DOCTOR OF PHILOSOPHY

Guided real time sampling using mobile electronic diaries

Morrison, Kenneth

Award date:
2010

Awarding institution:
University of Dundee

Guided real time sampling using mobile electronic diaries

Kenneth Morrison

2010

University of Dundee

Guided Real Time Sampling Using Mobile Electronic Diaries

Kenneth Morrison

Doctor of Philosophy

University of Dundee

March 2010

Table of contents

List of tables

List of figures

Acknowledgements

I would like to thank those who contributed to making this work possible:

Professor Ian Ricketts for his advice, encouragement and support over the past three years. Professor Nigel Pitts and Professor Frank Sullivan for their advice, support and endeavours that allowed this work to be possible.

Dr. Martyn Jones and Professor Derek Johnston for providing their expertise.

The School of Computing for providing a stimulating and supportive environment, especially the residents of Pod 2.15.

All the system users and study participants for giving so generously of their time.

My friends and family.

This thesis is dedicated to my Dad who passed away in 2008.

The work reported in this thesis has been funded in part by the UK Medical Research Council, the Social Dimensions of Health Institute (Universities of Dundee and St. Andrews), the Dental Health Services Research Unit and the School of Computing (University of Dundee).

Declaration by the candidate

I hereby declare that the work described in this thesis is my own; that I am the author of this thesis; that it has not previously been put forward in submission for any other higher degree or qualification; and that I have consulted the references listed in the table of references herein.

Kenneth Morrison

March 2010

Declaration by the supervisors

We certify that Kenneth Morrison has satisfied all the terms and conditions of the regulations made under Ordinances 12 and 39, and has completed the required nine terms of research to qualify submitting this thesis in application for the degree of Doctor of Philosophy.

Prof. Frank M. Sullivan

Prof. Nigel B. Pitts

March 2010

Associated Outputs and Awards

Publications

Pocket Interview – a handheld data collection tool, *Scottish School of Primary Care Conference*, Carnoustie, UK, January 2008.

Pocket Interview - a secure electronic data collection and diary tool, *HC2009*, Harrogate, UK, 28-30 April 2009. ISBN: 09546971 9 7

A secure electronic diary and data collection tool, *3rd International Conference on Pervasive Computing Technologies for Healthcare (2009),* London, UK, 31 March - 3 April 2009. ISBN: 9789639799 42 4

Pocket Interview - a secure electronic data collection and diary tool, *eHealth International Journal* (2009). 5(1), http://www.ehealthinternational.org. ISSN:14763591

Author list for publications:

K. Morrison, School of Computing, University of Dundee
I. W. Ricketts, School of Computing, University of Dundee
M. C. Jones, School of Nursing and Midwifery, University of Dundee
D. W. Johnston, School of Psychology, University of Aberdeen
N. B. Pitts, Dental Health Services Research Unit, University of Dundee
F. M. Sullivan, Clinical & Population Sciences & Education, Univ. of Dundee

Patent

Patent protection is currently being sought for Pocket Interview's Guiding mechanism: "Improvements In and Relating to Electronic Data Collection", filed with the Intellectual Property Office, September 2009

Software

Pocket Interview is freely available from the software website:

http://www.computing.dundee.ac.uk/acprojects/PocketInterview/

Grants

An Exploration of the use of handheld computers to determine the relative contribution of different cognitive, attitudinal, social and organizational factors in health workers' decisions to decontaminate hands, Dr Karen Lee (Nursing & Midwifery), Emma Burnett, (Nursing & Midwifery), Dr Martyn Jones (Nursing & Midwifery), Kenny Morrison (School of Computing), Prof. Ian Ricketts (School of Computing) £4301 funded by NHS Scottish Infection Research Network (SIRN), February - September 2009.

Evaluation of stress, well-being and clinical performance in recently graduated veterinary surgeons, Dr Richard Mellanby (Univ. Edinburgh), Dr Neil Hudson (Univ Edinburgh), Prof. Susan Rhind (Univ. Edinburgh), Prof. Ian Ricketts (Univ. Dundee) Kenny Morrison (Univ. Dundee), Dr Martyn Jones (Univ. Dundee), £5000 funded by MEDEV, the Higher Education Academy Subject Centre for Medicine, Dentistry and Veterinary Medicine, July 2008 – Sept 2010.

Abstract

This thesis describes the motivation and development of Pocket Interview, an easily configurable handheld electronic data collection and diary tool. The system can be used to apply 'experience sampling' methods that allow the collection of data in real-time and in the user's natural environment. Pocket Interview can prompt the user to make diary entries at fixed and/or random intervals. The system is configured via graphical user interfaces that are shown to be easily usable by non-computing users.

Pocket Interview includes an option that allows this sampling to be 'Guided' whereby inconvenient prompts are temporarily deferred until a more convenient time through the use of contextual audio information. Subjects participating in real-time studies require high levels of commitment and exhibit difficulties maintaining their motivation. Guiding can offer to reduce the perceived burden on the user, improve response rates, increase the quantity of replies and the quality of those replies.

This thesis describes a series of studies that investigate the following:

- What are the more convenient times for sampling and how can they be detected?
- Does Guided real-time sampling improve the data quality and participant compliance rates?
- Participants attitudes towards mobile devices automatically gathering their context information

Guiding is a strategy that could be applied to all context-aware computing, phone call or message delivery and indeed all other prompting. As computing power continues to expand and more powerful mobile devices become available we will see an increase in the quantity and sophistication of applications that interrupt their users. This will add to user's feelings of overload. To maximise user acceptability designers of computing systems require strategies, such as Guiding, to minimise the interruptions caused by proactive prompting.

1. Introduction

1.1 Motivation

Diary studies allow subjects to report on experiences and events during their daily lives. Computerised diary tools have been shown to be less susceptible to recall errors and encourage higher user-compliance than alternatives such as pen and paper. They can also provide greater flexibility to the study administrators and reduce errors during data-collection and entry.

Employing computerised diaries allows for a variety of protocols. As well as recording significant events, the user can also be prompted to provide entries at fixed and/or random intervals. However, the users may be interrupted during their daily activities that may prove irritating, disruptive and taxing especially when prompts are more frequent. Subjects participating in such studies require high levels of commitment and exhibit difficulties maintaining their motivation. This method of recording data in real-time and in the users natural environment has previously been referred to as 'experience sampling' and 'ecological momentary assessment'.

The use of contextual information could be used to help reduce the burden on data-providers. Is there a 'better' time to prompt these users when they are more interruptible? Guiding the sampling and allowing prompts to be deferred temporarily until a less disruptive time through the use of contextual information could potentially lessen the perceived burden on the user. Decreasing the annoyance and inconvenience may result in an increase in the volume and quality of the data collected, an improvement in compliance rates and potentially the quality of the data itself.

Guiding may present a viable strategy for all context-aware computing, message delivery and prompting. Designers of a wide range of computing systems require strategies for minimising interruptions of proactive messages. Computing power continues to improve which permits more powerful mobile devices and allows an increase in the amount and sophistication of applications. Many new and innovative mobile applications will proactively deliver information or prompt the user adding to feelings of information overload.

Pocket Interview is a new easily-configurable handheld electronic data collection and diary tool that can be used to apply ‘experience sampling’ methods. The system includes an option to allow the sampling to be guided through the use of environmental audio analysis. This approach uses ubiquitous mobile technologies and sensors that are embedded within the mobile device. The system does not require the use of wearable sensors as the participants may be resistant to systems that require them to wear additional equipment. There are no external sensors placed in the environment as this will tie the system to a static location. There is no development of new sensors as this will require design, development, deployment and evaluation which can be expensive, error-prone and can waste significant time and resources on something that may not even meet the desired requirements.

1.2 Thesis organisation

Chapter 2 – *Using contextual information to help predict computer users interruptibility and availability*

Chapter 2 presents a review of the current literature relating to Human Computer Interaction based interruptibility and availability. What desktop based systems have been reported that use predictive models? What contexts do they gather and strategies do they employ to model their user’s interruptibility and availability? What are the effects of interruptions on computer users, mobile users and surrounding people? More recent research focuses on mobile devices that seek user’s attention and assessing their availability. What progress has been made regarding mobile systems that aim to predict their users availability?

Chapter 3 – *Motivation and design of a handheld data collection and diary tool*

Chapter 3 describes the motivation behind the development of Pocket Interview, a new handheld data collection and diary tool. Background is provided relating to the use of handheld computers in healthcare and by researchers as data collection devices. As well as the benefits of using electronic diaries and recording data in real-time are described and the challenges involved with real-time data collection are also described e.g. low/decreasing compliance rates, participant motivation. What tools are currently available and what are their strengths and weaknesses?

Chapter 4 – *Development and evaluation of Pocket Interview*

Chapter 4 describes the development, evaluation and current uses of Pocket Interview. The 2 main parts of the system are described: the administration tool and the client. The novel features of the system are detailed, such as it's ease of use via graphical user interfaces, ease of installation, availability on multiple platforms e.g. PDA, smartphone, desktop/laptop. Other novel features include encryption, demo mode, the wide variety of question types available, audio recording, picture/thumbnail display and organisation of data through easy integration with Access. An evaluation is performed with administrator users.

Chapter 5 – *Description of Studies*

There are a number of studies reported and results analysed throughout Chapters 6, 7 and 8. To aid clarity the design and methodology of these experiments is described in this Chapter.

Chapter 6 – *Predicting availability of diary participants*

Chapter 6 describes a series of studies related to the availability of electronic diary users to make real-time data entries. What are the more convenient times for sampling and how can they be detected? Can the use of sound information gathered using mobile devices in work (and other) environments help predict availability of participants? The challenges using mobile devices such as microphone quality, processing power, battery life and working with a PDA's power modes. What are the effects of job role and gender when measuring availability?

Chapter 7 – *Guided sampling*

Chapter 7 details Guiding, a novel method for deferring inconvenient electronic diary interruptions to a better time and a series of studies that evaluate the method. What other methods for improving real-time sampling are there? A guided method that involves analysing the users' nearby environmental audio is implemented and evaluated. Does Guided sampling improve the user experience, compliance rates and the data collection overall?

Chapter 8 – *Privacy issues*

Chapter 8 describes a topical study of privacy attitudes related to sharing contextual information with mobile devices. Guided sampling will require the sampling of local contextual information. When microphones and other sensors are employed they may prevent systems being adopted through fears that the information gathered will be abused. What types of contextual information are participants willing to share with a mobile computer service that would help predict their availability and does not compromise their privacy concerns?

Chapter 9 – *Contributions and future work*

Chapter 9 summarises the lessons learned over the course of this thesis and provides suggestions for future research and system development.

2. Using contextual information to help predict computer users interruptibility and availability

2.1 Introduction

Interruptions are ubiquitous in people's lives today. This is partly due to the increasing number of computing devices that seek their user's attention and therefore interrupt them as they go about their day. It seems reasonable to state that an interruption is an event that breaks someone's attention from their current task to focus temporarily on another event. People can quickly decide when it is appropriate to interrupt one another and are able to interrupt each other without seeming rude or intrusive. Rarely will a person immediately interrupt another when they seem busy, rather they will wait for the person to finish or temporarily pause their current task. Computer systems, however, are unaware of social conventions and can either interrupt at inappropriate times or they can be set to be passive and then wait for user-interaction.

Experience sampling methods (ESM) have been employed by the research community for decades. These methodologies are designed to prompt research study participants to record their thoughts, feeling and actions throughout their daily lives. However, if the participants are interrupted during their daily activities this may prove irritating and disruptive especially when the prompts are more frequent. Deferring the prompts, or 'Guiding' during inappropriate times, temporarily until a more convenient time, may help decrease annoyance and result in an increase in the volume and quality of the data collected. Guiding, through the use of contextual information, may also prove to be a viable strategy for all for all systems that interrupt their users.

2.2 Interruptions

The study of interruptions has a long history dating back to the widely acknowledged classic work of Zeigarnik in the 1920s [1]. What has now been dubbed the Zeigarnik effect states that people can recall details of tasks that have been interrupted better than completed tasks after it was noticed that waiters had better recollections of orders that were still to be paid. This may be due to tension that is released once the task has been completed.

More recently there has been much research and study of interruption with regards to Human Computer Interaction (HCI). There are many new and innovative applications and computing devices that compete for their users attention e.g. desktop computers, mobile phones, personal digital assistants etc. These devices will proactively deliver information or prompt their users. Currently, designers of a wide range of computing systems require strategies for minimising the interruptions caused by their proactive messages.

2.3 Effects of interruptions

Computing power continues to improve which permits more powerful desktop and mobile devices and allows an increase in the amount and sophistication of their applications. This has added to what has been called "information overload", "interruption irritability" or "infomania". Infomania has been described as "the mental state of continuous stress and distraction caused by the combination of queued messaging overload and incessant interruptions" [2]. In recent years there has been an increased discussion of "attention economy" and "information overload". It has been suggested by Hudson, Christensen, Kellogg and Erickson that the "important commodity in the current economy is no longer money or other physical resources, it is individuals time and attention" [3].

Bailey et al demonstrated the disruptive effect of a computer interruption on a user's current task performance [4]. They found a peripheral task causes a greater increase in anxiety when it is presented during a primary task than when it is presented just after the completion of that task. Their experiments showed that interruptions have a disturbing effect on people's overall performance. Other research has shown that just the notification of an incoming interruption can affect task performance even if the message itself is ignored [5].

Human attention and memory also has limits that computer systems should take into account. Interruptions increase the load on human memory and can affect future tasks. O'Conaill and Frohlich found interruptions of tasks to be the most common reason for failure to remember future tasks, a phenomenon which has been referred to as *prospective memory* [6]. Cutrell, Czerwinski and Horvitz showed that instant messaging applications can decrease human performance and can have a negative effect on the memory required for resuming primary tasks [5].

Hudson et al propose that the focus should not be on reducing interruptions but making them more effective [3]. While computer systems can be temporarily disabled from giving interruptions their users often forget to re-enable the system afterwards therefore it may be better to introduce strategies that allow computer to interrupt at better times. The work presented in this thesis reflects this by assessing availability and deferring intrusive interruptions until what can be considered a more convenient time therefore reducing the effects of 'infomania' and the disruptive nature of interruptions.

2.4 Work-based studies

Estimating human interruptibility or a person's availability is known to be a very hard task. When using contextual information to model interruptibility a specific user-group or application is normally adopted e.g. office-workers, instant messaging or email. The work described in this thesis will focus on electronic methods of experience sampling and office-based university workers. The majority of recent work researching interruptibility and availability has also focused on office-based workers. Research has shown that interruptions are disruptive in the workplace and affect productivity.

However they may also be crucial as they can provide information that may be needed to work effectively. They may also be necessary (e.g. bathroom breaks) or self-initiated (e.g. coffee-breaks).

Perlow studied the time usage of a group of 45 software engineers in an information technology corporation [7]. She found a crisis mentality where a person up against a critical deadline feels free to interrupt colleagues which delays their work causing them to fall behind which in turn leads them to interrupt a colleague. Perlow also found a lack of management recognition of this issue. To help address this Perlow introduced quiet times where workers worked alone without interruption that seemed to improve group productivity but the effect proved to be short-lived.

Other researchers have shown that presence is not enough to determine interruptibility with some corporate research managers reporting a tension between the value of an appropriate interruption versus the disruption caused by in inappropriate one. The can lead to them being forced to physically move away from their computer or even their office when they do not want to be interrupted [3].

Hudson, Fogarty, Atkeson, Avrahami, Forlizzi, Kiesler, Lee, and Yang presented a number of experiments on modelling user interruptibility. They collected 600 hours of video and audio data of four office-based research managers [8]. Experience sampling was used to obtain 672 samples of the subjects interruptibility rated on a 1-5 scale. They trained models on simulated features and obtained scores of up to 78% when the data was grouped as a 2-class problem, highly non-interruptible vs. the rest. A two-class problem of the two least interruptible against the other three reduced the score while recognition of all five classes reduced the score further to 45%. They argue that five-way classification allows for greater user control although they did not ask the users themselves if the five-point scale was more meaningful than a binary scale. In a follow-up study they collected another 975 self-reports from 10 subjects and investigated different user-groups (managers, researchers, interns) and sensor sets (full-set, laptop only and without audio) [9]. They compared their models performance with human observers and demonstrated it is possible to create sensor-based models and estimates of human interruptibility within office workers that perform better than human observations. A more thorough analysis of the data is performed using ROC curve analysis [10].

Further work by Fogarty, Ko, Aung, Golden, Tang and Hudson consisted of observing subjects during a programming task [11]. Reaction time after an interruption was used as a measure of interruptibility. Three levels of interruptibility were identified from the data: an 'interruptible' state, an 'engaged' state and a 'non-interruptible' state. The statistical model for the two-class problem ('interruptible' against the other two) correctly identified 74.8% of the interruptible and 67.5% of the non-interruptible samples.

Another study of office workers investigated building models of their interruptibility [12]. The study demonstrated that the cost of interruption of a user was determined with a 73% accuracy using Outlooks calendar, the ambient acoustics in the office, visual analysis of the user's pose and desktop activity. The classifier was restricted to a particular physical space and therefore could not be extended to be more general.

2.5 Laboratory studies

There have been a number of reported studies that have been performed in laboratory settings. It can be argued that results obtained in these controlled environments can produce results that can only be applicable in those circumstances. This may be especially true when related to the development of applications that will be used in real-world situations such as tools for real-time data collection. However, studies in the lab may still provide useful insight.

Lab studies have included interrupting users during computer based navigation and object-collection tasks. Hess and Detweiler show that users become more expert in handling interruptions and not the actual task they are performing [13]. Mcfarlane explored different interruption strategies in a lab setting. He introduced and explored 4 different ways in which computer systems can present interruptions to users

- Immediately
- Scheduled – on a regular basis
- Negotiated – in cooperation with user
- Mediated – automatically at the 'best' time

The experiments involved a primary task (playing a handheld computer game) being interrupted by secondary task (a matching problem). He found that the negotiated method worked best and when implemented in such a way that the user initially receives small and unobtrusive notification and can decide by their self when to deal with it [14]. Drugge, Nilsson, Liljedahl, Synnes, and Parnes described similar studies with the same results that suggests that the notification method and cooperation with the user can be important factors [15] [16]. Scope is an awareness based system which uses these results by presenting summaries of interruptions from multiple sources such as email, instant messaging and appointments and then allowing the user decide how to deal with them [17].

2.6 Desktop systems

There is a wide range of desktop systems reported in the last decade that attempt to infer their user's availability. These tend to be systems that use predictive models of user-interruptibility constructed from earlier contextual sensor-based data to inform the decision-making regarding the best time to prompt the users.

Busybody is a desktop based system that can build decision-theoretic models of interruptibility by asking users to assess their own interruptibility intermittently during a training phase [18]. The system can log events such as typing, mouse movements, windows in focus, recent sequences of applications and window tiles and also more high-level statistics about switching among applications and windows. Contextual information such as time, day of week, computer name and electronic calendar and wireless signals can be also be recorded as well as a conversation-detection system using the Sphinx speech recognition module. During training the system can create large vectors of computer-activity based on a combination of these features to constructs models of the uses interruptibility. At run-time the system provides other applications with a current expected cost of interruption or information when the expected cost of interruption exceeds a user-set threshold. However, users may be averse to a system such as this due to the training phase that learns by continuously prompting (Figure 2-1).

Figure 2-1 BusyBody training mode
Users asked to assess their interruptibility [18]

The Ambush system predicts meeting attendance based on calendar information. The users label which meetings they have already attended and the system builds a predictive Bayesian model based on this information. No formal evaluation of the system was reported although the authors acknowledged there was a substantial burden placed on users by keeping an attendance diary [19].

Horvitz, Jacobs and Hovel discussed the issue of attention-sensitive alerting and the trade-off between value of information versus the cost of interruption and introduced the Priorities system [20]. The Priorities system infers the user's attention through sound and computer-activity and modelled the net value of alerts. Alerts are then presented differently depending on this value e.g. opening messages automatically or using specific sounds. Different audio notifications can be played for different levels of criticality. Text analysis of incoming email can be performed and critical messages forwarded to mobile phones. It also includes a smart out-of-office system where the sender of a critical email can be given an estimate of when recipient is likely to read the message.

The Coordinate system expanded on the Priorities system to develop a notification platform [21]. Coordinate is a multi-device system including a software development kit and considers computer activity, electronic calendar info, audio and video and location to detect the users last sensed location. This information is used to reason about a person's availability on multiple computing devices.

Further work by Horvitz, Apacible, Subramani, Sarin, Koch, Cadiz, Narin, and Rui has included the BestCom [22] and Bayesphone [23] systems. BestCom monitors incoming phone calls and intelligently reroutes them based on factors such as the receiver's availability or the caller's identity. The system allows the users themselves to define rules for their own interruptibility and importance of messages and will provide callers an indication of the receiver's availability or allow them to leave messages. Bayesphone allow the integration of mobile phones into systems such as BestCom. Bayesphone runs on a desktop computer which performs analysis offline and then determines how a phone should behave during scheduled times in the calendar. These precompiled policies are synchronised with the phone and at the same time data is collected from the phone (e.g. whether they attended a particular meeting) that can be used to update the policies.

MyVine is a context aware instant messaging communication client that uses built-in microphones on laptop computers to sense nearby speech [24] and combined with location, computer and calendar information the system models each users availability for communication. The system follows a socially translucent approach where contacts are greyed out depending on their availability and allows the user to choose how and when they should contact (Figure 2-2). The work involving the MyVine system introduces a distinction between a person's presence and their actual availability for interruption. However, it was found the senders focused only on the presence of the receiver as an appropriate time for communication rather than availability. This finding raises the question of whether sharing an indication of unavailability does actually reduce the number of inappropriate interruptions when the initiator is another person.

Figure 2-2 MyVine instant messaging
Contacts are greyed out depending on their availability [24]

Other systems designed to connect people at the right time and the right place include the Connector Service and MyConnector [25, 26]. They also aim to identify the best medium for socially appropriate communication and include personal agents that selectively facilitate some calls while blocking others.

2.7 Automated sampling

By extending previous work on the BusyBody system that uses desktop activity, calendar information and contextual information such as ambient audio and wifi signals Kapoor and Horvitz performed comparative analysis of four different methods of experience sampling [27] in an effort to maximise the value of each prompt.

Recently researchers have begun exploring automated experience sampling to collect data to build models of user interruptibility. Most of these procedures rely on random sampling or simple heuristics to trigger the sampling such as predefined locations being used to flag an appropriate time for a prompt. This approach requires that cues are known in advance and that there are reliable sensors in place that can be used for detecting these cues [28, 29]. Variants of random sampling have also included methods which allows users to adjust the parameters of the sampling such as the amount and distribution [18].

The four methods compared by Kapoor et al. were random sampling (four times an hour) and three other dynamic methods that employed inferential mechanisms seeking to be selective and identify the best times and therefore maximize the value of each probe. These dynamic methods were:

- uncertainty probe, a predictive model constructed with data collected so far
- a decision-theoretic probe based on probabilities related to users internal state and the expected value of information
- a dynamic decision theoretic probe which extends the previous method for addressing potential dynamics of context e.g. changes in location and introduces a 'value of forgetting' variant

This was a 2-week study with 44 participants with different roles within a research environment such as software developer, researcher, project manager and group manager. The software was installed on the participant's desktop or laptop computers. Some users experienced technical difficulties so 37 subjects remained after these issues.

The dynamic methods proved to be less annoying than random sampling though this may be as the number of prompts was actually reduced. Random sampling could produce an average of 21.35 prompts per day while the dynamic methods could produce on average as few as 4.65 prompts per day. However, this may not be preferable for an administrator/data-collector as they may require a certain number of samples within a predefined time period.

Chapter 3 includes a closer look at real-time data collection methods such as Experience Sampling and the tools and methods currently available to facilitate these methodologies.

2.8 Human activity

There are large amounts of work investigating human activity recognition that may be out with the scope of this work though there are systems of interest and worth mentioning as they make use of common sensor information to infer activity and could be used to inform availability.

Oliver, Garg and Horvitz describe the SEER system that employs the use of video, audio and desktop information to train models of human activity [30]. After training, the system was shown to identify typical office activities e.g. phone conversation, face-to-face conversation, giving a presentation and other office work. The system can also be selective where the value provided by a sensor may not justify the computational cost of enabling (e.g. vision based sensors). However, like other desktop systems SEER is restricted to the environment it was trained and will require at least partial retraining whenever this environment changes.

Horvitz, Breese, Heckerman, Hovel and Rommelse present the Lumiere system which monitored users application activity and working from a model of likely goals within the application predicts a users current goal and whether they need assistance [31]. The system created a personal competency profile that is updated when a person views documentation or successfully completes task and this is also considered when providing assistance. Lumiere served as the prototype for the Microsoft Office assistant. The assistant was represented by a variety of animated characters that popped up when it was decided the user might require assistance. It was widely considered among users as intrusive and annoying and was removed from more recent versions altogether. This highlights computer users annoyance when being interrupted during the middle of tasks and the need for designers to adopt appropriate strategies and methods acceptable to these users.

Begole, Tang and Hill record computer activity every minute and model and visualise patterns of a persons presence at their computer [32]. Periods of inactivity are clustered to model trends and therefore recurring events such as meetings and the length of their commute can be detected. Participants in their study, however, had some reservations regarding their privacy and had concerns about the information being taken out of context. They would like some control over how the person would view this information to avoid projecting a negative impression e.g. consistently being late for work.

Avrahami and Hudson model a computer users availability and responsiveness to instant messages [33]. Statistical models are created for users based on the person's computer activity and whether or not they have responded to previous instant messages. A weakness of the system however is that it is unaware of the content of messages sent and received and user responsiveness can incorrectly reflect availability when the user may simply be responding with a quick message indicating unavailability. Other work by the group includes the QnA system which did include some simple message analysis with a method for searching for the appearance of questions in instant messages and alerts user when they receive a non-trivial question [34]. A busy person can respond to identified questions and defer other messages until later.

2.9 Mobile systems

The majority of prior work on HCI based interruptions has focused on desktop computing applications, office environments and utilise a mixture of sensors. These prior studies suggest that only a small set of sensors may actually be needed to provide valuable information regarding the user's interruptibility. More recent research has focused on studies investigating interruptions for mobile users. There is an increasing number of mobile devices that seek users attention therefore it is essential to minimise interruptions and distractions caused to the users and the surrounding environment.

2. Using contextual information to help predict peoples interruptibility and availability

Context awareness can be particularly important for mobile computing as the user's situation may change quickly. Mobile devices present an increased challenge compared with desktop systems. They cannot rely on mouse or keyboard use to determine a person's presence or current activity. When designing for mobile devices it is also important to recognise the restricted processing power and limited battery life available.

Mobile phones are currently the most ubiquitous communication device the world over. There are many benefits offered by mobile phones such as their roaming ability i.e. they do not require a fixed location to make/receive calls, the ability to be in constant contact, the ability for emergency contact and the added sense of security that comes with this. However these benefits also come with a cost of increased interruptions and interaction demands. These demands can also be constrained by the devices capabilities and the user's current activity. They may even be in a situation where it is dangerous e.g. driving.

While mobile technologies have allowed more information and people to reach us than ever before though it is estimated that 50 percent of phone communications fail because they do not happen at a convenient time [26]. Therefore a number of mobile systems are being developed that monitor their owners context as a means of assessing their availability e.g. they can let callers make more informed decisions whether or not to make a call based on the receiver's contextual information. These are techniques that may make mobile phones seem less intrusive to those others nearby. It has been shown that most people consider use of mobile phones in public to be annoying [35, 36]. Wei and Leung conducted a large study and asked about contexts in which people find mobile use annoying and found that 81% responded restaurants/cafes, 80% classes or libraries and 79% airport/train stations [37]

ContextPhone is a smart phone framework that allows developers to build mobile based applications that can share their context with others using the same system [38]. Awarenex [39] and Live Addressbook [40] are systems that allow users to see other user's locations and availability status set by their owner in a similar way to buddy lists in instant messaging applications.

Hinckley and Horvitz modelled interruptibility by considering the users likelihood of response and their previous and current activity [41]. They used a PDA prototype and attached three sensors. Linear accelerometer and touch sensors that detect when a user is holding the device and an infrared proximity sensor that detects range to nearby objects so the phone can attract users attention using an appropriate notification modality e.g. if the user if holding the phone then it may not be necessary for the phone to ring as a vibration will be enough to alert the user.

Siewiorek, D., A. Smailagic, J. Furukawa, A. Krause, N. Moraveji, K. Reiger, J. Shaffer, and F. Wong present Sensay [42] a prototype context-aware mobile phone which can dynamically change alert method (e.g audio, vibration) and volume based on the users activity and the surrounding environmental information (Figure 2-3c). The system first checks whether the user is interruptible by checking their calendar and then by analysing audio information. The user is required to wear two microphones mounted on their throat (to detect speech) and chest (ambient audio). Their model focuses mainly on conversation and suggests that users may not want to be interrupted during face-to-face or phone conversations. However, the mobile phone was only tested on lab researchers and the accuracy of the sensors detecting the context of the user was simulated under laboratory settings.

2.10 Wearable sensors

Kern, Antifakos, Schiele and Schwaninger propose the use of wearable sensors which can include accelerometers to model users interruptibility [43, 44]. They introduced and validated a model for interruptibility wherein they distinguish between interruption to user's environment (social interruptibility) and to him/herself (personal interruptibility). They define 'personal' interruptibility as "the interruptibility of the user" and 'social' interruptibility as "the interruptibility of the user's environment" and provide the example that an audio notification might be acceptable to the person receiving it but may be disruptive to the meeting that they are attending. They model these concepts based on wearable sensors however only limited data based on actual fieldwork has been reported.

Ho and Intille introduced an approach in which interruptions are timed at transitions between physical activities [45]. They reasoned that a change in physical activity can, in many cases, correlate with a mental transition which in turn will make the interruptions less disruptive. They used a wearable accelerometer-based activity recognition system (Figure 2-3a) used by 25 participants who carried the sensor-enabled device through one working day. An experience sampling technique was used requesting self-reports at random times and at times triggered when an activity transition (sitting to standing, sitting to walking, walking to sitting, standing to sitting) was detected. While this solution may not address social interruptibility their study does show that the owner's were more receptive to an interruption triggered at an activity transition compared to random triggers. They also describe situations where this definitely was not true such as a doctor standing up to examine a patient.

Marti and Schmandt presented a solution to the social interruptibility of mobile phones that allowed the surrounding people to decide whether a mobile phone should ring. These people wear finger rings with attached sensors that vibrate when a call for anyone in the group is incoming. Any of the group can veto the call anonymously by touching a small button in his finger ring. This solution, while addressing social interruptibility, may in fact increase the personal interruptibility of each participant and the ring finger vibration may prove to be just as disruptive as a mobile ringing. The requirements and infrastructure may also make this approach impractical [46].

Figure 2-3 Systems using wearable sensors

(a) Ho et al's accelerometer-based activity recognition system [45] (b) Kern and Schiele's sensor network [43] (c) SenSay: a prototype context-aware mobile phone which can dynamically change alert method [42]

Kern and Schiele used a sensor network (Figure 2-3b) to estimate a persons interruptibility [43]. The network included a two-axis accelerometer attached to the user's thigh to measure activity, a wearable microphone that detected auditory context and wireless LAN (local area network) to determine the users location within the building. It was found in a pilot study that this model could determine interruptibility with a 94.6% accuracy, however these interruptible times were annotated manually afterwards using the experimenter as the subject.

Liu used input from an accelerometer, a heart-rate monitor and a pedometer to trigger interruptions at “non-stressful” moments. This method was used to study the impact of an emotionally-friendly interface in relation to interruptibility. Subjects in the study were shown to be most receptive to an interruption when the system was emotionally friendly and triggered during non-stressful activities [47].

2.11 Discussion

This chapter presents a review of the current literature relating to Human Computer Interaction based interruptibility and availability. The literature includes descriptions of desktop-based systems that use predictive models required to be trained with users self-reported data. However, users will be reluctant to adopt desktop systems that rely on trained classifiers and will be averse to systems with burdensome training phases. This will be especially true when the training requires them to be continuously prompted. Desktop systems will also be restricted to a particular physical space or environment and will require at least partial retraining when the environment changes.

