controlling interruptions: Awareness displays and social motivation for coordination.

Dabbish, L. A. (2006). *Coordinating Initiation and Response in Computer-Mediated Communication*. PhD thesis, School of Computer Science, Pittsburg, PA, USA.

Dabbish, L. A. and Baker, R. S. (2003). Administrative assistants as interruption mediators. In *CHI '03 extended abstracts on Human factors in computing systems*, CHI '03, pages 1020–1021, New York, NY, USA. ACM.

Davenport, T. and Beck, J. (2001). *The Attention Economy: Understanding the New Currency of Business*. Harvard Business School Press.

De Guzman, E. S., Sharmin, M., and Bailey, B. P. (2007). Should I call now? understanding what context is considered when deciding whether to initiate remote communication via mobile devices. In *GI '07: Proceedings of Graphics Interface 2007*, pages 143–150, New York, NY, USA. ACM.

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. *Hum.-Comput. Interact.*, 16(2):97–166.

Dey, A. K. and Häkkilä, J. (2008). Context-awareness and mobile devices. In Lumsden, J., editor, *Handbook of Research on User Interface Design and Evaluation for Mobile Technology*, volume 1 of *Information Science Reference*, pages 205–217. Hershey.

Dix, A. (1993). Pace and interaction. In *Proceedings of the conference on People and computers VII*, pages 193–207, New York, NY, USA. Cambridge University Press.

Dix, A., Ramduny, D., and Wilkinson, J. (1996). Long-term interaction: Learning the 4 Rs. In *Conference companion on Human factors in computing systems*, CHI '96, pages 169–170, New York, NY, USA. ACM.

Dix, A., Ramduny, D., and Wilkinson, J. (1998). Interaction in the large. *Interacting with computers*, 11(1):9–32.

Dix, A. J. (1990). Non-determinism as a paradigm for understanding the user interface. In Thimbleby, H. W. and Harrison, M. D., editors, *Formal Methods in Human-Computer Interaction*, chapter 4, pages 97–127. Cambridge University Press.

Donaldson, S. I. and Grant-Vallone, E. J. (2002). Understanding self-report bias in organizational behavior research. *Journal of Business and Psychology*, 17:245–260. 10.1023/A:1019637632584.

Dourish, P. (2001). *Where the action is: The foundations of embodied interaction*. MIT Press.

Dourish, P. (2004). What we talk about when we talk about context. *Personal Ubiquitous Comput.*, 8:19–30.

Dourish, P. (2006a). Implications for design. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, CHI '06, pages 541–550, New York, NY, USA. ACM.

Dourish, P. (2006b). Re-space-ing place: "Place" and "space" ten years on. In *Proceedings of the ACM conference on Computer supported cooperative work*, CSCW '06, pages 299–308, New York, NY, USA. ACM.

Dourish, P. and Bellotti, V. (1992). Awareness and coordination in shared workspaces. In *Proceedings of the ACM conference on Computer-supported cooperative work*, CSCW '92, pages 107–114, New York, NY, USA. ACM.

Durrant, A., Rowland, D., Kirk, D. S., Benford, S., Fischer, J. E., and McAuley, D. (2011). Automics: Souvenir generating photoware for theme parks. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '11, New York, NY, USA. ACM.

Ebner-Priemer, U. W. and Trull, T. J. (2009). Ecological momentary assessment of mood disorders and mood dysregulation. *Psychological Assessment*, 21(4):463–475.

Ellis, C. A., Gibbs, S. J., and Rein, G. (1991). Groupware: Some issues and experiences. *Commun. ACM*, 34:39–58.

Fällman, D. (2003). *In Romance with the Materials of Mobile Interaction : A Phenomenological Approach to the Design of Mobile Information Technology*. PhD thesis, Umeå University, Informatics.

Fischer, J. E. (2009). Experience-sampling tools: a critical review. In *Mobile Living Lab Workshop Proceedings available at http://mll09.novay.nl*. MobileHCI 2009, Bonn, Germany.

Fischer, J. E. and Benford, S. (2009). Inferring player engagement in a pervasive experience. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '09, pages 1903–1906, New York, NY, USA. ACM.

