THE UNIVERSITY *of* EDINBURGH

Usability Design of Short Message Service (SMS) Mobile Phone Banking

Gareth James Peevers

Ph.D Thesis

The University of Edinburgh

2010

Declaration of originality

2nd March 2010

This thesis is submitted for the Degree of Doctor of Philosophy. I declare that it has been composed by myself and that the work described was my own research.

Gareth James Peevers

Acknowledgements

I would like to thank my supervisor Professor Mervyn Jack and second supervisor Dr Gary Douglas for their advice and support. I would also like to thank Dr Fergus McInnes for kindly answering many questions on statistics and experimental design. I would also like to thank everyone in CCIR for all the help they have given me. Finally, I would like to thank Carol for telling me to just sit down and write.

Abstract

The financial services sector is investing considerable sums of money into mobile banking services, but the uptake by customers has been low. The cost to benefit ratio of mobile banking is highly unsatisfactory when the costs of developing and managing the channel are considered. Many of the advantages of Internet banking are shared by mobile banking e.g. control and time saving. Mobile banking also offers higher convenience with the ability to carry out banking whenever and wherever you are. It is hoped that mobile banking can be as successful as Internet banking. A major factor in the low adoption of mobile banking is usability, and there is a need for research on the issues surrounding mobile banking as so far little has been conducted. This thesis seeks to investigate the usability issues surrounding Short Message Service (SMS) banking. It identifies three general functions of SMS in electronic banking: transactions, communication/CRM and security. Three empirical usability evaluations are presented that explored customers' perceptions and attitudes of using these functions of SMS banking. The research presented here provides empirical evidence for the thesis that usability is a significant factor in the low customer adoption of SMS banking. It also shows that related to usability issues are customer concerns over the security of SMS as a banking channel. Older users will find SMS banking less usable than younger users and are more ambivalent regarding SMS in general. It recommends the most usable message input format to use in SMS banking and contributes insights on how best to realise the practical application of SMS banking and services. The findings from these studies will help improve usability in mobile banking services.

List of Publications

Peevers, G., Douglas, G., Jack, M. A., 2008. A usability comparison of three alternative message formats for an SMS banking service. International Journal of Human-Computer Studies 66, (2) 113-123.

Peevers, G. and McInnes, F., 2009. Laboratory studies. In: Love, S. (Ed.), Handbook of Mobile Technology Research Methods. Nova, Hauppauge NY.

Peevers, G., McInnes, F., Morton, H., Matthews, A., Jack, M. A., 2009. The mediating effects of brand music and waiting time updates on customers' satisfaction with a telephone service when put on-hold. International Journal of Bank Marketing 27, (3).

List of Figures

Figure 4.1: Sony Ericsson K700i

Figure 4.2: Customer Card for Abbreviations Service

Figure 4.3. Scatterplot of total task completion time (in seconds, vertical axis) for each message format against user's age (in years, horizontal axis), with regression lines

Figure 5.1: Example SMS Balance-only response format

Figure 5.2: Sony Ericsson K800i

Figure 5.3: Example of SMS Transaction Confirmation

Figure 6.1: The Sony Ericsson K-700 as a Hardware Authentication Token

Figure 6.2: An SMS Sends the Access Code to the Mobile

Figure 6.3: Instructions – Text Message Version, All Uses

Figure 6.4: Transaction Confirmation Screen with Password Entry

Figure 6.5: VASCO DIGIPASS 830 APACS Card Reader

Figure 6.6: Payment Confirmation Screen for APACS – First Use

Figure 6.7: The OTP Device as a Hardware Authentication Token

Figure 6.8: Instructions and Passcode Entry Field for the OTP Device – First Use

Figure 6.9: Mean Task Times for Authentication Methods at Tasks 1, 2 and 3

List of Tables

Table 4.1: Participant Demographics
Table 4.2: Participants' SMS Experience Levels
Table 4.3: Abbreviations based SMS Banking Service
Table 4.2: Numbers based SMS Banking Service
Table 4.5: Usability Questionnaire Items
Table 4.6: Mean Times (St. Dev.) in seconds by Task and SMS Version
Table 4.7: Completion Rates with Confidence Intervals for participants completing the tasks within three attempts
Table 4.8: Successful first time Completion Rates with Confidence Intervals
Table 4.9: Preference Rankings
Table 4.10: Participant Requirements for SMS Banking Services
Table 4.11: Participants Thoughts on the Abbreviations Used
Table 5.1: Participant Demographics
Table 5.2: Reasons for *not* using Mobile Phone Banking Services
Table 5.3: Participants' SMS Experience Levels
Table 5.4: Usability Questionnaire Items
Table 5.5: Completion Rates with Confidence Intervals for participants completing the tasks within three attempts
Table 5.6: Successful first time Completion Rates with Confidence Intervals
Table 5.7: Usability Scores for the Balance-on-demand versions
Table 5.8: Quality Ratings for the Balance-on-demand Designs
Table 5.9: Rank Order for Designs (General)
Table 5.10: Rank Order for SMS Variants vs. Automated service
Table 5.11: Rank Order for SMS Variants vs. Advisor service
Table 5.12: Potential Usage Frequency
Table 5.13: Circumstances where SMS may be used instead of Telephone
Table 5.14: Usage Intentions – Balance requests via service channels
Table 5.15: SMS Usage Intentions c.f. Advisor and Automated
Table 5.16: Participant Requirements for SMS-on-demand services
Table 5.17: Use/Cost Issues
Table 5.18: Usability questionnaire items

Table 5.19: Usability Scores for Funds Transfer Processes
Table 5.20: Quality Ratings for the Alternative Funds Transfer Processes
Table 5.21: Rank Order for Designs (Split between three experiences)
Table 5.22: Rank Order for Designs (Split between three experiences)
Table 5.23: Potential Frequency of Use per Week
Table 5.24: Circumstances where SMS may be used instead of Telephone Banking
Table 6.1: Participant Demographics
Table 6.2: Frequency of Use of Internet Banking
Table 6.3: Usability Questionnaire Items
Table 6.4: Mean times (S.D) in seconds by task and Authentication Method
Table 6.5: Pair-Wise Comparisons of Mean Task Times for Each Method
Table 6.6: Successful first time completion rates
Table 6.7: Pair-Wise Comparisons of Mean Usability Score for Each Method
Table 6.8: Individual Usability Scores and Overall Comparisons for each Method
Table 6.9: Pair-Wise Comparisons of Quality Ratings for Each Method
Table 6.10: Rank Order of Preference for the Alternative Authentication Methods
Table 6.11: Pair-Wise Comparisons of Security Ratings for Each Method
Table 6.12: Are Passwords Secure Enough to Confirm Transactions?
Table 6.13: Mandatory Device Usage to Confirm Transactions?
Table 6.14: Participant to choose Device?
Table 6.15: Device impact on IB usage Frequency?
Table 6.16: Previous Experience with the Three New Authentication Methods.

Abbreviations

CPU	Central Processing Unit
CRM	Customer Relationship Management
GSM	Global System for Mobile Communication
HCI	Human-Computer Interaction
HTML	Hyper Text Markup Language
IVR	Interactive Voice Response
J2ME	Java 2 Micro Edition
MIDP	Mobile Information Device Profile
OTP	One Time Passcode
SMS	Short Message Service
SMSC	Short Message Service Centre
UI	User Interface
W3C	World Wide Web Consortium
XHTML	Extensible Hyper Text Markup Language

Table of Contents

CHAPTER 1 ..... **1**

1. INTRODUCTION **2**

1.1. Thesis Outline 4
1.2. Contribution 5

CHAPTER 2 ..... **7**

2. BACKGROUND **8**

2.1. SMS and Banking 10
2.2. Functions of SMS Banking 12
2.3. Low Adoption of Mobile Banking 15
2.4. Usability 17

CHAPTER 3 ..... **25**

3. METHODOLOGY **26**

3.1. Usability Engineering Methods 26
3.2. Usability Experiments 28
3.3. Design of a Usability Experiment 29
3.4. Usability Metrics 31
3.5. Statistical Analysis 37
3.6. Data Collection and Quality Control 40
3.7. Advantages and Disadvantages of Experiments 41

CHAPTER 4 ..... **44**

4. A USABILITY COMPARISON OF THREE ALTERNATIVE MESSAGE FORMATS FOR AN SMS BANKING SERVICE **45**

4.1. Introduction 45
4.2. Method 48
4.3. Results 59
4.4. Discussion 74

CHAPTER 5 ..... **80**

5. EXPERIMENTAL ASSESSMENT OF CUSTOMER ATTITUDE TO THE INTEGRATION OF AN SMS CHANNEL INTO A MULTICHANNEL BANKING ENVIRONMENT AND AS A STRATEGY OF 'NEXT CALL AVOIDANCE' IN TELEPHONE BANKING **81**

5.1. Introduction 81
5.2. The SMS *Pull* Service Experiment 88
5.3. The SMS *Pull* Service Experiment: Results 97
5.4. Discussion 116
5.5. The SMS *Push* Confirmation Experiment 122
5.6. The SMS *Push* Confirmation Experiment: Results 129
5.7. Discussion 139

5.8. Discussion of Chapter 5 142

CHAPTER 6 146

6. THE USABILITY OF SMS AS A 2-FACTOR AUTHENTICATION PROCESS FOR INTERNET BANKING TRANSACTIONS 147

6.1. Introduction 147
6.2. Method 152
6.3. Results 167
6.4. Discussion 186

CHAPTER 7 195

7. CONCLUSIONS AND SUGGESTIONS FOR FURTHER WORK 196

7.1. Summary of Evidence 197
7.2. SMS Banking and Older Users 208
7.3. The Security Factor 210
7.4. Limitations 211
7.5. Further Work 212
7.6. Conclusions 213

BIBLIOGRAPHY 215

APPENDIX A: MATERIALS FROM CHAPTER 4 228

APPENDIX B: MATERIALS FROM CHAPTER 5 242

APPENDIX C: MATERIALS FROM CHAPTER 6 255

Chapter 1

1. Introduction

The thesis expounded here is that usability is a significant factor in the low customer adoption of SMS banking services. Related to these usability issues are customer concerns over the security of SMS as a banking channel. The spirit of this thesis is a contribution to the emerging area of mobile banking. The work contained here will offer advancement in this area by investigating the usability issues surrounding SMS banking. The thesis identifies three general functions of SMS in electronic banking: transactions, communication/CRM and security. Three empirical usability evaluations were carried out to explore customers' perceptions and attitudes of using these functions of SMS banking.

Internationally, banks are investing considerable sums of money into mobile banking services (Lee and Chung, 2009; Laukkanen, 2007; Luarn and Lin, 2004), but the uptake by customers has been disappointingly low (Laukkanen, 2007; Pousttchi and Schurig, 2004; Suoranta and Mattila, 2004). The cost to benefit ratio of mobile banking is highly unsatisfactory when the costs of developing and managing the channel are considered. Mobile banking applications have been available for some time but customers are just not using them. The presumed utility of mobile banking applications stems from the success of Internet banking, which has been the major success story in electronic banking and significantly changed the face of the retail banking sector. Internet banking gives customers control of their bank accounts and a major factor in its success is the 24/7, time saving convenience it affords. For banks the benefits of Internet banking are large manpower and cost savings. If Internet banking was a success then with the current improvements in mobile technologies

the next logical progression would seem to be mobile banking. Banks are continually searching for innovative services to offer customers as a way of differentiating themselves from their competitors (Lee and Kim, 2002). They are also quick to respond to new technology (Stamoulis et al., 2002). Many of the advantages of Internet banking are shared by mobile banking e.g. convenience and time saving. The most optimistic supporters of mobile banking claim it is cheaper, safer and more convenient compared with Internet banking (Lee and Chung, 2009; Laurn and Lin, 2004). Mobile banking appeals because the customer is able to access it on the move, regardless of time or place. A major factor in Internet banking success is also due to its ease of use (Hudson, 2002; Karagaluoto, 2002). A major factor in the low adoption of mobile banking applications is the usability problems inherent in these smaller devices. There is a need for research on the usability issues surrounding mobile banking (Laukkanen, 2007), as little research of this type has been carried out. The motivation for this thesis is to make a contribution to knowledge on the low adoption of mobile banking applications, specifically SMS banking. The thesis will advance reasons for the low adoption, and provide insights on how best to realise the practical application of SMS banking and services. The usability engineering experiments described here were performed in the context of strategic planning for the SMS banking channel of the Case Bank, one of the UK's leading High Street banks. The findings from these studies would improve usability in these mobile banking services.

1.1. Thesis Outline

Chapter two contains a full exploration of the research motivations behind this work in order to acknowledge why this research is needed and why it is an important contribution to knowledge. It reviews SMS and text messaging literature, the banking industry, electronic banking and the role of mobile banking in this industry. Finally it contains a review of usability engineering and discusses usability in relation to handheld devices and mobile banking.

Chapter three describes the usability engineering methodology, experiment design and analysis. It presents the usability methods employed in this thesis.

The next three chapters cover the empirical evaluations carried out to investigate the usability issues regarding the three functions of SMS banking: transactions, communication/CRM and security.

Chapter four presents the SMS format experiment comparing three possible message input formats that could be used in SMS banking: abbreviations, numbers and free-form. The participants were asked to complete three transactions using each format. The chapter will discuss the design, methodology and results of the experiment. The findings from this experiment will serve to define which message format should be used in SMS banking, and in the subsequent experiment described in chapter five.

Chapter five presents a two part experiment evaluating integrating an SMS channel into a banking multichannel environment as a strategy for 'next call avoidance.' The

chapter will discuss the design, methodology and results of the two part experiment. Part one, the SMS *Push* service experiment, compares an SMS channel using the abbreviations message format against a telephone banking channel for balance request transactions. The details of this experiment design are based on the findings from chapter four. The second part, the SMS *Pull* service experiment, evaluates participants' attitudes to the communication/CRM function of SMS banking. It compares a telephone service with SMS confirmation after a funds transfer against a telephone service without an SMS confirmation. It delivers qualitative data on the types of SMS CRM customers would want provided by the bank

Chapter six presents the SMS security experiment. SMS as a one time banking passcode (OTP) experiment. It evaluates a specific security function of SMS in banking, which is authentication of funds transfers in Internet banking. The SMS authentication method was compared against the existing password method, the OTP device method and the card reader method. The chapter will discuss the design, methodology and results of the experiment.

Chapter seven details and discusses the main findings and contribution that this research provides and highlights future work in this area.

1.2. Contribution

In the UK there is a large financial services industry and many banks have, or are considering, introducing mobile banking services in the future. The contribution to knowledge this thesis will make is on understanding how usability impacts the low

adoption of mobile banking applications, specifically SMS banking. It will contribute reasons for the low adoption and insights on how best to realise the practical application of SMS banking. It will show that the security of SMS as a channel for banking is a major concern for customers and an indicator of low usage intention in the future. It will also contribute knowledge to the body of research on text messaging in society and the debate revolving around the low usage of SMS and mobile phones by older users.

Chapter 2

2. Background

The continual widespread advances in computer technologies have encouraged many businesses to adopt new methods of interacting with customers to improve service, to lower costs and to maintain competitive advantage. The banking sector is continuously searching for ways to use technology for these purposes, and for the customer, create more convenient methods of banking. Banking has altered significantly since the time when it was conducted in a customer's local branch with the advent of electronic banking. With electronic banking a business can offer customers access to banking services via multiple electronic channels. Electronic banking is seen as one of the most successful areas of electronic commerce (Prousttchi and Schurig, 2004).