More recent research focuses on mobile devices that seek user’s attention and assesses their availability. However these mobile solutions can also place a burden on their users who may be required to ensure calendars are up to date or asked to continually complete location and attendance diaries. Users of mobile systems may also be required to don wearable sensors. However this approach will not always address social interruptibility as sensors such as accelerometers or heart monitors do not monitor the users surrounding environment. Other sensor-based approaches have acknowledged this and attempt to tackle social interruptibility but these can be impractical approaches as they require the surrounding people to also wear sensors.

Many of the systems described in this chapter have had no formal evaluations while some other systems have been evaluated under laboratory settings using researchers as the subjects and report limited data based on actual fieldwork. The accuracy of their sensors will be simulated under laboratory settings and not tested in real world scenarios. Interruptible times may be annotated manually afterwards by the researchers.

Other strategies aim to reduce the number of inconvenient interruptions however reducing interruptions, especially when sampling, may not be ideal. It may be these times that the data collector is most interested in. They may also require a number of samples within a period of time. Interruptions, while disruptive or inconvenient may also be necessary. It may be better to introduce strategies that allow computer to interrupt at better times and turn the focus to more effective interruptions.

The research described in this chapter has motivated some of the work presented in this thesis. Guided real-time sampling is a novel strategy for predicting availability for users of mobile electronic diaries to make real-time diary entries. People adopt strategies where they will wait for the other person to finish or temporarily pause their current task, Guiding adopts a similar strategy by deferring inconvenient interruptions temporarily until a ‘better’ time by incorporating the analysis of contextual information.

2.12 Summary

This chapter describes some of the recent research and development regarding computer-based interruptions, their user’s availability and the current context. This includes desktop systems that use predictive models and multiple static sensors to assess their users availability. More recent research focusing on mobile systems that assess their owner’s availability usually through the use of wearable sensors such as hi-specification microphones or accelerometers.

2. Using contextual information to help predict peoples interruptibility and availability

The next chapter will look at the motivation behind the development of Pocket Interview, the use of handheld computers in healthcare and by researchers as data collection devices and particularly as electronic diaries and for recording data in real-time.

3. Motivation and design of a handheld data collection and diary tool

3.1 Introduction

Mobile technologies are being used increasingly to gather data in clinical and other research studies. Computerised data-collection tools have been shown to be more accurate, less susceptible to recall errors and encourage higher user-compliance than pen and paper alternatives. This chapter details the background, motivation and requirements of Pocket Interview, a new configurable electronic diary and data collection tool that provides the ability to design and deliver questionnaires and collect data securely using handheld devices and also desktop computers. The client software can be set to prompt the data contributor, via audible prompts, to provide entries at fixed times and preset intervals with an optional element of randomness to prevent predictability. The device also allows the users to record personally significant incidents out with the schedule at any time point.

3.2 What is a handheld computer?

In 1993 Apple launched the Newton which is widely regarded as the first Personal Digital Assistant (PDA) (Figure 3-1a). At the time it wasn't quite as successful as was hoped but it did herald the start of a new industry. These days the term handheld computer can cover a wide range of devices, not only PDAs, but also enterprise digital assistants which are the more rugged version of the PDA favoured by industry (Figure 3-1b). However smart phones are now outselling PDAs [48]. Smart phones are mobile phones that incorporate personal computer functionality such as email, personal organizers and internet connectivity (Figure 3-1c). Smartphones sold 40 million units in Quarter3 2008, according to the latest estimates from leading analyst firm Canalys. This means smart phones now represent around 13% of the total mobile phone market [49].

There are many other forms of handheld and mobile computing. These can include tablet Personal Computers (PC), notebook computers with detachable keyboards and a touch screen with a stylus for input. More recently we have seen the introduction of ultra-mobile PCs and notebooks. These are small and lightweight devices capable of running desktop operating systems and applications. Other handheld devices may have more specific uses and can include handheld games consoles and travel aids.

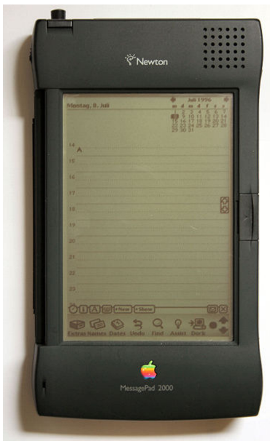
(a)  (b)  (c) 

Figure 3-1 A selection of handheld computers
(a) Apple Newton (b) Enterprise Digital Assistant (c) Smartphones

3.3 Using handheld computers in healthcare as data collection devices

Since their introduction handheld computers have become increasingly popular within the healthcare community as their benefits have become more appreciated and the devices have become cheaper and more ubiquitous. These are devices that are portable, unobtrusive and discreet in nature and provide the potential to view, store, organise and synchronise large amounts of data. At present they are used mainly for administration or organisational tasks [50]. Though, also, increasingly at the point of patient care as medical guidelines reference or drug references [51]. They have also provided an effective means to access patient information [52] and provide decision support at the point of care [53].

3.3.1 Barriers

Not everyone has embraced the technology and some barriers have been reported by medical users and during clinical trials. Users can lack experience and therefore may be intimidated and feel uncomfortable with the technology [54]. Due to small screen sizes the interface can appear quite busy and users with poor vision may have particular problems. Some users find text entry to be error prone and slow, and can have difficulties using the touch screen and stylus. Some users also report that the buttons are too small or predictive text too complex. There can be problems with some devices due to their limited memory, short battery life and fragile construction. Software problems and hardware malfunctions can also deter users. Cost may be a barrier with current prices of £150-£200 for a Pocket PC PDA and up to £500 for a smartphone.

Projects involving mobile technologies may be regarded as high-maintenance and expensive due to some or all of these issues and these aspects should be considered before committing to the technology.

3.3.2 Data collection

Mobile technologies are being used increasingly to gather data in clinical and other research studies. Traditionally, researchers have used pen and paper and subsequently entered the data into statistical software or subcontracted the task to a data entry company. The latter approach produces fewer errors but can be costly [55].

Direct data entry into a handheld computer can have several advantages compared to the traditional methods such as pen & paper. Capturing data using paper forms and then entering the results into a computer doubles the data entry required and increases the opportunity for introducing errors. Using mobile computers can save time by avoiding the double entry of data and reduce cost by reducing the need required of skilled staff. Direct entry to a computer also offers the potential to improve the quality of the data, for example by range checking. Additionally, researchers can analyse the results immediately once the data is entered [56]. Accurate results are assured every time as most input is point and click via menus and multiple choice questions so data capture is quick and accurate. This consistent input means that data can be collated easily and summarised reports can be produced.

There are commercial solutions available (e.g. Invivodata's eDiary, Pendragon Forms, PocketPC Forms and Satellite Forms) that can be used but these can be expensive and a separate license may need to be purchased for each client device that is used.

The Pocket Interview software can be used as a data collection tool where an administrator can configure the forms and the data to be collected. Currently, Pocket Interview only creates new data records and does not support updating existing records.

3.4 The use of handheld computers as electronic diaries

While PDAs can be used for many data collection tasks they have been used as a configurable electronic diaries (ED) in many research studies. EDs allow researchers to collect a user's self-reported data and thereby to monitor states, events and behaviour. Once again, the traditional methods of self-report tend to be pen and paper. Clinical researchers have used paper diaries for self-report data since the 1940's however more recently, EDs are being used more often [56]. There have been many applications reported using EDs with some of the more recent work involving patient self report including studies relating to smoking [57], alcohol intake [58], dietary programs [59], chronic pain [60] and allergy treatment [61].

3.4.1 Pen & paper vs. Electronic

Comparisons of pen and paper with electronic alternatives have been the subject of recent research. Studies report more complete entries using EDs with one particular study reporting approximately seven times fewer missing values when using EDs compared to a paper questionnaire [62]. Generally much fewer errors are reported while using EDs. The ED option is also more popular with users of all ages [63, 64]. It has also been shown that users may answer more forthrightly and honestly when using computer-assisted self-report [65] so EDs may be a more effective means of collecting sensitive information.

User entries can be date and time stamped so a precise record of when entries are made and how long entries take (even for individual questions) can be gathered. EDs are more efficient as less time is required to complete them and there are no problems or issues with deciphering hand written entries or losing paper forms. A wide variety of question formats can be made available and the requirement for a response can be enforced. Users can also be provided with examples to assist their understanding and shortened training times of five minutes have been reported in a recent review [66].

However, one area that has provoked recent debate is the issue of user-compliance with the data entry. This is the extent to which users comply with instructions particularly, with regard to the timing of the diary entries. Stone, Shiffman, Schwartz, Broderick and Hufford published research in 2002/2003 [67, 68] that proved alarming for researchers using paper diaries. They developed a novel paper diary fitted with a hidden photosensor that detected and recorded when the diary was opened and closed (Figure 3-2). The diary cards could not be removed and the users of the diary were not aware of the photosensor or the purpose of the study. They reported that patients completed 90% of the entries but only 11% were actually completed within 30 minutes of the scheduled assessment times. They found that 75% of users submitted entries for days that the diaries had not even been opened and 45% of users actually forward-filled entries. Large batches of entries were completed in the car park immediately before handing in. It was suggested that users may have missed diary entries due to lack of motivation, forgetfulness, or lack of opportunity and they didn’t want the embarrassment of missed entries. They perhaps did not realise they were actually defeating the purpose of the diaries.

A comparison group was run using EDs that prompted users with an audible signal and only allowed entries within a limited window of time and thereby prevented backfilling and forward-filling. Using EDs, 94% of diary entries were within 30 minutes of the scheduled times.

These studies stirred debate with other researchers arguing that motivating and educating participants about the importance of compliance was a bigger factor than whether studies are administered using pen & paper or EDs [69].

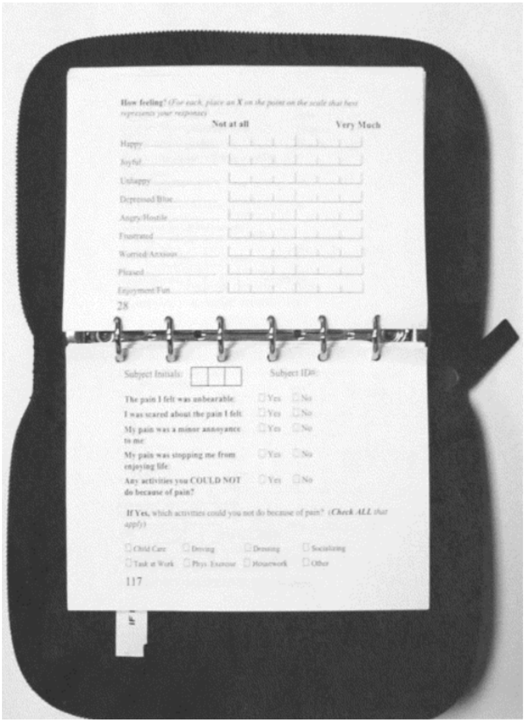


Figure 3-2 Stone, Shiffman et al.s' novel paper diary

The diary was fitted with a hidden photo-sensor that detected and recorded when the diary was opened and closed [70]

3.5 Experience Sampling - using electronic diaries to collect data in real-time

3.5.1 Brief history

In diary studies the participants are responsible for providing the information, however, they may forget to provide entries or only provide entries at the times they have time or inclination to do so rather than the ones of interest to the data-collector. As well as data loss this may lead to systematic biases in the data. For these reasons, the use of EDs to record data contemporaneously, also termed real-time and in-situ, in the user's natural environment is becoming increasingly popular. This method has previously been referred to as 'experience sampling' [71]. Experience sampling methods (ESM) have been employed by the research community for decades. These methodologies are designed to allow research study participants to record their thoughts, feeling and actions throughout their daily lives. Research has shown that a user's recall is unreliable and rife with inaccuracies and bias. People have also been shown to exhibit selective memory, especially for the most recent or the most severe incidents during a time period [72]. Rational for using these methods has been reviewed thoroughly (e.g. Beal & Weiss [73], Bolger, Davis & Rafaeli [72], Reis & Gable [74]) and more recently Conner-Christensen, Feldman-Barrett, Bliss-Moreau, Lebo and Kaschub have paid special attention to the use of handheld devices such as PDAs [75].

Although no single person or program can be credited with inventing ESM, among the very early pioneers were Sorokin and Berger [76], Wessman and Ricks [77] and most notably Larson and Csikszentmihalyi [71] who are credited with coining the term 'The Experience Sampling Method'. Originally the term referred to a particular technique where subjects were prompted at random however these days the term is used more generally to refer to any method where people report their experiences in real-time whether they be prompted or following a specific event. These experiences are therefore reported before a person has had the time to reflect and reconstruct their experiences after the fact. Researchers can also investigate the effect of time since events that have been reported on the diary e.g. the duration and decay of emotional response after a particular episode. Researcher involvement during data collection is minimal which can support longer data collection, more subjects and at more places. ESM is ideal for tracking data changes over time rather than requiring participants to collectively gather experience over a time-span and may provide richer data with considerably less retrospective bias [78]. These records may have increased validity as they represent experiences with the context of everyday life outside of laboratory settings.

In recent years ESM methods have increased dramatically in popularity with researchers. This may be in part due to technical advances and the growing affordability and availability of the tools currently adopted by ESM that make it more attractive as a research methodology. At present ESM involves having participants carry a device (usually a PDA or smartphone) that signals them to answer questions and record details of events in the moment, rather than at a later time retrospectively. The cost of PDAs has been dropping substantially recently and mobile phones are more ubiquitous even to the point of saturation in some western countries. Many people who own mobile phones will habitually take their mobiles with them at all times and places.

This method of recording data in real-time and in the users natural environment has also been referred to as ‘ecological momentary assessment’ (EMA) in health-related fields. EMA tends to be a broader term and can incorporate self-report as well as ambulatory monitoring of physical states such as blood pressure.

3.5.2 Applications of ESM and EMA

ESM has becoming increasing popular across a wide variety of research areas and fields. Social psychologists commonly use ESM as it can provide valuable insight into areas such as marital and family processes in the context of daily life which is not easily possible with more traditional methods [79]. ESM has been shown to be used with success in addiction studies where events and changes throughout the day may affect cravings and the likelihood of relapse [80]. ESM has also been shown to be valuable to record self-reported stress levels [81, 82], investigating the associations between physical symptoms and psychological variables [83], occupational psychology [84] and for emotion assessment [85].

ESM has recently been adopted and now growing in popularity in the field of HCI. ESM is being used to evaluate technologies that are used throughout everyday life and the unpredictable factors associated with using them in the real world. Zhang and Adipat review of new mobile technologies demonstrates the necessity of using field studies over laboratory settings to capture unexpected behaviours and situations [86]. Consolvo, Smith, Matthews, LaMarca, Tabert and Powledge used ESM to evaluate privacy attitudes when disclosing location to other people using mobile devices [87]. Froehlich, Chen, Smith and Potter used ESM to explore the relationship between location visit frequency and place ratings [88]. Other HCI studies include system evaluations [89] and exploring availability attitudes in project managers [3].

3.5.3 Protocols

Using an ED can enable a variety of diary protocols to be employed. These can be time-based protocols such as fixed intervals where entries are requested at the same time every day or random intervals where the user is prompted at random throughout the day. The users will be interrupted during their daily activities that, especially if the prompts are more frequent, may prove irritating. Diaries can also use an event-based protocol where the user decides when to make the entries, normally after some noteworthy incident.

3.5.4 Other ESM methods

ESM allows you to measure experiences as they occur, in the context that they occur and in their natural environments. Using PDAs is currently the most widely used method though other standard methods of capturing naturalistic information include:

- Pen and paper plus prompting from an electronic pager or programmable watch. The participant can be asked to write down what they are doing as they are doing it which may be a burden and also affect the activity itself [74]
- Lab settings - it may be difficult to simulate realistic behaviour in a lab and also fail to capture the influence of the setting on the participant's behaviors. It may also be unethical to administer in the lab a situation that induces stress
- Direct observation methods such as ethnography, time-motion analysis that can be costly, time-consuming and the presence of an observer may affect the behaviour of the participant
- Time diaries where the user is asked to write done what they are doing as they are doing it which can not only prove a burden but also affect the activity itself [90]
- Self-report using surveys which may be susceptible to recall and selective reporting bias [91]
- Interviews conducted individually or in focus groups
- Interactive voice response (IVR) - data-providers provide data using touch-tone phones. IVR is often criticized as being unhelpful and difficult to use if poorly designed and therefore should be designed with the minimum of complexity
- Internet - more recently researchers are using surveys that are administered over the web

3.6 Challenges with ESM

Computerised data-collection tools have been shown to be less susceptible to recall errors and encourage higher user-compliance than pen and paper alternatives [70]. However, if the user is interrupted during their daily activities this may prove irritating and disruptive especially when the prompts are more frequent [75]. The alarms will disrupt the participant's activities, work and conversations and may not only annoy them but also the surrounding people. Interruptions during inappropriate or inopportune moments may also lead to missed entries or negatively biasing the participants' responses. The participant will be repeatedly asked the same questions which may also negatively affect their responses. There may be difficulty in entering data in some situations e.g. in meetings, when driving.

ESM should probably not be used in all circumstances. Recent research suggests that people make important decisions about their future based on how they remember their experiences, not necessarily what actually happened in the moment [92]. ESM can also be costly and require time to design and implement. ESM does not give you control over variables, which is crucial for determining causal relationships [75]. ESM requires greater commitment and resources but provides greater understanding and more detail than a one-off questionnaire. The nature of the research questions should guide the decision whether to use ESM.

As with all self-report methodologies, ESM will rely on the participants answering the questions honestly and it may suffer from some unrealistic responses. Also, participants will sometimes be reacting to hypothetical situations i.e. they may be asked to imagine they are about to receive a mobile phone call but will not actually have to deal with the results of their response.

ESM studies may also be susceptible to the "Hawthorne Effect" which is well documented and is a form of reactivity whereby people change their behaviour due to the simple fact that they are being observed and/or their actions being recorded. Reactivity can also be especially problematic when repeatedly asking participants to assess their own behaviour or internal states as this may lead to behavioural changes e.g. they may become more aware that they are being inactive and over-compensate.

Compliance rates between PDAs and paper diaries do not always differ significantly. Recent research suggests that compliance with pen and paper diaries can be improved dramatically by establish good working relationships with participants [69]. It is fair to point out that paper diaries often can be effective and are easy to deploy.

3.7 Review of available ESM software tools

Related to the increasing popularity of ESM is the increasing number of software tools available freely and readily to the research community that can be used to design and run ESM studies.

3.7.1 Software for the Palm PDA operating system

ESP: The Experience Sampling Program [93] – the most popular and well-established with the latest version 4.0 released in 2005. There have not been any recent updates so it will require older versions of platforms. The software has not changed much since its introduction and so does not include many of the features available in more recent tools. There are limited question types with only buttons available. There is no option for open-ended questions, check boxes and no option for branching.

iESP: Intels modification of the ESP open-source tool (Figure 3-3a) [94]. Some of the modifications include: increased notification methods (e.g. if supported can be made to vibrate) and wider range of question types e.g. open-ended, sliders and check-boxes. Branching was also introduced and freeform entry is included i.e. graffiti with text recognition. A lack of current information suggests that it may be discontinued.

PMAT: Purdue Momentary Assessment Tool [95] – newer Palm and Windows desktop based tool which is available as open source. Allows one-level of branching. Includes a Java-based GUI for creating questionnaires and schedules of prompts, however the lack of a preview facility could make design and testing difficult.

A more detailed comparison and list of all features of the above 3 palm-based tools has been performed by Le, Choi and Beal [96]. They recognize the advancements these tools have made for diary research but also make some recommendations for new features. These include an increased flexibility in item presentation and format e.g. grouping and randomization, allowing multiple items on the screen and the ability to go back and edit previous responses within the same survey i.e. to correct an error. The ability to capture other types of data such as digital photos/video and sound recording would also be welcomed.

3.7.2 Software for Windows Pocket-PC based PDAs

C.A.E.S.: Context Aware Experience Sampling Tool [97] (Figure 3-3b). As well as prompting the data-provider this tool allows acquisition from other data input such as GPS and heart-rate monitors which could also be used to trigger sampling (e.g. when a subjects heart rate exceeds a certain level) and can record audio and take pictures [98]. This program has been optimised for multimedia inputs and may not be ideal for the flexible presentation of different questions and their formats. The software monopolises the device therefore does not allow normal usage of the device. Again, this tool is also no longer supported.

All of the tools mentioned above are restricted to older versions of the Palm and Windows CE operating systems and therefore will generally only operate on outdated hardware. ESP, iESP and CAES require users to modify configuration text files. A process that can be susceptible to errors and requires extended knowledge of the programs syntax. Although they are open-source it appears that ongoing development has ceased.

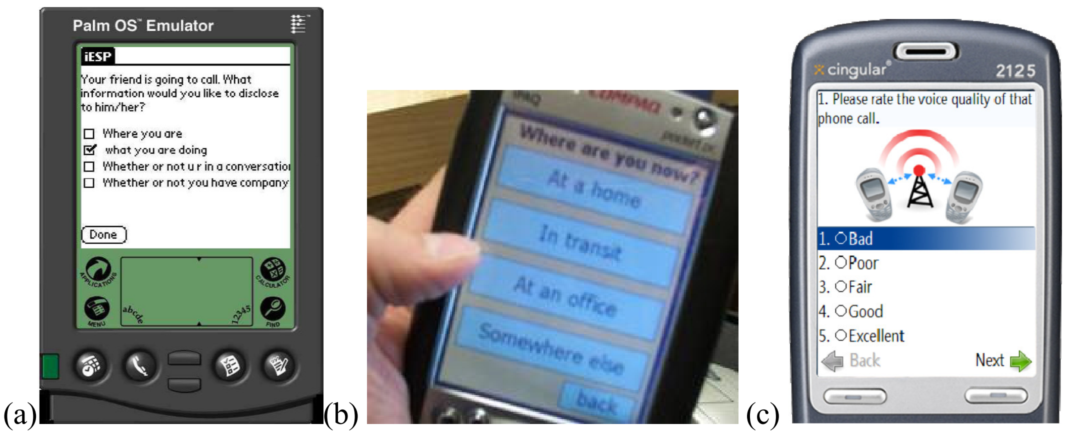


Figure 3-3 Screenshots of Experience Sampling Tools clients (a) iESP [94] (b) CAES [97] (c) MyExperience [29]

3.8 Smartphone based systems

More recently researchers have developed configurable smart phone based systems that can be used in Experience Sampling studies.

MyExperience: MyExperience extends previous work on the technique pioneered by MIT's Context-Aware Experience Sampling (CAES) toolkit for the PocketPC which employs sensors (such as GPS and accelerometers) to infer context, which then triggers a self-report questionnaire [29] (Figure 3-3c). At the time of writing MyExperience is still in early beta development stages. Configuration requires the users to modify XML-based configuration text files although assistance is provided in the form of a large user manual. MyExperience is open source software and runs on Windows Mobile Standard smartphones. A report in 2008 estimated Windows Mobile only account for 13.6% of the market [49] therefore it will be unlikely that administrators will find a large user-group with their own phones.

SocioXensor: Another mobile-based tool in ongoing development is the 'SocioXensor' system [99]. This is a Windows Mobile Standard based system. As well as an Experience Sampling tool, SocioXensor can record both mobile phone context and usage sensor data. The developers have reported issues with stability and overheating of the hardware and state that the system will be released as open source when technical issues have been resolved.

Momento is an open source application built around a client/server architecture [100]. Text and media messaging (SMS & MMS) technology is used to share information such as video, pictures and audio between the participants and the administrators. The system can be configured to automatically collect context data and in-situ qualitative information and does provide some tools to allow easier configuration.

Campaignr: Campaignr [101] is a mobile phone based Symbian system which was inspired by previous work on the Raccoon and StarScape systems. While not an experience sampling tool, as such, the system is designed to allow data capture from hardware sensors such as camera, microphone and also the location, time and usage data. A 'campaign' is defined by an XML file that contains information about what information to collect from what sensors and when to collect it. This can be automatic data collection or initiated by the user. The developers have identified many of the difficulties of developing for mobile phones and Symbian in particular such as getting the application signed to allow use and also highlight hardware problems such as limited battery life which can prove problematic when gathering large amounts of continuous data.

There may be other handheld-based data collection tools developed within the academic community but are not readily available for download. At the time of writing the author is aware of at least two other tools developed within the Scottish Academic community. Pocket Questionnaire developed at Aberdeen University which is a Pocket PC based configurable electronic diary tool and requires manual editing of XML questionnaires and schedules. Glasgow KelvinConnect's KCEngine and KCBuilder is a more general customisable handheld-based data collection toolset.

3.8.1 Smartphone considerations

Recently researchers have been developing configurable smart phone based systems as a means of gathering data. There are many potential advantages of using mobile phones in ESM studies. Costs can be kept down as users may already own a phone. Purchasing PDAs can be costly as can their maintenance, repairs and replacements. Participants are more likely to keep phone with them and this should help compliance rates. Mobile systems can also utilise mobile phone networks to remotely monitor participants and upload participants' data immediately.

However, as well as these benefits, there are still factors related to smartphone use in ESM studies that should be considered before committing:

- Smartphones are expensive and phones capable of running ESM applications are still a small percentage of the mobile phone market
- Battery life may be an issue especially if sensor data is being collected and monitored. The use of more processing power, increased user-interaction and display use will increase battery depletion. It may require people to change their charging habits to accommodate this
- They are not suitable for some situations such as hospitals where staff may not be permitted mobile phones with them
- Great care should be taken when deploying applications on participant's own devices. Early-stage applications may be fragile and cause crashes leading to the phone hanging and may prevent their owners from making or receiving phone calls
- Incorporating SMS & MMS should be used with caution as sending and receiving messages is not always immediate
- If using the participant's own 'sim' card they should be reimbursed for service charges. If using a study phone then copying the participants address book and calendar information onto the phone may raise some privacy issues
- Interaction with remote databases can take a few seconds which can lead to user frustration

PDAs are still widely used in electronic diary studies and should still be considered when planning new studies.

3.9 Pocket Interview – novel features

Pocket Interview is a new data collection and electronic diary tool that is designed to be easily usable by both non-technical administrators and data providers. Pocket Interview contains a number of novel features not featured on the systems described earlier in this chapter.

- Multiple platforms – desktop/PDA/mobile phone
- Graphical User Interfaces - allows ease of use via user-friendly graphical user interfaces that have been developed by adopting usability methods. Design of questionnaires and schedules.
- Demo mode - allows easy prototyping testing of questionnaires without the need to copy to the device
- Help Videos – demonstrate the more complex parts of the system administration such as creating questionnaires and schedules
- Encryption – data that is collected can be stored securely on the mobile device
- Kiosk mode – client can be run in fullscreen mode preventing the users from accessing other functionality of the device
- Integration with Access – allowing data to be easily organised

Pocket Interview also has a wide variety of question types and allows for audio recording and picture display. The system is easy to install on both desktop and client with step-by-step instructions. The system is described more fully in Chapter 4

3.10 Discussion

There are PDA-based data collection tools already freely available to the academic community, but administering these systems is not always as straightforward as it could be. They tend to rely on manual editing of the configuration files that requires knowledge of the syntax. All of the PDA-based tools are restricted to older versions of the Palm and Windows CE operating systems, and therefore may only operate on outdated hardware. Their protocol definition process can require users to modify configuration text files, a process that can be susceptible to error and requires extended knowledge of the syntax. More recently smart phone systems are becoming available though extra care is needed when developing for these devices.

Pocket Interview is a new data collection and electronic diary tool that is designed to be easily usable by both non-technical administrators and data providers. It allows ease of use via user-friendly graphical user interfaces that have been developed by adopting usability methods. It includes a demo mode to allow easy prototyping testing of questionnaires without the need to copy to the device, it has a wide variety of question types and allows for audio recording and picture display. The system is easy to install on both desktop and client with step-by-step instructions. It allows data to be organised through easy integration with Microsoft Access. Pocket Interview is the first configurable diary that is capable of running on multiple platforms that are currently available e.g. PDAs, smartphones and desktop computers.

These novel features allow the widest range of users to adopt the system and use efficiently. They will help to continue to increase the adoption rates and encourage continued use of the Pocket Interview system.

3.11 Summary

Electronic diaries can be used to collect a user's self-reported data and to record data in real-time and in the user's natural environment and have been shown to be more accurate, less susceptible to recall errors and encourage higher user-compliance than pen and paper alternatives. They can also apply Experience Sampling Methodologies that allow data to be recorded contemporaneously in the user's natural environment. ESM can help avoid data loss and systematic biases in the data due to unreliable user recall.

ESM is becoming increasingly popular and directly related to this is the increasing number of software tools available freely and readily to the research community that can be used to design and run ESM studies.

The next chapter will describe the development, evaluation and current uses of Pocket Interview.

4. Development and evaluation of Pocket Interview

4.1 Introduction

The previous chapter detailed the background and motivation behind Pocket Interview, a configurable electronic diary and data collection tool that provides the ability to design and deliver questionnaires and collect data using handheld devices and desktop computers. This chapter describes the development, evaluation and some of the current research studies using the Pocket Interview system.

4.2 System structure

After initial requirements gathering two user groups were identified: the administrators of the system and the data providers. Two subcomponents of the system were also identified: a PDA client used for prompting and data collection and an administration tool run on a desktop PC used for installing the client software, defining data collection schedules, defining the forms/questionnaires and collecting and organising the data from the devices in a suitable manner. Expert users offered their expertise during the requirements gathering and throughout the development process therefore ensuring user-involvement throughout the software's lifecycle.

The software was developed using the Visual Basic and C# programming languages and the Microsoft .net environment. The .net Compact Framework 2.0 was utilised for the PDA client. The desktop software requires Windows XP/Vista and the client runs on Pocket PC 2003, Windows Mobile 5 and 6. The implementation is approximately 20,000-25,000 lines of computer code.

4.3 Client software

The client software, used for data collection, is run on a Pocket PC handheld computer or smartphone (Figure 4-1). The interface to the questionnaire is created dynamically

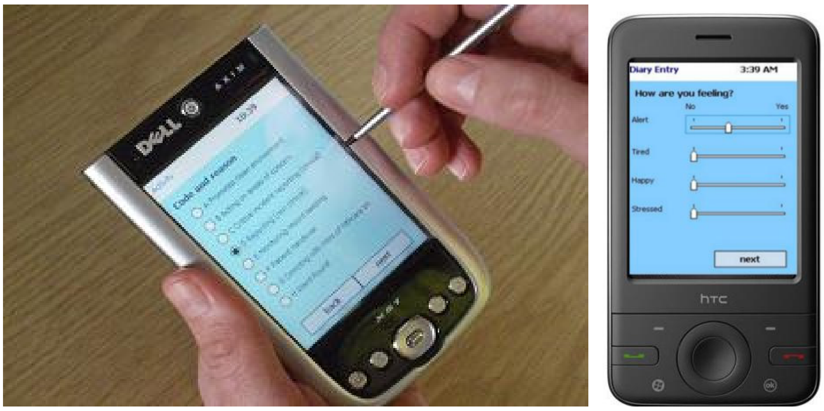


Figure 4-1 Pocket Interview client

by the client subcomponent. It reads the questions from a configuration file generated by the administration component. The software displays questionnaires, which can extend to multiple pages where necessary, and thus avoids the use of lengthy scrollbars where possible. Text-entry is performed using an onscreen keyboard. Questionnaires can include *back* and *next* buttons allowing the user to backtrack and review or change answers. Including the ‘back’ button is optional as set by the administrator.

4.3.1 Branching

The system allows for branching through the questionnaires. The questions that are displayed and the path through the questionnaire may depend on respondent’s answers to previous questions. Electronic diaries can be very effective at branching particularly when compared with a paper diary where participants can frequently became confused and fail to complete the questionnaires.