Flintham, M., Giannachi, G., Benford, S., and Adams, M. (2007). Day of the figurines: Supporting episodic storytelling on mobile phones. In Cavazza, M. and Donikian, S., editors, *Virtual storytelling - using virtual reality technologies for storytelling*, volume 4871 of *LNCS*, pages 167–175. Springer.

Fogarty, J., Hudson, S. E., and Lai, J. (2004). Examining the robustness of sensor-based statistical models of human interruptibility. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '04, New York, NY, USA. ACM.

Fogarty, J., Ko, A. J., Aung, H. H., Golden, E., Tang, K. P., and Hudson, S. E. (2005). Examining task engagement in sensor-based statistical models of human interruptibility. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '05, pages 331–340, New York, NY, USA. ACM.

Franklin, M. and Zdonik, S. (1998). "data in your face": push technology in perspective. In *Proceedings of the 1998 ACM SIGMOD international conference on Management of data*, SIGMOD '98, pages 516–519, New York, NY, USA. ACM.

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 *Proceedings of the 5th international conference on Mobile systems, applications and services*, MobiSys '07, pages 57–70, New York, NY, USA. ACM.

Garfinkel, H. (1967). *Studies in ethnomethodology*. Prentice-Hall, Englewood Cliffs, N.J.

Garson, D. G. (2010). Linear mixed models. In *http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm*. North Carolina State University.

Gillie, T. and Broadbent, D. (1989). What makes interruptions disruptive? a study of length, similarity, and complexity. *Psychological Research*, 50:243–250.

Glaser, B. G. and Strauss, A. L. (1967). *The discovery of grounded theory: strategies for qualitative research*. Aldine, Chicago, Illinois, USA.

Gluck, J., Bunt, A., and McGrenere, J. (2007). Matching attentional draw with utility in interruption. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '07, New York, NY, USA. ACM.

Gonzalez, V. M. and Mark, G. (2004). "Constant, constant, multi-tasking craziness": Managing multiple working spheres. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '04, pages 113–120, New York, NY, USA. ACM.

Grandhi, S. A. (2007). Human interruptibility: A relational perspective. In *Group '07 Doctoral Consortium papers*, Sanibel Island, Florida. ACM.

Grandhi, S. A. (2009). Conceptualizing interpersonal interruption management: A theoretical framework and research program. In Quentin, J., editor, *Proceedings of International Conference on System Sciences*, pages 1–10, Hawaii, USA.

Greenberg, S. (2001). Context as a dynamic construct. *Hum.-Comput. Interact.*, 16:257–268.

Harrison, S. and Dourish, P. (1996). Re-place-ing space: the roles of place and space in collaborative systems. In *Proceedings of the ACM conference on Computer supported cooperative work*, CSCW '96, pages 67–76, New York, NY, USA. ACM.

Heidegger, M. (1927, first English edition in 1962). *Being and Time*. Basil Blackwell, Oxford, UK.

Hektner, J. M., Schmidt, J. A., and Csikszentmihalyi, M. (2007). *Experience Sampling Method - Measuring the quality of everyday life*. Sage, Thousand Oaks, London, New Delhi.

Ho, J. and Intille, S. S. (2005). Using context-aware computing to reduce the perceived burden of interruptions from mobile devices. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '05, New York, NY, USA. ACM.

Horvitz, E. (1999). Principles of mixed-initiative user interfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '99, pages 159–166, New York, NY, USA. ACM.

Horvitz, E. and Apacible, J. (2003). Learning and reasoning about interruption. In *Proceedings of the 5th international conference on Multimodal interfaces*, Vancouver, British Columbia, Canada. ACM.

Horvitz, E., Apacible, J., and Subramani, M. (2005a). *Balancing Awareness and Interruption: Investigation of Notification Deferral Policies*, volume 3538 of *Lecture Notes in Computer Science*, pages 433–437. Springer, Berlin/Heidelberg, Germany.

Horvitz, E., Kadie, C., Paek, T., and Hovel, D. (2003). Models of attention in computing and communication: from principles to applications. *Commun. ACM*, 46(3):52–59.