Electronic banking can be traced back to the introduction of Automated Teller Machines (ATMs) (BBC News, 2007a), telephone banking, and the introduction of 'home banking', or online banking, in the first part of the 1980s (Cronin, 1997), with the first subscriber dial-up services on the primitive home computers of the time. Now there is the introduction of mobile banking. There has been research on user attitudes to ATMs (Rugimbana and Iversen, 1994; Davies et al., 1996), telephone banking (Lockett and Littler, 1997; Peevers et al., 2009b) and in recent years much work on Internet banking (Polatoglu and Ekin, 2001; Sathye, 1999; Shaw et al., 1997; Weir et al., 2006; Weir et al., 2007). The major success story in this area has been to Internet banking, which significantly changed the face of the retail banking sector. Now research has started on issues surrounding mobile banking (Laukkanen and Lauronen, 2005; Lee and Chung, 2009; Kim, Shin and Lee, 2008; Mallet, Rossi

and Tuunaninen, 2004; Pousttchi and Schurig, 2004), but there has been little work on the usability of mobile banking services (Peevers, Douglas and Jack, 2008).

The success of Internet banking has been attributed to its convenience and compatibility with the modern lifestyle (Black et al., 2002; Gerrard and Cunningham, 2003; Lichtenstien and Williams, 2006; Centeno, 2004), time saving attributes and low fees (Karajaluoto et al., 2002) and also because it gives customers control (Shih and Fang, 2004). A very important driver of the adoption of Internet banking has been ease of use (Hudson, 2002, Karjaluoto et al., 2002). The main driver and rewards for banks from Internet banking, and electronic banking in general, is cost savings (Jayawardhena and Foley, 2000).

With the introduction of third-generation (3G) services, improved wireless abilities, increased memory and central processing unit (CPU) speeds in handheld devices, a seemingly logical progression from desktop Internet banking is mobile banking or mBanking. Mobile banking services allow customers to carry out banking transactions such as balance checks, ordering cheque books and funds transfers using a mobile device. Mobile banking is a general term that can be applied to a number of methods for enabling customers to use their mobile device to perform banking actions, these are: the Mobile Internet, SMS banking and downloadable applications such as the Monilink[1] solution offered by the Monitise group.

[1] Monilink <http://www.monilink.co.uk/>

Mobile phones offer banks enormous potential as a service channel because of their ubiquity. Mobile banking can help banks to retain existing technology-savvy customers with value-added, innovative services and attract new customers from corresponding sections of society (Tiwari and Buse, 2007; Tiwari, Buse and Herstatt, 2007). Mobile banking offers customers many of the advantages of desktop Internet banking but with context-specific services, in time critical situations, with spontaneous decisions and needs, and offers an efficiency increase (Tiwari and Buse, 2007).

2.1. SMS and Banking

This thesis is concerned with one application of mobile banking: Short Message Service (SMS) banking. It could be argued the main long term application of mobile banking will be in the form of the mobile Internet, and the browsing of Internet banking sites on mobile phones, being made possible by an improvement in mobile technologies with browser support for HTML and XHTML, the use of technologies such as Opera's Small Screen Rendering (SSR), the increasing popularity of Cascading Style Sheets (CSS), along with standards such as the W3Cs Mobile Web Best Practices 1.0. This argument would though, overlook the huge popularity of SMS and its advantages such as a flat rate charge, and the fact that customers will not need an expensive smartphone, iPhone or Personal Digital Assistants (PDA) style handset.

Short Message Service (SMS), or text messaging, as it is commonly known, is still a tremendous growth area in mobile communications. It is estimated (Martin, 2010)

that worldwide 4.1trillion (UK) text messages were sent in 2008. The Mobile Data Association (MDA) reports that in 2008 a total of 78.9 billion text messages were sent in the UK alone (MDA, 2009), 216 million per day, and this was up 22 billion on the annual total in 2007. Research has found that text messaging is most commonly used as an effective one-to-one method of communication between friends (Sillence and Baber, 2004), but businesses have also realised that there is huge potential in SMS for carrying out business activities, and for individual communication with customers. It was estimated by market research group Radicati that in 2004, 55% of text messaging was for business use, with much further growth to come (Faulkner and Culwin, 2005). SMS banking services have already been successfully implemented by banks in Asia, the Middle East and South Africa, with both *Push* (automatic) and *Pull* (customer initiated) services offered to customers (Rumpa, 2005). At the time of writing the services offered by banks in the UK are limited to *Push* only e.g. the bank sends the customer a weekly account balance, and basic *Pull* services e.g. ordering a new cheque book. The popularity of SMS banking in markets such as India (Ahmed, 2004; BBC News, 2005a) is due in part to the low cost of mobile handsets compared to desktop computers. This may also be true for some socio-economic groups in the UK.

The Short Message Service (SMS), a Global System for Mobile Communications (GSM) service, allows the user to send text messages up to 140 bytes. The transmission of a message is carried out by the network operator's Short Message Service Centre (SMSC), which receives the message and routes it to the destination device. A bank can run its own SMSC and in this way generate SMS messages from

its own data on its customers' accounts. A weakness with SMS banking is that the messages are not automatically encrypted when they are transmitted (Pousttchi and Schurig, 2004). Encryption is possible though, and various software products have been developed for this purpose which would allow businesses to carry out more complicated financial transactions via SMS banking.

There is much current research on handheld devices carried out in the ubiquitous and mobile computing fields. Examples of recent work specifically on text messaging in the UK are studies by Faulkner and Culwin (2005) and Kurniawan et al. (2006b), Outside the UK there has been considerable research on SMS in society (Kasesniemi and Rautiainen, 2002; Ling, 2000, 2004, 2007). One major finding of this research is that SMS seems to be a medium favoured by the young. Related to this is the rise of 'textish' or 'text-speak', which is a form of abbreviations and has been defined as English with the vowels removed (Faulkner and Culwin, 2005). Textish has had much interest focussed on it recently in the UK media. An example of this is the widespread reporting of a 13 year old pupil who wrote an examination essay in textish (Carrington, 2005; Ward, 2004). Unsurprisingly, the use of textish is most prominent among young people (Faulkner and Culwin, 2005). Other work relevant to this thesis includes research on text entry and mobile phone user interfaces (Curren at al., 2006; James and Reischel, 2001; Lee et al., 2003).

2.2. Functions of SMS Banking

It is proposed in this thesis that the types of services a bank can offer under the umbrella of SMS banking can be divided into three general functions: transactions,

communication/CRM and security. There will be some overlap between these three, but banks could use SMS for each of these purposes separately, or in combination.

2.2.1. Transactions

Ordering a new cheque book or PIN number, requesting a mini statement, transferring money or making a payment, these are all types of banking transactions that could, and are, offered by an SMS service. Examples of such services offering SMS payments are the 'Mobile Wallet' service from T-Mobile and 'm-pay' from Vodafone[2]. Mobile payments are a form of payment combining elements from other methods of payments such as credit/debit cards, prepaid cards, telephone bills and premium SMS messages (SMS messages that cost a fixed, predetermined amount). The viability of mobile payments has been generally proved to be acceptable to consumers (Khodawindi et al., 2003). In one study, over 80% of participants were willing to accept mobile payment (Khodawindi et al., 2003), with 96% giving "privacy of personal data", and 93% giving "simplicity of the method" as reasons for their decision. SMS payment schemes are currently in development by Anam and TR2[3], and one is already used by the PayPal[4] service. The Anam scheme uses SMS text messages to make third party payments (The Register, 2007), overcoming the necessity for the customer to download software, such as the Monolink solution, to their mobile phone, and also takes advantage of the widespread usage and familiarity with text messaging. In the UK for example, Colchester Borough Council[5] has set up a scheme whereby residents can pay their council tax by sending an SMS text

[2] Vodafone and T-Mobile launch mobile wallet: <http://www.out-law.com/page-2448>
[3] TR2: <http://www.tr2.ie>
[4] PayPal: <http://www.paypal.com/mobile>
[5] Colchester Borough Council: <http://www.colchester.gov.uk>

message. These types of services would generally be *Pull* (customer initiated), but will also be *Push* (automatic) for confirmation.

2.2.2. Communication/CRM

SMS can be used as a one-to-one business to customer communication channel and offers massive potential for customer relationship management (CRM). SMS can be used for marketing of a bank's services and products, confirmation of transactions made by the customer with the bank via another channel (e.g. Internet, telephone banking), confirmation of contact with the bank via another channel, confirmation of appointments, complaints etc. Dealing with customer complaints is an important issue for businesses (Fornell and Wernerfelt, 1997; Johnston and Mehra, 2002) and for preventing customer switching behaviour. Banks need effective channels and procedures for resolving complaints, as this could lead to a customer switching to another bank. An SMS banking service could offer a useful channel for this purpose and there has been some research in this area (Richardson, 2005). There has been little research on mobile CRM so far, but one study (Lijander, Polsa and Forsberg, 2007) focussed on an airline using mobile CRM. They argue that customers are not ready for this type of mobile application yet, though they did find that participants who already used the mobile Internet had a more positive attitude. These types of service will generally be of the *Push* type.

2.2.3. Security

SMS can be used as method of adding 2-factor authentication to online transactions, and potentially to telephone transactions. SMS can be used to generate one time passcodes (OTP). An OTP is a password (usually a string of digits) that is valid only for a single online session or transaction that is made available to the customer either by a physical hardware device with a small display that the customer carries; by a Chip and PIN card reader device; or by using an "out-of-band" channel like SMS. To authenticate their transaction a customer must type in the OTP rather than a static password. SMS OTP generation has been implemented in Asia (Yeo, 2006) and mobile phone based 2-factor authentication has been researched (Aloul, Zahidi and El-Hajj, 2009) and proven to work, but there has been little usability research (Weir et al., 2009b) on using SMS as an authentication method.

2.3. Low Adoption of Mobile Banking

The advantages of mobile banking appear as convenience, access to banking no matter the location or time, and efficiency (Laukkanen, 2007; Jarvenpaa et al., 2003; Suoranta 2003; Tiwari and Buse, 2007). In spite of these advantages some authors (Lee and Chung, 2009; Pousttchi and Schurig, 2004) have remarked on the slow development of mobile banking. In countries such as Korea, Finland, and Taiwan studies have shown the usage levels of mobile banking are small (Laukkanen, 2007; Lee and Chung, 2009; Luarn and Lin, 2005) compared to what would be hoped for. What are the factors preventing large scale adoption of mobile banking? Research

has shown that customers worry about how much it will cost (Luarn and Lin, 2005), along with the security of the service (Brown et al., 2003; Luarn and Lin, 2005). Though some (Laukkanen, 2007; Laukkanen and Lauronen, 2005; Suoranta, 2003) have argued that security concerns are not a prohibitive issue. The perceived complexity of mobile banking is also argued to be a cause of low usage levels (Lee at al. 2003). Recent research (Lee and Chung, 2009) has argued that trust is one of the most important factors in the low adoption of mobile banking, and is the factor that most impacts on customer satisfaction with this banking channel. Trust has an impact on level of adoption in all forms of electronic banking (Aladwani, 2001), and has been researched extensively (Grabner-Krauter and Kaluscha, 2003; Kim and Prabhakar, 2000; Kim and Moon, 1998; Suh and Han, 2002).

In their survey study on mobile banking in Taiwan, Lee and Chung (2009) used a research model based on Delone and Mclean's (1992) information system (IS) success model with the three quality measures of system quality, information quality and interface design quality. They found that system quality and information quality affected trust more than user satisfaction, and argue that these factors are important in building trust in a mobile banking channel. System quality is defined as the quality manifested in the system's overall performance (Delone and McLean, 1992) as measured by a customers' perception. Information quality, including accuracy, is obviously of major importance to any electronic banking channel and will influence customer satisfaction (Kim et al., 2008). Lee and Chung argue that interface design quality may be an important factor in building trust, but it is not as important as system quality and information quality. For a bank offering an SMS banking

channel, interface design is something that can not be controlled, as it is dependent on the type of mobile phone the customer owns. With many of the studies described in this chapter there are still questions over how generalizable the findings are, because they are specific to individual countries and cultural factors may play a part.

Another major factor in the slow adoption of mobile banking is due to the limitations of mobile devices: tiny screen size, small keypads, reception and network problems and slow connection speeds. Writing a text message is not the easiest thing to do due to the variety of methods of text entry available on mobile devices, and the lack of a standard user interface, or even a standard layout of the keypad. The usability of mobile devices is an important factor in the low adoption of mobile banking (Laukkanen, 2007). The distinct lack of empirical research evaluating the usability issues surrounding implementing SMS services for either transactions, communication or security in banking leads to the work in this thesis (Peevers and McInnes, 2009; Peevers et al., 2008).

2.4. Usability

The history of usability as a concept can be traced back to the early 1980s with the papers published by Shackel (1981) and Bennett (1984). There are a number of definitions as to what usability is, with the first attempted by Miller (1971) and based on the concept of *ease of use* (Shackel, 1991), but the most often quoted is the one defined by the ISO as "the efficiency, effectiveness and satisfaction with which specified users can achieve specified goals in a particular environment" (ISO 9241-11, 1998). As the ISO definition suggests, usability is a multi-dimensional concept.

There are often compromises to be made, trading off different goals to achieve a usable product. An alternative definition from the Usability Professionals Association (UPA) states: “Usability is an approach to product development that incorporates direct user feedback throughout the development cycle in order to reduce costs and create products and tools that meet a users needs.” The first definition gives measurable dimensions with which to evaluate how usable a system/product is, and the second definition suggests a process of how to do this when developing a new system/product, along with the benefits of the approach.

Efficiency, effectiveness and satisfaction are independent qualities of the system (Frokjaer et al., 2000). Efficiency is concerned with the amount of effort required in usage. It is typically measured as the time taken or the number of clicks to complete a task. Effectiveness is indicative of application robustness and transparency; task completion, accuracy, prevention and easy recovery from errors are typical measures. Satisfaction relates to the degree to which users react positively to their experience whilst completing tasks. There is still not total agreement on what user satisfaction is (Lindgaard, 2009), but it can be proposed that it is related to measuring user attitude (Betsch et al., 2001). It may also concern perceived usefulness, attractiveness and other emotional responses to the system. The ISO definition concentrates on the attributes of efficiency, effectiveness and satisfaction, but there are other components of usability not included in this definition such as learnability and memorability (Nielson, 1993; Preece, Rodgers, Sharp, 2002). How easy the system is to learn, and how easy is it to remember how to use it again, are very important factor for a casual user. There is also the question of what makes a product useful and successful

(Kuniavsky, 2003) to a customer or user. All of these definitions suggest that usability has both subjective and objective components that can be measured.

The other usability standard is Human-Centred Design Processes for Interactive Systems (BS-EN ISO 13407: 1999). This standard is guidance for anyone who wants to follow a user-centred design process. The standard describes four principle activities that should be followed that will lead to a design that is of high usability. That is, an effective, efficient and satisfying design. The four activities are:

- Understand and specify the context of use
- Specify the user and socio-cultural requirements
- Produce design solutions
- Evaluate designs against requirements

The standard can be adapted and applied to the design of any product, and the level of effort and the sequence of the activities can vary depending on the type of product being designed.

Recently the there has been a trend towards the term ‘user experience’, and some practitioners distinguish between usability and user experience (Tullis and Albert, 2008). Those in agreement with this opinion consider that the concept of usability is too limited, or narrow, to explain the choices people make, and that a user’s whole experience with a product should be considered e.g. their thoughts, feelings, perceptions and even their interaction with the company who make/sell the product.

2.4.1. Usability Engineering

The process by which a usable product or computer system is achieved is called usability engineering. The goal of usability engineering is to engineer a quality product that does what it is meant to do, and fulfils a customer's actual needs. It does this by considering the user and following rigorous software engineering methods (Faulkner, 2000). The usability engineering process is well established and can apply to all products with a user interface (Nielsen, 1993). It follows a cyclic process of design and evaluation followed by redesign and evaluation. The methods used in usability engineering have developed from the fields of ergonomics, human factors and human-computer interaction (HCI), and also use more formal experimental methods favoured by psychologists (Preece et al., 1994).