4.3.2 Kiosk mode

The software interface can be displayed in ‘kiosk’ mode. This is where the questionnaire fills the entire screen. This mode can make the device seem less intimidating for the inexperienced user as the interface appears less cluttered and it also prevents the user from accessing the other functionality of PDA. Kiosk mode can be switched on/off by the administrator as they may not want to take over the entire device. They may be using the participants own device and therefore still allow them to run other applications. If this is the case Pocket Interview will automatically start and be brought to the foreground at the time of each prompt.

(a) (b) (c)

(d) (e) (f)

Figure 4-2 Screenshots of the Pocket Interview Client

(a) main screen (b) trackbars (c) incident report using text entry and onscreen keyboard (d) radio buttons (e) Spanish questionnaire (f) system area

4.3.3 Keyboard

To facilitate kiosk mode Pocket Interview's own custom onscreen keyboard is included (Figure 4-2c). The Windows Mobile onscreen keyboard permits exit from a full screen application when accessing the keyboard's settings and so is not used.

4.3.4 Demo mode

Included is a 'demo' mode within the client. While in demo mode all entry types are visible on the main screen and data is not stored. This can be useful for demonstration purposes and for training users as all questionnaires are accessible and no data is stored.

4.3.5 Prompts

The software can prompt the user to make diary entries, using audible prompts/alarms that are triggered according to a schedule set by the administrator. The administrator may also choose the sound of the alarm, the volume and the length or number of rings. The client will record whether the user has missed entries. If the data-provider does not respond to an alarm they can be reminded. The administrator can specify the number of reminders before each entry is recorded as missed. A 'snooze' facility is also provided. There may be situations where the user may not want to be disturbed (e.g. lectures, seminars, church, theatre etc.) so they may 'snooze' the device for up to one hour.

The client includes a system area, which is password protected, and allows administrators to edit settings such as volume of the prompt/alarm and the type of sound used (Figure 4-2f). The administrator can also edit the number and length of time (in minutes) between reminders.

4.3.6 Protocols

Multiple protocols of prompts may be run simultaneously. As well as fixed schedules or random intervals, the device also supports spontaneous data entry that might be triggered by an external event occurring e.g. a post-emergency report (Figure 4-2c).

4.3.7 Security

The system allows the data collected to be encrypted using the Microsoft Windows CE and Windows Mobile Enhanced Cryptographic Provider (RSAENH). This is a FIPS approved general-purpose, software-based, cryptographic module for Windows CE and Windows Mobile and conforms to the recent NHS Scotland standard for data held on laptops, memory sticks and all other mobile devices [102]. The standard, introduced in September 2008, outlines the minimum requirements for the protection of mobile data in NHS Scotland. It states that all mobile data should be protected through the application of appropriate security controls and encryption regardless of the sensitivity of the information and was introduced in response to recent incidents that involved the discovery of unencrypted medical data on portable computing devices.

Users may have concerns about a third-party getting access to the information, being spied on or someone getting hold of the device and viewing their information. If a device is lost or stolen then by employing the encryption features the data will be more secure and the participant's privacy will be protected. Participants themselves may even answer more openly knowing that their data is encrypted.

4.3.8 Audio recording

Voice recording is included to support the users responding to open-ended questions. These often require longer answers so instead of making a lengthy text entry the user could dictate their response.

4.3.9 Pictures

Pictures can be incorporated into questionnaires. Due to a PDAs small screensize, they can be displayed as thumbnails that the user can then click, enlarge and view fullsize via the use of scrollbars.

4.4 Administration software

The second component of the system is the administration tool that is run on a desktop or laptop PC and a main menu page allows access to the various sections (Figure 4-3a).

(a) (b)

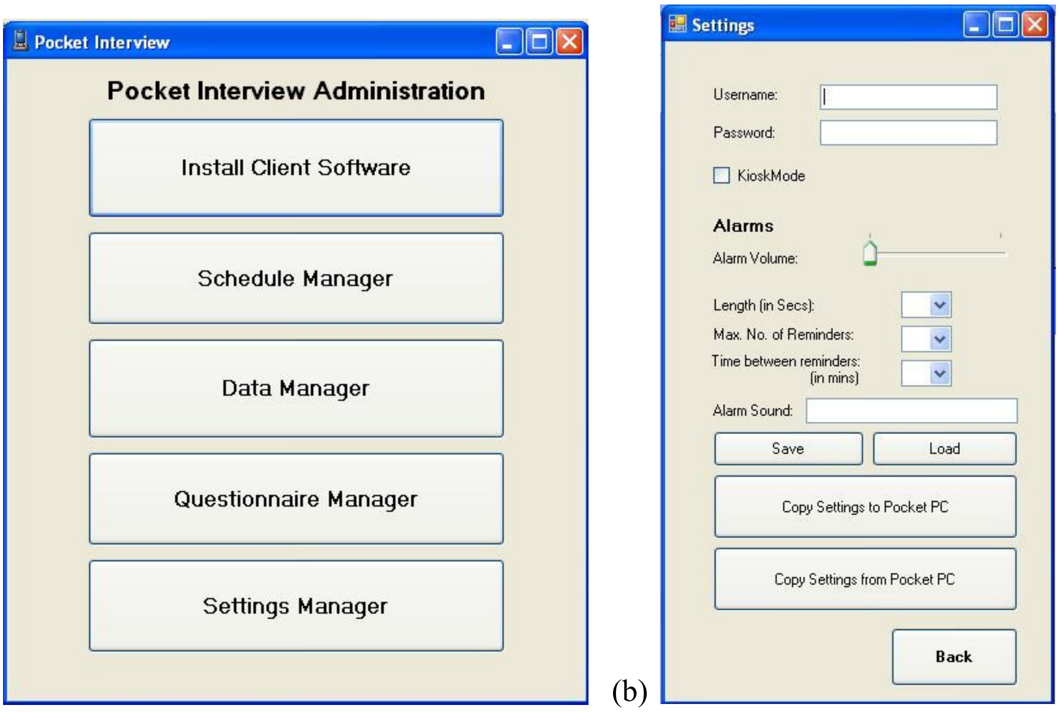


Figure 4-3 Screenshots of the Pocket Interview Administration tool (a) Menu (b) Settings Manager

4.4.1 Client manager

The Client manager is used to install the client software on the Pocket PC together with the Microsoft .net compact framework, the code-library that is utilised. Use of the Client Manager avoids extensive and complex installation procedures and is easy and quick to use by non-technical administrators.

4.4.2 Schedule manager

The Schedule manager consists of a graphical user interface that is used to create the schedules of prompts. The administrator can add individual prompts or the prompts can be organised in groups. The administrator can allocate the start and end times and also the number of alarms. The alarms will be evenly spaced. To prevent predictability and user-anticipation, an element of randomness can be added to the alarms. Alarms can also be deleted or edited individually. Once the schedule has been finalised it is saved to file before copying to the data-collection device. The schedule manager also allows the administrator to edit schedules already on the device by copying them from the device to the administrator's PC, making the required changes and then copying them back to the client.

4.4.3 Data manager

ESM can provide massive amounts of data. Collection and organisation of this data is an extremely important aspect of any study. The data manager allows the administrator to collect and organise the data in a suitable manner. When the clients' devices have been returned from the data-providers the data will be retrieved. The data is in XML format when copied from the client to desktop. The data manager allows the creation of a Microsoft Access database based on the questionnaire. Data can subsequently be entered into the appropriate Microsoft Access database where it is then easier to view and work with. It can also be imported and understood by analysis and statistical tools such as SPSS. The databases are created and the data is entered using SQL statements. The data manager constructs these statements dynamically though this process is hidden from the user.

4.4.4 Settings manager

The Settings manager allows the administrator to adjust the client settings and administer a master settings file. An administrator may be dealing with a number of clients and to avoid editing settings on multiple devices they can make any changes here and then copy the settings on to each client (Figure 4-3b). The system allows translation of the common user messages from English to other languages enabling Pocket Interview to be used by non-English speaking participants (Figure 4-2e).

4.4.5 Questionnaire manager

The Questionnaire manager is used for creating and editing existing questionnaires and consists of a graphical user interface (GUI) (Figure 4-4).

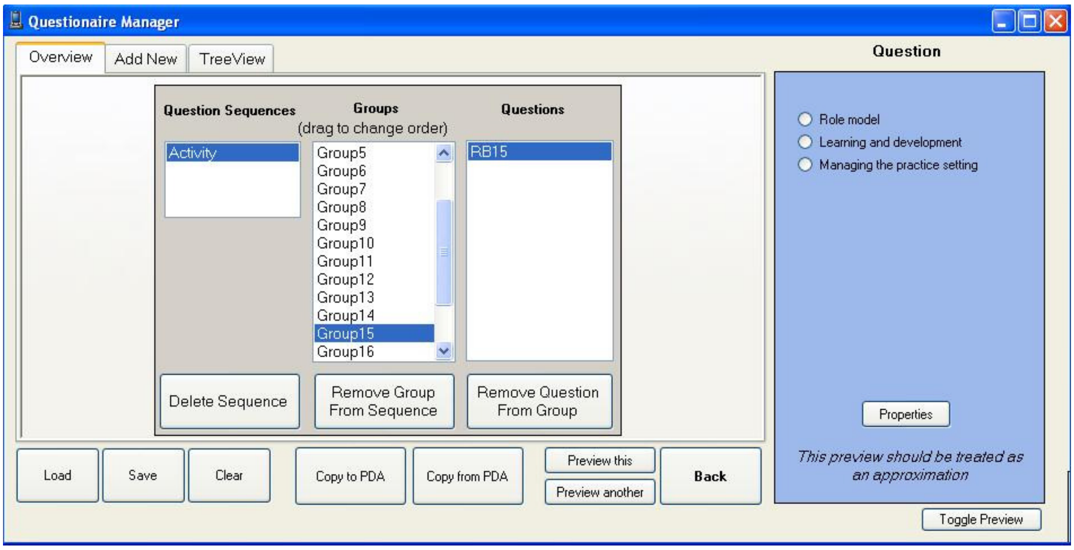


Figure 4-4 Questionnaire manager

Questionnaires are stored as configuration files in XML format. This may appear intimidating to the non-computer programmer and can be difficult to edit. Although XML is in many ways easier to use than programming in computer languages such as C# or VB it still requires expertise in mark-up languages and like many computer languages the syntax may appear fussy and small changes and elements incorrectly placed can cause malfunctions (Figure 4-5). Using the GUI shields the user from the complexities of XML.

```xml
<?xml version="1.0" ?>
- <Questionnaire xmlns="SchemaQuestionnaire">
  - <QuestionSequences>
      <QuestionSequence ID="DiaryEntry" MainPage="False" AlarmPage="True" Caption="Diary Entry" PreSeq="" PostSeq="" QuestionGroups="Group1|Group2|" />
      <QuestionSequence ID="ShiftEnd" MainPage="False" AlarmPage="True" Caption="Shift End" PreSeq="DiaryEntry" PostSeq="" QuestionGroups="Group3|Group4|Group5|Group7|Group9|Group10|Group14|Group15|Group11|Group12|" />
      <QuestionSequence ID="IncidentReport" MainPage="True" AlarmPage="False" Caption="Record an Incident" PreSeq="" PostSeq="" QuestionGroups="Group1|Group6|Group9|" />
    </QuestionSequences>
  - <QuestionGroups>
    - <QuestionGroup ID="Group1" Caption="How are you feeling?">
        <Question ID="Alert" Caption="Alert" ForceNewScreen="False" Type="TrackBar" Min="0" Max="100" MinCaption="No" MaxCaption="Yes" TickFrequency="0" />
        <Question ID="Tired" Caption="Tired" ForceNewScreen="False" Type="TrackBar" Min="0" Max="100" MinCaption="" MaxCaption="" TickFrequency="0" />
        <Question ID="Happy" Caption="Happy" ForceNewScreen="False" Type="TrackBar" Min="0" Max="100" MinCaption="" MaxCaption="" TickFrequency="0" />
        <Question ID="Stressed" Caption="Stressed" ForceNewScreen="False" Type="TrackBar" Min="0" Max="100" MinCaption="" MaxCaption="" TickFrequency="0" />
        <Question ID="Angry" Caption="Angry" ForceNewScreen="False" Type="TrackBar" Min="0" Max="100" MinCaption="" MaxCaption="" TickFrequency="0" />
        <Question ID="Energetic" Caption="Energetic" ForceNewScreen="False" Type="TrackBar" Min="0" Max="100" MinCaption="" MaxCaption="" TickFrequency="0" />
        <Question ID="Sad" Caption="Sad" ForceNewScreen="False" Type="TrackBar" Min="0" Max="100" MinCaption="" MaxCaption="" TickFrequency="0" />
        <Question ID="Frustrated" Caption="Frustrated" ForceNewScreen="False" Type="TrackBar" Min="0" Max="100" MinCaption="" MaxCaption="" TickFrequency="0" />
        <Question ID="Nervous" Caption="Nervous" ForceNewScreen="False" Type="TrackBar" Min="0" Max="100" MinCaption="" MaxCaption="" TickFrequency="0" />
      </QuestionGroup>
    - <QuestionGroup ID="Group2" Caption="Rate last 10 minutes">
        <Question ID="Group2Caption" Caption="Think about mental and physical activity in the past 10 minutes." ForceNewScreen="False" Type="Label" />
        <Question ID="WorkHard" Caption="Working hard" ForceNewScreen="False" Type="TrackBar" Min="0" Max="100" MinCaption="No" MaxCaption="Yes" TickFrequency="0" />
        <Question ID="CtrlWork" Caption="Had control over your work" ForceNewScreen="False" Type="TrackBar" Min="0" Max="100" MinCaption="" MaxCaption="" TickFrequency="0" />
        <Question ID="WorkFast" Caption="Working fast" ForceNewScreen="False" Type="TrackBar" Min="0" Max="100" MinCaption="" MaxCaption="" TickFrequency="0" />
        <Question ID="MoreCtrl" Caption="Would have liked more control" ForceNewScreen="False" Type="TrackBar" Min="0" Max="100" MinCaption="" MaxCaption="" TickFrequency="0" />
        <Question ID="WorkAppr" Caption="Work has been appreciated" ForceNewScreen="False" Type="TrackBar" Min="0" Max="100" MinCaption="" MaxCaption="" TickFrequency="0" />
      </QuestionGroup>
    - <QuestionGroup ID="Group3" Caption="Please think over the shift and identify its worst event for you.">
        <Question ID="Inc_Time" Caption="Time of incident" ForceNewScreen="False" Type="Time" />
        <Question ID="Serious" Caption="Seriousness of event" ForceNewScreen="False" Type="TrackBar" Min="0" Max="100" MinCaption="" MaxCaption="" TickFrequency="0" />
      </QuestionGroup>
    - <QuestionGroup ID="Group4" Caption="How did you feel?">
        <Question ID="EventAlert" Caption="Alert" ForceNewScreen="False" Type="TrackBar" Min="0" Max="100" MinCaption="No" MaxCaption="Yes" TickFrequency="0" />
        <Question ID="EventTired" Caption="Tired" ForceNewScreen="False" Type="TrackBar" Min="0" Max="100" MinCaption="" MaxCaption="" TickFrequency="0" />
        <Question ID="EventHappy" Caption="Happy" ForceNewScreen="False" Type="TrackBar" Min="0" Max="100" MinCaption="" MaxCaption="" TickFrequency="0" />
        <Question ID="EventStressed" Caption="Stressed" ForceNewScreen="False" Type="TrackBar" Min="0" Max="100" MinCaption="" MaxCaption="" TickFrequency="0" />
        <Question ID="EventAngry" Caption="Angry" ForceNewScreen="False" Type="TrackBar" Min="0" Max="100" MinCaption="" MaxCaption="" TickFrequency="0" />
        <Question ID="EventEnergetic" Caption="Energetic" ForceNewScreen="False" Type="TrackBar" Min="0" Max="100" MinCaption="" MaxCaption="" TickFrequency="0" />
        <Question ID="EventSad" Caption="Sad" ForceNewScreen="False" Type="TrackBar" Min="0" Max="100" MinCaption="" MaxCaption="" TickFrequency="0" />
        <Question ID="EventFrustrated" Caption="Frustrated" ForceNewScreen="False" Type="TrackBar" Min="0" Max="100" MinCaption="" MaxCaption="" TickFrequency="0" />
        <Question ID="EventNervous" Caption="Nervous" ForceNewScreen="False" Type="TrackBar" Min="0" Max="100" MinCaption="" MaxCaption="" TickFrequency="0" />
      </QuestionGroup>
    - <QuestionGroup ID="Group5" Caption="Recovered?">
        <Question ID="Group5Caption" Caption="Have you recovered from the incident?" ForceNewScreen="False" Type="Label" />
        <Question ID="Recovered" Caption="Recovered" ForceNewScreen="False" Type="TrackBar" Min="0" Max="100" MinCaption="Not at all" MaxCaption="Completely" TickFrequency="0" />
      </QuestionGroup>
    - <QuestionGroup ID="Group6" Caption="Who else was involved?">
        <Question ID="WhoElseWasInvolved" Caption="" ForceNewScreen="False" Type="CheckBoxes" Choices="Nobody Else|Colleague|Patient|Relative|Other|" ChoiceIDs="Nobody*|Colleague|Patient|Relative|Other|" Branches="No|No|No|No|No|"
          Required="True" />
        <Question ID="Group6Caption" Caption="More than 1 item can be ticked" ForceNewScreen="False" Type="Label" />
```


Figure 4-5 Example questionnaire file in XML format

There is a wide range of question types available that include:

- labels - text boxes
- combo boxes – a drop-down list of possible answers
- track bars/visual analogue scales - can detect exactly where the data-provider clicks. The administrator can set minimum and maximum values, labels, units and if/how many ticks are present (Figure 4-2b)
- check boxes – the user can select multiple answers from a selection of grouped choices. Items can set to be mutually exclusive i.e. if selected all other items in that group will be deselected (Figure 4-2e)
- radio buttons – the user can only select one item from a list (Figure 4-2d)
- likert scales – a custom control that allows the administrator to select the maximum value of the scale

The questionnaires have a hierarchical structure. Sequences contain groups that contain questions. Groups of questions may appear in different sequences. The Questionnaire manager includes a preview facility that emulates the client software and allows the administrator to view questionnaires without the need to copy to a PDA. The software includes user help in the form of documentation and videos covering the more complex parts of the system administration such as branching within questionnaires.

The scripts created by the Schedule and Questionnaire managers are stored in the XML format. XML is a language that allows the storage of structured information and can be read by multiple applications across multiple platforms. This will allow integration with new versions of the client on different platforms i.e. the same questionnaire script file can be used on both mobile and desktops and across multiple platforms.

4.5 Desktop version

A version of Pocket Interview has been developed to run on a desktop or laptop PC (Figure 4-6a). It may not always be necessary to provide users with a handheld computer as there may be situations where the data-provider is located in front of a PC or has easy access to one. The Pocket Interview client runs in the background on the PC. It's presence is indicated by an icon in the taskbar and a user will be prompted by a balloon appearing in the bottom right corner of the screen. The user clicks on the balloon to open the software and make an entry (Figure 4-6b). Therefore the system will not appear intrusive. The interface itself looks very similar to the handheld version in terms of size and layout. The questionnaires and schedules are downloaded and the data collected is automatically uploaded using an internet connection if available.

(a) (b)

Figure 4-6 Pocket Interview Desktop Version
(a) main screen (b) reminder prompt

4.6 Website

There are potentially many applications and to encourage its use the software is available, at no cost, to the research community via the project website at: http://www.computing.dundee.ac.uk/acprojects/PocketInterview/

The software is initially offered as a trial version. It can be started ten times before being required to register for continued use. The registration process is an automated online system whereby the user completes an online submission form. They are then sent a registration number automatically which can be used to fully unlock the Pocket Interview administration tool. At the time of writing (October 2009) there are 12 registered users of Pocket Interview. According to the 'Google Analytics' web statistics service [103] the website had 121 unique visitors from September 2008 to September 2009.

4.7 Evaluations

There has already been great interest in the Pocket Interview software from a broadening range of research groups. At the time of writing several groups are using the software to further their current work and are administering their own user-created questionnaires and schedules of prompts. Their interests include:

1) Using computerised ambulatory diaries for the assessment of job characteristics and work-related stress in nurses

Most research into work-related stress is based on retrospective self-reports whereas records made at the time could be more valuable [82]. In these studies models of work stress are assessed in trained nurses using ambulatory diaries in which recordings are made in real time in the working environment.

Current stage: data collection using Spanish version of questionnaire that was administered using earlier Palm-based diary software

2) Use of hand-held computers to determine the relative contribution of different cognitive, attitudinal, social and organisational factors on healthcare workers' decision to decontaminate hands.

This project explores the use of PDAs to examine factors affecting the decision-making process surrounding hand decontamination and glove use. Despite continued efforts, healthcare professionals' compliance with hand hygiene guidance remains sub-optimal [104]. One reason for this is the complexity of the decision-making process. It is envisaged that the ability to gather anonymous, valid data at the point of action will provide a step forward in understanding hand hygiene compliance and develop new ways of changing health professionals' behaviour.

Current stage: data collection is complete and data is being analysed

3) Evaluation of stress, well-being and clinical performance in recently graduated veterinary surgeons

There is an urgent need to understand the challenges that new graduates meet in their transition from student to practising vet, how this transition impacts on their well-being and performance in clinical practice and ultimately, based on this knowledge, to develop strategies to optimise the transition from training to employment [105]. Electronic diaries can provide rapid and reliable assessment of stress and well being together with a time-stamped record of clinical activity and performance in the form of computerised ambulatory diaries.

Current stage: questionnaire design and pilot testing

4) Determining the impact of the Senior Charge Nurse Review: A Study of Activity, Work Stress Levels and Well being

An investigation to see if Senior Charge Nurse (SCN) behaviour changes over time following the introduction of a new NHS policy (the 'National Role Framework') and a comparison of self-report methods with observer methods to see if the self-report method is valid and reliable. SCNs will be observed using a behaviour checklist delivered by Pocket Interview on PDA electronic diaries and collected by trained observers. Self reports of SCN activity and work environment perception will be recorded using Pocket interview on two separate occasions, Jan 2010 and Aug 2010.

Current stage: questionnaire design, observer training and pilot testing

5) Exploring prostate cancer patients' self-management demands and social support experiences using questionnaire and behavioural diaries

This study will test the stress buffering hypothesis in a sample of prostate cancer patients using both questionnaire and behavioural diaries. This study will identify the actual self-care demands and social supportive experiences for the prostate cancer patient, as these events happen, and will examine the methodological challenges of using behavioural diaries in this patient group.

Current stage: questionnaire design and pilot testing

To gather feedback on the design and use of Pocket Interview a questionnaire was sent to members of the research projects who are currently using the system. There were five responses, all from users responsible for configuring the Pocket Interview system.

None of the respondents had any previous computer programming experience although all of them could be considered experienced computer users as they all stated using a computer at least 21-30 hours a week.

4.7.1 Task performance

The users were asked about specific tasks, performed using the Pocket Interview administration tool, and the ease the tasks could be performed on a 5-point scale (1- unable, 5 – easily able) (Table 4-1). All of the participants answered 3 or above for all tasks except for one user who answered '2' for 'Create a new questionnaire'.

Table 4-1 Evaluation questionnaire ease of use variables

Variable (1- unable, 5-easily able)	Mean	Standard Deviation
How comfortable were you using the Pocket Interview system? (1-not comfortable, 5-very comftble)	4	0.71
Install the Pocket Interview software on your desktop or laptop	4	1
Install the Client Software onto a PDA	4.2	0.5
Create a new database	4	0.84
Copy data from client	4	0.71
Enter data into database	4	0.71
Create a new questionnaire	3.4	1.14
Edit a previously created questionnaire	3.6	0.89
Use the preview facility	3.8	0.83
Create a new schedule	3.5	0.58
Edit a previous schedule	3.75	0.95
How helpful were the help videos? (1-not helpful, 5-very)	4.75	0.5

When asked about the most beneficial aspects and the features that they like some of the responses included:

"*ease of use*"

"I like the flexibility of the software and the use of a graphical user interface, rather than editing xml script"

"having a GUI rather than having to use text files for setting schedules is quicker and should cut down on errors"

"the link with ACCESS"

"the tracker bar is extremely useful"

"visual analogue scales"

"the keyboard is good and does not offer any possibility of getting behind the running programme."

"being able to use this programme to meet the particular needs of data collection for my study"

All of the respondents indicated that they would recommend the software to other interested researchers.

4.8 Summary

Electronic diaries can be used to collect a user's self-reported data and to record data in real-time and in the user's natural environment. Pocket Interview provides the ability to design and deliver secure data collection using handheld devices and desktop computers. Through the use of an easy to use software tool, administrators can quickly build, test and distribute data forms. The responses subsequently provided by users can then be collected, organised and analysed in a timely fashion. The system has shown to be easily usable by both administrators and data-providers.

5. Description of studies

5.1 Introduction

There are a number of studies reported and results analysed throughout Chapters 6, 7 and 8. These studies have adopted Pocket Interview as a vehicle. To aid clarity the design and methodology of each experiment is described in this Chapter.

5.2 Study 1

Six users (three male, three female) were recruited from within the University of Dundee. Each user was provided with an electronic diary, in this case a Dell Axim PDA running the Pocket Interview system. An experience sampling method was adopted whereby each participant received 16 audio prompts from the electronic diary, spread over one working day. Eight of these prompts were random and eight were guided prompts each appearing within half-hour windows. The Guiding method adopted for this study was Guiding1 (Figure 7-2) using a 3-second audio sample, described more fully in Chapter 7. The type of prompt (i.e. guided or random) was randomised and it was ensured there was en equal number of each prompt type. Participants were asked after each prompt to rate on a 5-pt scaled how convenient an interruption would be at that time using the following question.

How convenient would you rate an interruption at this time?
1 (not convenient) - 5 (Very convenient)

Participants were asked to carry the PDA with them at all times and to answer the question based only on their current situation and not consider previous answers. Participants were unaware that there were two types of prompts i.e. random and guided.

5.3 Study 2

20 participants (eleven female, nine male) were recruited from within the University of Dundee. Recent context and sensor-based ESM studies have used between 10 and 25 users e.g. [45, 81, 88]. An experience sampling method was adopted. Each participant was provided with an electronic diary, a Dell Axim PDA running the Pocket Interview system. A total of 16 prompts per participant were issued spread over one working day consisting eight random prompts and eight guided prompts appearing once within half-hour windows. The Guiding method adopted for this study was Guiding1 (Figure 7-2) using a 3-second audio sample and is described more fully in Chapter 7. The type of prompt (i.e. guided or random) was randomised and it was ensured there was en equal number of each prompt type. Participants were asked after each prompt to rate using a binary (yes/no) scale if an interruption was convenient at that time using the following question:

Is an interruption convenient at this time?

Yes No

Participants were asked to carry the PDA with them at all times and were asked to answer the question based only on their current situation and not consider previous answers. Participants were unaware that there would be two types of prompts. 3-second audio samples were sampled just before the alarm was triggered for all of the entries and were then analysed to provide a reading of current sound levels.

5.4 Study 3

20 users (eleven female, nine male) were recruited from within the University of Dundee. An experience sampling method was adopted. Each user was provided with an electronic diary, a Dell Axim PDA running the Pocket Interview system. A total of 32 prompts per participant were issued spread over two working days consisting 16 random prompts and 16 guided prompts appearing once within forty minute windows. The Guiding method adopted for this study was Guiding2 (Figure 7-5) using a 18-20 sliding

window of audio samples and is described more fully in Chapter 7. The type of prompt (i.e. guided or random) was randomised and it was ensured there was en equal number of each prompt type. Participants were asked to carry the PDA with them at all times and were asked to answer the question based only on their current situation and not consider previous answers. Participants were unaware that there would be two different types of prompts.

After each prompt participants were asked to complete a short questionnaire, using the electronic diary, relating to the contextual information they would be happy to disclose automatically to their mobile phone and also the contextual information they would be happy for the phone to store. In addition they were also asked questions about their current location, whether they were in company, number of surrounding people and whether they were in conversation. At the end of the questionnaire they were asked whether they considered it a convenient time to complete the questionnaire.

The participants were first asked a question relating to accessing contextual information:

'Imagine you are about to receive a mobile phone call. Which of the following would you be happy for the phone to collect automatically about your current context in order to inform the caller of your availability (on a scale of 1- 10) - more than 1 item can be selected'

Options:

Location
Activity
Company
Conversation

Participants were then asked a further question relating to storing contextual information:

'The accuracy of predicting your availability to receive a call could be enhanced by storing a history. What information about your current context would you be happy for the device to store. (this information would not be available to anyone else) - more than 1 item can be selected'

Options:

Location

Activity

Company

Conversation

The participants were then asked the questions in Table 5-1.

Table 5-1 Questions asked after each prompt by the PDA client

Question	**Options**
What is your current location?	Work, Home, In transit, Other
Were you in conversation?	Yes, No
Are you in company?	Yes, No
How many people are there nearby (other than yourself)? *(only asked if previous answer is Yes)*	0, 1, 2, 3, 4+
Was this a convenient time to answer these questions?	Yes, No

At the conclusion of the study and immediately following their final diary entry participants were asked two further questions using the electronic diary:

Question:

'Would you be happy to use a computer service that would help predict your availability? Your availability could be inferred by collecting current information such as location, activity and whether you were in conversation?'

Options:

I would use if I can manage preferences

I would possibly use the service
I would avoid due to privacy concerns

Question:

How useful would you rate a computer service that would help predict your availability? Your availability could be inferred by collecting current information such as location, activity and whether you were in conversation? (1 - not useful, 5 - very useful)

5.5 Data analysis

Much of the data takes the form of repeated measures over a period of time using the same individuals. Tests that assume independent samples are not appropriate in cases such as these as the repeated measures are usually correlated. After receiving expert guidance from Dr. Victoria Bourne (School of Psychology, University of Dundee) it was decided to adopt the repeated paired t test for statistical analysis. This test is commonly used when comparing two sets of measurements where they have been gathered from the same participants [106].

When analysing the data collected in the studies it was noted that the number of observations varies between participants therefore the measurements have been averaged. By doing this the repeated observations have been reduced to two single measurements for each participant e.g. convenience rates when there is sound present compared with when it is quiet (Chapter 6), guided and non-guided convenience rates (Chapter 7) and guided and non-guided disclosure rates (Chapter 8).

6. Predicting the availability of participants in diary studies

6.1 Introduction

The study of interruptions has a long history. Recently there has been much research and study of availability with regards to Human Computer Interaction (HCI). The majority of prior work on HCI based interruptions has focused on desktop computing applications, office environments and utilise a mixture of static and wearable sensors. These studies suggest that even a small set of sensors can be enough to provide valuable information regarding the user's interruptibility. More recent research focuses on studies investigating interruptions for mobile computing users.

There are an increasing number of mobile devices, such as phones and PDAs, that seek their users attention. These devices act proactively to alert their owner of incoming phone calls, reminders, instant messages and can also provide services based on location such as friend-finders. Each time the device acts proactively it is competing for the users attention, possibly interrupting ongoing tasks and contributing to feelings of information overload.

The Pocket Interview client, described in Chapter 4, will usually run on a mobile device. By its nature, Pocket Interview will interrupt the data-providers when they are prompted to make a data entry. This chapter describes a series of studies, using Pocket Interview, that explore the availability of users of electronic diaries to provide data entries.

6.2 Challenges using mobile devices

When assessing their user's availability mobile devices present an increased challenge, especially compared with desktop systems. They cannot rely on mouse or keyboard use to determine a person's presence or current activity. When designing for a mobile environment it also is important to recognize the limits both in sensory capacity and in the computational power that is available on such devices. Constant analysis will consume the devices processing ability and quickly drain their battery.

6.2.1 Using Audio

It has been demonstrated by many researchers that continuous audio can be a rich source of information for activity recognition. When modelling human interruptibility previous research has suggested that the presence of human speech (e.g. someone is currently talking) can have a high predictive value using static or laptop microphones. When human speech is identified these times have been reported to be regarded as 'non-interruptible' by the participant [8, 9].