Horvitz, E., Koch, P., Sarin, R., Apacible, J., and Subramani, M. (2005b). Bayesphone: Precomputation of context-sensitive policies for inquiry and action in mobile devices. In Ardissono, L., Brna, P., and Mitrovic, A., editors, *User Modeling 2005*, volume 3538 of *Lecture Notes in Computer Science*, pages 251–260. Springer Berlin/Heidelberg.

Hudson, J. M., Christensen, J., Kellogg, W. A., and Erickson, T. (2002). "I'd be overwhelmed, but it's just one more thing to do": Availability and interruption in research management. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '02, New York, NY, USA. ACM.

Hudson, S., Fogarty, J., Atkeson, C., Avrahami, D., Forlizzi, J., Kiesler, S., Lee, J., and Yang, J. (2003). Predicting human interruptibility with sensors: a wizard of oz feasibility study. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '03, New York, NY, USA. ACM.

Hulley, S. B., Cummings, S. R., Browner, W. S., Grady, D. G., and Newman, T. B. (1988). *Designing clinical research*. Wolter Kluwer, Philadelphia, PA, USA, 3rd edition 2007 edition.

Hurlburt, R. and Akhter, S. (2006). The descriptive experience sampling method. *Phenomenology and the Cognitive Sciences*, 5:271–301.

Iachello, G., Truong, K. N., Abowd, G. D., Hayes, G. R., and Stevens, M. (2006). Prototyping and sampling experience to evaluate ubiquitous computing privacy in the real world. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, CHI '06, pages 1009–1018, New York, NY, USA. ACM.

Ihde, D. (1986). *Experimental Phenomenology*, volume Reprint. Originally published 1977. State University of New York Press, Albany, New York, USA.

Intille, S., Kukla, C., and Ma, X. (2002). Eliciting user preferences using image-based experience sampling and reflection. In *CHI '02 extended abstracts on Human factors in computing systems*, CHI '02, pages 738–739, New York, NY, USA. ACM.

Iqbal, S. T., Adamczyk, P. D., Zheng, X. S., and Bailey, B. P. (2005). Towards an index of opportunity: Understanding changes in mental workload during task execution. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, CHI '05, pages 311–320, New York, NY, USA. ACM.

Iqbal, S. T. and Bailey, B. P. (2005). Investigating the effectiveness of mental workload as a predictor of opportune moments for interruption. In *CHI '05 extended abstracts on Human factors in computing systems*, CHI '05, New York, NY, USA. ACM.

Iqbal, S. T. and Bailey, B. P. (2006). Leveraging characteristics of task structure to predict the cost of interruption. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, CHI '06, New York, NY, USA. ACM.

Iqbal, S. T. and Bailey, B. P. (2007). Understanding and developing models for detecting and differentiating breakpoints during interactive tasks. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, CHI '07, pages 697–706, New York, NY, USA. ACM.

Iqbal, S. T. and Bailey, B. P. (2008). Effects of intelligent notification management on users and their tasks. In *Proceeding of the SIGCHI conference on Human factors in computing systems*, CHI '08, New York, NY, USA. ACM.

Iqbal, S. T. and Horvitz, E. (2007). Disruption and recovery of computing tasks: Field study, analysis, and directions. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '07, pages 677–686, New York, NY, USA. ACM.

Iqbal, S. T., Ju, Y.-C., and Horvitz, E. (2010). Cars, calls, and cognition: Investigating driving and divided attention. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '10, pages 1281–1290, New York, NY, USA. ACM.

Kalyanaramen, S., Ivory, J., and Maschmeyer, L. (2005). Interruptions and online information processing: The role of interruption type, interruption content, and interruption frequency. In *Paper presented at the annual meeting of the International Communication Association*, pages 1–32, New York, NY, USA.

Kapoor, A. and Horvitz, E. (2008). Experience sampling for building predictive user models: a comparative study. In *Proceeding of the SIGCHI conference on Human factors in computing systems*, CHI '08, pages 657–666, New York, NY, USA. ACM.

Khan, V.-J., Markopoulos, P., Eggen, B., IJsselsteijn, W., and de Ruyter, B. (2008). Reconexp: a way to reduce the data loss of the experiencing sampling method. In *Proceedings of the 10th international conference on Human computer interaction with mobile devices and services*, MobileHCI '08, pages 471–476, New York, NY, USA. ACM.