Usability engineering can mean different things to different practitioners. At its simplest, usability engineering is the process of applying usability metrics (Whiteside, Bennett and Holtzblatt, 1988, cited in Faulkner, 2000). To others, usability engineering is not just about evaluation, but is involved in the whole development process from the very beginning to the release of the product (Culwin and Faulkner, 1997; Faulkner, 2000). For the purposes of this thesis, usability engineering is to be defined by the approach suggested by Preece et al. (2002), which involves using the documented usability results and feedback from early versions of a system to make changes to subsequent versions. As Nielsen (1993) states, usability engineering is a "discipline aimed at enhancing the usability of products."

A typical usability engineering lifecycle (Mayhew, 1999) consists of three basic stages: requirements analysis, design/testing/development and Installation (Preece et al., 2002). The lifecycle is a useful tool for integrating usability into software engineering practises. Each of these three stages has many subtasks within, with the middle set of tasks requiring the most attention. Not all projects will require the same level of complexity of lifecycle, and the usability methods employed will be dependent on the timeframe, resources and goals related to that specific project (Preece et al., 2002).

2.4.2. SMS and Mobile Banking Usability

Handheld devices offer many challenges for the user. Physically, they are by their very nature, small: they are typically meant to be held in one hand. They have tiny screens, and have small, fiddly keypads and challenging user input methods. They have less memory, CPU power and their connectivity can be slow and unreliable (though this is improving every year). But compared to desktop computers, or even laptops and notebooks, they offer great portability for on the move access to information, and their great power is communication.

Weiss (2002) proposed three tests which a device must pass to be considered a handheld device:

- General operation without cables (except for charging, connecting to desktop).

- Easily used while being held in one's hands.
- Addition of applications, or support Internet connectivity.

The popularity of SMS has led to a body of usability research on text entry methods (Curren et al., 2006; Friedman et al., 2001; James and Reischel, 2001), mobile phones (Balakrishnan et al., 2005; Soriano et al., 2005) and there has been general research on mobile phone user interfaces (Lee et al., 2006). One explanation for the usability problems encountered by users of mobile devices is a lack of extensive usability evaluations, due to the manufactures' rush to get their products out into the current competitive market place (Lee et al., 2006). The tiny key pads on mobile phones have been found to pose usability problems (Kurniawan et al., 2006; Soriano et al., 2005). Thumb size has been shown to cause usability problems with texting (Balakrishnan and Yeow, 2008). It has been found that older user have usability problems with texting on mobile phones (Kurniawan, 2008; Peevers et al., 2008), and there has been research into producing mobile phones aimed at the older generation (BBC News, 2005b). One study (Ornella and Stephanie, 2006) showed that when keys are placed to close together they cause problems for older users. Older user have been found to be passive users of mobile phones, and can find the process of text messaging intimidating (Kurniawan, 2008).

Much of the previous research in mobile banking has been based on surveys. Usability studies emphasise hands on usage and the collection of performance and qualitative data, and there are few studies (Hyvarinen, Kaikkonen and Hiltunen, 2005, Peevers et al., 2008) observing actual user performance with mobile banking

services, and SMS banking in particular. It is argued (Laukkanen, 2007) that to bring more understanding to the low adoption of mobile banking the usability issues surrounding this channel need to be researched. Usability, along with functionality, both influence real world usage (Whiteside et al., 1988).

It is important that any new product or service is based on user needs and requirements rather than being driven by technology. This is an important issue to consider with SMS and mobile banking. Research related to the usability design issues surrounding online banking are relevant to SMS banking. Customers want electronic banking designs that are secure (Furnell, 2005) and have good error prevention functionality (Liao and Chung, 2002). They also want user-friendliness, speed, accuracy and control (Jayawardhena and Foley, 2000). If a new banking channel such as SMS banking is to succeed it must also be satisfying to use. This is extremely important as the new channel will be in competition with the other channels offered by the bank, e.g. Internet banking, telephone banking, ATMs and branch based banking. It will be important to compare the new channel of SMS banking to existing channels as a bench mark and to discover any usability issues. Very little (Laukkanen, 2007) of this kind of comparison work has been carried out with mobile banking. Many of the existing SMS banking services use abbreviations. An example would be that to request an account balance the users sends an abbreviation as ACBAL, but is abbreviations the most usable format? Another issue to consider will be the demographic variables in the customer base, such as age and gender, and how this will affect the accessibility of an SMS banking service. It has already been discussed how text messaging is more popular amongst the young, and

the older generation find mobile phones and texting more difficult. It will be important to decide if the marketing of an SMS banking channel should be focussed on a specific group of users and how this will affect its cost/benefit ratio to the service provider. These issues will be addressed in this thesis, with the main line of research being concerned with the relationship between usability and the low adoption of SMS banking.

Chapter 3

3. Methodology

This chapter will detail the methodology used in the experiment evaluations described in this thesis.

3.1. Usability Engineering Methods

The usability engineering lifecycle model uses a number of methods that are applied at an appropriate time during the lifecycle. Examples of the established methods are listed below:

- Requirements analysis: task analysis, requirements capture and user profiling with scenarios. One of the basic tenets of usability engineering is "Know the user, know the task" (Faulkner, 2000).
- Apply guidelines and heuristics: expert usability practitioners evaluate the design using heuristics and guidelines such as those suggested by Nielsen, (1993).
- Formative study: prototyping and re-test using discount (Nielsen, 1993) style usability testing with a small group of users (5) trying the design in an iterative process. Problems can be fixed as they are found.
- Summative study: experimental evaluation with the collecting of usability metrics. To evaluate the design against the usability goals using more formal methods of experimental design, measurement and statistical analysis, as applied in experimental psychology

3.1.1. Formative verse Summative methods

Formative (also referred to as informal or discount) usability methods are now very popular in the usability community to the degree that it sometimes seems one or the other must be used, rather than using each one when appropriate, with formative methods being the preferred choice for a number of reasons. The reasons are that formative usability is cheaper, faster and the testing is less complicated to run. The major proponent of formative discount usability is Nielsen (1993), who acknowledges that there are weaknesses in formative testing, but practically, it is more likely to be carried out compared to summative testing. The real strength of formative usability is its iterative nature. It should be used to make improvements in the design by following the sequence of testing and finding problems, making recommendations, and testing again. Formative methods are a powerful technique and can find the most significant usability issues in a design iteration (Tullis and Albert, 2008) to improve the next iteration of the design. As the usability engineering lifecycle suggests (Faulkner, 2000), formative methods are very appropriate to use at the beginning of the design process.

Summative usability methods, unlike formative, can evaluate how well a product meets its objectives (Faulkner, 2000; Tullis and Albert, 2008), and they can answer the question of how usable a product is. Summative methods also produce scientific data which allows practitioners to test hypotheses, and so are more popular in the academic community. Summative methods are very appropriate to use at the end of the usability engineering lifecycle to experimentally evaluate the final product, but they are also the most useful way to compare new designs of a system to existing

ones, alternative designs providing the same functionality, or to compare a design against a competitors design (Preeces et al, 2002; Tullis and Albert, 2008). Summative methods do have their limitations (Weir, 2008), as they are best used in controlled environments where the metrics being measured are task-based and specific.

Each design project will have its own restraints on time, budget, complexity and specific requirements, so it is important to select the most appropriate usability engineering methods.

3.2. Usability Experiments

The summative methods used by usability engineering and employed in this thesis are based on the experiment design techniques used in experimental psychology. Controlled experiments are widely used in the field of human-computer interaction (HCI) to evaluate interface designs (McGuffin and Balakrishnan, 2005) and in usability engineering studies (Peevers et al., 2008; Weir et al., 2006; Weir et al., 2007) to compare between alternative designs. In a usability study there might be two alternative designs, A and B. The purpose of the study is to find which one has the highest usability. The experimenter will form hypotheses, or predictions, about the two designs. A and B are the experimental treatments, also referred to as the independent variables. Relevant tasks will then be designed for the user based on task analysis. The study will answer if changing the interface (A or B) will have an effect on the dependent variables (usability metrics), which could be completion rates, times on task and responses to attitude questionnaires. This will be answered

by performing statistical analysis on the usability metric data recorded during the controlled experiment. The aim of the experiment will be to refute the null hypothesis which is the proposition that manipulating the independent variable (changing the design: A or B) will not in fact impact on the measured dependent variables. The method used to design an experiment will be described in the next section.

3.3. Design of a Usability Experiment

3.3.1. Participants

It is very important that participants are appropriate to the goals of the study. For instance, if it is testing the usability of a new mobile phone aimed at older users then it is important to recruit participants from this target audience, e.g. older users. Whereas if the purpose is to test the usability of an online supermarket then it is advisable to recruit participants from a wider target audience of different ages and backgrounds. In reality, the type of participants that can be recruited is affected by such factors as time and cost restraints. In this case it is important to be aware of the limitations of the data generated and the meaningfulness of the results (Cairns and Cox, 2008). The number of participants needed depends upon the amount of segmentation required in the population. For robust statistical testing, larger numbers are needed in each key demographic group, i.e. of differing ages and genders to lessen the impact of individual differences (Faulkner, 2003; Landauer, 1988) .

3.3.2. Types of design

There are two types of design that are commonly used in usability experiments, they are designs that use different groups of participants (between-participants), or the same group of participants (within-participants). To compare two interfaces, A and B, the experiment could be between-participants, one group of users would use A and one group of users would use group B. Or it could be within-participants, the same group of participants will use both A and B in a randomised order, which is called a repeated measures design. The order the participants experience the interfaces must be randomised to minimise the effects of confounding variables, such as novelty and learning (Cairns and Cox, 2008), that may affect the usability metrics measured. If this type of extraneous variable can not be controlled for, then they should be identified in the planning stage of the experiment. One way to deal with confounding variables is to randomise them.

Each type of experiment design has its own strengths and weaknesses, but when participants are required to compare two interfaces then a within-participants study is the most appropriate (Cairns and Cox, 2008). An advantage of this type of design is that it requires fewer participants, because by using the same participants the influence of individual differences will be diminished. A between-participants design with different groups of users will have more variation to contend with so a larger cohort of participants will need to be recruited. The disadvantages of the within-participants design are possible carry over effects. If an experiment compares three interfaces a participant will possibly be required to complete the same set of tasks three times. This could be very repetitive and lead to carry over effects such as

fatigue and improving performance that could effect the validity of the experiments results. The way the experimenter can control for these effects is through counterbalancing. A experiment design can be counterbalanced by randomising the order the tasks are completed with each interface. Sometimes this is not possible to do as a set of tasks may lead on from each other in sequence.

A within-participants design can be balanced for individual groups within the participant cohort, for example by age, by gender or by experience (expert vs. novice) if they are considered relevant to the study. In this way the groups are treated as independent variables and can be compared in the statistical analysis to discover if one group finds an interface more difficult to use than another. The technique to enable this kind of comparison is to ensure that each group experiences all possible orders of treatments equal amounts of time. The downside is that with more variables included more participants will be required to balance the groups.

3.4. Usability Metrics

In usability experiments tasks will be carefully designed for the participants to attempt to complete using the interface or product. The experimenter has a choice of usability metrics, the dependent variables, which they can use to measure the participants' performance when completing the tasks. The data collected by these metrics can then be statistically analysed to judge how usable the product is. There are various usability metrics, such as performance metrics, self reported metrics and behavioural metrics. The usability metrics used in this thesis will be described below.

3.4.1. Performance Metrics

Performance metrics are the best way evaluate effectiveness and efficiency (Tullis and Albert, 2008). As discussed previously, these two dimensions of evaluation are cornerstones of the ISO definition of usability. These types of metrics produce an objective measure of usability.

Effectiveness is indicative of application robustness and transparency; task completion, accuracy, prevention and easy recovery from errors are typical measures. The main metric used for effectiveness throughout this thesis is task success. This is the most straightforward way of judging effectiveness in a usability study: can the participant complete the task? Binary success is the simplest way to measure task completion: 1 for success and 0 for failure. Once the completion rates have been collected confidence intervals can be calculated, which are essential because they reflect how confident the experimenter is in the data (Sauro and Lewis, 2005; Tullis and Albert, 2008). Confidence intervals can be calculated using the Adjusted Wald method (Sauro and Lewis, 2005; Lewis and Sauro, 2006). Confidence intervals are more accurate the larger the participant cohort. When using task success it is critical that the criteria for a successfully completed task is defined before the experiment starts. It is also critical to decide on how many attempts participants will be allowed to complete a task, and how this will affect the way the data from this metric is analysed and reported.

Efficiency is concerned with the amount of effort required in usage. It is typically measured as the time taken or the number of clicks to complete a task. In this thesis

efficiency will be measured by the time-on-task metric. Time-on-task can be measured using a stop watch, but it is more accurate to log the time using software and this is the approached used throughout the thesis. There are some instances where time-on-task is not appropriate to use as a usability metric such as when you are also using a think aloud methods. Also, some argue that timing is not appropriate in studies where the participants use a number of different interfaces for the first time (Weir et al., 2005; Ziefle, 2002), as this data will be subject to a learning curve. This effect though, could be useful for comparing how quickly users learn how to use an interface efficiently. Where appropriate, the number of attempts the participant took to complete a task was also noted by researcher and this too can be used as a pointer to efficiency.

3.4.2. Self Reported Metrics

Self reported metrics are the best way of evaluating the user satisfaction dimension of the ISO usability definition. The most efficient way of collecting self reported data is by using attitude questionnaires and rating scales (Tullis and Albert, 2008).

Usability Attitude Questionnaire

The design of the usability questionnaire used here followed standard practice (Likert, 1932) by using an equal number of negative and positive statements presented in a randomised order. The questionnaire used a 7-point Likert format that ranged from “Strongly Agree” (1) to “Strongly Disagree” (7). The use of questionnaires to evaluate user interfaces has a long history (LaLomia and Sidowski,

1990; Root and Draper, 1983). A number of usability questionnaires have been developed by research groups (Brooke, 1996; Chin, Diehl and Norman, 1998; Ryu and Smith-Jackson, 2006). A comparison of some common usability questionnaires (Tullis and Stetson, 2004) has shown them to give reliable results. The attitude questionnaire applied in the work in this thesis has been widely used and validated (Dutton et al., 1993; Love, 1997; Love et al., 1992; Love et al., 1994; Jack et al., 1993) in evaluating the usability of automated telephone systems. A core set of usability attributes were developed and the questionnaire has been adapted for use in evaluating Internet banking (Weir et al., 2006, Weir et al., 2007), online security (Weir et al, 2009a; Weir et al, 2009b), customer satisfaction with advisor telephone banking (Peevers et al., 2009a) and mobile banking (Peevers et al., 2008). An overall mean usability satisfaction score can be calculated for each interface, which is the unweighted mean of mean scores for all of the questionnaire items for that interface. The overall mean scores enabled direct comparisons of participants' attitudes towards the different products or interfaces tested. The individual items can also be analysed (Agarwal and Venkatesh, 2002) to gather further evidence.

In each of the studies described in this thesis the questionnaire was adapted with amendments and deletions made to the list of core usability items to suit the specifics of the devices and designs being evaluated. This was carried out in a rigorous manner with a panel of experts (Moore, 2001; Guttman and Suchman, 1967). The panel consisted of usability experts from the University of Edinburgh who all had extensive experience in questionnaire design. The majority of the attributes in the questionnaire did not change between studies as they are common to different types

of interfaces. The usability questionnaires were administered after direct user experience with the product or interface design.