Previous research regarding audio and availability tends to use high quality microphones. However the microphones that PDAs come equipped with tend to be of a lower quality and are primarily designed for dictation purposes, so therefore their ability to record audio out-with a range of a few feet is limited. If mobile devices and their inbuilt microphones could be used to assess their owner's availability this would allow the participants to be unencumbered. They will not be asked to wear or carry additional sensors such as accelerometers or monitors.

Much of the recent research regarding mobile audio is concerned with the classification of different environments. Environmental audio classification is a challenging problem due to the large quantity of different types of sounds and high variation between sounds that belong to the same categories. Some of the algorithms that are being explored include harmonicity ratio, spectral centroid, spectral spread and spectral flatness [107]. Much of this work represents solid progress, however, at present few would be appropriate for real world or real-time input due to the calculation time required by being computationally expensive.

However, event detection is simple and efficient. It is assumed that the exact value of a volume reading from a microphone is probably less predictive than a Boolean feature that is deduced from comparing the reading with a threshold. Therefore through a process of thresholding the audio samples' total energy value and incorporating constraints on sample length a Boolean value of whether sound activity is present can be deduced.

6.2.2 Audio Recording

For the experiments described later in this chapter a Dell Axim X51 PDA was used. Audio data is sampled using the PDA microphone at 22kHz with 16-bit depth in 3-second samples. Low-resolution audio samples of 3-second length recorded by PDAs have been demonstrated as sufficient to classify different environments such as offices, streets, bars and car-driving [108].

6.3 Audio analysis

For the analysis, a sound detection algorithm that uses amplitude thresholds, is performed on the sample to give an indication of sound activity. The threshold value is set during an initialisation phase whereby quiet audio is continuously sampled while the threshold is increased until an appropriate value is reached. The sample is recorded using the devices inbuilt microphone. Through informal user-testing the sound-detection method has been shown to be effective at detecting nearby speech while noise from televisions and radios goes undetected, computer activity such as typing and mouse clicks also go undetected.

6.4 Power management in mobile devices

In today's world many devices are never truly off unless their power supply is completely interrupted. A fully off state has been replaced with suspended or low power states. Mobile phones are typically always on, if a mobile phone is turned to it's true off state then the phone won't be able to receive phone calls or messages. Generally, when the device is passive, the screen will power off and the processor will be running at a reduced speed [109].

Windows smartphones typically have four power states –

- Fully Powered on
- Backlight off – backlight switches off after no interaction
- UserIdle – suspended state with reduced processing
- Off

Windows Pocket PCs add three additional power states -

- Unattended – screen, backlight and audio are all turned off. However, the programs on the device are still running
- Resuming – when the device wakes. Backlight and screen are still off though audio is active
- Suspended – device is asleep and the processor is halted

PDAs and smartphones will progress towards their 'UserIdle' or 'Suspended' states if there are no user actions. The PDAs additional power states allow them to be much more efficient in terms of managing and conserving power than mobile phones.

When sampling audio for the purposes of assessing availability it is preferable not to alert the user to the device's activity by switching on the screen. Ideally, the user should be unaware that the device is doing any work. When using a PDA this would require the device being in the 'Unattended' state, however by default the audio devices are off in this particular state. Therefore for the purposes of allowing the device to sample ambient audio while at the same time not alerting surrounding people of activity an additional customised user state has been added. This new state is similar to the unattended state while also allowing the microphone to be active.

6.5 Research question

- Can the presence of sound gathered using mobile devices in office (and other) environments help predict availability of participants to answer a short questionnaire delivered by an electronic diary?

The following studies had ethical approval from the School of Computing Ethics committee, University of Dundee.

6.6 Pilot study 1

6.6.1 Design

A pilot study was performed that asked participants to assess their own interruptibility at random times throughout the working day (Chapter 5 – Study 1).

Six participants (three male, three female) were recruited from within the University of Dundee. An experience sampling method was used. Each participant was provided with a Dell Axim PDA running the Pocket Interview system. Each participant received 16 prompts, spread over one working day, each occurring randomly within a half-hour window. Participants were asked after each prompt to rate their own availability on a 5-pt scale using the following question and were asked to not consider any of their previous answers:

How convenient would you rate an interruption at this time?
1 (not convenient) - 5 (very convenient)

Participants were asked to keep the PDA diary with them at all times.

6.6.2 Results and discussion

There were a total of 96 possible data points. 25 were not answered through a mixture of non-responses (19), the device being snoozed (2) and the participant not clicking the finish button [1] (4). This left a total of 71 answered data points.

After compiling the user responses it can be observed that the data distribution shown in Figure 6-1 is approximately an inverse normal distribution i.e. the data peaks at 1 and 5 and is lowest at 3. The data for this small sample size demonstrates a tendency to cluster at the 1 and 5 measures of the scale. This suggests that capturing self-reported interruptibility levels may be bimodal in nature and that interruptibility may be better captured using a dichotomous measure rather than a continuous scale. A finding that is similar to earlier research. Picard also used a binary yes/no when measuring interruptibility [81]. Fogarty et al. initially used a 5-point Likert scale when measuring interruptibility, though they found the most common response was highly non-interruptible. They then chose to distinguish between this and all the others hence their scale became bimodal also [9].

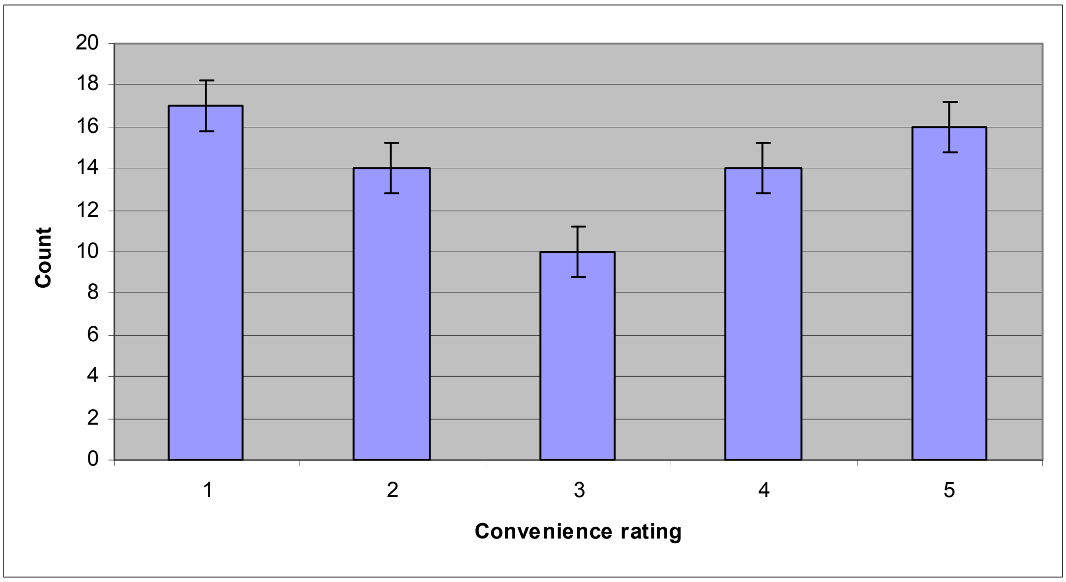


Figure 6-1 Pilot study 1 data distribution

---

[1] *Clicking the 'finish' button was necessary for the entry to be recorded. An issue that was resolved in future versions of the system.*

6.7 Pilot study 2

6.7.1 Design

A second pilot study was performed (Chapter 5 – Study 2). This time with an increased number of participants and a revised question that allowed a binary response (based on the results from Pilot Study 1). 3-second audio samples were sampled just before the alarm was triggered for all of the entries and were then analysed to provide a reading of current sound levels.

20 participants (11 female, 9 male) were recruited from the University of Dundee. Recent context and sensor-based ESM studies have used between 10 and 25 users e.g. [45, 81, 88]. Each participant in this study was provided with a Dell Axim PDA running the Pocket Interview system. A total of 16 prompts per participant spread over 1 working day (9am – 5pm) occurring randomly within half-hour windows. Participants were asked after each prompt to rate using a binary (yes/no) scale if an interruption was convenient at that time using the following question:

Is an interruption convenient at this time?

Yes No

6.7.2 Results and discussion

There were a total of 320 possible data points. 46 were not answered through a mixture of non-responses (38), the device being snoozed (5) and the participant not clicking the finish button [2] (3). This left a total of 274 answered data points.

---

[2] *Clicking the 'finish' button was necessary for the entry to be recorded. An issue that was resolved in future versions of the system.*

Overall 143 answered yes, it was convenient and 131 answered no, it was not a convenient time.

When using ESM non-responses can be treated in 2 ways:

1) They are simply dropped from the analysis, this is a common approach in ESM and is appropriate as it is unknown why there was no response [45]
2) They are treated as not convenient with the assumption that the subject was too busy to respond or to carry the PDA (as was this case in the study by Fogarty et al. [9])

The subjects in this experiment indicated that it was convenient

53.0% of the time (when not including non-responses and snoozed)

or 46.1% (when treating no response and snoozed as inconvenient)

The results and analysis reported in the rest of this thesis will treat non-responses as option 1, they are simply dropped.

There were a total of 88 out of 274 instances that were classed as containing sound using a 3-second sample. Due to this short sample length there may be some samples that are quiet within otherwise noisy times that may go undetected. This is an issue that is addressed in Chapter 7 with the introduction of a sliding window.

Table 6-1 Overall convenience counts from Pilot Study 2

	Yes	No	
Sound present	22	66	88 (32.1%)
Quiet	121	65	186 (67.9%)
	143 (52.2%)	131 (47.8%)	274 (100%)

Overall, the results suggest that the presence of sound significantly increases the likelihood that people working in an office-environment will not consider themselves to be interruptible (Table 6-1). Statistical analysis confirms this (Paired $t(19) = 4.112$, $p < 0.001$).

Out of the 38 samples that were missed only nine were classified as containing sound, a result consistent with the overall result (88/274 classed as containing sound).

6.7.3 Gender

Table 6-2 Convenience counts by gender from Pilot Study 2 comparing Male/Female and Sound/Quiet

	Female (N = 11)		Male (N = 9)		
	Yes	No	Yes	No	
Sound	18	33	4	33	88
Quiet	69	37	52	28	186
	87	70	56	61	274

Examining the responses by gender reveals a significant difference between male and females self-assessment of their availability (Table 6-2). The results indicate that males are much more likely to report themselves as unavailable when sound is present. Statistical analysis confirms this result (Paired $t(19) = 2.945$, $p < 0.01$).

The author is not aware of any previous work investigating or commenting on HCI based interruptibility based on gender. This could provide an area for further investigation.

6.7.4 Job Roles

The participants had roles within the workplace that could be categorised as project manager, researcher or administrator (Table 6-3).

Table 6-3 Convenience counts by job role from Pilot Study 2

	Researcher (N = 10)		Administrator (N = 4)		Manager (N = 6)		
	Yes	No	Yes	No	Yes	No	
Sound	13	31	6	15	3	20	88
Quiet	54	39	18	13	50	12	186
	67	70	24	28	53	32	274

Statistical significance that the presence of sound will affect convenience rates:

Researchers: Not significant

Administrators: Paired $t(3) = 5.133$, $p < 0.05$

Managers: Paired $t(5) = 6.807$, $p < 0.001$

On average administrators were found to work in environments that, compared to the other groups, contain more sound. Though more data will be required to back this up statistically. Statistical analysis does not show a significant difference between Administrators, Researchers and Managers in terms of their availability when sound is present. While researchers may tend to be located in quieter lab settings the presence of sound still suggest they consider themselves as not available though more data will be required to back this up statistically. Sound presence is most indicative for the manager category. This is a result that suggests that managers work activities and their availability is the most closely related to the presence of sound when compared with the other job roles.

6.8 Study 3 – self-reported attributes

Results from Pilot studies 1 and 2, reported earlier in this chapter, demonstrate that the presence of sound can be a good cue that a variety of office workers consider themselves as unavailable for interruption. As was the case in these studies, the sound information can be gathered using inbuilt microphones in mobile devices such as PDAs and mobile phones that the participants can carry with them throughout their day. Using inbuilt sensors allows the user to be unencumbered by additional equipment and the sensors to be mobile i.e. not tied to one location.

As well as sound, is there further contextual information that can also be automatically collected by the device and be used to predict the participant's availability?

6.8.1 Research question

- Is there contextual information that can be gathered by mobile devices and other mobile sensors that could help infer a person's availability?

6.8.2 Design

To help answer this question the self-reported data that was collected during an experience sampling study (Chapter 5 – Study 3) related to privacy attitudes (Chapter 8) can be analysed for this purpose.

For this study 20 participants (11 female, 9 male) were recruited from within the University of Dundee. An experience sampling method was used. Each participant was provided with a Dell Axim PDA running the Pocket Interview system and received a total of 32 prompts per participant spread over two full days (9am – 9pm) occurring randomly within 45-minute windows.

As well as questions related to the participants current context and privacy attitudes regarding mobile devices the participants were asked the questions in Table 5-1.

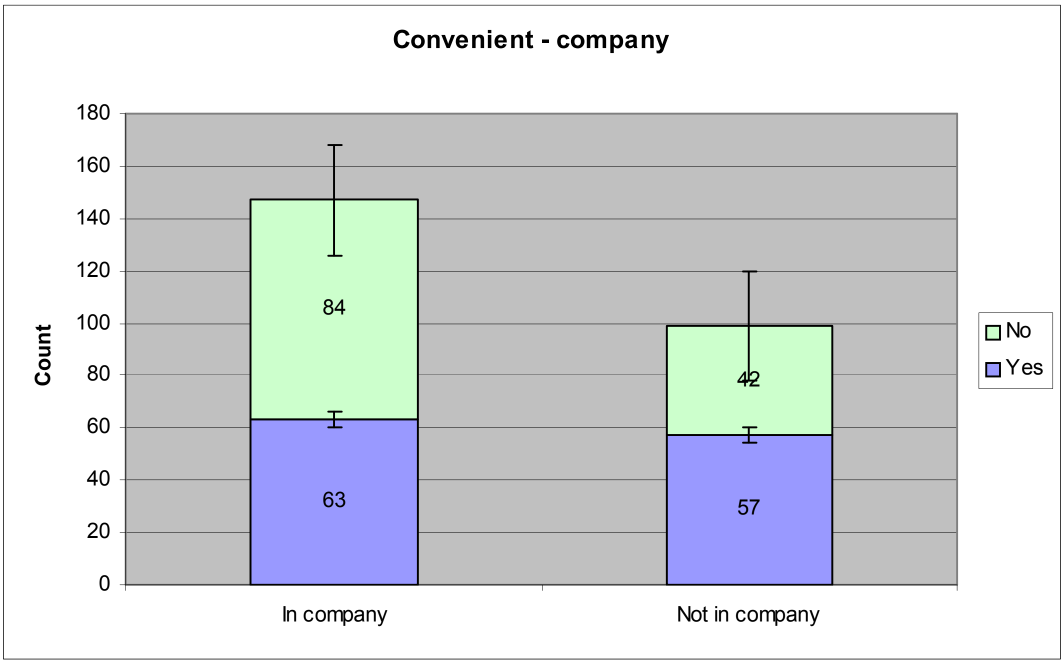
For this analysis data was not included that was collected using the Guided sampling method that is described in Chapter 7. This section is an exploration of availability using self-reported attributes. The guided data may be biased due to the fact that it is sampled at, what is targeted to be, more convenient times and so for these purposes may not be considered completely random. Therefore a total of 246 entries that occurred at non-guided times were analysed.

6.8.3 Results

This section will present an examination of the results for each attribute and examining the relationship between the attribute and the participant's convenience for completing the questionnaire.

6.8.3.1 Company – whether the participant was in the company at the time of the prompts

Overall the results suggest that whether the participant is in company affects the likelihood that people working in an office-environment will consider themselves to be interruptible (Figure 6-2). If they are not in company then they are more likely to be interruptible. Statistical analysis confirms this (Paired $t(19) = 2.501$, $p < 0.05$).



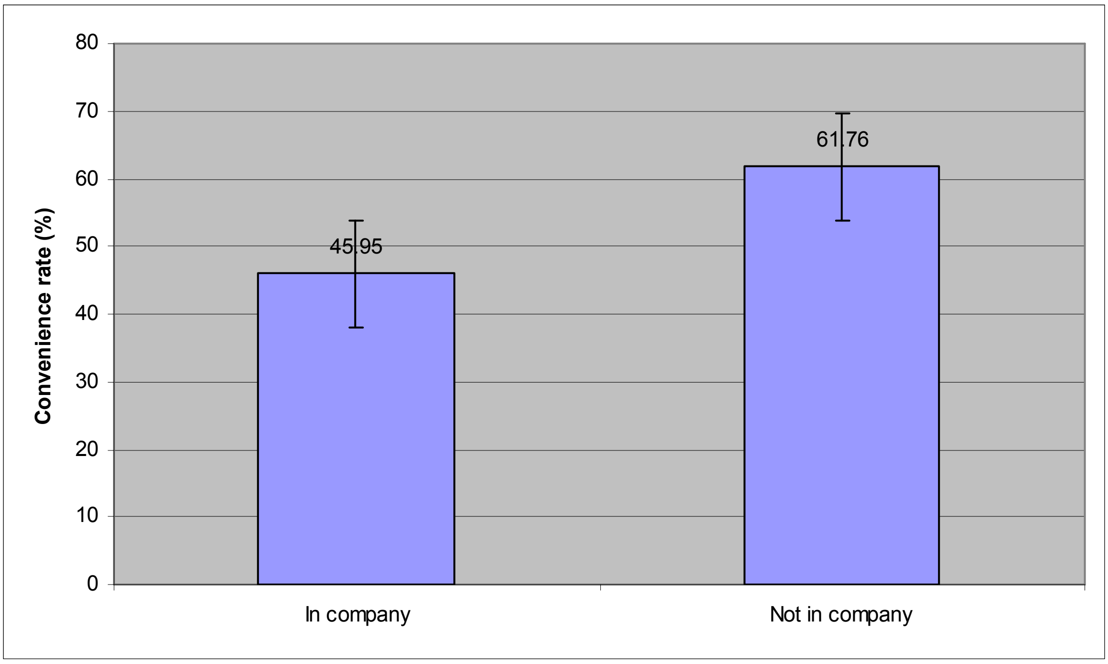
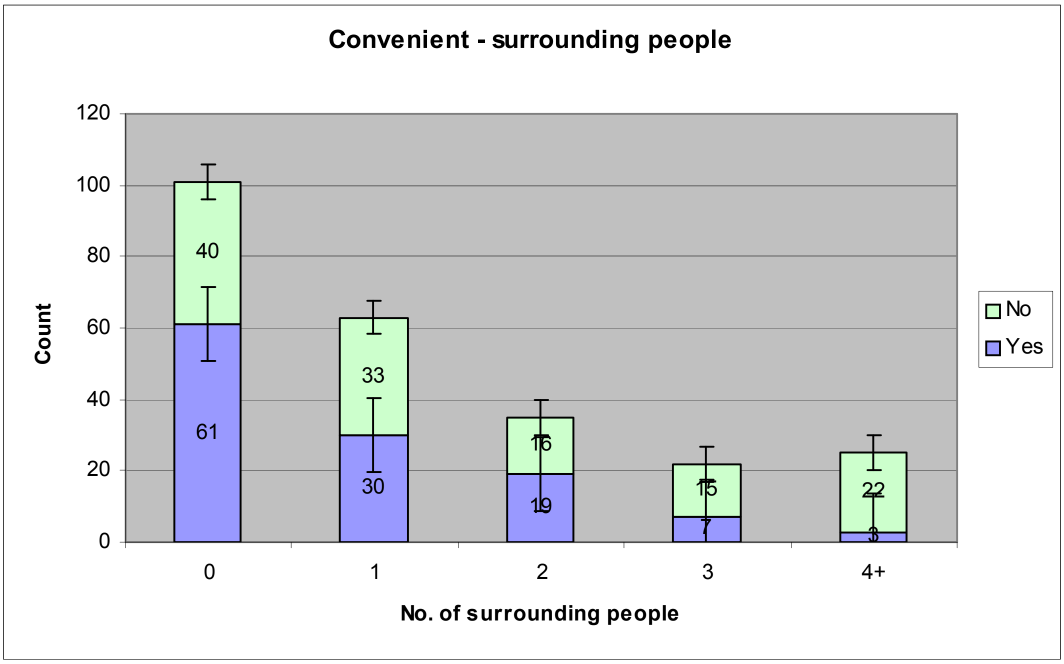


Figure 6-2 Data distribution with relation to convenience rates based on 'in company'

6.8.3.2 The number of surrounding people (not including the participant)



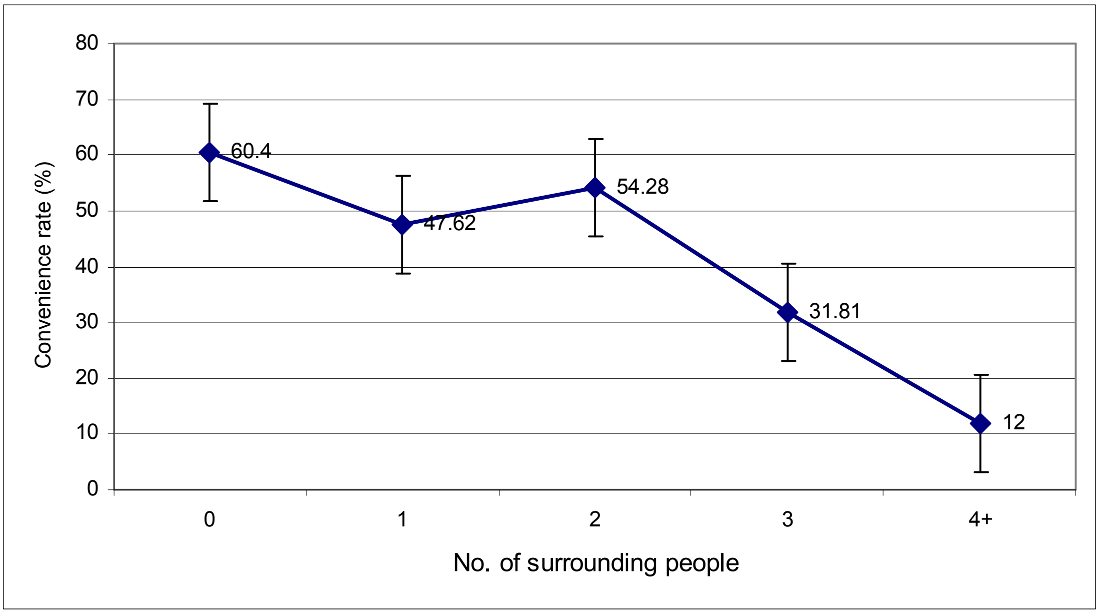
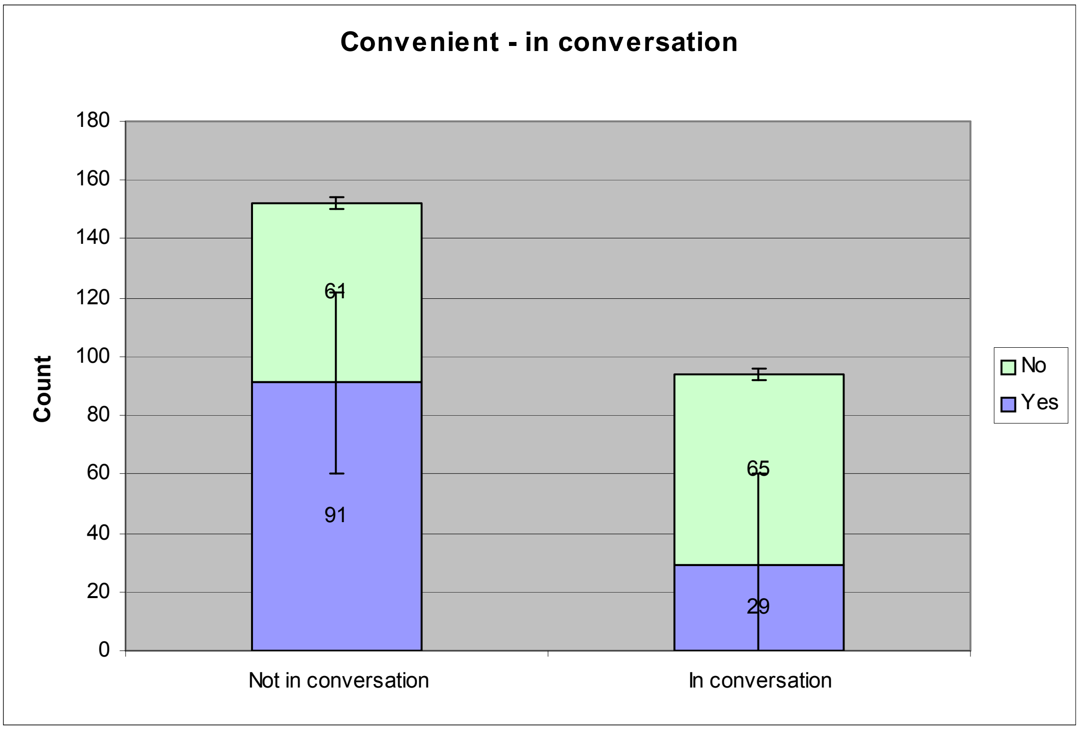


Figure 6-3 Data distribution with relation to convenience rates based on 'Number of surrounding people'

Figure 6-3 suggest that the 'No. of surrounding people' is related to the likelihood that office workers will not consider themselves to be interruptible. Further data will be required to back this statistically.

6.8.3.3 In conversation – whether the participant was in conversation at the time of the prompt



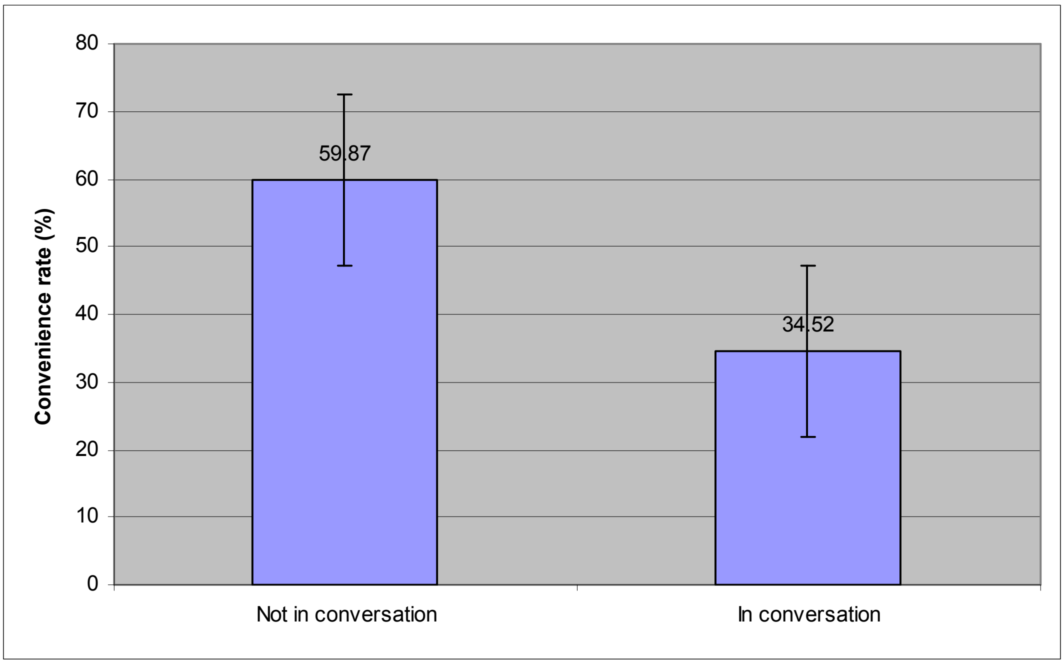
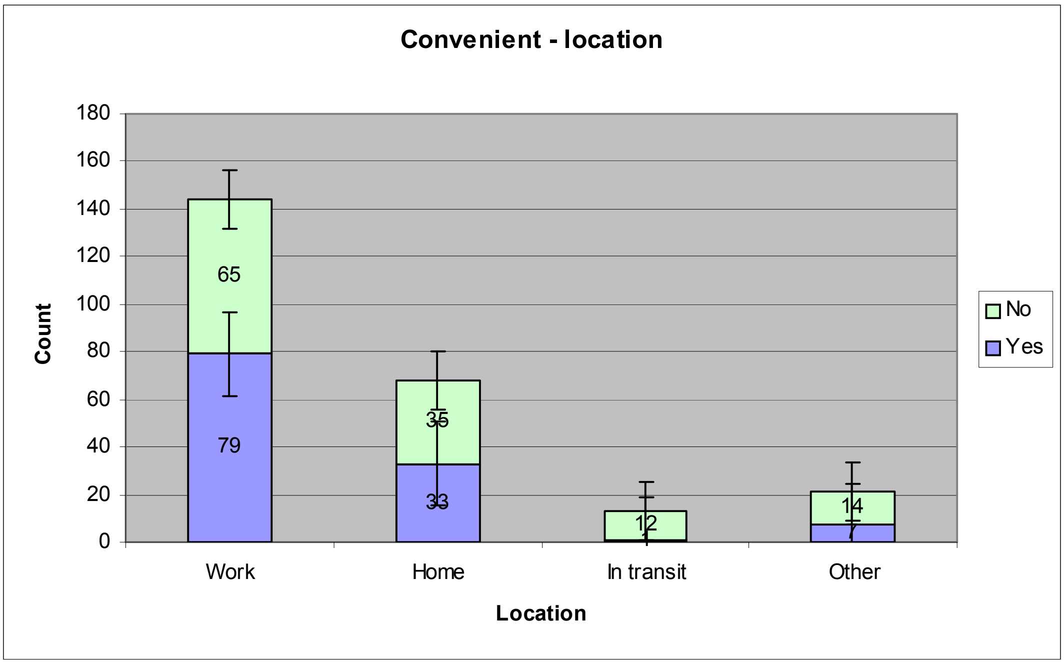


Figure 6-4 Data distribution with relation to convenience rates based on 'in conversation'

The results suggest that whether the participant is in conversation will affect the likelihood that office workers will not consider themselves to be interruptible (Figure 6-4). If the participant is in conversation then they are less likely to be interruptible. Statistical analysis confirms this (Paired t(19) = 6.492, $p < 0.001$).

6.8.3.4 The participants location at the time of the prompt



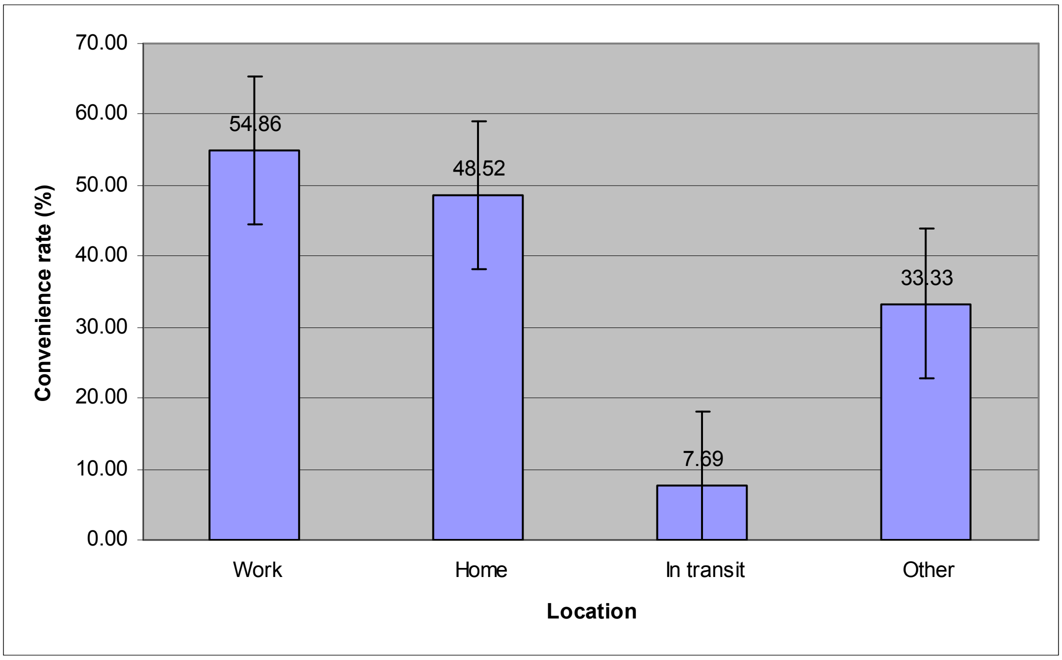


Figure 6-5 Data distribution with relation to convenience rates based on location

The results suggest that where the participant is at the time of the interruption will affect the likelihood that office workers will consider themselves to be interruptible (Figure 6-5).