Kinnear, P. R. and Gray, C. D. (2008). *SPSS 15 made simple*. Psychology Press, Hove, East Sussex.

Klasnja, P., Harrison, B. L., LeGrand, L., LaMarca, A., Froehlich, J., and Hudson, S. E. (2008). Using wearable sensors and real time inference to understand human recall of routine activities. In *Proceedings of the 10th international conference on Ubiquitous computing*, UbiComp '08, pages 154–163, New York, NY, USA. ACM.

Krause, A., Smailagic, A., and Siewiorek, D. (2006). Context-aware mobile computing: learning context-dependent personal preferences from a wearable sensor array. *IEEE Transactions on Mobile Computing*, 5(2):113–127.

Lasswell, H. (1948). *The structure and function of communication in society*. Harper.

Latorella, K. A. (1998). Effects of modality on interrupted flight deck performance: Implications for data link. Technical report, NASA Langley Technical Report Server.

Lazar, J., Feng, J. H., and Hochheiser, H. (2010). *Reasearch Methods in Human-Computer Interaction*. Wiley, Chichester, United Kingdom.

Lee, S., Seo, J., and Lee, G. (2010). An adaptive speed-call list algorithm and its evaluation with ESM. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '10, pages 2019–2022, New York, NY, USA. ACM.

Lei, S., Zhang, K., and Sha, E. (2007). Applying situation awareness to mobile proactive information delivery. In *Emerging directions in embedded and ubiquitous computing*, pages 592–603. Springer, Berlin, Heidelberg, Germany.

Light, A. (2008). Transports of delight? what the experience of receiving (mobile) phone calls can tell us about design. *Personal Ubiquitous Comput.*, 12:391–400.

Mark, G., Gonzalez, V. M., and Harris, J. (2005). No task left behind?: Examining the nature of fragmented work. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '05, New York, NY, USA. ACM.

Mark, G., Gudith, D., and Klocke, U. (2008). The cost of interrupted work: More speed and stress. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '08, New York, NY, USA. ACM.

Marti, S. and Schmandt, C. (2005). Giving the caller the finger: Collaborative responsibility for cellphone interruptions. In *CHI '05 extended abstracts on Human factors in computing systems*, CHI '05, New York, NY, USA. ACM.

Mazmanian, M., Orlikowski, W. J., and Yates, J. (2006). CrackBerrys: Exploring the social implications of ubiquitous wireless email devices. In *Proceedings of the EGOS 2006*.

McCrickard, D. S., Catrambone, R., Chewar, C. M., and Stasko, J. T. (2003). Establishing tradeoffs that leverage attention for utility: Empirically evaluating information display in notification systems. *Int. J. Hum.-Comput. Stud.*, 58(5):547–582.

McFarlane, D. C. (1997). Interruption of people in human-computer interaction: A general unifying definition of human interruption and taxonomy. NRL Formal Report NRL/FR/5510-97-9870, Naval Research Laboratory, Washington, USA.

McFarlane, D. C. (1999). Coordinating the interruption of people in human-computer interaction. In Sasse, A. and Johnson, C., editors, *Proceedings of Human-Computer Interaction*, INTERACT, pages 295–303, Amsterdam. IOS Press.

McFarlane, D. C. (2002). Comparison of four primary methods for coordinating the interruption of people in human-computer interaction. *Human-Computer Interaction*, 17(1):63–139.

McFarlane, D. C. and Latorella, K. A. (2002). The scope and importance of human interruption in human-computer interaction design. *Human-Computer Interaction*, 17(1):1–61.

Merleau-Ponty, M. (1962). *The Phenomenology of Perception*. Routledge, London, UK, 2005, originally published in French, 1945 edition.

Meyrowitz, J. (1985). *No sense of place: The impact of electronic media on social behavior*. Oxford University Press, New York, NY, USA.

Miettinen, M. and Oulasvirta, A. (2007). Predicting time-sharing in mobile interaction. *User Modeling and User-Adapted Interaction*, 17(5):475–510.

Miyata, Y. and Norman, D. A. (1986). Psychological issues in support of multiple activites. In Norman, D. A., editor, *User-Centered System Design*, pages 265–284. Lawrence Erlbaum Associates.