Likert questionnaire responses are quantitative and strictly are ordinal rather than interval, but it has been found that in practice many statistical tests designed for interval data are robust to departures from the interval data assumption and can thus be applied to Likert scale data (Garson, 1998; Tullis and Albert, 2008). The most appropriate method of analysing Likert data is the analysis of variance (ANOVA) (Cortson and Coleman, 2003). Attitude scales such as the Likert scale do have some disadvantages as they are open to position response bias. The scale used here has seven points explicitly labelled, so this risk is reduced. The scale employed here also includes a neutral response as it uses an odd number of points. There is a debate in the usability profession as to whether to have an odd or even number of points on a rating scale. One reasonable argument (Tullis and Albert, 2008) for an odd number is that a neutral reaction is a perfectly valid response in the real world. A consistently neutral response is also useful in that it can indicate questionnaire items that are not relevant to participants when judging the usability of a product.

Quality Preference Metric

An advantage of a repeated-measures design is that after participants have had direct experience with all of the alternative interfaces, they can specifically compare them for preference and quality. The quality metric has been used and validated in previous research (Peevers et al., 2008; Weir, 2008; Weir et al., 2006, 2009a, 2009b) and involves participants making a quality rating, recorded as a value on a 0-30

linear scale labelled “poor” at 0, and “excellent” at the 30 end. This quality rating involves evaluating all of the alternative interface designs against each other on the linear scale, and is also recorded to indicate an explicit preference for one version of the interface, or no preference. Each interface is represented by a marker, and the markers are placed on the scale simultaneously. The quality rating is a subjective satisfaction measurement, but unlike the usability questionnaire a participant is specifically asked to compare treatments against each other which can be a useful result to compare with the data collected from the questionnaire. If there is a discrepancy between the two, then the qualitative data from a one-to-one structured interview can be consulted for an explanation.

Qualitative Usability Data

It is important for a usability engineer to collect qualitative data along with quantitative data, so that they can get the best possible insight into any usability issues participants have with a product. The strength of qualitative research is in helping the experimenter understand the qualities of a particular technology and how people use it, and how they think and feel about it (Adams, Lunt and Cairns, 2008). Whereas the strength of quantitative data is that it can be compared statistically.

An exit questionnaire can be carefully designed and then used at the end of an experiment session. The questionnaire can be structured or semi-structured in form, consisting of both open and closed questions. The order and wording of questions is strictly adhered to so that for each participant it is the same, and they will be responding to the same stimuli. The closed questions can also have descriptive

statistic calculated and some analysis can be performed. The experiment facilitator will read aloud the questions and note down the participant's responses. It is important in an experimental session to de-brief the participant on their experience and let them comment on it.

3.5. Statistical Analysis

Both qualitative and quantitative data are collected in a usability experiment and so the data would be provided in several different forms. The general types of data that can be collected in a usability study are:

- Nominal data: unordered groups or categories e.g. male/female, apple/orange. The groups are simply different, one is not better than the other, as there is no order between them.
- Ordinal data: ordered groups or categories, but the interval between them is not meaningful e.g. top ten lists. The song at number one in the list is considered better than that at number two, but the size of the difference between them is not known.
- Interval and ratio data: these are both similar to ordinal data in that they are on a scale, but the difference between points on the scale is equal so it is meaningful. This type of data allows the experimenter to conclude that one unit it twice as large as another. The difference between the two types is that ratio data has an absolute zero that has an inherent meaning (Tullis and Albert, 2008). An example of ratio data is time-on-task.
- Binary data: only two options are possible, Yes or No.

With a controlled usability experiment the main focus is on the quantitative data which can be analysed to allow the experiment hypotheses to be tested. To determine if the differences in data observed between treatments is due to chance, or if it is actually due to the manipulations of the independent variable, a statistical test of significance must be performed. The test will result in a significance score, the *p* value. This is the probability of (falsely) obtaining data at least as extreme as these, assuming that there is in fact no effect caused by the experiment treatments. Conventionally, if the level is <0.05 then the result is 'significant' and the null hypotheses can be rejected. The 0.05 means there is a 5% probability that the result is due to chance. If this level is not reached then the hypothesis is rejected, as the differences were considered to be down to chance alone. The smaller the *p* value then the stronger the evidence that the experiment treatments have produced the effect.

The task timings will produced ratio data and are quantitative in nature. This type of data lends itself to analysis by statistical methods. First descriptive statistic can be computed and analysed by taking the means, standard deviations, median, mode and range for each experimental condition. The mean can be taken by individual task and can then be averaged for an overall score. Descriptive statistics are essential for interval data and are calculated to describe the data recorded from the experimental cohort, but they do not indicate anything about the larger population (Tullis and Albert, 2008). Inferential statistics are then performed, which are used by the experimenter to try and draw conclusion from the random sample data about the general population. For example, if the sample data suggests interface A has higher usability compared to B, by carrying out inferential statistics the experimenter can

infer if there is evidence to confidently allow them to predict about a larger population. In this thesis inferential statistics are performed using a repeated measures analysis of variance (ANOVA). If the ANOVA returns a significant result, pair-wise tests can be performed using Fisher's least significant differences (LSD) method to investigate further.

The task completion data is binary and quantitative in nature, so these data lent themselves to analysis by statistical methods in the same manner as the timing data. When analysing completion rates it is helpful to compute a confidence interval as this is more informative than a point estimate: e.g. 98.6% is the best estimate of the population's completion rate for a task using interface B. It is better to report the confidence interval (see Sauro & Lewis, 2005 for more details), which for a task could be 92-100%. This means somewhere between 92% and 100% of users will be able to complete this task. To work out the confidence interval with the Adjusted Wald method the online calculator[6] was used.

The quality ratings were interval data (30-point scale) and quantitative in nature, so again this type of data lends itself to analysis by statistical methods and by using a repeated-measures ANOVA. Likert questionnaire responses are quantitative and strictly ordinal rather than interval, but it has been found that in practice many statistical tests designed for interval data are robust to departures from the interval data assumption and can thus be applied to Likert scale data (Garson, 1998). Again,

[6] <http://www.measuringusability.com>

descriptive statistics and inferential statistic using ANOVAs can be performed on the data from this usability metric.

Responses to interview questions are free-form text. This type of data is qualitative in nature. Qualitative analysis can be used to 'tell a story' about what was observed and reported by the quantitative data. For example, if there is a discrepancy between the Likert questionnaire results and the quality ratings, then the qualitative data from the one-to-one structured interviews can be consulted for an explanation. The type of analysis that can be carried out on the participants' responses to the exit interview is coarse-grained categorization. The data can be converted to categorical (nominal) data by classification into broad categories such as comments on ease of use, speed and memorability. The analysis involved reading all of the participants' interview responses and placing them in the above mentioned categories in positive and negative fields. The responses were then counted and converted into percentages. The data is sorted and ordered to focus on certain groups of data e.g. age, gender. The data from the closed questions are of the categorical type (nominal). To analyse these data simple descriptive statistics such as the mean, range and simple counts and percentages are calculated.

3.6. Data Collection and Quality Control

Each participant is assigned a random ID so that they remain anonymous in the analysis. The ID is then the only way of identifying the different data collected and is used to connect the data when it comes to analysis. Folders (plus extras) marked with an ID number are provided, each containing all the paper work and materials

needed for the session e.g. task sheets, questionnaires and check sheets (each had the ID pre-printed on them). The ID is used in all the data collection techniques employed in the study, for example the researcher would enter the ID for the usability questionnaire and this would be stored in the database to identify the recorded data.

The researcher has a form (on paper) setting out the questions to be asked during the interview, with spaces to record the participant's responses. The responses written on the form are subsequently keyed in by the researcher, using an Excel spreadsheet with appropriate headings for all the questions. The spreadsheet also provides fields for entry of the researcher's observations on participant behaviour and performance, noted during the session. The data entered by the researcher into the spreadsheet are double checked for errors as a means of quality control. A random 10% of the data entered are checked against the original raw data. If there are errors these are corrected and another random 10% can be checked.

3.7. Advantages and Disadvantages of Experiments

The laboratory usability experiment as a method has several major advantages (Peevers and McInnes, 2009b). Experiments enable a high degree of control over the interfaces (hardware and software) used by participants, the tasks for which they are used and the context of use. This is helpful in reducing the variation and bias which are liable to arise in an observational study, due to extraneous and potentially confounding variables. For example, in a study of real-world use where users are free to choose what tasks they attempt with the system under study, those users who

are more experienced with the system may tend to attempt more complex tasks than the less experienced users. Then the variables of user experience level and task complexity will be correlated, so that it may be impossible to distinguish effects of experience level from effects of task complexity when analysing the results. These two variables are then said to be confounded. In a controlled experiment it is possible to avoid such confounding by assigning tasks to participants in a random or balanced manner rather than giving participants a choice of tasks.

Also, within a controlled experiment it is easy to implement a repeated measures design, in which the same participants use two or more interfaces in succession (the order of use being balanced across the participant cohort). This reduces the level of variation in the comparative results due to individual differences between users, relative to a design in which the different interfaces are used by different groups of participants. It also provides a context for explicit preference questions, to which all participants can give informed answers, since each participant will have experienced all the interface designs under comparison.

For applications or interfaces which do not yet exist in the real world, the experimental setting provides a convenient test-bed for a comparison based on prototypes or simulations – eliminating much of the cost in time and effort that would be required to implement the systems in full. Because the tasks are given as part of the experiment design, the interfaces under test need only be designed to accomplish these specific tasks, rather than the (typically wider) range of tasks for which real-world versions might be used.

Finally, an experiment provides a convenient opportunity to collect a large amount of data in a short time, since participants can be booked to attend the research location at agreed times (typically for a session of approximately one hour per participant) and a large number of measurements of performance and attitude can be collected from a participant within a single session.

The experiment research method also has some significant limitations. These are mostly related to realism – in the tasks being attempted, the contexts of use of the systems under test, and the users' motivation and background knowledge. For instance, mobile devices are used on the move and in various locations so it could be argued that laboratory tests do not simulate the context where mobile devices would be used in the real world. For usability experiments in general the experiment situation only approximates reality, but real world behaviour is inferred from measurements and observations collected in the laboratory. The investigation of banking applications is though, suited to evaluations by controlled experiments because the tasks are well established, and realistic and engaging scenarios can be designed for participants.

With the large number of participants often recruited in the experiment method the practical and time constraints often preclude a field study. A study by Kaikkonen et al. (2005) comparing a mobile application in a laboratory test and in a field test found the same problems in both environments. The only difference was in the frequency of the findings. They concluded that field testing may not be worthwhile when searching for interface flaws.

Chapter 4

4. A Usability Comparison of Three Alternative Message Formats for an SMS Banking Service

4.1. Introduction

This chapter discusses the results of an empirical investigation comparing the usability of three alternative message formats that could be adopted for an SMS banking channel. The aim of the experiment was to find the most usable message input format to use in a *pull* SMS banking service. To this end three formats were compared: Abbreviations, Numbers and Free-form.

4.1.1. Background

Many of the existing SMS banking services use abbreviations, for example for an account balance the user sends ACBAL, but are abbreviations the most usable format? In choosing a message input format banks must recognise that text messaging is not the easiest thing to do due to the variety of methods of text entry available on mobile devices, and the lack of a standard user interface, or even a standard layout of the keypad. The three formats tested here Abbreviations, Free-form and Numbers are on a general continuum and represent specific, relatively discrete, formats that could be used in SMS banking. They were tested for their efficiency by measuring the time taken, their effectiveness by measuring completion rates, and satisfaction by measuring responses to a usability attitude questionnaire and a quality rating. This would give an overall picture of their individual usability, and thus their suitability as the basis for an SMS banking service.

4.1.2. SMS Banking versions

Logically, as writing a text message can still be a slow and painful process, the ideal SMS banking input message format would be the simplest and most succinct. This would point towards the use of numbers. For example, to see an account balance text '1', to receive a mini statement text '2'. The problem with a numbers based system is the lack of mapping between the numbers and what function they represent, which would initially result in the service being difficult to learn to use. This could be alleviated by giving users some form of guide that is small enough size to carry in their wallet. Another solution could be a 'help' function built into the SMS banking system. Neither of these is ideal. SMS banking is proposed as a simple convenient way of banking. Customers may not want to carry around a guide that could be lost or damaged, or text for an explanation because they are unable to remember what commands the numbers represent.

The second message format is the one that has been implemented by a number of banks: abbreviations. If the abbreviations are short and intuitive then this system could be almost as fast and easy to use as a numbers system. The mapping between the abbreviation and function should be more obvious than with numbers. If the abbreviations chosen are not intuitive then this method may suffer from the same usability problems as numbers, e.g. users will at first find it difficult to remember them. The use of abbreviations may prove to be a sensible idea for SMS banking because of their existing widespread use in text messages, so it would benefit from being modelled on users existing behaviour. A question though remains: how widespread is the use of textish among all user groups?

The final message format (free-form) has one inherent advantage: the user can write SMS requests in their own words. This sidesteps the potential usability concern of requiring the user to memorise meaningless numbers, or potentially non-intuitive abbreviations. In an interaction with a bank users may also feel more comfortable using formal language rather than textish. A study in Finland (Kasesniemi and Rautiainen, 2002) found that teenagers wrote more formal messages when texting a person outside their social circle. One disadvantage of a free-form system would be the increased cost of development. It would need a more powerful grammar interpreter compared with an abbreviations or numbers system, which would result in a longer development time and may not prove to be cost-efficient. The second major disadvantage with a free-form system is that users will have to write longer messages, which will require more time and effort. A free-form system would have similarities to the first version of the search engine Ask Jeeves[7], where users were encouraged to phrase their search criteria as natural language questions. Users though, may be unsure of how to formulate their messages, and this may also contribute to the service taking longer to use.

4.1.3. Hypotheses

The experiment reported here investigated which message input format participants would find the most usable, and which they would prefer to use. Based on the condition that participants were provided with a guide to each of the three services,

[7] Ask Jeeves: <http://uk.ask.com/>

and would not be relying on their own memory, the following hypotheses were posited.

H1. There will be differences in the efficiency, based on speed, of the three message input formats, and Numbers will prove to be the most efficient.

H2. There will be differences in the effectiveness of the three messages input formats, based on completion rates, and Free-form will prove to be least effective.

H3. There will be differences in participant satisfaction with the three message input formats, based on the results of a usability attitude questionnaire, and Numbers will perform the best.

H4. There will be differences in the participant satisfaction with the three message input formats, based on 0-30 quality rating scale and Free-form will perform the worst.

4.2. Method

4.2.1 Experiment Design

To enable the participants to compare the three different message input formats the participants were given simple background scenarios, and three banking tasks to carry out using each format. Three different working versions of a *Pull* SMS banking application were created for use with a Sony Ericsson K700i handset (Figure 4.1.) with a 'standard' keypad. One worked with abbreviations for common transactions,

one allowed free-form messages and the third worked with numbers. The three SMS banking services developed were fully functioning and offered six different types of banking transactions/enquires. Customers received realistic confirmation messages back when using the services.

Fig. 4.1: Sony Ericsson K700i

It was decided to conduct this experiment in a usability laboratory, and not in the field. Mobile devices are used on the move and in various locations so it could be argued that laboratory tests do not simulate the context where mobile devices would be used in the real world. With the large number of participants recruited in this experiment practical and time constraints precluded a field study. Further to this, a study by Kaikkonen et al. (2005) comparing a mobile application in a laboratory test and in a field test found the same problems in both environments. The only

difference was in the frequency of the findings. They conclude that field testing may not be worthwhile when searching for interface flaws.

The experiment described in this Chapter was not specifically testing how easy the abbreviation and number commands are to remember. A longitudinal study of such nature is beyond the scope of this research, but the experiment would offer qualitative data from the participants recorded in structured one-to-one interviews, which would address the issue of memorability.