6.8.4 Discussion

All of the self-reported attributes (location, surrounding people, company, conversation) are shown to be significantly associated with the participant's self-reported availability. Further work should investigate whether these attributes can be captured accurately using mobile-based sensors. The remainder of this chapter will investigate the accuracy of predictive models of availability that have been built using the four self-reported attributes.

6.9 Predictive models of availability

6.9.1 Weka

The Weka machine-learning environment was used to build the availability models discussed in the remainder of this chapter. Weka (Waikato Environment for Knowledge Analysis) is a comprehensive computer application consisting of a collection of machine-learning algorithms and was developed at the University of Waiketo in New Zealand [110]. Weka is freely available under a GNU general public license and is open source. It comes packaged with a console version complete with graphical user interfaces and visualisation tools. It includes a wide range classification algorithms and attribute selection methods.

6.9.2 Naïve Bayes classifier

The models that were built using Weka were Naïve Bayes classifiers and utilised correlation-based attribute selection. To evaluate the models a standard 10-fold cross-validation approach was adopted which involved 10 trials of model construction. In each trial 90% of the samples were used to train the model and the remaining 10% used for testing therefore each sample was used for training nine times and testing one time.

The Naïve Bayes classifier is one of the most straightforward and widely used Bayesian methods. It is an independent feature model/probabilistic classifier that assumes that the presence or absence of a feature is unrelated to any other feature. In real-life situations this is not normally the case but it works well in practice. It can be trained very efficiently and demonstrates effectiveness in complex real world situations. It requires small amount of training data as only independent variances are calculated, not the entire covariance matrix. It is computationally very efficient so is feasible for real-time recognition.

Naïve Bayes classifiers have shown to be effective in a variety of research areas. Fogarty et al. adopted these models in their studies modelling availability and interruptibility in a small group of office workers. They also explored using other classifiers such as decision trees and Support Vector Machines although with no significant improvement. Ter Hofte used Naïve Bayes classifiers in his experiments regarding participant's availability to receive mobile phone calls [99].

6.9.3 Results

19 out of the total 20 participants were included. One participant was excluded, as they answered 'no' to being a convenient time for all entries. Therefore it was not possible to build a classifier for this participant. This left a total of 235 entries.

The model's prediction was compared with the actual response for each of the 235 entries. Overall there were 160/235 correctly classified instances. This demonstrates the models have an average predictive value of 68% using the participants' self-reported availability as ground truth. The percentage correctly identified for each participant can be seen in Figure 6-6.

Participants reported availability 119/235 of the time. Therefore a system that predicted participants' availability as positive at these times would be correct 50.6% of the time. Models that have been trained on a small set of context information would fare significantly better: on average 68% of the time (Figure 6-6). A result which can be demonstrated by statistical analysis: Paired t(18) = 4.140, $p < 0.001$.

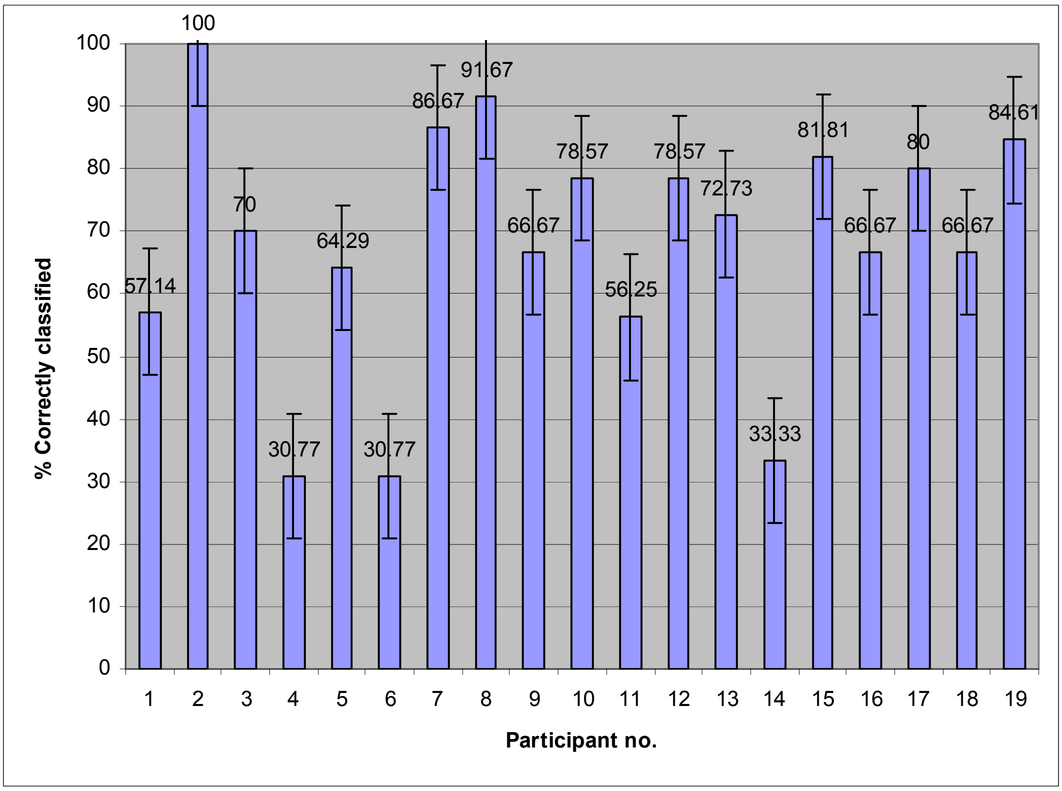


Figure 6-6 Correctly classified % for each participant when adopting Naïve classifiers trained using self-reported 'location', 'number of surrounding people', 'in company' and 'in conversation'

6.9.4 Attribute selection

Weka can be used to select the attributes that provide the most predictive power. In this case 'Location' (12/19 models) and 'Conversation' (14/19 models) are the most selected attributes. They are also, arguably, more technically feasible and easily implemented than detecting whether the participant is currently in company and the number of surrounding people.

Location is the most widely used context in context-aware computing which can be demonstrated by the wide range of location-aware applications available. Current technologies include GPS, GSM and Wifi that can make it possible to accurately locate a user to within a few metres.

Conversation: the audio processing community have been working on a variety of features that can be calculated efficiently and can be used to identify human speech from audio streams [111].

6.10 **Summary**

Pilot studies 1 and 2 reported in this chapter demonstrate that the use of sound information, gathered using mobile devices, can help predict the availability of participants to answer a short questionnaire delivered by an electronic diary. When it is quiet the participants were twice as likely to be available. When there was sound present the participants were three times more likely to be unavailable. These sound levels were calculated using low-resolution audio and 3-second length samples.

The characteristics of the participant such as gender and job role can have an effect, in some cases a statistically significant one, on the participants self-reported availability level when using sound as an availability cue.

The self-reported data in the third study show that additional contextual information can also be used to predict participant's availability. The contexts that were reported were 'location', 'number of surrounding people', 'in company' and 'in conversation'. Naïve Bayes classifiers constructed from these predictors were shown to have an average predictive value of 68%. A system that always predicted participants' availability as positive would be correct 50.6% of the time.

All of these contexts were demonstrated as being significantly related to the participant's self-reported availability. When constructing Naïve Bayes predictive models for each participant the most predictive contexts were shown to be location and conversation. In practice, these contexts could potentially be straightforward to capture through the use of GPS and sound analysis.

Chapter 7 will introduce Guided real-time sampling, a novel strategy for real-time prompting. A number of experiments are described that evaluate it's effectiveness at delivering real-time experience sampling prompts at more convenient times for electronic diary users.

7. Guided sampling

7.1 Introduction

Pocket Interview is a handheld electronic data collection and diary tool that can be used to apply experience sampling methods (ESM). Participants in real-time studies require high levels of commitment and exhibit difficulties maintaining their motivation. Therefore, Pocket Interview includes an option that allows this sampling to be 'guided' through the use of audio analysis of the user's immediate environment. Guiding the sampling attempts to automatically defer potentially inconvenient prompts, temporarily, until a less disruptive time.

7.2 Guiding

People can quickly decide when it is appropriate to interrupt one another and are able to interrupt each other without seeming rude or intrusive. Rarely will a person immediately interrupt another when they seem busy, rather they will wait for the person to finish or temporarily pause their current task. Computer systems, however, are unaware of social conventions and can either interrupt at inappropriate times or are set to be passive and wait for user-interaction. Even changing the notification method, for example, to a tactile prompt would still not allow them to respond immediately though an interruption that is placed at the end of a task will usually be less disruptive than one during a task [4].

Guiding is a novel strategy whereby a person's availability to complete a handheld computer based interview is assessed using contextual information. If appropriate, the interview is deferred temporarily until a less disruptive or 'better' time (Figure 7-1). The quality of ESM data is best if the participant answers immediately. If a large delay is allowed between signalling and answering this can introduce memory biases and defeat the purpose of the method. In the past when researchers have allowed for a delayed response they typically restrict this to no more than 30 minutes after the prompt [112]. Guiding allows prompts to be deferred for a short period and looks for a good time within a short predefined window, e.g. 15 or 20 minutes.

The Guiding method will attempt to adopt ubiquitous mobile technologies and sensors that are embedded within the mobile device. One of the aims of the method is to avoid the use of wearable sensors as participants will be resistant to technologies that require them to wear additional equipment. There should also be no external sensors placed in the environment as this will tie the system to a static location. There is no 'bottom-up' development of new sensors as this will require design, development, deployment and evaluation which can be expensive, error-prone and can waste significant time and resources on something that may not even meet the desired requirements.

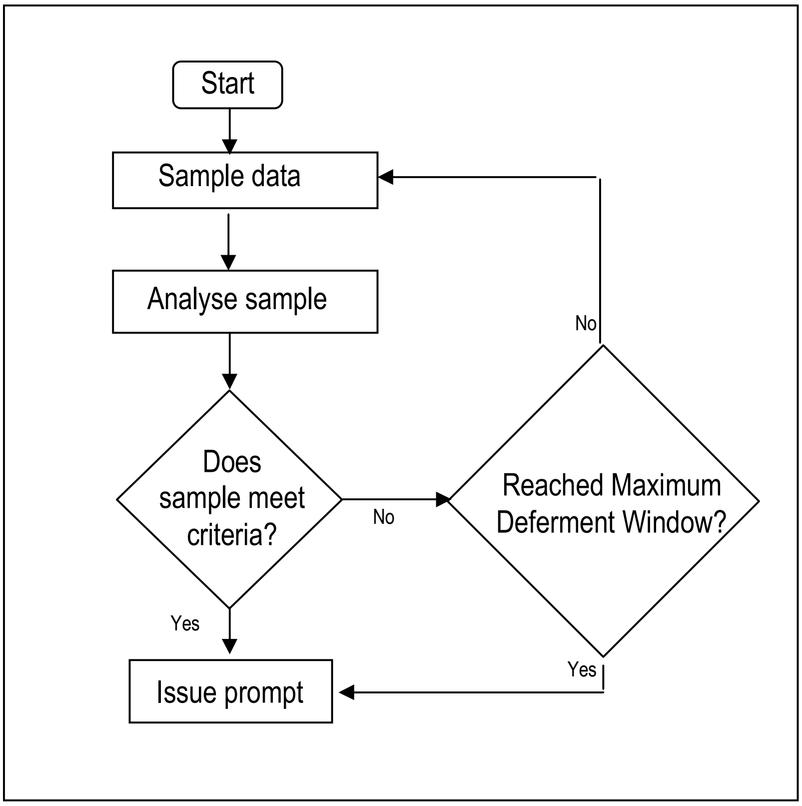


Figure 7-1 Guiding flow chart

7.3 Related work addressing experience sampling weaknesses

Guiding is a strategy that can be used to address weaknesses in ESM. Other recent research is also concerned with addressing some of the inherent shortcomings of ESM e.g. interrupting participants from their current activities or prompting at inappropriate times.

Lisa Feldman Barrett, an expert in experience sampling for emotion assessment, has been quoted as saying that “you cannot pay people enough to be in these studies” and advises that drop-out rates can be high [81]. Interrupting people more than a dozen times a day and asking them to stop what they are doing can prove to be very irritating. She also highlights the importance of carefully motivating the subjects.

ESM can be expensive in both hardware requirements and user-time and may provide little information about uncommon or brief incidents [113]. Participants may find it difficult or even dangerous (e.g. when driving) to make some entries.

One method of addressing these issues is context-triggered sampling. This is a strategy that is most closely related to event-based sampling. This is where the participant is asked to fill out a self-report as the result of some event or activity. The key difference, however, is that event-based sampling relies on the participant to recognise when a relevant event occurs and then initiate a diary-entry themselves. Context-triggered sampling, in contrast shifts this burden to the technology as it uses sensor data to automatically infer when an event of interest has occurred. This allows sampling to be more targeted, in that it only occurs after events of interest and will potentially reduce disruption caused by random prompting. The context data received from the sensors can also be saved continuously and used to test the effectiveness of the sensor technologies.

The Context aware experience sampling tool (CAES) [28] and, more recently, MyExperience [29] are systems that include the context-triggered sampling option. Though promising there is little data about how effective this sampling method may be. One study, using the MyExperience tool, investigated the relationship between place visit behavior and preference [88]. The smartphone based system used GSM signals to detect when the phone had been stationary for a period of 10 minutes after having been on the move. This was used to trigger the prompts. While the authors reported problems with both false positives and negatives they state the system was ‘quite successful’ in capturing data and resulted in 1,981 entries over 28 days with 16 participants. Though there is no reported comparison with the non-triggered method.

The earlier PDA-based CAES tool, along with a wearable accelerometer, was used in a study by Ho et al. This study explored triggering prompts at transitions between physical activities [45]. Their study indicated that people were more receptive to an interruption triggered at an activity transition compared to random triggers.

Context-triggered sampling is subject to some of the methodological issues that affect interval-contingent and signal-contingent sampling. For example, it is probably not the most appropriate technique for studying events where it is not possible to respond to a prompt. In this case, an event contingent based method or a hybrid between the two may be more appropriate because it would give the participant control of when to record an event.

Extra care needs to be taken when configuring trigger-based system as was highlighted in Picard and Liu’s stress study [81]. This study involved the use of a system using a PDA, pedometer and accelerometer and was designed to trigger when significant heart-rate changes were detected that would typically be associated with a change in activity. The study involved 10 subjects, however three were interrupted significantly more than the others with one participant experiencing over 50 interruptions a day. It was found that the prompts were triggered accidentally by a power level charge. A fourth participant also experienced problems due to being overweight and the extra effort needed to stand up it resulted in an increased number of triggers.

An alternative to ESM was proposed by Kahneman, called the Day Reconstruction Method (DRM) [113]. Subjects are asked to complete a detailed diary of events and activities throughout the day. This diary is then used as a memory aid to allow informants to recall thoughts and feelings during a follow-up interview. Kahneman produced evidence regarding the effectiveness of DRM however the method can suffer from low accuracy related to factual aspects e.g. timing of events and other factual details that are easily captured using ESM.

Reconexp is a smartphone based application which combined aspects of ESM and DRM. The participants data is uploaded daily to a website. The participant can then review their data at the end of the day using the website and the participants can then complete any omissions. The authors report an overall response improvement of 29.35% after data has been reviewed. However, due to technical difficulties regarding firewalls and other network security measures, nearly half of participants could not upload and therefore review their data. This method also introduces the possibility of 'postponing' data entry i.e. participants might not answer on the device as they have the option to complete the entry later which could lead to underperformance of the ESM aspect of the method.

7.4 ESM completion rates

Loss of data is a widely acknowledged problem with ESM. Froehlich et al. report completion rates of 80.5% in their context-triggered study [88]. Khalil report completion rates of 80% [114] and Consolvo et al. also report completion rate of 80% although in these studies completion can be as low as 28.5% [89]. The completion rate for the one-day availability studies reported in this thesis in Chapter 5 were 73.9% and 85.6%. All of these rates do not reflect the significance of the data lost as they may occur during times when the data providers are busiest and it may be these times that the researchers are most interested in studying. Reported completion rates can also be inflated as they do not include users who have dropped out of the studies or have been excluded due to low level of their completion.

7.5 Guided sampling using environmental audio

In an office environment, interruptions can range from emails, phone calls to impromptu meetings. Of the many types of messages even those that are time or location specific could be delayed for a short period of time. Previous studies suggest that a small set of sensors can provide valuable information about interruptibility in an office setting and there is evidence that the presence of sound is a useful indicator of a persons' availability to complete a short questionnaire (Chapter 6). None of the user subgroups (administrators, managers, researchers) from these studies indicate that the presence of sound increases the likelihood that the participant is not interruptible.

Handheld computing is currently the most widely used method of collecting data in real time. However, at present there are limited sensors available on a typical PDA though they generally come equipped with built in microphones and Bluetooth connectivity.

When a period of quiet follows a noisy period a subject may be in the process of finishing the current task e.g. a phone call or a meeting and may be more receptive to the interruption. In one experience sampling study participants indicated that they were particularly unreceptive to a sampling prompt while they were talking, meeting with supervisors or in board discussions [45].

A Guided sampling method has been implemented on the Pocket Interview system. When the sampling is guided the system will look for a quiet time within a predefined window in which to prompt the user.

Computer-based environmental audio classification of known events can be a challenging task as there can be a high quantity of different types of sounds and high variation between sounds that belong to the same category. This guided method avoids this challenge by choosing to distinguish between sound and no sound with the overall goal to create a system that is real-time and robust to real-world environments. The sound detection algorithm itself is simple and efficient. By thresholding the audio samples' total energy value and incorporating constraints on event length a Boolean value of whether sound activity is present can be deduced.

The Guiding method that was evaluated initially, and named Guiding1, involves sampling audio for three seconds, analyse and if not 'quiet' then wait 1 minute and then sample again (Figure 7-2). Low-resolution audio samples of 3-second length recorded by PDAs have been demonstrated as sufficient to classify different environments such as offices, streets, bars and car-driving [108]. They have also been shown to give a good indication that an electronic diary user is unavailable (Chapter 6).

Sampling once every minute avoids the need for continuous sampling by the device and the technical issues involved maintaining the device in an unattended state until prompts are issued.

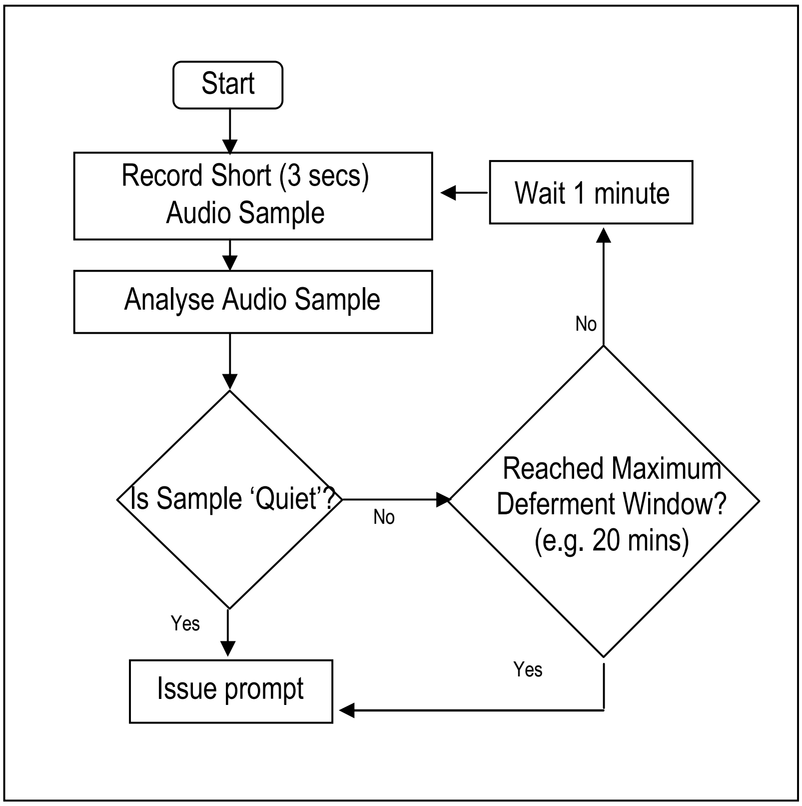


Figure 7-2 Guiding1 - guided method using audio

7.6 Research questions

- Can Guided sampling help trigger prompts at more convenient times for data-participants?
- Can Guided sampling help offer to improve completion rates?

The following studies had ethical approval from the School of Computing Ethics committee, University of Dundee.

7.7 Pilot study 1

A pilot study was performed (Chapter 5 – Study 1) comparing non-guided prompts with prompts using the Guiding1 strategy.

Six users (three male, three female) were recruited from within the University of Dundee. Each user was provided with a Dell Axim PDA running the Pocket Interview system. There were a total of 16 prompts per participant, spread over one working day and consisting eight random prompts and eight guided prompts each within a half-hour window. Participants were asked after each prompt to rate on a 5-pt scaled how convenient an interruption would be at that time (1- not convenient, 5- very convenient). Participants were asked to carry the PDA with them at all times and were asked to answer the question based only on their current situation and not consider previous answers. Participants were unaware that there were two different types of prompts i.e. random and guided.

7.7.1 Results and discussion

Comparing guided and random responses data showed a 0.5 difference in their means in favour of guided however when non-parametric tests are applied this difference was shown not to be statistically significant. This may be due to the wide distribution of responses and/or the small sample size. The data is not normally distributed for either group.

Collectively the data distribution is approximately the inverse of normal i.e. the data peaks at 1 and 5 and is lowest at 3 - separately the random group peaks at 1 and the guided group peaks at 5 (Figure 7-3). The data demonstrates a tendency to cluster at the 1 and 5 measures of the scale. This is promising though suggests that capturing self-reported interruptibility levels may be bimodal in nature and that interruptibility may be better captured using a dichotomous measure rather than a continuous scale.

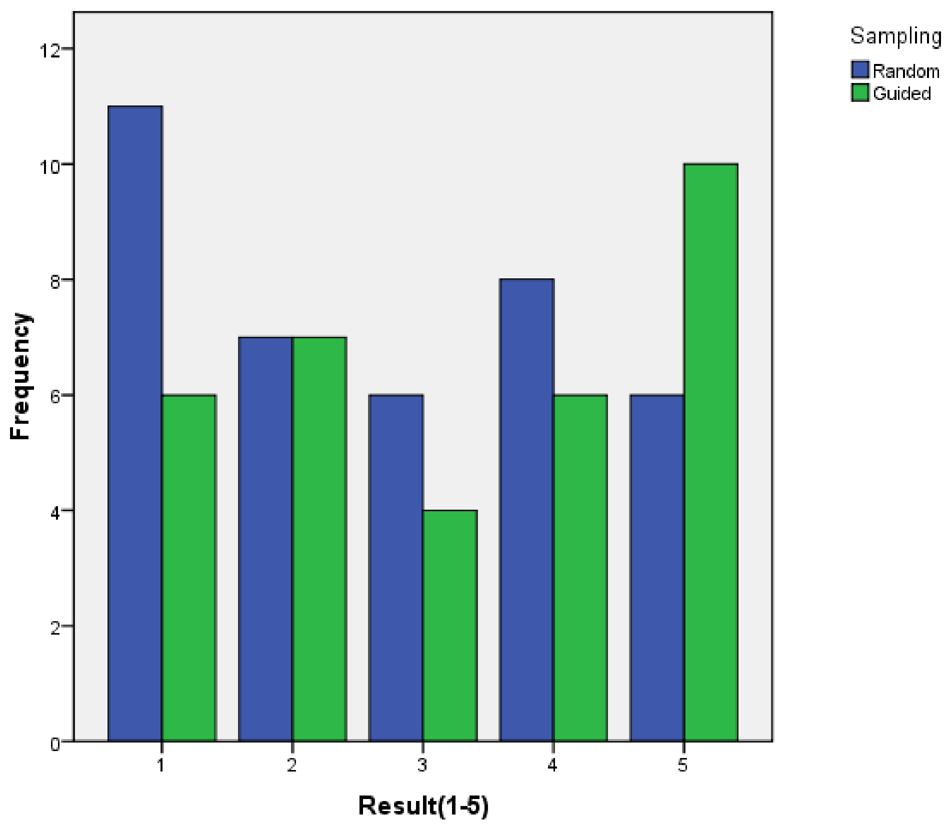


Figure 7-3 Pilot study 1 data distribution

7.8 Pilot study 2

A second pilot study was performed (Chapter 5 – Study 2) comparing non-guided prompts with prompts using the Guiding1 strategy.

20 users (11 female, 9 male) were recruited from within the University of Dundee. Each user was provided with a Dell Axim PDA running the Pocket Interview system. A total of 16 prompts per participant were issued spread over one working day consisting eight random prompts and eight guided prompts appearing once within half-hour windows. Participants were asked after each prompt to rate using a binary (yes/no) scale if an interruption was convenient at that time. Participants were asked to carry the PDA with them at all times and were asked to answer the question based only on their current situation and not consider previous answers. Participants were unaware that there would be two different types of prompts. Audio samples were also recorded and analysed for all entries, not only the guided entries.

7.8.1 Results and discussion

They were a total of 320 possible data points. 47 were not answered through a mixture of non-responses (38) the device being snoozed (5) and the participant not clicking the finish button [3] (4). This left a total of 274 answered data points.

Overall, 143 answered yes (it was convenient) and 131 answered no (it was not convenient a time)

Table 7-1 Pilot study 2 - Guided vs. Non-guided

	Yes	*No*	*Missed*
Non-guided	67	73	21
Guided	76	58	17

Subjects indicated that it was convenient only 47.9% of the time when the prompts were not guided (Table 7-1). A result similar with earlier research that indicated people only consider themselves available to receive mobile phone calls 53% of the time [114].

[3] *Clicking the 'finish' button was necessary for the entry to be recorded. An issue that was resolved in future versions of the system.*

Guiding was triggered 54 out of a possible 160. Only two (3.7%) entries were missed when Guiding was triggered. Overall 11.9% of entries were missed.

When prompts are guided the convenience rate increases to 56.7%. Although this improvement does show an 8.8% increase in accuracy, it is not statistically significant. When examining the prompts where Guiding was triggered, they reveal an accuracy of 50%, a result that suggests further work could be added to improve the Guiding method itself.

7.9 Study 3 - sliding window

Results so far indicate that the presence of sound has a high predictive value. When the data is analysed collectively or within subgroups (i.e. when categorised by gender or job role) the presence of sound will significantly increase the likelihood that the participants are not interruptible. While the Guiding method shows promise, with roughly 50% of triggered guided prompts being steered to a convenient time, these results also suggest that the method may need refining.

In the results presented so far the participant's audio environment is sampled using the PDA's microphone for three seconds per minute. However, due to the relative briefness of this window, it may be possible that the sampling may happen during quiet periods in a busy time e.g. the participant is on the phone and listening, the participant is speaking with someone who is speaking out with the microphones range.

Questions:

- Are there periods of quiet within busy times?
- How often do they occur and for how long?

To help answer these questions and build a model of these quiet periods a PDA application was developed which could be used to record and analyse characteristics of the environmental audio.

While it is possible for PDAs to 'listen' constantly, this will drain their batteries very quickly. As described in Chapter 6, PDAs are optimised to conserve power. Unlike smartphones that have three power states, Windows-based PDAs have a possible seven power states. The will PDA enter a customised version of the 'unattended mode' where the PDA is on and processing but the user is not alerted as the backlight and display are off.

7.9.1 Design

Five users were provided with a PDA running a software application that, when activated by the user, would 'listen' to the surrounding environment and record sound levels in 3-second windows. Sound activity levels for each window would be stored. The sound detection algorithm is the same as previously described in Chapter 6, however, for this purpose the device will now listen continuously until stopped by the user. The device records the audio to file, analyses (a process which takes just a few milliseconds), before recording the next 3-second sample. Users were asked to activate the 'listening' in situations they classed themselves as 'busy' which contained human interaction e.g. meetings, phone calls, lectures, seminars. The sound recordings are only stored on the device for the time taken to analyse and were then automatically deleted.

7.9.2 Results and discussion

A total of approximately 7 hours and 30 minutes of audio was recorded and analysed. When recording in 3-second samples it was found that there was a total of 89 minutes of quiet (minimum length 3 seconds) within this time. 464 periods of quiet were detected which ranged in length from '3-6' seconds to '51-54' seconds (Table 7-2, Figure 7-4).

Table 7-2 Frequency of quiet periods during 'busy' times

Period of Quiet (secs)	3	6	9	12	15	18	21	24	27
Count	313	78	40	17	6	3	2	1	0

Period of Quiet (secs)	27	30	33	36	39	42	45	48	51
Count	0	0	1	1	0	0	0	0	2

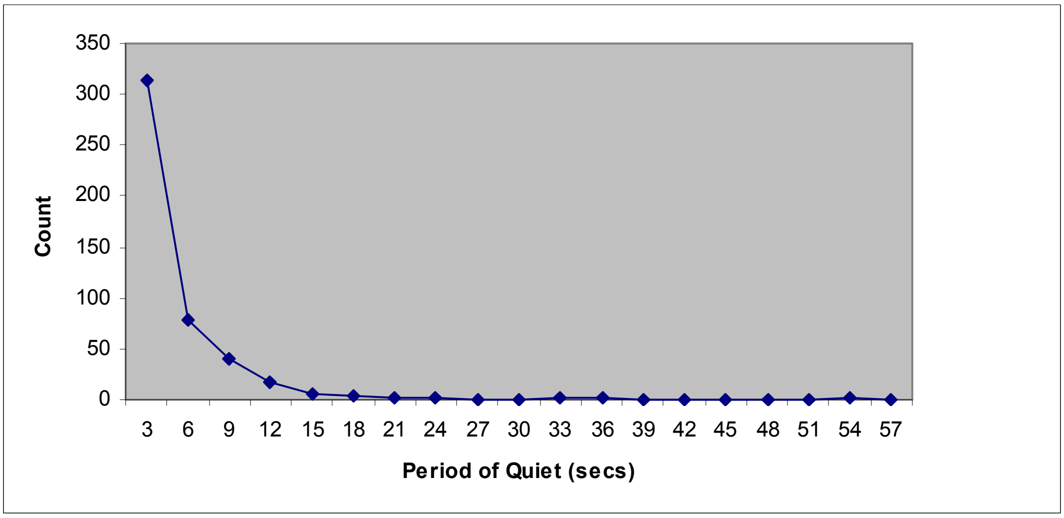


Figure 7-4 Quiet lengths count during 'busy' periods

It was found that 96.7% of the Quiet times were 18 seconds or less in duration therefore it can be inferred that if there is a period of quiet longer than 18 seconds it will have a 3.3% chance of occurring during a busy period containing sound.

7.10 Study 4

Results so far demonstrate that, when using 3-seconds samples, quiet times are considered convenient 65.1% of the time (Chapter 6). This result also suggests that the guided method could be further improved and refined through the introduction of a sliding window of samples.

7.10.1 Design

A Guided sampling method was implemented that incorporated a sliding window (Figure 7-5). This window is 18-20 seconds in length. The window length was informed by the results of the 'Sliding Window' study described earlier in this chapter. This Guiding method is named Guiding2.