Monk, Christopher A., B.-D. D. A. and Trafton, J. G. (2002). The attentional costs of interrupting task performance at various stages. *Human Factors and Ergonomics Society Annual Meeting Proceedings*, 46:1824–1828.

Moore, G. E. (1965). Cramming more components onto integrated circuits. *Electronics*, 38(8).

Mulhall, S. (1996). *Heidegger and Being and Time*. Routledge, 2005, second edition.

Nagel, K. S., Hudson, J. M., and Abowd, G. D. (2004). Predictors of availability in home life context-mediated communication. In *Proceedings of the 2004 ACM conference on Computer supported cooperative work*, New York, NY, USA. ACM.

Nardi, B. A., Whittaker, S., and Bradner, E. (2000). Interaction and outeraction: instant messaging in action. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work*, CSCW '00, pages 79–88, New York, NY, USA. ACM.

Norman, D. A. (1986). Cognitive engineering. In Norman, D. A., editor, *User-Centered System Design*, pages 31–61. Lawrence Erlbaum Associates.

O'Conaill, B. and Frohlich, D. (1995). Timespace in the workplace: Dealing with interruptions. In *Conference companion on Human factors in computing systems*, pages 262–263, Denver, Colorado, USA. ACM.

Oulasvirta, A. (2005). Interrupted cognition and design for non-disruptiveness: The skilled memory approach. In *CHI '05 extended abstracts on Human factors in computing systems*, CHI '05, New York, NY, USA. ACM.

Oulasvirta, A., Raento, M., and Tiitta, S. (2005a). ContextContacts: Redesigning smartphone's contact book to support mobile awareness and collaboration. In *Proceedings of the 7th international conference on Human computer interaction with mobile devices & services*, MobileHCI '05, pages 167–174, New York, NY, USA. ACM.

Oulasvirta, A., Sakar, T., Roto, V., and Kuorelahti, J. (2005b). 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, New York, NY, USA. ACM.

Paelke, V., Oppermann, L., and Reimann, C. (2007). *Mobile location-based gaming*. Springer.

Patel, S., Kientz, J., Hayes, G., Bhat, S., and Abowd, G. (2006). Farther than you may think: An empirical investigation of the proximity of users to their mobile phones. In Dourish, P. and Friday, A., editors, *UbiComp 2006: Ubiquitous Computing*, volume 4206 of *Lecture Notes in Computer Science*, pages 123–140. Springer Berlin / Heidelberg.

Pedersen, E. R. (2001). Calls.calm: Enabling caller and callee to collaborate. In *CHI '01 extended abstracts on Human factors in computing systems*, CHI '01, pages 235–236, New York, NY, USA. ACM.

Perry, M., O'hara, K., Sellen, A., Brown, B., and Harper, R. (2001). Dealing with mobility: Understanding access anytime anywhere. *ACM Trans. Comput.-Hum. Interact.*, 8:323–347.

Picard, R. W. and Liu, K. K. (2007). Relative subjective count and assessment of interruptive technologies applied to mobile monitoring of stress. *Int. J. Hum.-Comput. Stud.*, 65(4):361–375.

Prasopoulou, E., Pouloudi, A., and Panteli, N. (2006). Enacting new temporal boundaries: The role of mobile phones. *European Journal of Information Systems*, 15(3):277–284.

Riley, P. (1976). Discursive and communicative functions of non-verbal communication. In *Proceedings of the Annual meeting of the British Association for Applied Linguistics*, Universite de Nancy II, Nancy, France. C.R.A.P.E.L.

Robertson, T. J., Prabhakararao, S., Burnett, M., Cook, C., Ruthruff, J. R., Beckwith, L., and Phalgune, A. (2004). Impact of interruption style on end-user debugging. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '04, pages 287–294, New York, NY, USA. ACM.

Rodenstein, R., Abowd, G., and Catrambone, R. (1999). OwnTime: a system for timespace management. In *CHI '99 extended abstracts on Human factors in computing systems*, CHI '99, pages 200–201, New York, NY, USA. ACM.

Rouncefield, M., Hughes, J. A., Rodden, T., and Viller, S. (1994). Working with "constant interruption": CSCW and the small office. In *Proceedings of the ACM conference on Computer supported cooperative work*, CSCW '94, pages 275–286, New York, NY, USA. ACM.