The experiment had three conditions: (1) Abbreviations, (2) Free-Form, and (3) Numbers. Participants experienced all three conditions in a repeated measures design with complete counterbalancing using all six possible orders of these three conditions. The order of the tasks was also randomised for each participant. The usability of each version of the service would be measured by the following independent variables: timings for *efficiency*, completion rates for *effectiveness*, responses to individual items in a usability questionnaire and quality rating with deduced preferences for *satisfaction*. The two between-participants factors were age (two groups) and gender. The experiment was designed to have the order of experience for the three conditions balanced across the participants of differing ages and genders. This was to ensure that each group would experience all possible orders equal amounts of time. Due to no-shows and incomplete data sets the final recruited cohort was reasonably well balanced for these factors (Table 4.1).

4.2.2. Participants

A cohort of 74 participants was recruited in Edinburgh. They were all customers of the Case Bank and all mobile phone owners. There were 36 male participants and 38 female participants (Table 4.1). The age range of the participants was 20 to 70 and the mean age was 40. The participants' length of time as a customer ranged from five months to 68 years, with a mean of 16 years. 53% of the participants were Internet banking customers and 88% of them had sent a text message before, with 45% of those using predictive text when writing text messages. Recruiting enough users who had experience of a Sony Ericsson K700i handset from the customer base would have been impossible, so the participants recruited were owners of various handsets.

	Age 18-35	**Age 36+**	**Total**
Males	17 (23%)	19 (25.7%)	36 (48.7%)
Females	18 (24.3%)	20 (27%)	38 (51.3%)
Total	35 (47.3%)	39 (52.7%)	74 (100%)

Table 4.1: Participant Demographics

The different experience levels that the participants had in writing text messages can be seen in Table 4.2.

Frequency	**Number (Percentage)**	**18-35**	**Age 36 +**
A few times a day or more	29 (44.7%)	19 (59.4%)	10 (30.3%)
Daily	13 (20%)	8 (25%)	5 (15.2%)
A few times a week	18 (27.7%)	5 (15.6%)	13 (39.4%)
Weekly	2 (3%)	0 (0%)	2 (6.1%)
Monthly	3 (4.6%)	0 (0%)	3 (9.1%)
Less often	0 (0%)	0 (0%)	0 (0%)

Table 4.2: Participants' SMS Experience Levels

4.2.3. The SMS banking service

The six different banking tasks available using the SMS banking service were: request a mini statement, order a full statement, transfer funds from savings account, top-up mobile phone, order a new cheque book and find the nearest ATM. There was also a 'help' function that returned a list of the services with the relevant abbreviation or number commands. These six services were available with each of the three versions of SMS banking. The commands used by the different services are detailed below in Tables 4.3 and 4.4, but with Free-form the participants were free to write their request in any manner they wished. For each mode participants were given the relevant credit card sized SMS banking card (see Figure 4.2.).

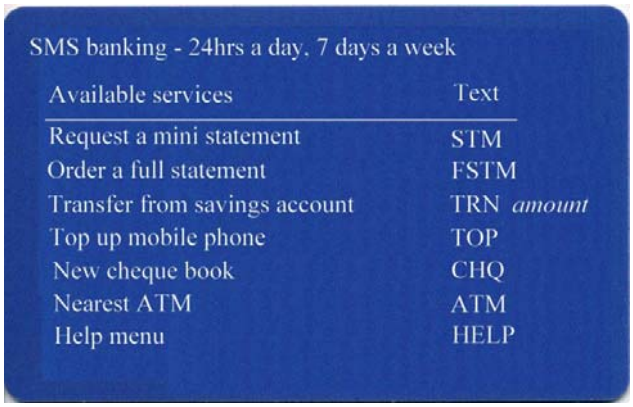


Figure 4.2: Customer card for Abbreviations Service

The six services available through SMS banking were listed on the back of the card with the relevant abbreviation, or number, alongside. The front of the card detailed the phone number to send the text messages to. The Free-form version of the card just listed the available services and the number to text to. There were A4 laminated versions of the cards available for any participants with visual difficulties. For

example, to request a mini-statement using Abbreviations the participant would send ‘STM’; using Numbers they would send ‘1’; for Free-form an acceptable message could be ‘please send me a mini statement’.

Abbreviation based design	
Request a mini-statement	**STM**
Order full statement	**FSTM**
Transfer from savings account	**TRN** amount
Top-up mobile phone	**TOP**
New cheque book	**CHQ**
Nearest ATM	**ATM**
Help menu	**HELP**

Table 4.3: Abbreviations based SMS Banking Service

Numbers based design	
Request a mini-statement	**1**
Order full statement	**2**
Transfer from savings account	**3** amount
Top-up mobile phone	**4**
New cheque book	**5**
Nearest ATM	**6**
Help menu	**HELP**

Table 4.4: Numbers based SMS Banking Service

In order that all participants received the same experience during each session, it was vital that each text message reply sent from the bank was delivered in the same time interval. Thus, in order to guarantee this, an emulator was used to recreate the required network environment. A J2ME application using MIDP2.0 was created for the Sony Ericsson K700i handsets used in the experiment. This J2ME application accurately emulated the Graphical User Interface (GUI) of the host device, in particular the 'stand-by' screens and all screens associated with the composition and

delivery of a standard text message. By using such an application, it was also possible to record accurately the length of time taken to compose each message and the content of these messages using the inbuilt record store function of the J2ME platform. The additional functionality of vibration and sound alerts for the incoming messages employed the Nokia UI extensions.

The Free-form service worked using simple keyword matching. For instance, if the words 'mini' and 'statement' were included in a message in the right order then this SMS message was accepted as correct. The service was case insensitive. The contents of each of the participants' messages were logged by the application. In this way if a participant failed the task it was possible to use the log files to diagnose what had gone wrong.

4.2.4. Text input methods

The two most common ways of entering text on mobile phones are AOL Mobile's T9[8] predictive text input, and Multi-tap, the default for all mobile phones. When using Multi-tap a user has to press a number key several times in order to input a single text character, for instance, to write the letter f the '3' key is pressed three times (d, e, f, 3). As Weiss (2002) comments, Multi-tap is not very intuitive, and is thus difficult to use. T9 works by comparing the sequence of keystrokes to words in a large database to determine the user's intended word. The problem with T9 is the often confusing results displayed by the predictive interpreter as the user inputs each

[8] T9: <http://www.t9.com/eu/index.html>

letter of the word. A study on novice users (Weiss, 2002) found they were no quicker with T9 than with Multi-tap. Some users find T9 confusing, or too difficult and subsequently resort to the familiar Multi-tap. It has been claimed that more than half of all users who have T9 do not use it (Eatoni Ergonomics, 2006), a claim that was given credence by analyses of our randomly recruited cohorts texting habits. For the purposes of experimental control it was decided to restrict the participants to the use of the default Multi-tap text entry method. The reasoning behind this was that all participants would have experience of using the Multi-tap method, while Multi-tap users would find it difficult to learn to use T9 for the purposes of the experiment[9].

4.2.5. Scenario and tasks

Participants were told to imagine that they had recently registered for the bank's new SMS text messaging banking service. They were told that when they registered they had set their current account as the primary account for SMS banking, which meant all transactions would be carried out on their current account. When they registered they also received an SMS banking customer card, as described above. The participants had to perform three tasks. They had to order a new cheque book, transfer £200 from their savings account and request a mini statement to be sent to their phone.

[9] In the design of the experiment, two formats for free-text entry were considered, T9 and multi-tap. T9 usability is impacted by user choice and prior usage patterns; use of multi-tap is more prevalent and more amenable to user familiarisation in an experiment. Including both formats in the experiment would have required a much larger sample. A decision to run with only multi-tap in the experiment was validated in that usability scores for multi-tap were significantly lower than scores for Abbreviations and Numbers, obviating the need to further examine T9 usability.

4.2.6. Attitude Questionnaire

The design of the usability attitude questionnaire used for this experiment followed standard practice (Likert, 1932) by using an equal number of negative and positive statements presented in a randomised order. The use of questionnaires to evaluate user interfaces has a long history (LaLomia and Sidowski, 1990; Root and Draper, 1983). The questionnaire used a 7-point Likert format that ranged from "Strongly Agree" (1) to "Strongly Disagree" (7). Following reversal of the polarity of positive questionnaire statements, a score of 7 consistently indicates a strongly positive attitude and 1 a strongly negative attitude. The questionnaire used consisted of 18 statements that address a range of usability attributes pertaining to human-computer interaction (HCI): *cognitive attributes* (level of concentration and degree of confusion), *the fluency and transparency of the service* (knowledge about what is expected, ease of use, degree of complication), and *quality attributes* (efficiency of service, amount of improvement service is felt to require, reliability of service). The 18 questionnaire items are shown in Table 4.5.

1	I thought this service was too complicated
2	When using this service I didn't always know what I was expected to do
3	I thought the service was efficient
4	I liked the service
5	I would be happy to use this service again
6	I found this service confusing
7	I found this service user-friendly
8	I felt under stress using this service
9	I could accomplish my goals quickly using this service
10	I enjoyed using this service
11	I found this service frustrating to use
12	I felt flustered using this service
13	The service was easy to use
14	Using the service took a lot of concentration
15	I did not feel in control using this service
16	I feel this service needs a lot of improvement
17	I felt confident in the security of this service
18	I found it easy to compose messages using this service

Table 4.5: Usability Questionnaire Items

All participants ($N = 74$) completed the usability questionnaire following exposure to each of the three experimental conditions. The usability questionnaire was presented to the participants on a laptop computer and the order of questions was randomised for each participant. The mean scores for these usability attributes enabled direct comparisons of participants' attitudes towards the three different message formats.

Overall usability was determined by determining the mean for all the usability questions by all participants. Individual statements were also analysed separately to identify any specific usability issues that arose from the hands-on usage sessions.

4.2.7. Quality and exit interviews

After experiencing all three SMS banking services, and completing the usability questionnaires, participants made a quality rating, recorded as a value on a 30-point

linear scale labelled “poor” at the 0 end and “excellent” at 30. This quality rating involved evaluating all three services against each other on the linear scale, and was also recorded to indicate an explicit preference for one version of the SMS service, or no preference. The quality rating is a subjective satisfaction measurement, but unlike the usability questionnaire a participant is specifically asked to compare the three services against each other which can be a useful result to compare with the questionnaire scores. If there is a discrepancy between the two, then the qualitative data from the one-to-one structured interviews can be consulted for an explanation. In the exit interview participants were given an opportunity to comment on their experiences and give their opinions on a range of related issues.

4.2.7. Procedure

The participants were given an A4 blow-up diagram of the layout of the keypad on the handset to demonstrate the use of the buttons and their mapping. They were given a demonstration of how to use the keypad to get a desired character by repeated pressing of a key. They were also specifically shown that if they held down the desired key, rather than repeatedly pressing it, they could get the number assigned to that key. They were then given two practice tasks. The first was to write their own name and the second was to write the message “Meet me at 10am”, which was chosen for its mixture of letters and numbers.

The participant was then introduced to the first randomly assigned version of the SMS banking service with its specific message format. They were then given the

SMS banking customer card with the instructions on, and given a verbal example of how to use the service, which was not one of the experiment tasks. The example was how to order a postal statement and this was the same for all participants. The participant was then given the first of the three randomly ordered tasks. They were allowed three attempts to complete the task. If they failed three times the researcher stepped in and helped them to complete the task. They were then given the second and third tasks to complete in the same way. After completing all three tasks the participant was given a usability questionnaire to complete. The participant was then introduced to the second version of SMS banking and the procedure was repeated. Finally they were introduced to the third version and again the same procedure was followed.

At the end of the session the participant completed a one-on-one 'exit' interview and were also asked to rate the overall quality of the three different version on the 30-point linear scale. The experiment was concluded by completion of a demographic questionnaire to collect factual data about the participant and their banking and mobile phone usage.

4.3. Results

4.3.1. Efficiency – Task Timings

The time that participants took on each task was logged. Table 4.6 shows the mean times for the individual tasks in seconds, and also the total combined mean time for all three tasks. The participants were allowed up to three attempts to complete a task.

The times shown in the table reflect this, and are derived from the total time participants took on a task, notwithstanding how many attempts were made.

Task	Abbreviations	Free-from	Numbers
Mini statement	27.11 (15.28)	76.81 (63.34)	25.99 (31.02)
Cheque book	31.51 (32.6)	89.88 (86.98)	21.70 (18.42)
Funds transfer	58.8 (38.68)	163.93 (114.78)	63.82 (51.64)
Total mean time	117.42 (61.77)	330.62 (226.33)	111.51 (76.32)

Table 4.6: Mean Times (St. Dev.) in seconds by Task and SMS Version

A repeated measures analysis of variance (ANOVA) was performed on the mean total times, with version as the within-participants factor, and age group and gender as between-participants factors. Deviation from Sphericity was taken into account by applying the Greenhouse-Geisser adjustment in the analysis. The analysis showed a very highly significant difference between versions [$df = 1.23$, $F = 84.6$, $p<0.001$], and a significant interaction of version and age [$df = 1.23$, $F = 4.36$, $p = 0.032$], with the older age group taking much longer to complete the tasks with Free-form. There was also a between-participants effect for age [$df = 1.0$, $F = 14.8$, $p<0.001$], with the older age group taking a longer time on the tasks across all three versions. Pair-wise tests were performed using Fisher's least significant differences (LSD) method and these showed that the differences in total mean time between Free-form and Abbreviations, and Free-form and Numbers, were very highly significant, both at the $p<0.001$ level. There was no significant difference in total mean time between Abbreviations and Numbers.

A repeated measures ANOVA was run on the mean times for each task (Table 4.6), with the same within-participants factors as used in the combined mean time test. For the mini statement task there was a highly significant difference between versions [$df = 1.48$, $F = 46.5$, $p<0.001$], and a highly significant difference for the between-participants factor of age [$df = 1.0$, $F = 12.2$, $p<0.001$], with the older age group taking longer on the tasks across all three versions. Pair-wise tests (LSD) showed that the differences in time taken between Free-form and both the other versions on the mini statement task were highly significant ($p<0.001$). There was no significant difference between Abbreviations and Numbers.

For the cheque book task there was a highly significant difference between versions (d.f. = 1.27, $F = 38.5$, $p<0.001$), and a significant interaction of version and age (d.f. = 1.27, $F = 5.17$, $p = 0.018$), with older customers taking longer to complete this task with Free-form. There was also a significant difference for the between-participants factor of age (d.f. = 1.0, $F = 10.62$, $p = 0.002$), with the older age group again taking longer on the tasks across all three versions. Pair-wise tests (LSD) showed that the differences in time taken between Free-form and both the other versions on the cheque book task were highly significant ($p<0.001$). There was also a significant difference ($p = 0.018$) in time taken between Numbers and Abbreviations.

For the funds transfer task there was a highly significant difference between versions (d.f. = 1.41, $F = 56.5$, $p<0.001$). There was also a significant difference for the within-participants factor of age (d.f. = 1.0, $F = 12.0$, $p<0.001$), with the older age group again taking longer on the tasks across all three versions. Pair-wise tests (LSD)

showed that the differences in time taken between Free-form and both the other versions on the transfer task were highly significant ($p<0.001$). There was no significant difference between Abbreviations and Numbers.

To analyse if there was a relationship between the age of the participant and time it took them to complete tasks with each version the data were check for correlations using Pearson's correlation coefficient (r). A scatterplot of the relationship is shown in Figure 4.2. Here the user's age is plotted on the horizontal axis, and the total task completion time in each version of the service is plotted on the vertical axis. A linear regression line is also shown for each version: this summarises the trend in task completion time across the age range. It can be seen that task completion time tends to increase with the age of the user for each message format, but more so (i.e. the slope of the regression line is steeper) for Free-form than for Numbers or Abbreviations. The correlations were significant for each version: Abbreviations (r =.527, p = .000), Free-form (r =.527, p = .000) and Numbers (r =.503, p = .000).