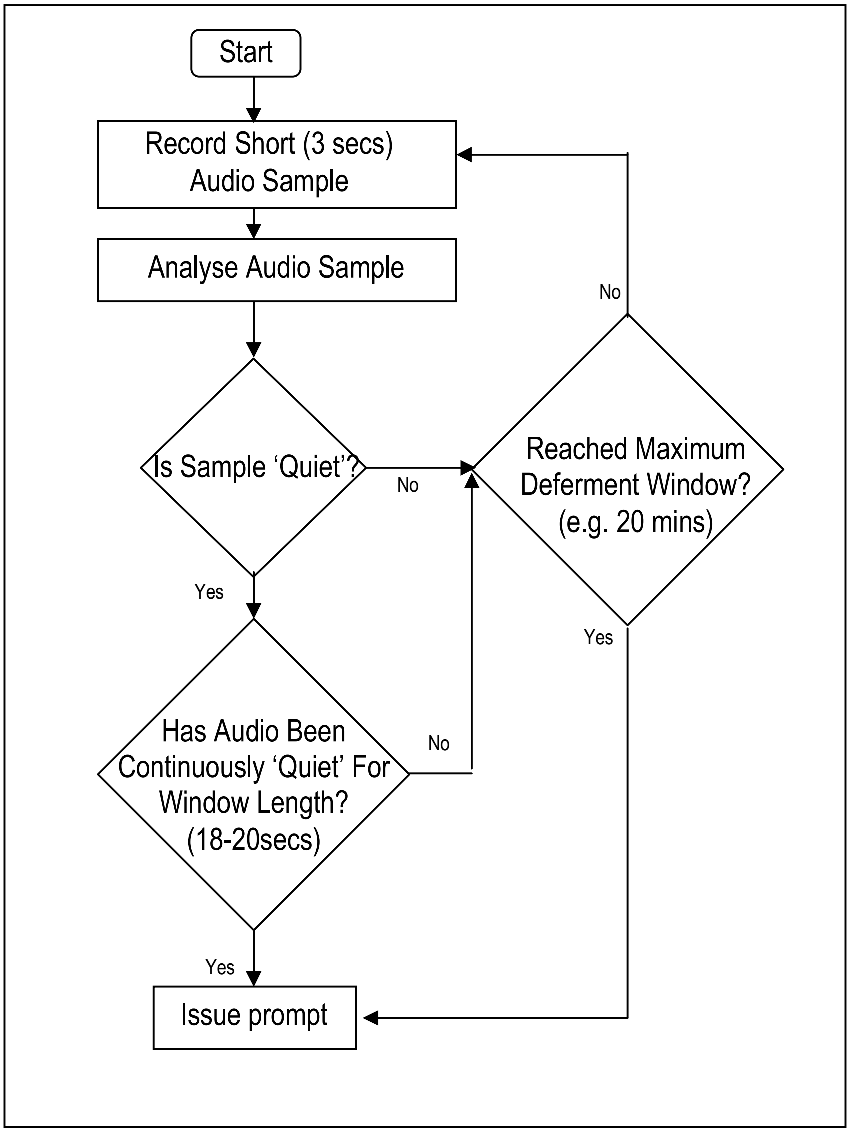


Figure 7-5 Guiding2 - guided method using audio and sliding window

The refined ‘sliding’ method was evaluated using a topical study related to mobiles and privacy attitudes (Chapter 8).

A study was performed (Chapter 5 - Study 3) comparing non-guided prompts with prompts using the Guiding2 strategy.

20 users (11 female, 9 male) were recruited from within the University of Dundee. Each user was provided with a Dell Axim PDA running the Pocket Interview system. A total of 16 prompts per participant were spread over one full day (9am-9pm) and consisted of eight random prompts and eight guided prompts. Each prompt occurred within a forty-minute window.

Participants were asked to answer a short questionnaire related to privacy attitudes and were asked at the end of the questionnaire to rate, using a binary (yes/no) scale, if it was a convenient time to complete the questionnaire using the following question:

Was this a convenient time to answer these questions?
Yes, No

Participants were asked to carry the PDA with them at all times and to answer the question based only on their current situation and not consider previous answers. Participants were unaware that there would be two different types of prompts.

7.10.2 Results and discussion

7.10.2.1 Day 1

Out of the total of 160 guided samples, guiding was triggered a total of 77 out of a possible 155 times (5 were snoozed).

Noticeably, when guiding is triggered (i.e. not quiet) compliance rates increase to 71/77 (92.2%) when compared with 61/78 (78%) when guided (quiet) and 134/159 (84%) when not guided.

Overall when the participants were asked if it was a convenient time 59% replied yes when the prompt was guided and 50% when not guided. This result was not statistically significant, however, when a statistical power analysis was performed using the web-based DSS Researchers toolkit [115] it suggested that between 50% and 100% more samples were required for this result to be shown to be statistically significant ($p < 0.05$). The same 20 participants were therefore asked to participate 1 further day to add significance to the results.

7.10.2.2 Day 1 and Day 2

Table 7-3 Study 4 - Guided vs. Non-guided convenience counts

	Non-guided			*Guided*		
	Yes	**No**	**Missed**	**Yes**	**No**	**Missed**
Day1	69	65	25	77	54	24
Day2	51	61	48	72	53	26
Total	120	126	73	149	107	50

Compiling the results for both days, 120/246 non-guided prompts occurred at convenient times (48.7%) and 149/256 guided prompts (58.2%) (Table 7-3). This is shown to be a significant difference (Paired $t(19) = -2.298$, $p < 0.05$). This suggests that guided prompts will occur at more convenient times for the participants in an experience sampling study.

7.10.2.3 When Guiding is triggered

Table 7-4 Convenience counts when Guiding is triggered

	Guided	Y	N	Missed	Guided to end	Y	N	Missed
Day1	77	41	30	6	21	1	17	3
Day2	97	54	32	11	33	8	20	5
Total	174	95	62	17	54	9	37	8

Guiding, by it's nature, is passive. In this case it will only be triggered when sound is present within the sliding window. As can be seen in Table 7-4 Guiding was triggered 174/320 times. Out of these 174 prompts, 95 were convenient and 62 were not convenient. This indicates that the introduction of the sliding window improves the convenience rate of triggered guided prompts, 60.5% compared with 50% in Pilot Study 2 reported earlier in this chapter.

17/174 entries were missed when Guiding was triggered, just less than 10% of the total when guided. Compared with the non-guided missed entries (73/320) this is a significant difference (Paired $t(19) = -3.814$, $p < 0.001$).

This suggests that although the participants may not consider these prompts particularly convenient, the triggered samples do seem to give an indication of their presence. The participants are more likely to respond when Guiding has been triggered. This is similar to participants actions as reported in a study by Fogarty et al. Using a context-aware instant messaging system as the test system, if the recipient was speaking the sender took that as an indication that they are present while not considering if they were available [24].

7.10.2.4 When guided to end

When guided to end (i.e. Guiding reached the end of the 20 minute deferment window and then prompted), 9/54 (16.7%) of these prompts were convenient. 37/54 (68.5%) prompts were not convenient. Comparing this result with the non-guided prompts shows that 126/246 (51.2%) were not convenient. Further work will be needed to explore this effect although it does indicate, along with earlier results, that during the Guiding process participants are more likely to be present though less likely for it to be a convenient time.

Eight prompts were missed when guided to the end of the deferment window (Table 7-4). Some participants indicated that they noticed that prompts occurred while driving or cycling and therefore they could not answer. In this particular case guided prompts would reach the end of the deferment window and, while the participant is present, it would explain why the may have not responded to some of these eight missed prompts.

7.10.3 Compliance

Compliance is extremely important in ESM studies. Data loss can occur when people are busy or engaged in social/professional activities and may be the ones researchers are interested in. If a participant is only compliant under certain circumstances e.g. when they have nothing else to do or they are in a good mood then the studies results can be undermined. ESM can be demanding for the participants in terms of time and, by its nature, will interrupt and ask the participants to stop what they are doing to complete a questionnaire. This can lead to irritation and the case where participants may even begin to ignore the prompts.

One solution to the compliance problem is to compensate the participants. Monetary rewards have been shown to improve compliance when completing questionnaires [116] although high-value cash incentives may also lead to poorer data quality (e.g. faked data) due to the participant's financial interest and lack of personal involvement [117].

Pocket Interview includes an option to include reminder alarms. If the participant does not respond after a short period additional alarms may sound. However, this could lead to an increase in participant's irritation levels and may actually contribute to them being less likely to respond to future prompts. Pocket Interview also includes a snooze facility, however this relies on the user remembering to initiate a snooze and pre-empt any inappropriate interruptions.

ESM studies generally don't report compliance trends though it is commonly acknowledged that compliance rates will drop off as the study progresses especially in longer studies (e.g. over two weeks) [112]. One study that investigated compliance for participants in an ESM study was a 25-day study where the control group compliance dropped sharply (by over 20%). In this case, the researchers demonstrated maintaining compliance by giving the users a visualisation of the data collected so far [118].

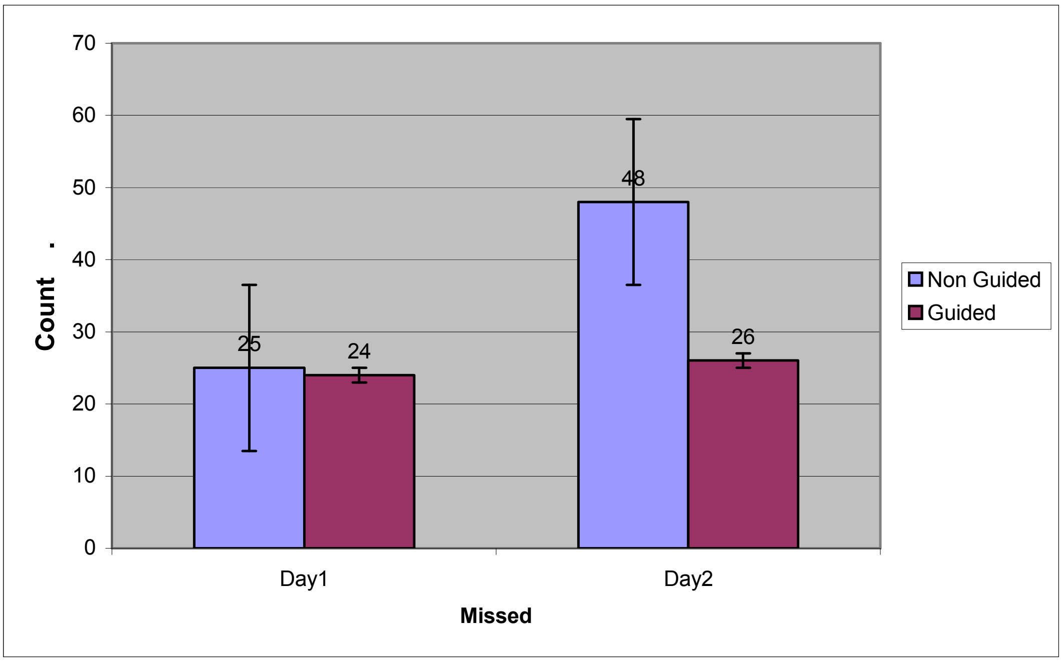


Figure 7-6 Study 4 - number of missed entries organised by day

For this purpose of examining the data collected in Study 4, compliance rate is defined as the number of completed responses divided by the total number of prompts issued.

Table 7-5 Study 4 - compliance rates

	Day 1	Day 2	Paired t test
Non guided	134/159 (84.3%)	112/160 (70%)	Paired $t(19) = -2.334$, $p < 0.05$
Guided	131/155 (84.5%)	125/151 (82.3%)	Not significant

When the compliance rates are examined for the 2-day experience sampling study it can be seen that while the number of missed entries is roughly the same for the first day, the second day demonstrates a large increase in the non-guided missed entries (Figure 7-6). The non-guided random prompts demonstrate a significant decrease in compliance for the second day of the study. The Guided prompts show a small decline in compliance though not a significant one (Table 7-5). This is a result that suggests that Guided real-time sampling can help maintain completion rates.

Studies over longer periods of time will be required to explore this effect further. Is the effect short-term or does the adoption of Guiding help maintain completion rates for longer periods? Compared with recent studies that examine compliance in Experience Sampling studies, two days can be considered a short period. Also, in this case the second day of data collection did not immediately follow the first. There was approximately one month between the two days of data collection though Guiding is still shown to have an effect.

7.11 **Discussion**

Results indicate that Guiding real-time sampling can lead to prompts at more convenient times. Guiding is also shown to maintain compliance rates in a 2-day study. This may be related to the more convenient aspect of Guiding and/or may be due to the triggering of Guiding indicating a participant's presence.

While the results reported in this chapter suggest that the Guiding method using audio is a viable strategy for deferring prompts there are some weaknesses that should be acknowledged. It will not always be the case that quiet times are good times. Prompts during quiet times could disrupt tasks that require concentration such as computer programming, reading and writing. Results presented in Chapter 6, however, suggest that prompts at quiet times are perceived as less disruptive than those provided completely at random. Guiding does seem to give an indication of presence so while they may not be always be guided towards a convenient time they are still more likely to provide a response from the recipient.

There may be times when the device's microphone is obscured e.g. when the device is carried in a bag or the user's pocket. In this case it may be likely that the user will not hear the audio prompt, so will not respond. Guiding will likely be triggered in this situation as the microphone will detect sound through being enclosed. Guiding will defer until quiet as normal. Therefore it could be argued that Guiding will help in this particular case.

7.12 Summary

Guiding is a novel strategy whereby computer-based prompts can be deferred temporarily until a less disruptive time through the sampling of contextual information. In this chapter Guiding methods that involve sampling electronic diary users' nearby environmental audio have been implemented and evaluated. A sliding window is used when sampling to identify periods of quiet that are longer than those that are shown to commonly occur in meetings or other situations involving people interaction.

Data gathered in a 2-day experience sampling study demonstrates that guided real-time sampling can lead to a significant increase in completion rates compared with random sampling ($p < .05$). This may be due to the fact that Guided sampling can result in prompts that are shown to occur at more convenient times for the participants compared with random sampling ($p < .05$). Participants in the experience sampling studies in this thesis, when asked using a bimodal scale, report that unguided prompts happen at convenient times on average 50.1% while Guided sampling leads to 59%.

Chapter 8 will describe a novel study of privacy attitudes related to sharing contextual information with mobile devices.

8. Privacy issues

8.1 Introduction

Systems that use microphones and other sensors run the risk of not being adopted through user's privacy fears that they may be abused. Guided sampling, using the Pocket Interview system, does not require the device to be switched on therefore the user, in fact, may not even be aware that the device may be active and gathering contextual data. To expand Guided sampling further to consider more contextual information such as location and conversation will raise more privacy concerns. These should be addressed before the system can be judged feasible and practical. While the Pocket Interview client device does not store any audio files (only numerical data about the environmental audio levels) participants may be discouraged from providing accurate information when prompted if they feel their privacy is being compromised.

These privacy issues are not unique to Guided sampling but can be linked all forms of context-aware computing and telephony. If people feel that they do not benefit from allowing access to context information they may not agree to share this information.
These systems should also respect the privacy and comfort of others nearby who may be monitored. Privacy concerns will vary greatly depending on what the sensors are recording, the context in which the participants work and live and the perceived value of the information.

8.2 Brief history

In the past, surveys of privacy concerns generally focus on information disclosure to online retailers. The P&AB Harris Interactive privacy classification survey groups users according to their level of privacy concern: privacy fundamentalists, pragmatists and unconcerned [119]. Previous research in consumer privacy has shown that one fourth of consumers are 'privacy unconcerned', one fourth are highly concerned 'privacy fundamentalists' and half fit in between the extremes as 'privacy pragmatists' [120]. Khalil [114] and Consolvo [87] both discovered that while the distribution of participants in their studies related to mobile phone context privacy was the same, many of their participants were categorised differently. Therefore this index may not be suitable as a global measure and different indexes may be needed for different areas such as context-aware computing and telephony.

Recent research has been conducted with the aim of studying privacy issues related to context-aware computing. The common finding from these studies tends to suggest that privacy preferences tend to be complex and can depend on a variety of contextual attributes e.g. the users social relationship with the requester, the time of day and location. One factor commonly observed is that even though users may not react favourably to a new technology they may adopt it anyway. The may be due to peer pressure or possibly an attitude change once they realise the value that it may provide.

The introduction of GPS, cameras in mobile devices and other such technological advances are continually forcing researchers and designers to re-evaluate the boundaries of privacy.

8.3 Disclosing location

Many researchers have identified location context as important and particularly indicative while, at the same time, raising privacy concerns relating to sharing this information.

Sadeh, Hong, Cranor, Fette, Kelley, Prabaker and Rao in a combination of field and lab studies, using a mobile phone based PeopleFinder application, generally found that users were apprehensive about the privacy implications associated with location tracking. Throughout their studies, they also found that most users are not good at articulating these preferences. The accuracy of the policies they define increases only marginally over time unless they are given tools that help them better understand how their policies behave in practice [121].

Consolvo et al. described a study that explored user's location disclosure to social relations [87]. They used an experience sampling technique whereby participants were asked to disclose their location to people they knew. They found that the disclosure of location depends mainly on the recipient, the degree of detail and the reason for disclosure. Location disclosure would appear to be correlated with the strength of the social relationship.

Olson, Grudin and Horvitz focused on personal information and studied the sharing of private information with the aim to create a simple and efficient privacy management system by identifying clusters of information and recipients [120].

Lederer et al presented a mechanism that allows people to control disclosure of their context information [122]. They introduced the 'face' concept as a metaphor to what the user is willing to show others in certain situations. Results from their questionnaire study suggest that who is requesting the information is the primary factor in choosing whether to disclose information or not.

Recently, many person-finder applications have started appear. Some of these have a consumer-orientated slant such as the UK-based mapAmobile that allows users to track a mobile phone to the precision allowed by mobile phone tower coverage [123]. In February 2009 Google introduced Latitude. This is a Google Maps feature that allows most smartphone users to share their location (or more accurately their phones location) with friends. Amidst concerns over location-based privacy Google have claimed that the device won't log data therefore preventing a location history being recorded [124].

Other studies explore the factors that people use to make decisions regarding information disclosure. Work by Adams [125] and Lederer, Mankoff and Dey [122] indicates that 'who' is inquiring is more important than the participant's context in deciding what to disclose. With Consolvo et al. [87] adding that 'why' they wanted to know was also important. Consolvo and Lederer agree that user location and activity are of lesser importance.

8.4 Methods of capturing privacy attitudes

Many recent studies have found discrepancies between people's privacy attitudes and their actual behaviour. For example, Jensen, Potts and Jensen [126] found that the privacy practices of internet users did not match their written questionnaire statements and concluded that questionnaires are best suited to study attitudes and opinions rather than behaviours or experiences and called for a re-evaluation of the role of questionnaires when studying privacy behaviour.

Harper and Singleton [127] have also questioned the validity of using written questionnaires and polls when measuring privacy concerns and preferences. They listed many problems such as questionnaires suffering from the 'talk is cheap' issue whereby it does not cost respondents to be very stringent or unrealistic. Questionnaires can suffer from not being able to replicate the different factors that are considered when participants are making decisions in real life.

Meanwhile, ESM has been shown to provide an effective method of gathering and exploring user's privacy attitudes. Experience sampling studies provide real-time data that is inherently affected by the situations of the participant that cannot be predicted in advance. Connelly, Khalil and Liu extend previous findings from the field of commerce and market research. They compared in-situ techniques with surveys when measuring privacy concerns when sharing context information and found a significant difference when participants gave answers to the same questions, both more and less conservative [128]. Their results suggest that participants will overestimate their privacy concerns associated with company and talking contexts and underestimate privacy concerns related to location and activity.

8.5 Related work

Palen and Dourish point out that the lack of real-world cues makes it more difficult to know who might have access to information collected by technology and how they might use it [129]. These concerns could prevent systems using microphones or cameras being adopted as there may be fears that they might be abused. They suggest it may be worth considering whether more limited sensors are useful.

Khalil et al. present a study that investigated the context information mobile phone users were willing to disclose to potential callers from six different types of social relations [114]. They identified four different types of technically feasible contextual information: location, activity, company (whether the user was with other people) and conversation (whether the user was participating in a conversation). Using an experience sampling method 20 participants were triggered randomly throughout the day (10am – 10pm) for 10 days. They report that on average 'company' was most the most disclosed (74.3%), closely followed by conversation (69.4%). Location (47.4%) and activity (46.4%) proved to be much lower. There was shown to be a significant relationship between location and activity and also a relationship (although not significant) between company and conversation suggesting that the categories could be further split into two groups. 70% reported they would use a service that publishes context information if provided with a tool to manage preferences, 20% would possibly use and 10% would avoid due to privacy concerns. Sharing patterns were found to vary greatly across different social relations. However it was also found that gender had the most significant effect on the sharing rate of context types with females much less likely to share contextual information.

It was found throughout their experience sampling study that overall there was only a 53% availability rate of participants to receive mobile phone calls. This suggests that roughly half of all mobile phone calls are received at inappropriate times and further highlights the usefulness of detecting mobile computing users availability.

In a similar experience sampling study by ter Hofte with 10 subjects over 10 days, 784/980 samples were gathered using 14 samples per day (8am – 10pm). It was reported high availability for mobile phone calls only 51% of the time [99]. Again they found that sharing patterns can vary greatly across different social relations. On average across all social relations 64% coarse location, 25% detailed location, though only 34% released conversation and 20% in company, a finding which be differs from Khalil et als earlier work.

8.6 Motivation

The choice of sensing technology can be critical with regards to user acceptance of a technology. Also the choice of sensing technology will be important with regards to deployment e.g. sensing a person's location is easily possible, however, sensing their exact position indoors will be more difficult. Sensing sound is straightforward while detecting speech can be more challenging.

While studies of users' privacy and identification preferences have been reported, none of these studies focus solely on automatic disclosure and storage of contextual information by mobile devices as a means to assess and estimate their owner's availability. This is a novel study that aims to better understand people's attitudes and behaviors towards privacy as they interact with an application that uses contextual information to assess their availability.

8.7 Research questions

- Will gathering contextual information by mobile devices, and other mobile sensors, compromise their user's privacy concerns?
- Does the user's context itself have an affect on their willingness to disclose context information to a mobile device?

8.8 Design

When a mobile phone call is made the receivers can see who is calling them. The receiver can choose not to answer the phone call though the call itself is still disruptive for the receiver and also for those people nearby. These are interruptions that are disruptive even if they are not acted on.

The test application employed in this study is a context-aware service aimed at reducing inappropriate mobile phone call interruptions by providing a potential caller with a rating of the receiver's availability. This would help the caller make a more informed decision about whether this was an appropriate time to call. The availability rating would be calculated through contextual information such as their current activity, location, company and conversation.

An experience sampling method was used whereby participants received prompts from the Pocket Interview device throughout the day (Chapter 5 – Study 3). They were asked to complete a short questionnaire relating to what contextual information they would be happy to disclose automatically to their phone and also what contextual information they would be happy for the phone to store. In addition they were also asked questions about their current location, whether they were in company, number of surrounding people and whether they were in conversation.

The questionnaires were triggered randomly throughout the day. This helps to minimise their predictability and therefore decrease potential biases in the answers. The questionnaires were triggered over two working days between the hours of 9am and 9pm.

The questions were based on earlier studies by Khalil and Connelly [114] and Barkhuus and Dey [130]. In these previous studies participants were asked to indicate what context information they would share and what people from a group of social relations ranging from strangers to family members they would share it with. For every questionnaire the participants could select any combination of the four different types of context information they were willing to allow access, including the option to store 'none'. Compared to those studies a further question about storing the context information on the device was asked.

Note that the participants were not asked about their attitudes regarding Guiding and the context information that could be used to guide the prompts. Doing so would inform them of the nature of the guided samples and may bias the data related to their convenience. However, the questions asked are directly related to a Guided Sampling application in that they are both context-aware applications that sense their participants surroundings and use this information to infer their availability.

8.9 Context features

The four different types of context information that were deemed relevant to mobile telephony were location, activity, company and conversation. These are context features that have been previously classed as 'high risk' factors (location and activity) and 'low risk' factors (company and conversation) [114].

Capturing these contexts can also be deemed technically feasible given the state of the art in sensing technologies. The results from Chapter 6 demonstrate that these contexts can provide valuable cues about user's availability.

Location: the most widely used context in context-aware computing. The wide range of location-aware applications available can demonstrate this. Current technologies include GPS and GSM and make it possible to accurately locate a user to within a few metres.

Activity: activity can be inferred from location information (e.g work, school) and other cues could also be used such as calendar information, computer activity and real-time analysis of video and audio. IPhones and some other smartphones have built-in accelerometers that can give an indication of activity levels.

8. Privacy issues

Conversation: has already shown to be a very good indicator when predicting interruptibility and availability [24]

Company: the most technically challenging of the four contexts though may be inferred through the use of heat sensors and audio/video analysis

None of these contexts was actually gathered. Future work could investigate capturing these contexts and the accuracy of doing so. Sound levels were monitored to inform the Guided sampling.

8.10 Questionnaire structure

The participants were first asked a question relating to accessing contextual information:

> *'Imagine you are about to receive a mobile phone call. Which of the following would you be happy for the phone to collect automatically about your current context in order to inform the caller of your availability (on a scale of 1- 10) - more than 1 item can be selected'*

Options:

> *Location*
> *Activity*
> *Company*
> *Conversation*

Participants were then asked a further question relating to storing contextual information:

> *'The accuracy of predicting your availability to receive a call could be enhanced by storing a history. What information about your current context would you be happy for the device to store. (this information would not be available to anyone else) - more than 1 item can be selected'*

Options:

> *Location*

Activity

Company

Conversation

The participants were also asked the questions in Table 5-1 related to their location, in company, the number of surrounding people, in conversation and convenience.

8.11 Results

On average answering a questionnaire took 32 seconds to complete, ranging from 13 seconds to 1 min 26 seconds.

In total the study covered 480 hours during which 625 samples occurred. (32 per subject over two days 9am – 9pm). 15 samples were not triggered due to the participant choosing to snooze the device for a short time. 123 were not reacted to within 2 minutes and so therefore were timed out and classed as ‘missed’. The response rate per subject varied from 59.3% to 90.6%. There were a total of 502 data entries, an overall response rate of 80.3%

Participants chose to disclose all their context information 48.6% of the time. Participants chose to disclose at least two of the options 75.4% of the time. Participants chose not to disclose any contexts 4% of the time. These results suggests that participants are willing to share as much information as possible and may go through a process of removing contexts rather than adding them to their list based on their current situations.

5 out of 20 people chose to disclose all of the contexts each time they responded.
One person chose to reveal none of the contexts when asked to ‘access’ while no users chose to disclose none of the contexts when asked to ‘store’.

8.11.1 Correlations

After examining the disclosure rates for all participants and for each context there are strong correlations between disclosure of company and conversation (0.78, 0.76) and moderate correlations between activity and location disclosure (0.52, 0.54) as can be seen in Table 8-1. These results suggest that the participants are grouping these particular contexts. When disclosing one they are likely to disclose the other also.

Table 8-1 Correlation matrices for the 'Access' and 'Store' questions

	Access				Store			
	location	activity	company	conversation	location	activity	company	conversation
location	1				1			
activity	0.52	1			0.54	1		
company	0.23	0.31	1		0.14	0.31	1	
conversation	0.22	0.45	0.78	1	0.17	0.35	0.76	1

8.11.2 Disclosure Rates

Table 8-2 Disclosure rates for each of the contexts: 'location', 'activity', 'company' & 'conversation'

Contextual Info	Access	Store	Total Frequency	Percentage
Location	381	343	724/1004	72.1
Activity	323	319	642/1004	63.9
Company	352	361	713/1004	71.0
Conversation	311	334	645/1004	64.2

The overall disclosure rate was 67.7% (Table 8-2).

Unlike previous research by Khalil and Connelly [114], when disclosing context to social relations, in this study participants are just as willing to share their location and activity. Participants are willing to share these contexts with their device as much as the other contexts.

Also, unlike the same previous research, the location context has the highest disclosure rate. However, location disclosure decreases significantly when the participant is asked to store this information (Paired t(19) = 2.501, $p < 0.05$). This indicates that while participants are likely to allow access to their location, they are more uncomfortable with a history of where they have been being stored on their device. This is unlike the other context disclosure rates that remain approximately the same or actually increase (a result that seems curious as the device will need to access the information before it can be stored).

8.11.3 Disclosure rates based on the context

Does the participant's context itself affect their willingness to disclose their context information? Figure 8-1 presents the overall disclosure rate as related to the number of surrounding people. Figure 8-2 displays the average disclosure rate as related to their location. Table 8-3 presents the disclosure rates for Gender, In company, In conversation, Convenient and guided.

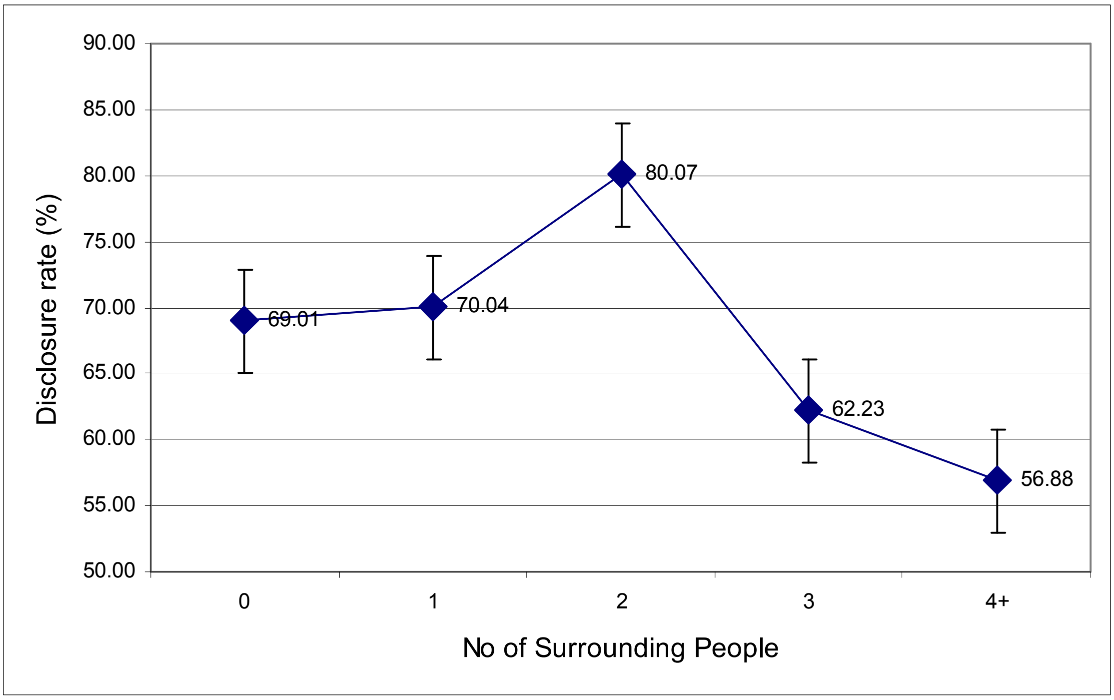


Figure 8-1 The average disclosure rate based on 'number of surrounding people'

Participants were alone 40.2% of the time, surrounded by one person 26.3%, two people 15.5%, three people 9.2% and four or more people 8.4% of the time. Though peaking at 2 the disclosure rate trend is shown to decrease as the number of people increases (Figure 8-1).