Rowland, D., Flintham, M., Oppermann, L., Marshall, J., Chamberlain, A., Koleva, B., Benford, S., and Perez, C. (2009). Ubikequitous computing: Designing interactive experiences for cyclists. In *Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services*, MobileHCI '09, New York, NY, USA. ACM.

Sadler, K., Robertson, T., and Kan, M. (2006). "It's always there, it's always on": Australian freelancer's management of availability using mobile technologies. In *MobileHCI '06: Proceedings of the 8th conference on Human-computer interaction with mobile devices and services*, pages 49–52, New York, NY, USA. ACM.

Sawhney, N. and Schmandt, C. (2000). Nomadic radio: Speech and audio interaction for contextual messaging in nomadic environments. *ACM Trans. Comput.-Hum. Interact.*, 7(3):353–383.

Schegloff, E. A. (1968). Sequencing in conversational openings. *American Anthropologist*, 70(6):1075–1095.

Schilit, B. N., Adam, N., and Want, R. (1994). Context-aware computing applications. In *Proceedings of the IEEE Workshop on Mobile Computing Systems and Applications*, pages 85–90.

Schutz, A. (1932). *The phenomenology of the social world*. Northwestern University Press, Evanston, Illinois, USA.

Scollon, C. N., Prieto, C.-K., and Diener, E. (2009). Experience sampling: Promises and pitfalls, strength and weaknesses. In Diener, E., editor, *Assessing Well-Being*, volume 39 of *Social Indicators Research Series*, pages 157–180. Springer Netherlands.

Shannon, C. E. (2001). A mathematical theory of communication. *SIGMOBILE Mob. Comput. Commun. Rev.*, 5:3–55.

Shen, X. and Chen, H. (2008). An empirical study of what drives consumers to use mobile advertising in china. In *GPC-WORKSHOPS '08: Proceedings of the 2008 The 3rd International Conference on Grid and Pervasive Computing - Workshops*, pages 158–163, Washington, DC, USA. IEEE Computer Society.

Sohn, T., Li, K. A., Griswold, W. G., and Hollan, J. D. (2008). A diary study of mobile information needs. In *Proceeding of the SIGCHI conference on Human factors in computing systems*, CHI '08, pages 433–442, New York, NY, USA. ACM.

Sproull, L. S. (1984). The nature of managerial attention. *Adv. in Information Processing in Organizations*, 1:9–27.

Stone, A. and Shiffman, S. (2002). Capturing momentary, self-report data: A proposal for reporting guidelines. *Annals of Behavioral Medicine*, 24:236–243.

Stone, A. A., Kessler, R. C., and Haythomthwatte, J. A. (1991). Measuring daily events and experiences: Decisions for the researcher. *Journal of Personality*, 59(3):575–607.

Streefkerk, J. W., van Esch-Bussemakers, M. P., and Neerincx, M. A. (2006). Designing personal attentive user interfaces in the mobile public safety domain. *Computers in Human Behavior*, 22(4):749–770. Special issue: Attention aware systems.

Su, N. M. and Mark, G. (2008). Communication chains and multitasking. In *Proceeding of the SIGCHI conference on Human factors in computing systems*, CHI '08, New York, NY, USA. ACM.

Suchman, L. (1985). Working relations of technology production and use. In MacKenzie, D. and Wajcman, J., editors, *The social shaping of technology - second edition*, pages 258–268. Open University Press, Buckingham, UK and Philadelphia, PA, USA, 1999, second edition.

Suchman, L. A. (1987). *Plans and situated actions: the problem of human-machine communication*. Cambridge University Press, New York, NY, USA.

Svanaes, D. (2001). Context-aware technology: A phenomenological perspective. *Hum.-Comput. Interact.*, 16:379–400.

Szóstek, A. M. and Markopoulos, P. (2006). Factors defining face-to-face interruptions in the office environment. In *CHI '06 extended abstracts on Human factors in computing systems*, CHI '06, New York, NY, USA. ACM.