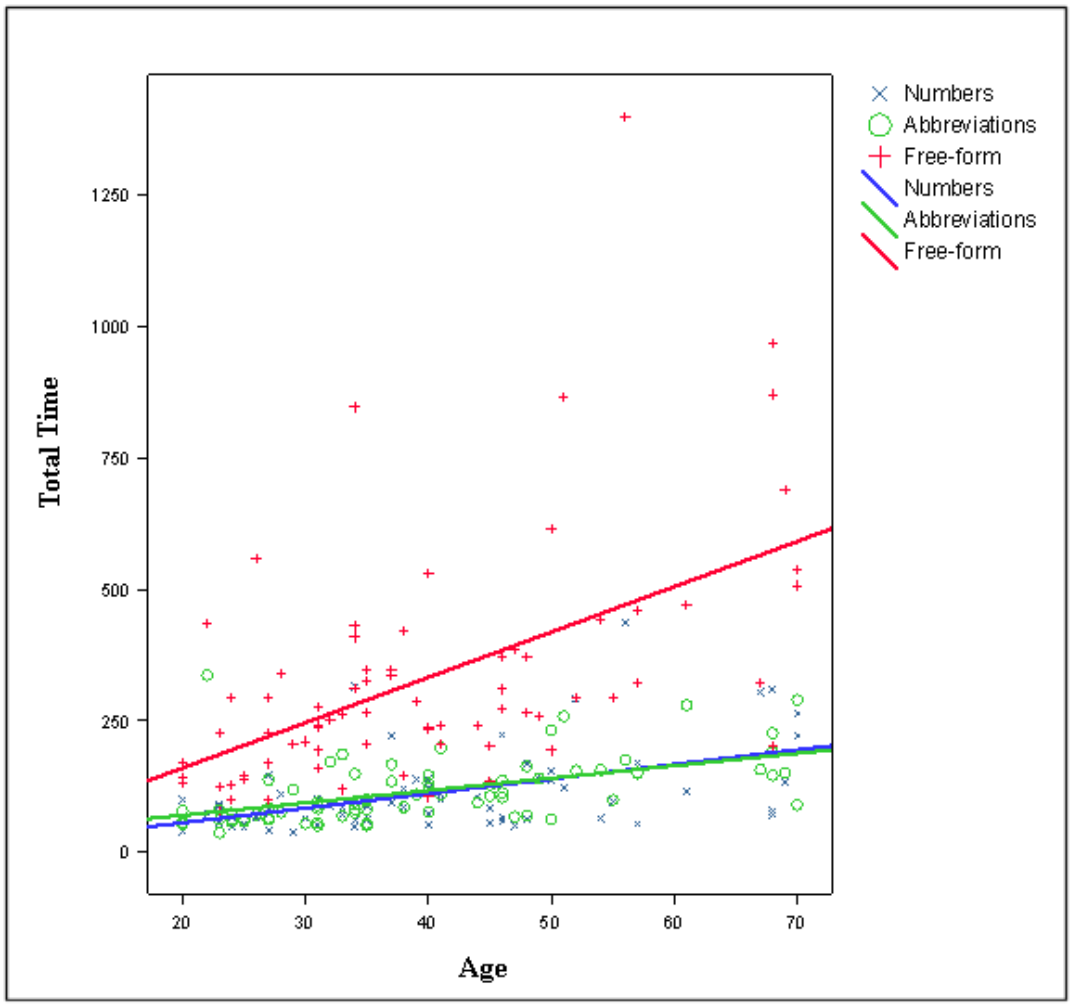


Figure: 4.3. Scatterplot of total task completion time (in seconds, vertical axis) for each message format against user's age (in years, horizontal axis), with regression lines

4.3.2. Effectiveness – Completion Rates

The mobile phone software was used to log if the participants completed each task. Table 4.7 shows the completion rate (not taking into account the number of attempts) for the individual tasks and the overall task completion rate, which is obtained by averaging the data across all participants and all tasks for each SMS banking version. Using the Adjusted Wald method (Lewis and Sauro, 2006; Sauro and Lewis, 2005)

the confidence intervals for each completion rate were calculated. These are also shown in Table 4.7. No ANOVAs were carried out on this data due to the ceiling effects caused by the lack of variability in scores.

	Abbreviations		**Free-form**		**Numbers**	
Task	Mean	Interval	Mean	Interval	Mean	Interval
Mini statement	98.6%	92% - 100%	100%	96% - 100%	98.6%	92% - 100%
Cheque book	100%	96% - 100%	98.6%	92% - 100%	100%	96% - 100%
Funds transfer	100%	96% - 100%	100%	96% - 100%	97.3%	90% - 100%
Overall score	99.5%	92% - 100%	99.5%	92%-100%	98.6%	88% - 99%

Table 4.7: Completion Rates with Confidence Intervals for participants completing the tasks within three attempts

First time Completion Rates

Using the data, the percentage of customers to complete each task first time was calculated, again the overall completion rate was obtained by averaging the data across all participants and all tasks for each SMS banking version. Table 4.8 shows the successful first time completion rates with their confidence intervals.

	Abbreviations		**Free-form**		**Numbers**	
Task	Mean	Interval	Mean	Interval	Mean	Interval
Mini statement	94.6%	87% - 98%	79.7%	69% - 87%	91.9%	83% - 97%
Cheque book	94.6%	87% - 98%	75.7%	65% - 84%	95.9%	88% - 99%
Funds transfer	86.5%	77% - 93%	91.9%	83% - 97%	69.0%	57% - 79%
Overall score	92.0%	68% - 86%	82.4%	51% - 72%	86.9%	53% - 75%

Table 4.8: Successful first time Completion Rates with Confidence Intervals

A repeated measures analysis of variance (ANOVA) was performed on the overall successful first time completion rates, with version as the within-participants factor,

and age group and gender as between-participants factors. Deviation from Sphericity was taken into account by applying the Greenhouse-Geisser adjustment in the analysis. The analysis showed a slightly significant difference between versions (d.f. =1.0, $F = 4$, $p = 0.049$). There was also a within-participants interaction between SMS version and gender (d.f. =1.0, $F = 4$, $p = 0.049$) with males having a lower overall completion rate for tasks using Numbers compared to females. The data also revealed a between-participants effect for age (d.f. =1.0, $F = 14.04$, $p<0.001$) with older participants having a lower overall completion rate across all three SMS versions. Pair-wise tests using Fisher's (LSD) method showed no significant differences in the overall completion rates.

A repeated measures analysis of variance (ANOVA) was performed on the completion rates for the individual tasks, with version as the within-participants factor, and age group and gender as between-participants factors. For the mini statement task only one significant effect was found, which was a between-participants effect for age (d.f. =1.0, $F = 6.83$, $p = 0.011$), with older participants having a lower overall completion rate across all three SMS versions. No other significant within or between subject effects were found. For the cheque book task no significant effects were found. For the funds transfer task there was a significant difference between versions (d.f. =1.0, $F = 6.15$, $p = 0.016$), and a significant interaction (d.f. =1.0, $F = 4.24$, $p = 0.043$) between SMS version and gender with males having higher completion rates using Abbreviations and Free-form compared to females, but having a lower completion rate for Numbers. Pair-wise tests (LSD) showed significant differences between Abbreviations and Numbers, and Free-form

and Numbers (both $p<0.05$). There were no other significant pair-wise differences found.

4.3.3. Attitude

The scores for the 18 usability attributes were averaged to obtain an overall usability score for each version of SMS banking. The mean usability scores for the three version of SMS banking were as follows, Numbers scored highest with a mean score of 5.73 (SD=0.91), second was Abbreviations with 5.62 (SD=0.75) and Free-form had the lowest mean usability score of 4.39 (SD=1.29).

A repeated measures analysis of variance (ANOVA) was performed, with version as the within-participants factor, and age group and gender as between-participants factors. Deviation from Sphericity was taken into account by applying the Greenhouse-Geisser adjustment in the analysis. The ANOVA revealed a very highly significant difference between versions (d.f. = 1.61, $F = 63.1$, $p<0.001$), but there were no significant effects for interactions between version and age or version and gender. There was a significant between-participants effect (d.f. = 1.0, $F = 6.6$, $p<0.01$) found for age with the older age group giving a lower mean score across all three versions compared to the younger age group.

Further pair-wise statistical analysis on the overall means using Fisher's (LSD) method revealed that the differences in overall mean scores between Abbreviations and Free-form, and between Numbers and Free-Form, were very highly significant

(p<0.001). There was no significant difference in the overall mean scores of Abbreviations and Numbers.

Pair-wise tests (LSD) were used to analyse the data for each of the 18 individual usability attributes, with age and gender as the between-participants variables. The pair-wise results were only consulted when the overall ANOVA test on each individual attribute had already shown a significant effect. The scores for Free-Form on all 18 usability attributes were significantly (p<0.01) lower than those for Abbreviations and Numbers, and in almost all cases they were all highly significant (p<0.001). There were significant differences (p<0.05) between Abbreviations and Numbers for four of the usability attributes, all with the Numbers scheme scoring higher than Abbreviations. Compared to Abbreviations, the Numbers versions of SMS banking was less *complicated*, *quicker* to use, more *user-friendly* and the customers *liked* it more.

Significant within-participants effects were found for the interaction of version and age for the attributes *complication* and *concentration*: the older age group gave Free-form much lower usability scores (p<0.05). A between-participants effect (p<0.05) of age was found for nine of the usability attributes, *complication*, *know what to do next*, *liked*, *happy to use again*, *stressful*, *frustrating*, *flustered*, *concentration* and *ease of composing message;* with the older age group giving lower scores across all three versions. A between-participants interaction for age and gender (p<0.01) was found for the attribute *confusing* with young men giving a noticeably higher score across all three versions. A between-participants effect (p<0.05) for gender was

found for the attribute *concentration* with females giving a lower score across all three versions. There was also a between-participants effect ($p<0.01$) for the interaction of age and gender for this attribute with younger males giving a higher score across all versions.

4.3.4. Quality

To collect a quality rating for each of the SMS banking versions participants were asked to order and rate each version by preference by placing magnetic markers on a scale marked from 0 (poor) to 30 (excellent). The mean ratings scores for the three version of SMS banking were as follows, Numbers scored highest with a mean score of 23.61 (SD=6.26), closely followed by Abbreviations with 22.39 (SD=6.37) and Free-form had a much lower mean rating score of 12.03 (SD=8.03).

A repeated measures analysis of variance (ANOVA) was performed on the means, with version as the within-participants factor, and age group and gender as between-participants factors. Again, the Greenhouse-Geisser adjustment was applied in the analysis. This revealed a very highly significant difference between versions (d.f. =1.8, F =64.7, $p<0.001$), and a highly significant effect (d.f. =1.8, F =6.1, $p<0.01$) for the interaction between SMS version and age, with the older customers giving Free-form a lower rating. There were no between-participant effects.
Pair-wise (LSD) tests revealed differences between Free-form and the other two versions, which were both very highly significant ($p<0.001$), with no significant difference between the quality ratings for Abbreviations and Numbers.

For each participant, a ranking of the three versions by perceived quality was derived from the three ratings. From this an overall picture of preference for versions was collected. The rankings are given in Table 4.9 (Not all columns add up to 74, as when a participant rated two versions as equal best on the quality scale, this has been counted as half a vote for a ranking of 1 and half a vote for a ranking of 2 for each of them. Similarly, when two versions were rated equal worst this has been counted as half a vote for a ranking of 2 and half a vote for a ranking of 3 for each of these two. Participants who rated all three versions equally have been counted as giving each version a ranking of 2). It can be seen from Table 4.9 that Numbers received the top rating by the highest amount of participants, Free-form received the lowest rating by the highest amount of participants, and Abbreviations received the highest amount of second place ratings.

	Abbreviations	Free-form	Numbers
1st	29	7	39
2nd	49.5	7.5	27
3rd	4.5	59.5	8

Table 4.9: Preference Rankings

4.3.5. Qualitative exit interview results

When asked what they liked about the Abbreviations message format many participants commented on how *easy/simple* it was to use (mentioned by 35% of participants in the experiment), as one participant commented *"Easy to use, they make sense so are easy to remember"*. How *quick* it was to use was an aspect cited by 23% of participants. A further 12% commented that it would be easy to remember

the Abbreviations. As one participant commented *"It's more specific than the other (versions), only simple abbreviations from the actual words so easy to remember"*.

When asked what they disliked about the Abbreviations 20% of the participants focussed on the difficulty in *remembering*, or *memorising* them. A further 14% commented on their dislike of the actual abbreviations chosen. As one participant commented *"Didn't like abbreviations, couldn't guess them, would need card with you"*.

In general the participants did not offer many positive comments on the Free-form message format, but 10% of participants did comment that they would not need to *memorise*, or *remember*, any abbreviations or numbers. A further 5% commented on the fact that they could *"type what you liked"*. As one participant commented *"Really liked this as you could type in exactly what you want"*.

When asked what they disliked about Free-from many participants cited how *slow* it was to use (mentioned by 46% of participants in the experiment), as one participant commented *"Too long, had to concentrate more, more room for error"*. Another aspect cited by participants was that they were *unsure* what to type, or what was acceptable to type, commented on by 12% of the participants. A further 19% of the participants commented on their attempts not being accepted, due to spelling mistakes and unsuitable attempts at abbreviations.

As with the Abbreviations format many participants focussed on how *quick* the Numbers format was to use (mentioned by 39% of participants in the experiment). The most comments gathered cited how *easy/simple* the Numbers format was to use, which was mentioned by 53% of the participants. One participant commented *"Perfect for a lazy person like me. You don't have to be a text expert to do this one."* Only 4% commented that numbers would be easy to remember.

When asked what they disliked about Numbers many of the comments focussed on the difficulty in *remembering*, or *memorising* the numbers (mentioned by 26% of participants in the experiment), *"Have to memorize numbers as they bear no relation to services"*. A further 12% of participants commented on the *transfer* task being difficult to complete with Numbers. These comments correspond with the completion data, as 31% of participants failed to complete this task first time using Numbers. As one participant commented *"Could confuse numbers, especially with transfers"*.

Customers were given a questionnaire to complete that listed 13 services (the six used in the experiment and seven further suggestions) that could be made available through SMS banking. Reponses to each service were given on a seven-point response scale from "definitely would use" to "definitely would not use", and were coded as values from 1 to 7, with a score of 7 representing they would use it, and 1 that they would not use it, and 4 representing the neutral point. Table 4.10 shows the 13 statements and the mean score for each in descending order of popularity, with the percentage of participants who indicated that they would definitely use each service

shown in brackets. The most popular services are mini statement, account balance, list of recent debits and list of recent credits. There are no services that they would definitely not use.

Service	Mean
Account balance	6.49 (75.7%)
Mini statement	6.07 (64.9%)
List of recent debits	5.86 (58.1%)
List of recent credits	5.81 (55.4%)
Order cheque book	4.98 (40.5%)
Find nearest branch	4.74 (35.1%)
Find nearest ATM	4.65 (36.5%)
Transfer to savings account	4.62 (29.7%)
Transfer from savings account	4.54 (33.8%)
Full statement	4.51 (33.8%)
Top-up mobile phone	4.18 (37.8%)
Stop cheque	3.89 (28.4%)
Pay Bills	3.88 (20.3%)

Table 4.10: Participants Requirements for SMS Banking Services

The customers were given a questionnaire listing the abbreviations used in the Abbreviations SMS banking experiment and asked to indicate whether they liked the abbreviations used. If they did not they were asked if they could suggest alternatives abbreviations. Table 4.11 below lists the services and corresponding abbreviations and shows the numbers of participants that liked each abbreviation.