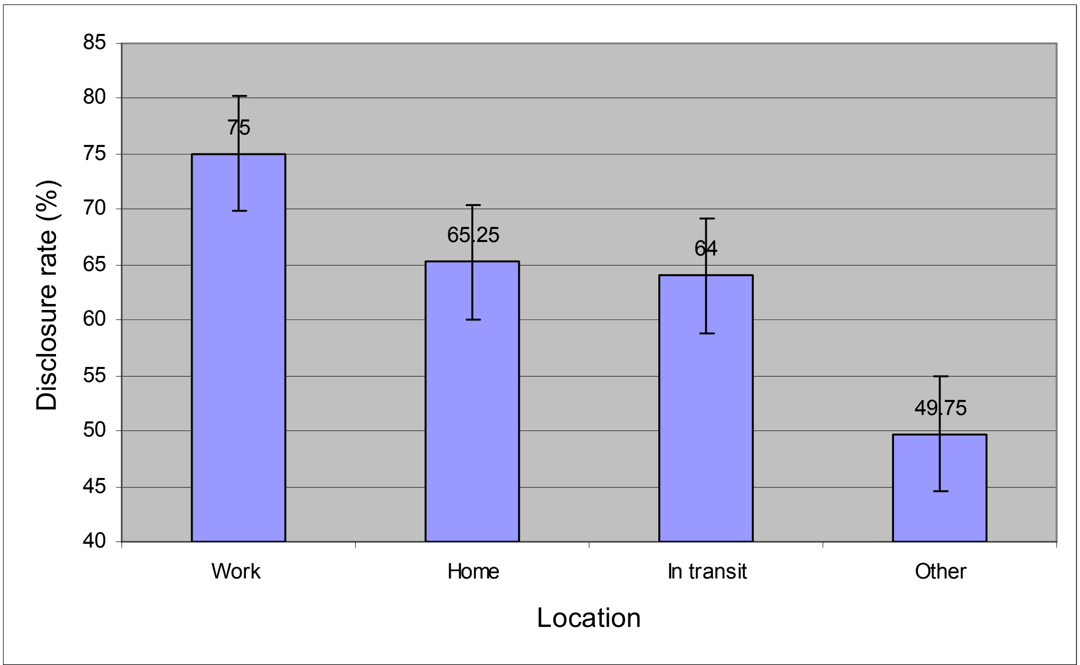


Figure 8-2 The average disclosure rate based on 'location'

Participants disclose most context information when they are at work and the least when they class their location as 'Other' (Figure 8-2).

Table 8-3 Disclosure rates for 'gender', 'company', 'conversation', 'convenient' and 'Guided'.

Variable	Disclosure rates	t test
Gender	M 68.9% F 66.9%	Not significant
In Convstn	Y 65.7% N 69.1%	Not significant
In Company	Y 63.7% N 73.8%	Paired $t(19) = 2.218$, $p < 0.05$
Convenient	Y 70.9% N 63.5%	Paired $t(19) = -2.240$, $p < 0.05$
Guided	Y 70.0% N 64.3%	Paired $t(19) = -2.206$ $p < 0.05$

Unlike previous research assessing privacy disclosure to people of different social relations [114] the participants gender does not have a significant effect on their willingness to disclose context information to their device alone. Statistical analysis demonstrates that there is a significant decrease in disclosure rates when the participants are in company and when they are in conversation. There is also a significant increase in disclosure rates when the participants report that time as convenient and when the prompts have been guided (Table 8-3).

8.12 End of study questions

At the conclusion of the study and immediately following their final diary entry participants were asked two further questions using the electronic diary:

Question:

> *'Would you be happy to use a computer service that would help predict your availability? Your availability could be inferred by collecting current information such as location, activity and whether you were in conversation?'*

Options:

> *I would use if I can manage preferences*
> *I would possibly use the service*
> *I would avoid due to privacy concerns*

68.4% answered they would use if they could manage preferences, 10.5% answered they would possibly use and 21.1% answered they would avoid (19/20 participants answered this question).

Question:

> *How useful would you rate a computer service that would help predict your availability? Your availability could be inferred by collecting current information such as location, activity and whether you were in conversation? (1 - not useful, 5 - very useful)*

The participants answered with a mean 3.2 (standard deviation of 0.79) that suggests that the users in this study would rate this service as useful.

8.13 **Discussion**

The results of this study are relevant to context-aware applications in general. The study explores how much context information a mobile devices user is prepared to disclose to the device and whether they would be comfortable for the device to store context information.

'Location' and 'number of surrounding people' were shown have an effect on disclosure rates. There was found to be a significant difference in disclosure rates when the participants reported they were 'In company', 'In conversation' and also when it was 'Convenient' to answer the questionnaire. Disclosure rates significantly increase when the participants state that it was a convenient time and also when prompts were guided. This is an interesting result as it suggests that if a guided system can find a better time to prompt the user then the user may either provide higher quality richer responses. It may be that disclosure rates are actually related to convenience. This provides an area for further investigation.

Context-aware applications, such as a service that could help predict their user's availability, do not only seem desirable, as is reflected with the high acceptance and usefulness rates, but also feasible due to the high disclosure rates. Similar studies related to mobile phone context disclosure report significant differences based on the types of context information. This is not the case in this study where all context attributes are highly disclosed. However, the participants do show a significant decrease when asked to store their location information. The participants are willing to let the device access their current location although they are less comfortable when a history of their previous locations is stored on the device.

8.14 Summary

This chapter describes a novel study that explores the types of contextual information could potentially be gathered by a mobile device. Would people be willing to allow their mobile device to automatically capture and store this information if it can help a computer service predict their availability? An experience sampling method was used and participants were asked throughout the day which contexts out of four they would be happy for their device to capture at that time. Disclosure rates were shown to be high, 67.7% and also directly related to the participant's context itself. 68.4% of users said they would use a service that could help predict their availability if they could manage the preferences, a result that suggests such services would be desirable to owners of mobile devices.

9. Conclusions and future work

9.1 Lessons learned

This thesis describes the motivation and development of Pocket Interview, a configurable electronic diary and data collection tool that provides the ability to design and deliver questionnaires and collect data using handheld devices and desktop computers. Compared with similar tools the system includes novel and attractive features such as administration through the use of graphical user interfaces, a demo mode, easy installation procedures, integration with Access database, secure data collection and the ability to operate on multiple platforms. The system was shown to be easily usable by both administrators and data providers.

A number of studies have been reported throughout this thesis that have adopted Pocket Interview as a vehicle. Chapter 6 included analysis regarding electronic diary users and their availability to make diary entries. What are the more convenient times for prompting and how can they be detected? Can the use of sound information gathered using mobile devices in work (and other) environments help predict availability of participants?

Chapter 7 introduced a novel Guiding method that involved sampling the users nearby environmental audio in an effort to temporarily defer inconvenient interruptions to a more convenient time. Studies relating to implementation and evaluation were reported. Does Guided sampling improve the user experience, compliance rates and the data collection overall?

Chapter 8 described a novel study of privacy attitudes related to sharing contextual information with mobile devices.

9. Conclusions and future work

The experimental results reported in this thesis demonstrate that:

- Guided real-time sampling using audio can help maintain completion rates when compared with random sampling as was demonstrated in a 2-day study ($p < .05$)

- Participants in the experience sampling studies in this thesis, when asked using a bimodal scale, report that unguided prompts happen at convenient times an average of 50.1% of the time

- Guided sampling using audio can result in prompts that are shown to occur at more convenient times for the participants (59%) compared with random sampling ($p < .05$)

- Self-reported context information 'location', 'conversation', 'company' and 'surrounding people' are all associated with a person's self-reported availability

- Self-reported 'location', 'conversation', 'company' and 'surrounding people' were used as features in Naïve Bayes classifiers and were shown to have an average predictive value of 68%. A system that always predicted participants' availability as positive in this case would be correct 50.6% of the time.

- When real-time sampling occurs at self-reported convenient times there is shown to be a significant increase in contextual information disclosure rates ($p < .05$)

- Context disclosure rates at Guided times increase significantly ($p < .05$)

- Context disclosure rates were shown to be high, 67.7%, and directly related to the participant's context itself. 68.4% of users said they would use a service that could help predict their availability if they could manage the preferences. 21.1% stated they would avoid due to privacy concerns.

Like other studies, it has been demonstrated that only a small set of context information is needed to significantly improve the performance of an availability predictor. This is also shown to be context information that participants are willing to disclose to their mobile devices.

Disclosure rates increase when the participants state that it was a convenient time. This is an interesting result as it suggests that if a Guided system can find a more convenient time to prompt the user, then the user may either provide richer, higher quality responses i.e. less likely to skip quickly through a questionnaire. It may also be that disclosure rates are related to convenience. This provides an area for further investigation.

9.2 Contributions of Guiding

Guiding is a novel strategy whereby a person's availability to complete a handheld computer based interview is assessed using contextual information. Guiding is also a strategy that can be used to tackle the increasing problem of infomania. At present, computing system designers require strategies for minimising the interruptions caused by their proactive messages. Guiding can offer such a strategy as well as a number of other attractive features.

Guiding does not require a training stage which participants may be resistant to. Audio-based Guiding demonstrates that a default model can provide reliable results for many people and situations.

Guiding uses ubiquitous mobile technologies and sensors that are embedded within the mobile device. There are no external sensors placed in the environment that would tie the system to a static location. The system does not require the use of wearable sensors, the participants may be resistant to systems that require them to wear additional equipment. There are no additional costs when using inbuilt microphones making large-scale deployment easily possible.

Development of new sensors is avoided as this will require design, development, deployment and evaluation which can be expensive, error-prone and can waste significant time and resources on something that may not even meet the desired requirements.

Guiding does not require the device to be switched on. A PDA's power management includes an 'unattended' mode whereby the PDA can be active and run programs without alerting their owners. Efficiency is especially important when deploying applications on mobile devices. A mobile system that constantly monitors will quickly drain the devices battery therefore power consumption is a concern for both manufacturers and users of mobile devices. Given these considerations efficiency can be a major factor that affects whether a new mobile technology will be adopted. Guiding is a mechanism where environmental context information is gathered only for the length of the maximum deferment window and has been demonstrated to perform efficiently over a 2-day experience sampling study.

9.3 Future work

The work presented in this thesis suggests a variety of future work.

9.3.1 Larger studies

Larger studies with more participants, over longer periods of time would be necessary to strengthen the observations that have been reported in this thesis. In terms of study length, the sample size may be considered small compared to other recent context-based ESM studies. There should be additional investigations with more subjects and over longer periods.

UK based office-workers were recruited to take part in theses studies. More job and social roles could be included and participants from different countries and cultures.

Results suggest that gender and job role can have effects on self-reported interruptibility and availability. Further studies could explore these differences.

9.3.2 New features for Pocket Interview

Pocket Interview provides an excellent vehicle for exploration. Pocket Interview has shown to be an effective and easily usable system for collecting not only real-time experience sampling data but also sensor-gathered context information.

As the Pocket Interview software evolves and continues to develop there are a number of new features that could be added.

9.3.2.1 Mobile phone version

Recently researchers have been developing configurable smart phone based systems as a means of gathering data [29, 99]. At present the Pocket Interview software is compatible with smart phones running the Windows Mobile Professional operating system and with some further development will run on all Windows-based mobiles. However a report in 2008 estimated that Windows Mobile only accounted for 13.5% of the global smart phone market [131]. Symbian is the platform adopted by Nokia, Sony Ericsson and Panasonic amongst others and it accounts for 46.6% of the global market. The Apple iPhone has 17.3% and the Blackberry has 15.3%. Therefore a mobile phone version of Pocket Interview that is compatible with all mobile platforms should be developed.

There are many potential advantages of using mobile phones. These are devices that many users already own and are comfortable operating. As well as being more convenient for the participants as they will not be required to carry another device (as well as their personal one), mobile phone users are more likely to keep the devices with them and are less likely to lose or forget them. There is also the potential for project savings as additional equipment may not be required.

Using mobile networks allows the data to be downloaded instantly or as soon as network coverage is available. Therefore it is possible to remotely monitor participants while the experiment is still running. The data that is being collected could be amended in real-time based on the responses that have been received so far. Using participant's own phones would also allow the studies to be more widespread and may even aid in recruitment of participants.

Extra care does need to be taken when adopting a participant's own device. Software will require to be thoroughly tested to avoid crashing the device and thus preventing the owner from being able to use it. Phones also have limited memory and processors and so therefore the system should not noticeably impact the phones performance.

Fischer argues for ESM tools that utilise browsers, mobile communication such as SMS, web browsers and, wherever possible, avoids altogether the need for additional client software [132]. While this may help with the issues involved with deploying software on mobile devices it may also lead to a system with reduced flexibility and scope for evolution.

9.3.2.2 Additional sensors

The system should allow for the connection of additional sensors. Researchers may be interested in the participant's activity levels, heart-rate monitoring or other readings from wearable technologies. These sensors may be connected with wired connections or wireless protocols such as Bluetooth.

9.3.2.3 Multimedia input

At present Pocket Interview includes an audio question type. This allows participants to dictate lengthy answers to open-ended questions. This could be further extended to allow participants the most convenient method of answering questions or allow them to complement their answers with multimedia. When appropriate the user may want to include a photograph with their response or record a video. If the device is capable, the facility to allow more multimedia input, such as photographs or videos, could be added. This would greatly enrich the data for the researcher.

9.3.2.4 Roaming

At present, Pocket Interview can run on multiple devices e.g. desktop, laptop and mobile devices. The introduction of a roaming feature could be explored. The participants could be prompted on the most convenient or easily available device for them at that particular time e.g. their desktop when they are at their desk, their mobile device when they are on the move.

9.3.2.5 Visualisation tools

The participant's data is stored in XML format. Pocket Interview provides tools that allow this data to be easily integrated into an Access database application. From there the data is in a format that can be easily manipulated and exported to many statistical analysis tools. Pocket Interview itself does not provide any analysis capabilities. It would be useful to provide visualisation tools for the data that has been gathered. As a study is in progress researchers could quickly view and visualise the data.

9.3.2.6 Notification methods

Ideally, context-aware services should not be constantly intrusive or attention demanding. Attention demands will play a major role in shaping user experience and ultimately the level of acceptance and adoption. Adopting PDAs for ESM studies will restrict the notification methods available to audio prompts that will interrupt surrounding people also. Adopting other mobile devices such as smartphones will allow for different notification methods such as tactile vibrations. These alerts could then range from private to public, subtle to intrusive.

What is the right balance between the need for user attention and subtleness? Different notification may be appropriate for different situations and the user preferences may also vary. Some users may want their notification methods to change automatically while others may want complete control.

9.4 Wider applications

The use of Guiding is not limited to this system or ESM in general. As well as making a participant's involvement in research easier and more effective the Guiding strategy, whereby interruptions are deferred temporarily, could be applied to any system that interrupts.

9.4.1 Phone calls

Wider applications of Guiding could include mobile (and other) phone call delivery and delivery of SMS/MMS messages. In today's society mobile phone are becoming more popular which leads to many mobile phone calls arriving at inconvenient times. It was found throughout one experience sampling study that participants reported themselves as available to receive a mobile phone call 53% of the time [114]. Another experience sampling study reported 51.1% [99]. These are results that suggest that roughly half of all mobile phone calls are received at inappropriate times. The receiver of a mobile call can choose not to answer a phone call though the call itself is still disruptive for the receiver and also for those people nearby as well. These are interruptions that are disruptive even if they are not acted on. If someone forgets to switch off a mobile phone and it interrupts a meeting or lecture then they blame themselves, as do the other people attending. If they forget to switch the phone back on then they won't be able to receive incoming calls or messages and will then be blamed by the sender.

Results from a field study of home telephone usage show that callers and receivers come to learn each others daily patterns and this influences the call initiation process [133] e.g. a caller will avoid making a call when they know the receiver is busy or receiver will not answer unknown callers when they are busy. If a sender is provided with an indication the receiver's availability before making a call then they may defer to a more convenient time.

9.4.2 Computer-based interruptions

There are a wide variety of ways that desktop and laptop computers will interrupt their users. Hodgetts and Jones demonstrated that even brief computer interruptions such as emails, instant messages, office assistants, 'save' reminders, notifications e.g printing, internet pop-ups and software update alerts are all disruptive [134]. Even if distracted for a few moments the overall effect is not just the time needed to deal with interruption but also the time needed to refocus and resume the current task.

9.4.3 Other applications

Guiding could be adopted by consumer surveys and marketing companies that target the goods and services that they are selling. Systems that do not try to follow social politeness and convention (e.g. sales calls) can lead to high levels of irritation and could benefit strongly from such strategies.

Further studies would be needed to explore whether Guiding provides an effective strategy for these wider applications.

9.5 Extending Guiding

The context information adopted by the Guiding method reported in Chapter 7 is a Boolean value based on the presence of sound. Results in Chapter 6 suggest that other context information such as location, conversation, company and surrounding people are all significantly associated with a persons self-reported availability.

Results in this thesis and other previously reported studies suggest that a person's location and whether they are in conversation are particularly indicative of their availability. These may be technically feasible given current technologies.

9.5.1 Location

Embedded sensors within personal mobile devices are becoming more pervasive. Sensors such as GPS units are available even in midrange phones and so therefore moving into the mainstream.

9.5.2 Conversation

In Fogarty et als previous work examining robustness of sensor-based models they report that sensing speech can have a predictive value of 75.9% [9]. There are a variety of audio features that could be calculated and tested for their effectiveness in detecting speech using a mobile devices microphone. Lu et al. 2002 presented a set of computationally efficient features: high zero crossing rate ratio, low short-time energy ratio and spectrum flux. They have been shown to be effective for recognising the difference between speech, music and environmental sounds and have been reported to be effective in prior deployment of a context-aware instant messaging client [24].

9.5.3 Activity

Context attributes that have not been examined in this thesis include the relationship between the participant's activity, levels of activity and their availability. At present, iPhones and some other smartphones have built-in accelerometers. There should be further work regarding the participant's activities and how to capture this information. This should not be restricted to types of activity but also the amount of activity could be explored.

9.5.4 Device activity

Another attribute for possible inclusion in a Guiding method is the activity on the device itself. This could include, but not limited to, keypad use, receipt and/or transmission of calls, receipt and/or transmission of SMS/MMS messages, use of software applications.

Each new attribute will require development, testing and evaluation.

9.6 Response rates and missed entries

Keeping participants motivated over the course of longer ESM studies is an area that requires further investigation. Utilising mobile network coverage could make it easier to monitor participant's motivation levels. Are all the participants still engaged? Do some of them need assistance? Researchers could identify these participants and easily contact them by sending an email or SMS.

Providing the participants a means to review the data they have provided could be added. This could be on the device itself or via a project website. Allowing participants to review data that has been collected has been shown to help participants motivation levels is ESM studies [118].

It has been demonstrated in this thesis that Guiding offers to maintain compliance rates and reduce the number of missed entries in ESM studies. Though at present, it is not clear why some individuals may respond more than others. Are there particular characteristics of these people that can explain this and what are these characteristics?

By their nature, it is not straightforward to deduce why entries are missed. A mechanism could be added whereby participants are asked at a later point to state why they have missed a previous entry. While not ideal as this will rely on user's accurate recollection of earlier events or feelings it will aim to provide an insight into the most common reasons why participants don't respond.

There is little reported data investigating the relationship between response rate and participant's irritation levels. How does annoyance affect compliance rates and the quality of the data itself, if at all? Does Guiding help reduce irritation levels in longer studies? If participants are less irritated will they provide richer and more accurate responses?

9.7 Learning

The desktop version of Pocket Interview allows access to more sensors and indications of computer-activity such as typing, mouse movements, windows in focus, current and recent applications. Additional static sensors could easily be integrated such as cameras and sensors that could monitor phone activity and office activity e.g. door use. These additional contexts could be used as features in user models. Chapter 6 included an investigation regarding the use of multiple contexts as features in predictive models of availability. In the past systems that adopt adaptable predictive models will normally be deployed with default settings that have been learned from a general user model. These models could then adapt to each user over time. This avoids the need for a training phase which users will be resistant to. However, sensors that are predictive for one user or for a particular setting such as an office may be less predictive for other people and environments. The system could adopt an approach whereby it learns and adapts to each user through the user's self reported availability. Alternatively, the learning could be performed automatically where the system learns directly from user reaction times and question response rates.

It may be difficult to know what available contexts will be most useful in a statistical model for individual users as was demonstrated by the large variation in user accuracy in Chapter 6. It will prove difficult to manually select the best possible features. It will also be difficult for static models to accommodate unexpected situations. Extra care should be taken so that attributes are relevant. If the available features have little relationship to the concept being modeled then no amount of processing will be able to extract a relationship.

References

1. Zeigarnik, B., "On finished and unfinished tasks", *A source book of Gestalt psychology*, 1967: p. 300-314.
2. Zeldes, N., D. Sward, and S. Louchheim, "Infomania: Why we can't afford to ignore it any longer", *First Monday*, 2007. **12**(8-6).
3. Hudson, J.M., J. Christensen, W.A. Kellogg, and T. Erickson, "" I'd be overwhelmed, but it's just one more thing to do": availability and interruption in research management", in Human Factors in Computing Systems: Proceedings of CHI'02. 2002, ACM New York, NY, USA. p. 97-104.
4. Bailey, B.P., J.A. Konstan, and J.V. Carlis, "Measuring the effects of interruptions on task performance in the user interface", in IEEE International Conference on Systems, Man, and Cybernetics. 2000. p. 757-762.
5. Cutrell, E., M. Czerwinski, and E. Horvitz, "Notification, Disruption, and Memory: Effects of Messaging Interruptions on Memory and Performance", in Interact. 2001. p. 263-269.
6. O'Conaill, B. and D. Frohlich, "Timespace in the workplace: dealing with interruptions", in Conference companion on Human factors in computing systems. 1995, ACM: Denver, Colorado, United States. p. 262-263.
7. Perlow, L., "The time famine: Toward a sociology of work time", *Administrative Science Quarterly*, 1999. **44**(1): p. 57-81.
8. Hudson, S., J. Fogarty, C. Atkeson, D. Avrahami, J. Forlizzi, S. Kiesler, J. Lee, and J. Yang, "Predicting human interruptibility with sensors: a Wizard of Oz feasibility study", in ACM Conference on Human Factors in Computing Systems (CHI'03). 2003, ACM Press New York, NY, USA. p. 257-264.
9. Fogarty, J., S.E. Hudson, and J. Lai, "Examining the robustness of sensor-based statistical models of human interruptibility", in SIGCHI Conference on Human Factors in Computing Systems. 2004, ACM New York, NY, USA. p. 207-214.

10. Fogarty, J., R. Baker, and S. Hudson, "Case studies in the use of ROC curve analysis for sensor-based estimates in human computer interaction", in Graphics Interface 2005. 2005, Canadian Human-Computer Communications Society: Victoria, British Columbia. p. 129-136.
11. Fogarty, J., A. Ko, H. Aung, E. Golden, K. Tang, and S. Hudson, "Examining task engagement in sensor-based statistical models of human interruptibility", in SIGCHI conference on Human factors in computing systems. 2005, ACM: Portland, Oregon, USA. p. 331-340.
12. Horvitz, E. and J. Apacible, "Learning and reasoning about interruption", in 5th international conference on Multimodal interfaces. 2003, ACM: Vancouver, British Columbia, Canada. p. 20-27.
13. Hess, S.M. and M.C. Detweiler, "Training to Reduce the Disruptive Effects of Interruptions", in Proceedings of the Human Factors and Ergonomics Society Annual Meeting. 1994, Human Factors and Ergonomics Society. p. 1173-1177.
14. McFarlane, D.C., "Comparison of Four Primary Methods for Coordinating the Interruption of People in Human-Computer Interaction", *Human-Computer Interaction*, 2002. **17**(1): p. 63-139.
15. Drugge, M., M. Nilsson, U. Liljedahl, K. Synnes, and P. Parnes, "Methods for interrupting a wearable computer user", in Eighth International Symposium on Wearable Computers 2004. p. 150-157.
16. Nilsson, M., M. Drugge, U. Liljedahl, K. Synnes, and P. Parnes, "A Study on Users' Preference on Interruption When Using Wearable Computers and Head Mounted Displays", in Third IEEE International Conference on Pervasive Computing and Communications. 2005, IEEE Computer Society. p. 149-158.
17. van Dantzich, M., D. Robbins, E. Horvitz, and M. Czerwinski, "Scope: providing awareness of multiple notifications at a glance", in Advanced Visual Interfaces. 2002: Trento, Italy. p. 267-281.
18. Horvitz, E., P. Koch, and J. Apacible, "BusyBody: creating and fielding personalized models of the cost of interruption", in Conference on Computer Supported Cooperative Work (CSCW 2004). 2004, ACM New York, NY, USA. p. 507-510.

19. Mynatt, E. and J. Tullio, "Inferring calendar event attendance", in Proceedings of the 6th international conference on Intelligent user interfaces. 2001, ACM: Santa Fe, New Mexico, United States. p. 121-128.
20. Horvitz, E., A. Jacobs, and D. Hovel, "Attention-Sensitive Alerting", in 15th Conference in Conference on Uncertainty and Artificial Intelligence. 1999. p. 305-313.
21. Horvitz, E., P. Koch, C.M. Kadie, and A. Jacobs, "Coordinate: Probabilistic Forecasting of Presence and Availability", in 18th Conference in Uncertainty in Artificial Intelligence. 2002. p. 224-233.
22. Horvitz, E., J. Apacible, M. Subramani, R. Sarin, P. Koch, J. Cadiz, A. Narin, and Y. Rui, "Experiences with the Design, Fielding, and Evaluation of a Real-Time Communications Agent". 2003, Microsoft Research.
23. Horvitz, E., P. Koch, R. Sarin, J. Apacible, and M. Subramani, "Bayesphone: Precomputation of Context-Sensitive Policies for Inquiry and Action in Mobile Devices", in User Modelling. 2005, Springer. p. 251-260.
24. Fogarty, J., J. Lai, and J. Christensen, "Presence versus availability: the design and evaluation of a context-aware communication client", *International Journal of Human-Computer Studies*, 2004. **61**(3): p. 299-317.
25. Danninger, M., T. Kluge, and R. Stiefelhagen, "MyConnector: analysis of context cues to predict human availability for communication", in Proceedings of the 8th international conference on Multimodal interfaces. 2006, ACM: Banff, Alberta, Canada. p. 12-19.
26. Danninger, M., T. Kluge, E. Robles, L. Takayama, Q.Y. Wang, R. Stiefelhagen, C. Nass, and A. Waibel, "The Connector Service - Predicting Availability in Mobile Contexts", in Joint Workshop on Machine Learning and Multimodal Interfaces. 2006, Springer. p. 129.
27. Kapoor, A. and E. Horvitz, "Experience sampling for building predictive user models: a comparative study", in Twenty-sixth annual SIGCHI conference on Human factors in computing systems. 2008, ACM: Florence, Italy. p. 657-666.

28. Intille, S., J. Rondon, C. Kukla, I. Ancona, and L. Bao, "A context-aware experience sampling tool", in CHI '03 extended abstracts on Human factors in computing systems. 2003, ACM Press: Ft. Lauderdale, Florida, USA. p. 972-973.
29. Froehlich, J., M.Y. Chen, S. Consolvo, B. Harrison, and J.A. Landay, "MyExperience: a system for in situ tracing and capturing of user feedback on mobile phones", in MobiSys. 2007, ACM New York, NY, USA. p. 57-70.
30. Oliver, N., A. Garg, and E. Horvitz, "Layered representations for learning and inferring office activity from multiple sensory channels", *Computer Vision and Understanding* 2004. **96**(2): p. 163-180.
31. Horvitz, E., J. Breese, D. Heckerman, D. Hovel, and K. Rommelse, "The Lumiere Project: Bayesian User Modeling for Inferring the Goals and Needs of Software Users", in Conference on Uncertainty and Artificial Intelligence. 1998. p. 256-265.
32. Begole, J.B., J.C. Tang, and R. Hill, "Rhythm modeling, visualizations and applications", in ACM Symposium on User Interface Software and Technology. 2003. p. 11-20.
33. Avrahami, D. and S. Hudson, "Responsiveness in instant messaging: predictive models supporting inter-personal communication", in SIGCHI conference on Human Factors in computing systems. 2006, ACM: Montreal, Quebec, Canada. p. 731-740.
34. Avrahami, D. and S.E. Hudson, "QnA: augmenting an instant messaging client to balance user responsiveness and performance", in ACM conference on Computer Supported Cooperative Work. 2004, ACM New York, NY, USA. p. 515-518.
35. Lasen, A., "A Comparative Study of Mobile Phone Use in Public Places in London, Madrid and Paris", Digital World Resource Centre, University of Surrey.
36. Monk, A., J. Carroll, S. Parker, and M. Blythe, "Why are mobile phones annoying?" *Behaviour and Information Technology*, 2004. **23**(1): p. 33-41.
37. Wei, R. and L. Leung, "Blurring public and private behaviors in public space: policy challenges in the use and improper use of the cell phone", *Telematics and Informatics*, 1999. **16**(1-2): p. 11-26.

38. Raento, M., A. Oulasvirta, R. Petit, and H. Toivonen, "ContextPhone: A Prototyping Platform for Context-Aware Mobile Applications", in IEEE Pervasive Computing. 2005, IEEE Computer Society. p. 51-59.
39. Tang, J.C., N. Yankelovich, J. Begole, F. Li, and J. Bhalodia, "ConNexus to awarenex: extending awareness to mobile users", in Computer Human Interaction. 2001, ACM New York, NY, USA. p. 221-228.
40. Milewski, A.E. and T.M. Smith. "Providing presence cues to telephone users", in ACM conference on Computer supported cooperative work. 2000. Philadelphia, Pennsylvania, United States: ACM.
41. Hinckley, K. and E. Horvitz, "Toward more sensitive mobile phones", in 14th annual ACM symposium on User interface software and technology. 2001, ACM: Orlando, Florida. p. 191-192.
42. Siewiorek, D., A. Smailagic, J. Furukawa, A. Krause, N. Moraveji, K. Reiger, J. Shaffer, and F. Wong, "SenSay: A Context-Aware Mobile Phone", in 7th IEEE International Symposium on Wearable Computers. 2003, IEEE Computer Society. p. 248.
43. Kern, N. and B. Schiele, "Context-Aware Notification for Wearable Computing", in 7th International Symposium on Wearable Computing. 2003, ACM Press. p. 257-264.
44. Kern, N., S. Antifakos, B. Schiele, and A. Schwaninger, "A model for human interruptability: experimental evaluation and automatic estimation from wearable sensors", in Eighth International Symposium on Wearable Computers (ISWC'04). 2004. p. 158-165.
45. Ho, J. and S.S. Intille, "Using context-aware computing to reduce the perceived burden of interruptions from mobile devices", in SIGCHI conference on Human factors in computing systems. 2005, ACM: Portland, Oregon, USA. p. 909-918.
46. Marti, S. and C. Schmandt. "Giving the caller the finger: collaborative responsibility for cellphone interruptions", in CHI Extended Abstracts. 2005.
47. Liu, K.K., "A Personal, Mobile System for Understanding Stress and Interruptions". 2004, Massachusetts Institute of Technology.

48. etforecasts. "Worldwide PDA & Smartphone Forecast", [cited 2007 June]; Available from: http://www.etforecasts.com/products/ES_SP-PDA.htm.
49. Canalys. "Global smart phone shipments rise 28%", 2008 [cited 2009; Available from: http://www.canalys.com/pr/2008/r2008112.htm.
50. Garritty, C. and K. El Emam, "Who's using PDAs? Estimates of PDA use by health care providers: a systematic review of surveys", *Journal of medical Internet research*, 2006. **8**(2).
51. Fischer, S., T.E. Stewart, S. Mehta, R. Wax, and S.E. Lapinsky, "Handheld computing in medicine", *Journal of the American Medical Informatics Association*, 2003. **10**(2): p. 139-49.
52. Wu, R.C. and S.E. Straus, "Evidence for handheld electronic medical records in improving care: a systematic review", *BMC medical informatics and decision making*, 2006. **6**(26).
53. Ruland, C.M., "Handheld technology to improve patient care: evaluating a support system for preference-based care planning at the bedside", *J Am Med Inform Assoc.*, 2002. **9**(2): p. 192-201.
54. Wu, J.H., S.C. Wang, and L.M. Lin, "Mobile computing acceptance factors in the healthcare industry: A structural equation model", *International journal of medical informatics*, 2007. **76**(1): p. 66-77.
55. Weber, B.A. and B.L. Roberts, "Data collection using handheld computers", *Nursing research*, 2000. **49**(3).
56. Hufford, M.R. and A.L. Shields, "Electronic diaries: applications and what works in the field", *Appl Clin Trials*, 2002: p. 46-59.
57. Shiffman, S., J.A. Paty, M. Gnys, J.A. Kassel, and M. Hickox, "First Lapses to Smoking: Within-Subjects Analysis of Real-Time Reports." *Journal of Consulting and Clinical Psychology*, 1996. **64**(2): p. 366-379.
58. Todd, M., S. Armeli, H. Tennen, M.A. Carney, S.A. Ball, H.R. Kranzler, and G. Affleck, "Drinking to Cope: A Comparison of Questionnaire and Electronic Diary Reports*", *Journal of studies on alcohol*, 2005. **66**(1): p. 121-130.