Tang, J. C., Yankelovich, N., Begole, J., Van Kleek, M., Li, F., and Bhalodia, J. (2001). ConNexus to Awarenex: Extending awareness to mobile users. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '01, pages 221–228, New York, NY, USA. ACM.

ter Hofte, G. H. (2007). Xensible interruptions from your mobile phone. In *MobileHCI '07: Proceedings of the 9th international conference on Human computer interaction with mobile devices and services*, pages 178–181, New York, NY, USA. ACM.

ter Hofte, G. H., Otte, R. A. A., Kruse, H. C. J., and Snijders, M. (2006). What's your lab doing in my pocket? supporting mobile field studies with SocioXensor. In *Conference supplement of CSCW '06*, CSCW '06, pages 109–110, New York, NY, USA. ACM.

Tolmie, P., Crabtree, A., Rodden, T., and Benford, S. (2008). "Are you watching this film or what?": Interruption and the juggling of cohorts. In *Proceedings of the ACM conference on Computer supported cooperative work*, CSCW '08, New York, NY, USA. ACM.

Tsang, M. M., Ho, S.-C., and Liang, T.-P. (2004). Consumer attitudes toward mobile advertising: An empirical study. *Int. J. Electron. Commerce*, 8(3):65–78.

Vastenburg, M. H., Keyson, D. V., and de Ridder, H. (2004). Interrupting people at home. In *Proceedings of the 2004 IEEE conference on systems, man and cybernetics*, volume 1, pages 59–64. Omnipress.

Vertegaal, R., Dickie, C., Sohn, C., and Flickner, M. (2002). Designing attentive cell phone using wearable eyecontact sensors. In *CHI '02 extended abstracts on Human factors in computing systems*, CHI '02, pages 646–647, New York, NY, USA. ACM.

Wang, H.-C., Fussell, S. F., and Setlock, L. D. (2009). Cultural difference and adaptation of communication styles in computer-mediated group brainstorming. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '09, pages 669–678, New York, NY, USA. ACM.

Weiser, M. (1991). The computer for the 21st century. *Scientific American*, 265(3):94–104.

Weiser, M. and Brown, J. S. (1996). Designing calm technology. *PowerGrid Journal*, 1.

Whittaker, S., Frohlich, D., and Daly-Jones, O. (1994). Informal workplace communication: What is it like and how might we support it? In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '94, pages 131–137, New York, NY, USA. ACM.

Wiberg, M. and Whittaker, S. (2005). Managing availability: Supporting lightweight negotiations to handle interruptions. *ACM Trans. Comput.-Hum. Interact.*, 12(4):356–387.

Winograd, T. and Flores, F. (1985). *Understanding computers and cognition*. Ablex Publishing Corp., Norwood, NJ, USA.

Witt, H. and Drugge, M. (2006). HotWire: An apparatus for simulating primary tasks in wearable computing. In *CHI '06 extended abstracts on Human factors in computing systems*, CHI '06, pages 1535–1540, New York, NY, USA. ACM.

Xia, L. and Sudharshan, D. (2002). Effects of interruptions on consumer online decision processes. *Journal of Consumer Psychology*, 12(3):265–280.

Yoshiuchi, K., Yamamoto, Y., and Akabayashi, A. (2008). Application of ecological momentary assessment in stress-related diseases. *BioPsychoSocial Medicine*, 2(1).

Zacks, J. M., Braver, T. S., Sheridan, M. A., Donaldson, D. I., Snyder, A. Z., Ollinger, J. M., Buckner, R. L., and Raichle, M. E. (2001). Human brain activity time-locked to perceptual event boundaries. *Nature Neuroscience*, 4(6):651–655.

Zeigarnik, B. (1927). Das behalten erledigter und unerledigter handlungen. *Psychologische Forschung*, 9(1):1–85.

COLOPHON

This thesis was typeset with $\LaTeX\,2_\varepsilon$ using André Miede's `classicthesis` package inspired by Hermann Zapf's *Palatino* and *Euler* type faces (Type 1 PostScript fonts *URW Palladio L* and *FPL* were used, made available by Malte Rosenau and Ulrich Dirr).

The typographic style was inspired by Bringhurst's *The Elements of Typographic Style*. It is available for LaTeX via CTAN as "`classicthesis`".

Final Version as of 20th March 2012 at 13:31.