Service	Abbreviation	Liked
Mini statement	STM	43 (58.1%)
Full statement	FSTM	50 (67.5%)
Transfer from savings account	TRN	50 (67.5%)
Top-up mobile phone	TOP	70 (94.5%)
Order cheque book	CHQ	74 (100%)
Find nearest ATM	ATM	72 (97.2%)

Table 4.11: Participants Thoughts on the Abbreviations Used

For mini statement the most common alternative suggestions were MST, suggested by six participants, and MSTM, also suggested by six participants. For full statement the most common alternative suggestions were FST and FS, both suggested by five participants. For Transfer from savings account there was only one abbreviation suggested more than once, which was TRF suggested by five participants. There were no alternative suggestions for toping up a mobile phone, cheque book or find nearest ATM.

Participants were also asked to suggest abbreviations for the seven further services (Table 10). For a pay bills service 14 participants suggested BILL, and 14 suggested PB. A further eight participants suggested PAY as an abbreviation. Suggestions made for a stop cheque service were STOP, by 10 participants, STP CHQ by six, and STCHQ by four participants. For recent debits the most common suggestions were DEB, suggested by 13 participants, and DEBITS suggested by four participants. For recent credits the most popular suggestions were CRED with six suggestions, CREDITS with five and CR with four. For account balance the most common suggestions were BAL, suggested by 29 participants, and AB suggested by 12. For transfer to savings account the most common suggestions were TSA, by five participants, and TRANS and TRN SA both suggested by three participants. For a find nearest branch service the most popular suggestions were TSB with nine suggestions, FIND with five and BRANCH and BRAN with four each.

Participants were also asked what they thought a SMS text messaging banking service should be called. The most common suggestions were Text Banking (19

participants), Mobile Banking and SMS Banking (both with six participants) and Text Bank (three participants).

After completing the three tasks with each version of SMS banking, participants were asked various questions about SMS banking and what they liked and what they disliked about each of the versions in an interview questionnaire.

The participants were asked to comment on the text message replies they received from the bank. They were asked if each contained enough information. For the mini statement 83.8% said that it contained enough information. For the cheque book confirmation message 87.8% said that it contained enough information and for the transfer confirmation message 83.8% said that it contained enough information.

When asked if they whether they would use an SMS text messaging banking service like the one they had said they had preferred, 54.1% said they would use it, but only if it was free of charge, 28.4% said they would be willing to pay a small charge to use it and 17.6% said they were unsure if they would ever use it.

4.4. Discussion

The results from this experiment show that the Numbers and Abbreviations message input formats were the most beneficial to the user. The participants used each version three times over a short duration, and as it was the first time they had used each version they had access to an SMS banking card detailing the relevant abbreviation or number command to use. In this context, it had been predicted that Numbers

would be the most efficient version of the service to use, and it was perceived to be faster by the participants, but in actuality, participants only performed faster in the cheque book task compared with Abbreviations. It was also predicted that Numbers would be rated highest for *satisfaction*, but again there were no significant difference between the overall usability questionnaire scores for Abbreviations and Numbers, or the quality ratings, though the participants perceived numbers to be less *complicated*, more *user-friendly* and they *liked* it more.

True to predictions, Free-form in general performed the worst. The reasons for this can be deduced from its inherent limitations when applied to a mobile phone user interface. When analysing the participants' comments 46% of them gave its slowness as a reason for disliking the service. It should be noted that even though participants took longer to write messages and gave Free-form significantly lower overall mean questionnaire and quality scores for *satisfaction*, this did not impact on how successful they were in completing the tasks, as the task completion data testifies. Indeed, in the case of the funds transfer task Free-form actually had a significantly higher first time completion rate compared to Numbers. It could be suggested that because it took participants more effort to write messages in Free-form they actually thought more about what they writing. An alternative possibility is that because the mapping between what they were writing and what they wanted was very high this led to fewer mistakes. The performance of Numbers on this task could be due to a level of confusion caused by the command for a transfer being a number (3), and the amount also being a number (200), as one participant commented *"Could confuse numbers, especially with transfers."*

It could be argued that Free-form performed the worst because it forced people to write text messages in grammatically correct English. In reality many people use abbreviations, or textish. It was observed in the text logs that many users wrote very concise messages e.g. ‘new cheque book’, and some even used the service in a keyword type manner, but even so, it was still much slower to use. Performance may improve over repeated use, and Free-form theoretically has high learnability and memorability, but the first impressions of a Free-from service, as shown in these results, may put users off. This experiment used one type of form-factor and was limited to investigating the commonest text input method for mobile phones. Performance with Free-form may be improved with different text input methods, such as a stylus, but it is debatable if it would ever be as efficient, or as cost effective, as the use of abbreviations or numbers.

It did not prove to be the case in the context of this experiment, but it is still possible that the Numbers format is not as intuitive because of a lack of natural mapping between the commands and the services that they represent, and this factor could impact on its learnability and memorability. Providing the participants with the SMS banking card was realistic, as it is based on an existing automated telephone banking service, but over repeated use it would be expected that heavy users of the service would learn and hopefully memorize the commands. If after an extended period of use the majority of participants still had to rely on the SMS banking card then this would point to a problem in that versions memorability. It may be that the nature of this study has been more favourable on the Numbers format, and that in a longitudinal study Numbers would not perform as well. This study did not measure

participant performance over extended repeated use, but it did consider participants' comments on the problem of memorising the numbers and abbreviations, and 27% for the former, and 20% for the later, mentioned this as something they disliked about these versions. As one participant commented: *"Have to memorise numbers as they bear no relation to service."* This argument, of course, can also be applied to a service using abbreviations, but it seems plausible that abbreviations, if chosen well, may have higher learnability and memorability. These questions could be answered with a longitudinal study investigating user behaviour over an extended time period, or artificially tested by taking away access to the SMS banking card. Which format then to recommend as the model for an SMS banking service? In the context of this study, the results show that users would find an SMS banking service that used either numbers or abbreviations to be equally usable, but it is suggested that abbreviations may prove to be the most usable format based on an assumption of higher learnability and memorability, but further research should be carried out to test these assumptions.

Older users suffer from visual, cognitive and motor impairments to varying degrees (Christopher, 1999; Jagaacinski et al., 1995; Krampe and Ericsson, 1996; Kurniawan et al., 2006a; Walker et al., 1997). Each of these impairments could impact on how usable an SMS banking service will be to older users. The term 'older users' does not just apply to the over 60s, or the retired, as a decline in visual acuity is noticeable by the mid-40s (Kurniawan et al., 2006a), and this decline would make reading the text on a mobile phone's small screen much more difficult. Older users can also suffer from a reduced ability to cope with repetitive fast movements. These are the types of

movements required when entering text into a mobile phone. The result from this Chapter have shown a correlation between age and speed typing text messages, with time taken increasing with age. Older users can also have difficulty with finer movements and motor coordination, due to stiffening of joints and arthritis, which would cause difficulty when using the small keypads on mobile phones, and this has been shown in recent research (Ornella and Stephanie, 2006). In this study 16 of the participants were in their 40s, nine in their 50s, six in their 60s and two in their 70s, it is reasonable to assume that the issues discussed above have contributed to the differences in results found between the two age groups in this experiment. Some mobile phone manufacturers are trying to respond to these issues and have introduced models with less functionality, larger screens and larger keypads that are specifically aimed at the older population (BBC News, 2005b).

There is some debate in the HCI community as to how to explain the SMS usage patterns of older users (Ling, 2007), with one side arguing that it is down to usability problems, and the other arguing that there are more complicated sociological reasons, with the attitude of the older users playing a part in their low adoption of the technology. It has been argued that older people are passive users of mobile phones, and, interestingly, see them more as a safety line rather than a communication device (Kurniawan, 2008). The SMS usage patterns in this study show that the younger age group in general send SMS more frequently than the older group (see Tables 4.2), with nearly 60% sending them a few times a day or more, compared to only 30% in the older group. It is probable that the difference in experience levels between the two groups has also contributed to the difference in results. To answer the question

as to whether lower usage levels, and the consequent lower experience levels, are due to the usability issues surrounding mobile devices, or are caused by sociological explanations, is beyond the scope of this thesis.

Chapter 5

5. Experimental Assessment of Customer Attitude to the Integration of an SMS Channel into a multichannel Banking environment and as a Strategy of 'Next Call Avoidance' in Telephone Banking

5.1. Introduction

This chapter discusses the results of a two part empirical investigation into the integration of an SMS channel into a multichannel banking environment. The first experiment compares an SMS banking service to an existing telephone banking service for balance requests. Two SMS versions: SMS balance-only and SMS balance with Mini-Statement, and two telephone banking services: Automated and Advisor, are compared by measuring their usability. The second part investigates integrating SMS into an existing telephone banking channel by way of funds transfer confirmation messages: SMS as a bank to customer communication medium. The aims of the experiment were to compare the usability of a new SMS banking channel against the existing established telephone banking channel, to investigate the usability of SMS confirmations after telephone banking funds transfers and to gather data as to the acceptability of integrating an SMS channel with existing banking services.

The continual widespread advances in computer technologies have encouraged many banks to adopt new methods of interacting with customers to improve service, to lower costs and to maintain competitive advantage, and for the customer create more convenient methods of banking. Banks are now committed to transferring many

customer services that traditionally involve interacting with branch staff to different technology channels such as telephone banking or Internet banking. In some smaller branches in the UK there are now no traditional counter staff and customers are directed to cash/cheque deposit machines and dedicated telephone lines for access to banking agents, or automated services. Banks aim to encourage customers to contact them using direct channels such as automated telephone banking or mobile banking, rather than by person-to-person contact. They also wish to find strategies for 'next call avoidance' whereby the customer telephones the bank to enquire if a recent transaction has been processed.

Chapter 4 discovered that an SMS balance on-demand service was the one most likely to be used by customers. The capability to provide balance on-demand initiated by *Push* SMS message requests from customers would remove the need for the customer to call telephone banking for a balance, and would offer potential cost reductions and improved customer experience. The message format chosen to implement the SMS banking service was abbreviations. This is based on the findings of chapter 4, and as the subsequent format adopted by the Case Bank in real practice.

The introduction of any new banking channel will require research to compare it with available services in terms of usability and customer satisfaction with the service, because as has been pointed out (Anderson et al., 1996) if a user's first experience of a new service is unsatisfactory they may revert to the use of a more traditional form of service. There has been little research into comparisons of the usability of mobile banking and competing banking channels. One survey study by Laukkanen (2007)

compared customer value perceptions of Internet versus mobile banking for bill payments. The two most important differences found between the two channels were related to the location free access to the service and the display of the device, which resulted in perceived differences based on efficiency, convenience and safety. Laukkanen (2007) argue that by safety the participants meant a feeling of uncertainty with mobile banking due to a fear of making mistakes because of the size of the screen, and slow access to the service, rather than concerns with the actual data security of the system. They found that this difference in usability between the two channels due to the smaller screen size and keyboard on the mobile phone was a major inhibitor to the use of a mobile bill payments service.

If a bank can furnish transaction confirmation SMS alerts on transactions carried out using their channels and thereby remove the need for customers to call an advisor in order to enquire and double-check transaction details, this will also results in cost reductions for the bank and a better user experience for the customer with additional convenience, reassurance and confidence. This is SMS banking used as a one-to-one business to customer communication channel, and offers massive potential for customer relationship management (CRM). A bank's CRM strategy can employ many channels, such as direct mail, telephone, loyalty cards and emails. SMS can be used for this purpose and Mobile CRM has been defined as "*Customer relationship management of any kind including interactive communication between an organization and a customer using a mobile device*" (Liljander, Polsa and Forsberg, 2007).

There has been limited research into mobile CRM. One study (Richardson, 2005 PhD) on customer complaint management compared multi-channel electronic banking options such as ATM, SMS banking, email and a message function on Internet banking. The majority of customers (72%) thought that SMS banking was an appropriate channel for updating them on their complaint. Another study (Liljander et al, 2007), examined mobile CRM used by an airline and argue that customer are not ready for this type of mobile application yet, though they did find that participants who already used the mobile Internet had a more positive attitude to mobile CRM. With such little research so far there is a need for usability research on mobile CRM.

The key questions addressed by this experiment were:

1. How do customers view the role of *Pull* SMS balance on-demand requests in a multi-channel environment? How frequently would they use SMS balance on-demand requests? Would such a service change their contact habits?

2. How do customers view the role of *Push* SMS (confirmation) Text Alerts in a multi-channel environment? Do they gain better re-assurance that transactions have been successfully completed?

3. How does the enhanced SMS Text Alerts offering compare with existing contact points?

4. What are the impacts on perceived usability derived from use of these new SMS Text Alert services?

5.1.2. Hypotheses

The experiment described here had two parts. The first part investigated which channel participants would find the most usable for requesting their account balance, an SMS banking channel or a telephone banking channel. The second part of the experiment investigated if a SMS confirmation after completing a funds transfer using telephone banking would improve the usability of this channel.

Experiment 1

H1. There will be differences in the effectiveness of the two banking channels, based on completion rates.

H2. There will be differences in participant attitude to the usability of the two channels, based on the results of a usability questionnaire.

H3. There will be differences in the participant satisfaction with the two channels, based on a 0-30 point quality rating scale.

H4. There will be a difference in usability scores by age, with older participants giving lower scores.

Experiment 2

H1. There will be differences in participant attitude to the usability of the different versions of telephone banking, based on the results of a usability questionnaire.

H2. There will be differences in the participant satisfaction with the different versions of telephone banking, based on a 0-30 point quality rating scale.

5.1.3. Participants

A cohort of 116 participants was recruited in Edinburgh. They were all customers of the Case Bank and all mobile phone owners. There were 53 male participants and 63 female participants (Table 5.1). The age range of the participants was 20 to 76 and the mean age was 46. All 116 participants provided data for part 1, while 110 provided data for part 2 of the experiment.

	18 – 45	**46 and over**	**Total**
Males	26 (22.4%)	27 (23.3%)	53 (45.7%)
Females	29 (25.0%)	34 (29.3%)	63 (54.3%)
Total	55 (47.4%)	61 (52.6%)	116 (100%)

Table 5.1: Participant Demographics

The participants were asked about mobile banking (mBanking) services, and from their responses it was clear that not all of them were aware that these facilities were available. Some were receiving SMS text alerts such as weekly balances (13

participants, 11%) and when their cards were used abroad (2 participants, 2%). One participant mentioned requesting a cheque book using the SMS request service.

When asked why they did not use mobile banking services, most participants were unaware of these options or thought they had no need to use them, see Table 5.2. The other issues mentioned included: “thinking the services were too expensive”,” finding other services more convenient”, “thinking they would be difficult to use”, “confidence”, “don’t know how to use them”, “no time to learn”, “don’t like getting text messages”.

Reasons	**N**	**(%)**
Not aware	24	29%
No need	8	10%
Use Internet Banking	22	27%
Use other channels (PB, ATM, Branch)	7	9%
Security concerns	4	5%
Don’t use mobiles	4	5%
Other issues	13	16%
Total	82	*

Table 5.2: Reasons for *not* using Mobile Phone Banking Services

A total of 82% of the cohort had sent a text message before. Those who used text messaging were asked about their frequency of use (Table 5.3) and whether they used predictive text. Some 49 (52%) participants used predictive text (T9), while 40 (42%) did not, and 6 (6%) sometimes made use of predictive text when writing SMS messages.