59. Wang, D.H., M. Kogashiwa, and S. Kira, "Development of a new instrument for evaluating individuals' dietary intakes", *Journal of the American Dietetic Association*, 2006. **106**(10): p. 1588-93.

60. Jamison, R.N., S.A. Raymond, J.G. Levine, E.A. Slawsby, S.S. Nedeljkovic, and N.P. Katz, "Electronic diaries for monitoring chronic pain: 1-year validation study." *Pain*, 2001. **91**(3): p. 277-285.

61. Koop, A. and R. Mosges, "The use of handheld computers in clinical trials", *Controlled clinical trials*, 2002. **23**(5): p. 469-80.

62. Jaspan, H.B., A.J. Flisher, L. Myer, C. Mathews, C. Seebregts, J.R. Berwick, R. Wood, and L.G. Bekker, "Brief report: Methods for collecting sexual behaviour information from South African adolescents-a comparison of paper versus personal digital assistant questionnaires", *Journal of adolescence*, 2006 **30**(2): p. 353-359.

63. Palermo, T.M., D. Valenzuela, and P.P. Stork, "A randomized trial of electronic versus paper pain diaries in children: impact on compliance, accuracy, and acceptability", *Pain*, 2004. **107**(3): p. 213-9.

64. Kvien, T.K., P. Mowinckel, T. Heiberg, K.L. Dammann, O. Dale, G.J. Aanerud, T.N. Alme, and T. Uhlig, "Performance of health status measures with a pen based personal digital assistant", *Annals of the Rheumatic Diseases*, 2005. **64**(10): p. 1480-4.

65. Tourangeau, R. and T.W. Smith, "Asking sensitive questions: The impact of data collection mode, question format, and question context". 1996, AAPOR. p. 275-304.

66. Palmblad, M. and B. Tiplady, "Electronic diaries and questionnaires: designing user interfaces that are easy for all patients to use", *Qual Life Res*, 2004. **13**(7): p. 1199-1207.

67. Stone, A.A., S. Shiffman, J.E. Schwartz, J.E. Broderick, and M.R. Hufford, "Patient non-compliance with paper diaries", *British Medical Journal*, 2002. **324**(7347): p. 1193-1994.

68. Stone, A.A., S. Shiffman, J.E. Schwartz, J.E. Broderick, and M.R. Hufford, "Patient compliance with paper and electronic diaries", *Controlled Clinical Trials*, 2003. **24**(2): p. 182-99.
69. Green, A.S., E. Rafaeli, N. Bolger, P.E. Shrout, and H.T. Reis, "Paper or plastic? Data equivalence in paper and electronic diaries", *Psychological Methods*, 2006. **11**(1): p. 87-105.
70. Stone, A.A., S. Shiffman, J.E. Schwartz, J.E. Broderick, and M.R. Hufford, "Patient non-compliance with paper diaries", *BMJ* 2002. **324**(7347): p. 1193-4.
71. Larson, R. and M. Csikszentmihalyi, "The experience sampling method", *New Directions for Methodology of Social and Behavioral Science*, 1983. **15**: p. 41-56.
72. Bolger, N., A. Davis, and E. Rafaeli, "Diary methods: Capturing life as it is lived", *Annual Review of Psychology*, 2003. **54**: p. 579-616.
73. Beal, D.J. and H.M. Weiss, "Methods of Ecological Momentary Assessment in Organizational Research", *Organizational Research Methods*, 2003. **6**(4): p. 440-464.
74. Reis, H.T. and S.L. Gable, "Event-sampling and other methods for studying everyday experience", *Handbook of research methods in social and personality psychology*, 2000: p. 190–222.
75. Conner Christensen, T., L. Feldman Barrett, E. Bliss-Moreau, K. Lebo, and C. Kaschub, "A practical guide to experience-sampling procedures", *Journal of Happiness Studies*, 2003. **4**: p. 53-78.
76. Sorokin, P.A. and C.Q. Berger, "*Time-budgets of human behavior*". 1939: Harvard University Press.
77. Wessman, A.E. and D.F. Ricks, "*Mood and personality*". 1966: Holt, Rinehart and Winston.
78. Robinson, M.D. and G.L. Clore, "Belief and feeling: Evidence for an accessibility model of emotional self-report", *Psychological Bulletin*, 2002. **128**(6): p. 934-960.
79. Laurenceau, J.P. and N. Bolger, "Using diary methods to study marital and family processes", *Journal of family psychology*, 2005. **19**(1): p. 86-97.

80. Husky, M.M., C.M. Mazure, K.M. Carroll, D. Barry, and N.M. Petry, "Using the experience sampling method in the context of contingency management for substance abuse treatment", *Journal of Applied Behavior Analysis*, 2008. **41**(4): p. 635.
81. Picard, R.W. and K. Liu, "Relative subjective count and assessment of interruptive technologies applied to mobile monitoring of stress", *International Journal of Human-Computer Studies*, 2007. **65**(4): p. 361-375.
82. Johnston, D.W., A. Beedie, and M.C. Jones, "Using computerized ambulatory diaries for the assessment of job characteristics and work-related stress in nurses", *Work & Stress*, 2006. **20**(2): p. 163-172.
83. Burton, C., D. Weller, and M. Sharpe, "Are electronic diaries useful for symptoms research? A systematic review", *Journal of Psychosomatic Research*, 2007. **62**(5): p. 553-561.
84. Van Eerde, W., D. Holman, and P. Totterdell, "Diary studies in work psychology", *Journal of Occupational and Organizational Psychology*, 2005. **78**: p. 151-154.
85. Barrett, L.F., "Feelings or words? Understanding the content in self-report ratings of experienced emotion", *Journal of Personality and Social Psychology*, 2004. **87**(2): p. 266.
86. Zhang, D. and B. Adipat, "Challenges, methodologies, and issues in the usability testing of mobile applications", *International Journal of Human Computer Interaction*, 2005. **18**(3): p. 293-308.
87. Consolvo, S., I.E. Smith, T. Matthews, A. LaMarca, J. Tabert, and P. Powledge, "Location disclosure to social relations: why, when, what people want to share", in SIGCHI conference on Human factors in computing systems. 2005, ACM: Portland, Oregon, USA. p. 81-90.
88. Froehlich, J., M.Y. Chen, I.E. Smith, and F. Potter, "Voting with your feet: An investigative study of the relationship between place visit behavior and preference", in UbiComp 2006. 2006, Springer: Orange County, California. p. 333-350.

89. Consolvo, S., M. Walker, I. Res, and W.A. Seattle, "Using the experience sampling method to evaluate ubicomp applications", *IEEE Pervasive Computing*, 2003. **2**(2): p. 24-31.

90. Robinson, J.P., " The validity and reliability of diaries versus alternative time use measures", in *Time Goods and Well-Being*, F.T. Juster and F.P. Stafford, Editors. 1985. p. 33-62.

91. Stone, A.A. and S. Shiffman, "Ecological momentary assessment (EMA) in behavioural medicine", *Annals of Behavioral Medicine*, 1994. **16**: p. 199-202.

92. Wirtz, D., J. Kruger, C. Napa Scollon, and E. Diener, "The role of predicted, on-line, and remembered experience in future choice", *Psychological Science*, 2003. **14**(5): p. 520-524.

93. Barrett, E. and D. Barrett. "ESP The Experience Sampling Program", 2005 [cited 2007; Available from: http://www.experience-sampling.org.

94. "Intel Research Seattle and University of Washington Computer Science and Engineering Department, iESP", [cited 2009; Available from: http://seattleweb.intel-research.net/projects/ESM/iESP.html.

95. Weiss, H., D. Beal, S. Lucy, and S. MacDermid. "PMAT: Purdue Momentary Assessment ", 2004 [cited 2007; Available from: http://www.mfri.purdue.edu/pages/PMAT/.

96. Le, B., H.N. Choi, and D.J. Beal, "Pocket-sized psychology studies: Exploring daily diary software for Palm Pilots", *Behaviour Research Methods*, 2006. **38**(2): p. 325-332.

97. "Changing Places Consortium, MIT: CAES: Context Aware Experience Sampling Tool ", 2007 [cited 2009; Available from: http://www.media.mit.edu/~intille/caes.

98. Intille, S., E.M. Tapia, J. Rondoni, J. Beaudin, C. Kukla, S. Agarwal, L. Bao, and K. Larson. "Tools for Studying Behavior and Technology in Natural Settings ", in Proccedings of UbiComp. 2003: Springer.

99. ter Hofte, G.H., "Xensible interruptions from your mobile phone", in 9th international conference on Human computer interaction with mobile devices and services. 2007, ACM: Singapore. p. 178-181.

100. Carter, S., J. Mankoff, and J. Heer, "Momento: support for situated ubicomp experimentation", in Proceedings of the SIGCHI conference on Human factors in computing systems. 2007, ACM: San Jose, California, USA. p. 125-134.

101. Joki, A., J.A. Burke, and D. Estrin, "Campaignr: A Framework for Participatory Data Collection on Mobile Phones". 2007, Centre for Embedded Network Sensing, University of California, Los Angeles.

102. "NHS Scotland Mobile Data Protection Standard v1.0. 2008 ", [cited 2010; Available from: http://www.ehealth.scot.nhs.uk/wp-content/documents/mobile-data-protection-standard-pdf3.pdf.

103. "Google Analytics", 2009 [cited 2009]; Available from: http://www.google.com/analytics/.

104. Lee, K.E., M. Themessl Huber, and P. Davidson, "Research or audit? The benefits and limitations of structured observation of the hand hygiene practice of named staff", *British Journal of Infection*, 2008. **9**(2): p. 12-17.

105. Mellanby, R.J. and M.E. Herrtage, "Survey of mistakes made by recent veterinary graduates", *The Veterinary Record*, 2004. **155**(24): p. 761-765.

106. Field, A., "*Discovering statistics using SPSS*". 2009: Sage Publications Ltd.

107. Korpipaa, P., M. Koskinen, J. Peltola, S.M. Makela, and T. Seppanen, "Bayesian approach to sensor-based context awareness", *Personal and Ubiqitious Computing*, 2003. **7**(2): p. 113-124.

108. Ma, L., B. Milner, and D. Smith, "Acoustic environment classification", *ACM Transactions on Speech and Language Processing*, 2006. **3**(2): p. 1-22.

109. Johnson, J.I. "Windows Mobile Power Management", 2008 Mar 2009 [cited 2009]; Available from: http://www.codeproject.com/KB/mobile/WiMoPower1.aspx.

110. Witten, I.H. and E. Frank, "*Data Mining: Practical machine learning tools and techniques*", 2nd ed. 2005: Morgan Kaufmann, San Francisco.

111. Lu, L., H.J. Zhang, and H. Jiang, "Content analysis for audio classification and segmentation", *IEEE transactions on speech and audio processing*, 2002. **10**(7): p. 504-516.

112. Scollon, C.N., C. Kim-Prieto, and E. Diener, "Experience Sampling: Promises and pitfalls, strengths and weaknesses", *Journal of Happiness Studies*, 2003. **4**: p. 5-35.

113. Kahneman, D., A.B. Krueger, D.A. Schkade, N. Schwarz, and A.A. Stone, "A Survey Method for Characterizing Daily Life Experience: The Day Reconstruction Method", *Science*, 2004. **306**(5702): p. 1776-1780.

114. Khalil, A. and K. Connelly, "Context-aware telephony: privacy preferences and sharing patterns", in 20th anniversary conference on Computer supported cooperative work. 2006, ACM: Banff, Alberta, Canada. p. 469-478.

115. "DSS Researcher's toolkit", 2009 [cited 2009]; Available from: http://www.dssresearch.com/toolkit/.

116. Lynn, P., "The impact of incentives on response rates to personal interview surveys: Role and perceptions of interviewers", *International Journal of Public Opinion Research*, 2001. **13**(3): p. 326-336.

117. Stone, A.A., R.C. Kessler, and J.A. Haythomthwatte, "Measuring daily events and experiences: Decisions for the researcher", *Journal of personality*, 1991. **59**(3): p. 575-607.

118. Hsieh, G., I. Li, A. Dey, J. Forlizzi, and S. Hudson. "Using visualizations to increase compliance in experience sampling", in 10th international conference on Ubiquitous computing. 2008. Seoul, Korea: ACM.

119. P&AB, "Consumer Privacy Attitudes: A Major Shift since 2000 and Why", *Privacy & American Business Newsletter*, 2003. **10**(6).

120. Olson, J.S., J. Grudin, and E. Horvitz, "A study of preferences for sharing and privacy", in CHI '05 extended abstracts on Human factors in computing systems. 2005, ACM: Portland, OR, USA.

121. Sadeh, N., J. Hong, L. Cranor, I. Fette, P. Kelley, M. Prabaker, and J. Rao, "Understanding and capturing people's privacy policies in a people finder application", *Personal and Ubiquitous Computing*, 2008. **13**(6): p. 401-412.

122. Lederer, S., M. Mankoff, and A.K. Dey, "Who wants to know what when? privacy preference determinants in ubiquitous computing", in CHI '03 extended abstracts on Human factors in computing systems. 2003, ACM: Ft. Lauderdale, Florida, USA. p. 724-725.
123. "mapAmobile", 2009 [cited 2009]; Available from: http://www.mapamobile.com/.
124. Gilbertson, S. "Google Latitude Broadcasts Your Location", 2009 [cited 2009]; Available from: http://www.wired.com/epicenter/2009/02/google-latitude/.
125. Adams, A., "Multimedia information changes the whole privacy ballgame", in Computers, Freedom & Privacy. 2000, ACM New York, NY, USA. p. 25-32.
126. Jensen, C., C. Potts, and C. Jensen, "Privacy practices of Internet users: self-reports versus observed behavior", *International Journal of Human Computer Studies special issue on HCI Research in Privacy and Security*, 2005. **63**(1-2): p. 203-227.
127. Harper, J. and S. Singleton, "With a grain of salt: What consumer privacy surveys don’t tell us". 2001.
128. Connelly, K., A. Khalil, and Y. Liu, "Do I Do What I Say?: Observed Versus Stated Privacy Preferences", in Interact. 2007. p. 620-623.
129. Palen, L. and P. Dourish, "Unpacking "privacy" for a networked world", in SIGCHI conference on Human factors in computing systems. 2003, ACM: Ft. Lauderdale, Florida, USA. p. 129-136.
130. Barkhuus, L. and A. Dey, "Is context-aware computing taking control away from the user? Three levels of interactivity examined", in UbiComp. 2003, Springer. p. 149-156.
131. Canalys. "Smart mobile device shipments hit 118 million in 2007, up 53% on 2006", 2008 [cited 2009; Available from: http://www.canalys.com/pr/2008/r2008021.htm.
132. Fischer, J.E., "Experience-Sampling Tools: a Critical Review", in Mobile HCI '09. 2009: Bonn, Germany. p. 35-38.
133. Lacohée, H. and B. Anderson, "Interacting with the telephone", *International Journal of Human Computer Studes*, 2001. **54**(5).

134. Hodgetts, H.M. and D.M. Jones, "Reminders, Alerts and Pop-ups: The Cost of Computer-Initiated Interruptions", in Human-Computer Interaction, Part I. 2007, Springer. p. 818-826.

Appendix 1 – Study participant information sheet

University of Dundee

School of Computing

Guided Experience Sampling

Information

Experience sampling is a term that refers to real-time data collection and methods that allow people to record thoughts, feelings and actions as they happen. This usually involves having participants carry a device during the day that prompts them to answer questions.

The experience sampling method may be disruptive for the participant especially if the prompts are more frequent. Guiding the sampling in some way could potentially decrease this disruption and inconvenience.

The aim of this study is to investigate whether we can guide the sampling to detect a better time that may be more convenient for the participants.

Participants will be asked to sign a form saying that they are willing to participate in the study. The consent form will explain what will be asked and what will happen to any information that is collected. The participants will be given a copy of this form to keep.

If you would like to know more about this research and/or you have questions please feel free to contact Kenny Morrison, by phone at 388234 or by email at kmorrison@computing.dundee.ac.uk.

Instructions

You will be provided with an electronic diary (a Personal Digital Assistant) which will prompt you randomly by audio alarm on average once every half an hour throughout one working day (9am - 5pm). You will be asked to make a diary entry after each prompt which will consist of one yes/no question asking the participant if a that time it is convenient for an interruption.

The purpose of the diary is to collect information in real-time as it happens so it would be preferable if you keep the diary with you at all times e.g. lunchtime. You can snooze the device for different periods of time, up to 1 hour, when you won't be prompted.

Appendix 2 – Participant consent form

University of Dundee

School of Computing

Guided Experience Sampling

Consent Form

Dear Participant

Thank you for your interest in this study. This page describes what you will be asked to do for the study. Please read through it and then sign at the bottom to say that you understand and accept the conditions of this study. If you have questions, please feel free to ask the researcher.

You will be provided with an electronic diary (a Personal Digital Assistant) which will prompt you randomly, by an audio alarm, on average once every half an hour throughout one working day (9am - 5pm). You will be asked to make a diary entry after each prompt that will consist of one yes/no question enquiring if that time it is convenient for an interruption.

Your participation is voluntary and you can leave the study at any time without penalty or giving reasons. No undue risk arises from the participation in this study. All the information which you give us will be stored safely and kept separate from information about your identity. Access to your data is minimised to the people involved in this research. If information about you is used for publications or presentation, we will ensure that no reference to your identity is made. Please date and sign this page below to indicate that you understand and accept the conditions of this study.

Thank you for your help.

Name of Participant: ____

Signature: ____

W*itness of Researcher:* ____

Date: ___ /______ / 200 _

Appendix 3 - Ethics response

Kenny Morrison   Date: 1/06/09
School of Computing
University of Dundee

Dear Kenny

Full title of study: Guided Experience Sampling

SoCEC reference number: 09/001

Thank you for responding to my letter dated 30/01/09 received on 11/05/09.

Ethical issues arising from the proposed study
You have indicated that are no significant ethical issues arise from this project.

Conditions of approval
By submitting an application to the Ethics Committee you confirm that you have read and understand the University of Dundee Guidelines for Ethical Practices in Research and the School of Computing Code of Practice for Research involving Human Participants and undertake to abide by these guidelines. Permission is therefore granted for you to proceed with the study. The committee however reserves the right to fully review the application at a later date.

Please inform the committee of any change in project methodology which may have ethical implications.

Best wishes for your research,

Annalu

Dr Annalu Waller CSci MBCS
Convenor: School of Computing Ethics Committee

Appendix 4 – Administrator questionnaire

Pocket Interview Administrator Evaluation

How many hours a week on average do you use a computer?

Less than 10 ☐
11-20 ☐
21-30 ☐
31-40 ☐
More than 40 ☐

Do you have any computer programming experience?

Yes ☐ No ☐

If so, how would you rate your level of computer programming experience? (1 – novice, 5 – expert)

1 2 3 4 5

Have you used any other configurable diary software?

Yes ☐ No ☐

If so, which one/s?

Overall how comfortable were you using the Pocket Interview system?
(1 – not comfortable, 5 – very comfortable)

1 2 3 4 5

If applicable, please rate your ability to complete the following tasks using Pocket Interview (not including any software bugs you might have encountered)

	(1 – unable, 5 – easily able)				
Install the Pocket Interview software on your desktop or laptop	1	2	3	4	5
Client Manager					
Install the Client Software onto a PDA	1	2	3	4	5
Data Manager					
Create a new database	1	2	3	4	5
Copy data from client	1	2	3	4	5
Enter data into database	1	2	3	4	5
Questionnaire Manager					
Create a new questionnaire	1	2	3	4	5
Edit a previously created questionnaire	1	2	3	4	5
Use the preview facility	1	2	3	4	5
Schedule Manager					
Create a new schedule	1	2	3	4	5
Edit a previous schedule	1	2	3	4	5

Did you view the help videos for creating questionnaires and schedules?

Yes ☐ No ☐

How helpful did you find the videos?
(1 – not helpful, 5 – very helpful)

1 2 3 4 5

Are there any features of Pocket Interview that you particularly like?

Which aspects were most beneficial for your study?

Is there anything you would change about the software?

Are there any features that you would like to see added?

Would you recommend the software to other interested users?

Yes ☐ No ☐

Thank you for completing this questionnaire

Appendix 5 – Non-functional requirements

Development

Client software shall run on a Windows-based Pocket PC

Administration software shall run on a Windows-based desktop PC/laptop

Software shall be developed with Visual Studio incorporating the .net framework and .net Compact Framework

Application scripts and data shall be stored in XML format that allows for storage of structured information that can be read by multiple applications and platforms

Usability

The administration software shall include Graphical User Interfaces

There shall be simple installation and set-up procedures

Documentation and help shall be provided in the form of step-by-step videos

System shall be developed and evaluated with continual user involvement and iterative design process

The system shall easily recover from errors and provide informative error messages

The system shall provide users with feedback at all steps

Users

Administration users - assumed to have some computer experience though not computer programming

Data providers - client software shall be easily usable with the minimal amount of training

Appendix 6 – Functional Requirements

Client Software Functional Requirements

A1.1	Onscreen keyboard
DESCRIPTION	An onscreen keyboard shall be provided when text entry is required
RATIONALE	Allows the user to enter text
NEED	Essential
PRIORITY	High
STABILITY	Unlikely to change

A1.2	Multipage
DESCRIPTION	Questionnaires should be viewed on multiple pages and will avoid the use of scollbars when possible
RATIONALE	Allows the user to see all relevant information at once
NEED	Desirable
PRIORITY	Low
STABILITY	Unlikely to change

A1.3	Next button
DESCRIPTION	Onscreen 'Next' buttons shall be present to navigate questionnaires
RATIONALE	Allows the user to proceed to next question
NEED	Essential
PRIORITY	High
STABILITY	Unlikely to change

A1.4	Back button
DESCRIPTION	An optional 'Back' button shall be present to navigate questionnaires
RATIONALE	Allows the user to review answers. Administrators may decide not to allow this so should be optional
NEED	Essential
PRIORITY	High
STABILITY	Unlikely to change

A1.5	Branching
DESCRIPTION	Shall allow for branching through the questionnaires
RATIONALE	The questions that are displayed and the path through the questionnaire may depend on previous answers
NEED	Essential
PRIORITY	High
STABILITY	Unlikely to change

A1.6	Multiple questionnaires
DESCRIPTION	The client device shall be able to present multiple questionnaires
RATIONALE	Different questionnaires may be required to be answered at different times
NEED	Essential
PRIORITY	High
STABILITY	Unlikely to change

A1.7	Kiosk mode
DESCRIPTION	The client shall allow for Kiosk mode
RATIONALE	Kiosk mode prevents the users from accessing other functionality of the device
NEED	Essential
PRIORITY	High
STABILITY	Unlikely to change

A1.8	Demo mode
DESCRIPTION	The client shall include a demo mode that allows display of the questionnaires but does not store data
RATIONALE	Can be used for demonstration and user-training
NEED	Essential
PRIORITY	High
STABILITY	Unlikely to change

A1.9	Prompts
DESCRIPTION	The client shall prompt the user using audio alarms to make entries. The volume, sound and length can be set by the administrator.
RATIONALE	Allows data to be collected in real-time
NEED	Essential
PRIORITY	High
STABILITY	Unlikely to change

A1.10	Incident reports
DESCRIPTION	The client shall store allow questionnaires to be self-initiated e.g. incident reports
RATIONALE	Allows users to make spontaneous entries
NEED	Essential
PRIORITY	High
STABILITY	Unlikely to change

A1.11	Missed entries
DESCRIPTION	The client will keep a record of missed entries
RATIONALE	Provides the administrator with data regarding prompts that have been issued and not responded to
NEED	Essential
PRIORITY	High
STABILITY	Unlikely to change

A1.12	Snooze
DESCRIPTION	The client shall allow the user to snooze prompts for up to one hour
RATIONALE	There may be situations where the user may not want to be disturbed by an audio alarm
NEED	Essential
PRIORITY	High
STABILITY	Unlikely to change

A1.13	System area
DESCRIPTION	The client shall include a password protected settings area that allows the administrator to change system settings
RATIONALE	Device settings should be not be editable by the participants
NEED	Essential
PRIORITY	High
STABILITY	Unlikely to change

A1.14	Security
DESCRIPTION	The option to encrypt data entries shall be included
RATIONALE	If the device is lost or stolen this will protect the participants privacy
NEED	Essential
PRIORITY	High
STABILITY	Unlikely to change

A1.15	Picture display
DESCRIPTION	The client shall be able to display pictures within the questionnaires
RATIONALE	Pictures can provide additional information and complement the questions
NEED	Desirable
PRIORITY	Low
STABILITY	Unlikely to change

Administration Software Functional Requirements

Client Manager

B1.1	**Install client software on device**
DESCRIPTION	The Client manager shall install the software on the client devices
RATIONALE	Allows the administrator to easily distribute the client software
NEED	Essential
PRIORITY	High
STABILITY	Unlikely to change

B1.2	**Install .net frameworks**
DESCRIPTION	The Client manager shall install the .net compact framework on the client devices
RATIONALE	The .net compact framework is essential for the client to operate
NEED	Essential
PRIORITY	High
STABILITY	Unlikely to change

Schedule manager

C1.1	**Create prompt schedules**
DESCRIPTION	The Schedule manager shall allow the administrator to create a schedule of prompts
RATIONALE	Allows the administrator to define client audio prompts
NEED	Essential
PRIORITY	High
STABILITY	Unlikely to change

C1.2	**Randomisation**
DESCRIPTION	The Schedule manager shall allow the prompts to be randomised
RATIONALE	Prevents predictability when prompting the data-provider
NEED	Essential
PRIORITY	High
STABILITY	Unlikely to change

C1.3	Select questionnaire
DESCRIPTION	The schedule manager shall allow the administrator to define which questionnaire should be answered after each prompt
RATIONALE	There may be multiple questionnaires that may be displayed at different times throughout the schedule
NEED	Essential
PRIORITY	High
STABILITY	Unlikely to change

C1.4	Add/Edit Individual prompts
DESCRIPTION	The Schedule manager shall allow the administrator to add or edit individual prompts within the schedule
RATIONALE	Allows the administrator more control with the definition of the schedule
NEED	Essential
PRIORITY	High
STABILITY	Unlikely to change

C1.5	Save/Load schedules
DESCRIPTION	The Schedule manager shall allow the administrator to save and load schedules to/from file
RATIONALE	Allows the administrator to save and load schedules
NEED	Essential
PRIORITY	High
STABILITY	Unlikely to change

C1.6	Copy to/from client
DESCRIPTION	The Schedule manager shall allow the administrator to copy a schedule to a client device and copy a schedule from a client device
RATIONALE	Shields the administrator from the copying process
NEED	Essential
PRIORITY	High
STABILITY	Unlikely to change

Data Manager

D1.1	Create database
DESCRIPTION	The Data manager shall allow the administrator to create an Access database based on a previously created questionnaire file
RATIONALE	An appropriate database is required before data entry can proceed
NEED	Essential
PRIORITY	High
STABILITY	Unlikely to change

D1.2	Copy data from client
DESCRIPTION	The Data manager shall allow the administrator to copy a clients data from the device to their desktop
RATIONALE	The data is copied from the device to the administration PC before it is entered into the database
NEED	Essential
PRIORITY	High
STABILITY	Unlikely to change

D1.3	Enter data
DESCRIPTION	The Data manager shall allow the administrator to enter copied data into a previously created database
RATIONALE	Allows administrators to enter data into a database where it can be easily viewed
NEED	Essential
PRIORITY	High
STABILITY	Unlikely to change

Settings manager requirements

E1.1	Edit client settings
DESCRIPTION	The Settings manager shall allow the administrator to edit the client settings
RATIONALE	An administrator may be dealing with multiple clients and avoids editing settings on many devices
NEED	Essential
PRIORITY	High
STABILITY	Unlikely to change

E1.2	Edit user messages
DESCRIPTION	The Settings manager shall allow the administrator to edit or translate the common client user messages
RATIONALE	Allows the client software to be used by non-english speaking data providers
NEED	Essential
PRIORITY	High
STABILITY	Unlikely to change

E1.3	Copy settings
DESCRIPTION	The Settings manager shall allow the administrator to copy settings to/from a client device
RATIONALE	An administrator may be dealing with a number of clients and avoids editing settings on multiple devices
NEED	Essential
PRIORITY	High
STABILITY	Unlikely to change

Questionnaire Manager

F1.1	Create new questionnaire
DESCRIPTION	The Questionnaire manager shall allow the administrator to create a new questionnaire
RATIONALE	Allows the administrator to define a questionnaire
NEED	Essential
PRIORITY	High
STABILITY	Unlikely to change

F1.2	Edit existing questionnaire
DESCRIPTION	The Questionnaire manager shall allow the administrator to create edit an existing questionnaire
RATIONALE	Allows the administrator to make changes to existing questionnaires
NEED	Essential
PRIORITY	High
STABILITY	Unlikely to change

F1.3	Save/Load questionnaires
DESCRIPTION	The Questionnaire manager shall allow questionnaire files to be saved/loaded from file
RATIONALE	Questionnaire are stored on the computers hard drive
NEED	Essential
PRIORITY	High
STABILITY	Unlikely to change

Question types

F1.4	Create new questionnaire
DESCRIPTION	The Questionnaire manager shall allow for a selection of different question types
RATIONALE	Allows the administrator to define a more adaptable questionnaire
NEED	Essential
PRIORITY	High
STABILITY	Unlikely to change

F1.4.1	Label
DESCRIPTION	The Questionnaire manager shall allow for ‘Labels’
RATIONALE	Text labels that require no user entry
NEED	Essential
PRIORITY	High
STABILITY	Unlikely to change

F1.4.2	Combo boxes
DESCRIPTION	The Questionnaire manager shall allow for ‘Combo boxes’
RATIONALE	Drop-down lists where the user can make one selection
NEED	Essential
PRIORITY	High
STABILITY	Unlikely to change

F1.4.3	Trackbars
DESCRIPTION	The Questionnaire manager shall allow for ‘Trackbars’
RATIONALE	Slider scales that can detect exactly where the data-provider clicks.
NEED	Essential
PRIORITY	High
STABILITY	Unlikely to change

F1.4.4	Check boxes
DESCRIPTION	The Questionnaire manager shall allow for ‘Check boxes’
RATIONALE	The user can select multiple answers from a selection of grouped choices
NEED	Essential
PRIORITY	High
STABILITY	Unlikely to change

F1.4.5	Radio buttons
DESCRIPTION	The Questionnaire manager shall allow for ‘Radio buttons’
RATIONALE	The user can only select one item from a selection of grouped choices
NEED	Essential
PRIORITY	High
STABILITY	Unlikely to change

F1.4.6	Likert scales
DESCRIPTION	The Questionnaire manager shall allow for ‘Likert scales
RATIONALE	A numbered scale commonly used in research
NEED	Essential
PRIORITY	High
STABILITY	Unlikely to change

F1.5	Required to answer
DESCRIPTION	The Questionnaire manager shall allow the administrator to label questions as ‘required to answer’
RATIONALE	The data-provider will be required to provide an answer before being allowed to continue
NEED	Essential
PRIORITY	High
STABILITY	Unlikely to change