Frequency	Number (Percentage)	18 – 45	46 and over
A few times a day or more	42 (44%)	28 (56%)	14 (31.1%)
Daily	22 (23%)	14 (28%)	8 (17.8%)
A few times a week	20 (21%)	7 (14%)	13 (28.9%)
Weekly	5 (5%)	0 (0%)	5 (11.1%)
Monthly	0 (0%)	0 (0%)	0 (0%)
Less often	6 (6%)	1 (2%)	5 (11.1%)

Table 5.3: Participants' SMS Experience Levels

30 participants (26%) said they received text alerts from other banks and businesses, but only three examples were collected: *Tesco offers*, *Phone Bills*, and one other foreign banking service.

5.2. The SMS *Pull* Service Experiment

5.2.1. Experiment Design

To enable the participants to compare the two different banking channels they were given simple background scenarios and the task of getting their account balance. A working version of a *Pull* SMS banking application was created for use with a Sony Ericsson K800i handset with a 'standard' keypad. The SMS banking service developed was fully functioning. Participants received realistic confirmation messages back when using the service. For the telephone banking service 'mirror' versions of the actual telephone banking service employed by the bank were used. One was a fully automated interactive voice response (IVR) system; the second version employed an actual human advisor.

The experiment had two within-participants conditions: (1) SMS banking and (2) Telephone banking. Each of these independent variables had two between-participant levels. SMS banking's two levels were: (A) balance or (B) balance plus mini-statement, and Telephone banking's two levels were: (C) automated IVR or (D) advisor. Participants experienced both within-participants conditions (1) and (2) in a repeated-measures design, and for each condition 50% of the participants experienced level one of the variable and 50% level two in a balanced design. To divide the participants for the two versions of the telephone banking service half were given security numbers in their personas to allow them to use the automated version of the service.

The usability of each version of the service would be measured by the following independent variables: completion rates for *effectiveness*, responses to individual items in a usability attitude questionnaire and quality rating with deduced preferences for *satisfaction*. For this experiment it was not deemed appropriate to time the participants' performance as a measure of efficiency. In completing the task there would invariably be differences in time taken using the telephone banking service compared to the SMS service due to the nature of the two types of interaction. Indicators of *efficiency* then, would be measured by verbal reports and perceptions of efficiency as measured in the usability questionnaire.

The two between-participants factors were age (two groups) and gender. The experiment was designed to have the order of experience of the independent

variables balanced across the participants of differing ages and genders. This was to ensure that each group would experience all possible orders equal amounts of time.

5.2.2. The SMS banking service

The SMS service was designed to match the functionality of the Case Bank's existing SMS *Pull* service for new PIN and cheque/pay-in books as, **Text BAL to 61119**. According to research by Lee and Chung (2009) information quality significantly influenced customers' satisfaction with a mobile banking service, and their trust in the service. Both are important factors in the adoption of mobile banking. To explore the issue of message content in the balance on-demand message, half of the participants received an SMS message with balance only: the other half received a balance plus a mini statement. This would be a simple way of testing for information quality. An example SMS Balance-only response is shown in Figure 5.1.

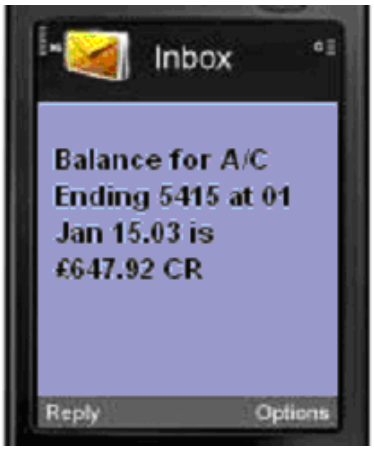


Figure 5.1: Example SMS Balance-only response format

The SMS banking service worked with abbreviations following the findings of chapter 4. The choice of a ‘balance request’ was due to this being a very common customer task for telephone banking, and it was found to score the highest in chapter 4 on ‘definitely would use’ as an SMS service. Mini-statement scored second highest and was included in the balance request message for half of the participants. Further to the research in chapter 4, it was decided to investigate the usability of including the last 4 digits of the customer’s account number in the text message sent to the bank, as a method of selecting which account they required, for example **‘BAL 6439’**. This is the format that the Case Bank has subsequently introduced to its customers.

In order that all participants received the same experience during each session, it was vital that each text message reply sent from the bank was delivered in the same time interval. Thus, in order to guarantee this, an emulator was used to recreate the required network environment. A J2ME application using MIDP2.0 was created for the Sony Ericsson K800i handsets used in the experiment (see Figure 5.2). This J2ME application accurately emulated the Graphical User Interface (GUI) of the host device, in particular the 'stand-by' screens and all screens associated with the composition and delivery of a standard text message. By using such an application, it was also possible to record accurately the length of time taken to compose each message and the content of these messages using the inbuilt record store function of the J2ME platform. The additional functionality of vibration and sound alerts for the incoming messages employed the Nokia UI extensions.

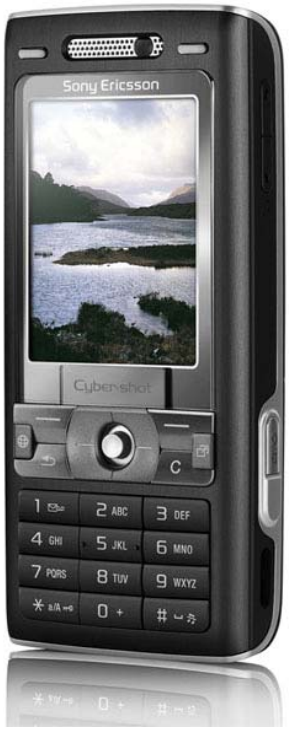


Figure 5.2: Sony Ericsson K800i

5.2.3. Text input methods

For the purposes of experimental control it was decided to restrict the participants to the use of the default Multi-tap text entry method based on the same reasoning for chapter 4.

5.2.4. Scenario and tasks

Participants were told to imagine that they were customer of the bank:

Scenario 1: Participants were told to find out the balance of their account by sending a text message to the bank using the SMS banking service.

Scenario 2: Participants were told to find out the balance of their account by phoning the bank.

The participants were given all of the account details and information they would need to complete the tasks.

Telephone Advisors

Two experienced telephone banking advisors were provided by the Case Bank for the duration of the experiment, to achieve total realism in experience for the participants.

5.2.5. Usability Attitude Questionnaire

The design of the usability questionnaire used for this experiment followed standard practice (Likert, 1932) by using an equal number of negative and positive statements presented in a randomised order. The questionnaire used a 7-point Likert format that ranged from "Strongly Agree" (1) to "Strongly Disagree" (7). Following reversal of the polarity of positive questionnaire statements, a score of 7 consistently indicates a strongly positive attitude and 1 a strongly negative attitude. The questionnaire used consisted of 17 statements that address a range of usability attributes pertaining to human-computer interaction (HCI): *cognitive attributes* (level of concentration and degree of confusion), *the fluency and transparency of the service* (knowledge about what is expected, ease of use, degree of complication), and *quality attributes*

(efficiency of service, amount of improvement service is felt to require, reliability of service). The 17 questionnaire items are shown in Table 5.4.

1	I thought this service was too complicated
2	When using this service I didn't always know what I was expected to do
3	I thought the service was efficient
4	I would be happy to use this service again
5	I found this service confusing
6	The service was friendly
7	I felt under stress using this service
8	I thought the service was polite
9	I enjoyed using this service
10	I found this service frustrating to use
11	I felt flustered using this service
12	The service was easy to use
13	Using the service took a lot of concentration
14	I did not feel in control using this service
15	I feel this service needs a lot of improvement
16	I felt confident in the security of this service
17	I felt that the service was reliable

Table 5.4: Usability Questionnaire Items

The questionnaire consists of the same core statements used in Chapter 4 with a few deletions and inclusions as adjudicated by a panel of experts (Moore, 2001; Guttman and Suchman, 1967). The panel consisted of usability experts from the University of Edinburgh who all had extensive experience in questionnaire design. The statement "I found it easy to compose messages using this service" was excluded as not being relevant when comparing the SMS service to a telephone banking service. The statement "I felt that the service was reliable" was included as being an important factor to compare between the two types of service. The statement "I thought the service was polite" was expected to favour the telephone banking services, but would

be interesting to compare between the two levels of telephone banking: automated compared to advisor.

All participants (N = 116) completed the usability questionnaire following exposure to each of the two experimental conditions. The usability questionnaire was presented to the participants on a laptop computer and the order of questions was randomised for each participant. The mean scores for these usability attributes enabled direct comparisons of participants' attitudes towards the different versions of the service.

Overall usability was determined by computing the mean for all the usability questions by all participants. Individual statements were also analysed separately to identify any specific usability issues that arose from the hands-on usage sessions.

5.2.6. Quality and exit interviews

After experiencing both services, and completing the usability questionnaires, participants made a quality rating, recorded as a value on a 30-point linear scale labelled "poor" at the 0 end and "excellent" at 30. This quality rating involved evaluating both services against each other on the linear scale, and was also recorded to indicate an explicit preference, or no preference. The quality rating is a subjective satisfaction measurement, but unlike the usability questionnaire a participant is specifically asked to compare both services against each other which can be a useful result to compare with the questionnaire scores. If there is a discrepancy between the

two, then the qualitative data from the one-to-one structured interviews can be consulted for an explanation. In the exit interview participants were given an opportunity to comment on their experiences and give their opinions on a range of related issues.

5.2.7. Procedure

The participant was introduced to the first randomly assigned service: SMS banking or telephone banking.

For the SMS banking the participants were given an A4 size diagram of the layout of the keypad on the handset to demonstrate the use of the buttons and their mapping. They were given a demonstration of how to use the keypad to get a desired character by repeated pressing of a key. They were also specifically shown that if they held down the desired key, rather than repeatedly pressing it, they could get the number assigned to that key. They were then given two practice tasks. The first was to write their own name and the second was to write the message '12345', send it and then open and read the reply message.

The participant was then given a task sheet with the instructions on, which the researcher also read out aloud to them. They were allowed three attempts to complete the task. If they failed three times the researcher stepped in and helped them to complete the task.

For telephone banking the participant was given a task sheet with the instructions on, which the researcher also read out aloud to them. They were allowed three attempts to complete the task. If they failed three times the researcher stepped in and helped them to complete the task.

At the end of the session the participant completed a one-on-one 'exit' interview and were also asked to rate the overall quality of the three different version on the 30-point linear scale. The experiment was concluded by completion of a demographic questionnaire to collect factual data about the participant and their banking and mobile phone usage.

5.3. The SMS *Pull* Service Experiment: Results

5.3.1 Effectiveness - Completion Rates

The software used for each version of the service was used to log if the participants completed each task. Table 16 shows the completion rates (not taking into account the number of attempts). Using the Adjusted Wald method (Lewis and Sauro, 2006; Sauro and Lewis, 2005) the confidence intervals for each completion rate were calculated. These are also shown in Table 5.5. No ANOVAs were carried out on this data due to the ceiling effects caused by the lack of variability in scores.

	SMS Banking		Telephone Banking	
Task	Mean	Interval	Mean	Interval
Balance	91.5%	85.7% - 96.0%	94.1%	88.9% - 97.8%

Table 5.5: Completion Rates with Confidence Intervals for participants completing the tasks within three attempts

Using the data the percentage of customers to complete each task first time was calculated. Table 5.6 shows the successful first time completion rates with their confidence intervals.

	SMS Banking		Telephone Banking	
Task	Mean	Interval	Mean	Interval
Balance	85.6%	78.7% - 91.4%	94.1%	88.9% - 97.8%

Table 5.6: Successful first time Completion Rates with Confidence Intervals

A repeated measures analysis of variance (ANOVA) was performed on the overall successful first time completion rates, with service as the within-participants factor, and age group and gender as between-participants factors. The analysis showed a significant difference between services (d.f. =1.0, $F = 5.15$, $p = 0.025$). There was also a within-participants interaction between SMS version and age (d.f. =1.0, $F = 7.61$, $p = 0.007$) with the older age group having a lower overall completion rate with the SMS service compared to the younger group. The data also revealed a between-participants effect for age (d.f. =1.0, $F = 10.45$, $p = 0.002$) with older participants having a lower overall completion rate across both versions of the service.

5.3.2. Usability Attitude

The scores for the 17 usability attributes were averaged to obtain an overall usability score for each service: SMS banking and Telephone banking. The mean overall usability scores for both services were SMS Banking: 5.36 (SD=0.90) and Telephone Banking: 5.78 (SD=0.82). The usability scores broken up by version of each service are shown in Table 5.7.

Version	Mean
Automated Telephone	5.76
Advisor Telephone	5.81
SMS Balance-only	5.49
SMS Balance + Mini-Statement	5.24

Table 5.7: Usability Scores for the Balance-on-demand versions

A repeated measures ANOVA was performed on the overall means with age, gender, SMS version (balance-only or balance + mini-statement) and phone version (automated or advisor) as the between-participant factors. The main effect of the ANOVA was highly significant (d.f. = 1.0, $F = 24.25$, $p < 0.001$) indicating significant differences overall in the attitudes to telephone banking compared to SMS banking. A significant age interaction was found (d.f. = 1.0, $F = 5.02$, $p < 0.05$) where younger customers scored the two methods (telephone banking or SMS banking) similarly, while older customers gave much higher scores to the traditional telephone banking methods. A between-participants gender effect was also noted (*p*

< 0.05) where male participants gave lower scores to all designs compared with their female counterparts.

Further pair-wise statistical analysis on the overall combined means using Fishers (LSD) method revealed differences between telephone and SMS for 11 of the 17 usability attributes: *confusion*, *flustered*, *stress*, *knew what to do next* at $p < 0.05$, and *concentration*, *would use again*, *reliability*, *improvement needed*, *friendly*, *polite* and *confidence in security* at $p < 0.01$, with telephone scoring higher for all 11.

The SMS balance-only design scored higher than the balance plus mini-statement version, but this was not a significant. Analysis of the individual attributes with independent samples t-tests did show significant differences between the two SMS formats on two attributes, with balance-only scoring significantly higher than the balance plus Mini-Statement for *confusion* and *stress* at $p < 0.5$. There were no significant differences in usability between the overall means for the automated and advisor telephone banking services, or on any of the individual usability attributes.

To test for a relationship between the usability attitude and the age of the participants, correlation analysis was carried out. Pearson's correlation coefficient (r) was used to analyse the overall mean for SMS and overall mean for telephone banking. The analysis showed there was a weak but significant relationship for SMS banking and age ($r = -0.247$, $p < 0.01$). This was a negative relationship, which meant that attitude decreased with the age of the participant. There was no

significant relationship between age and attitude scores for telephone banking (r = -0.055, p < 0.56).

5.3.3. Quality and Preference

To collect a quality rating for each version of the Balance-on-demand process participants were asked to rate them comparatively on a 30-point scale. Table 5.8 shows the mean and standard deviation of the quality ratings awarded to each.

Version	Mean (SD)
Automated Telephone	22.56 (5.23)
Advisor Telephone	22.65 (5.64)
SMS Balance-only	20.93 (6.05)
SMS Balance + Mini-Statement	20.71 (7.65)

Table 5.8: Quality Ratings for the Balance-on-demand Designs

The advisor version obtained the highest mean score, closely followed by the automated version. The two SMS versions were scored slightly lower, with the Balance only design scoring higher than the longer text message containing mini-statement items.

A repeated measures ANOVA was performed on the overall means with age, gender SMS version and phone version as the between-participant factors. The SMS versions were found to score statistically significantly lower than the phone-based services, p = 0.036 (SMS average 20.82 vs. Telephone average 22.60). There was also a significant interaction (p < 0.05) for age with younger customers scoring both

types of service (phone or SMS) equally, but older customers giving higher scores to the traditional phone services.

Comparing only the two SMS variants (between-participants), there were no significant differences between the two ratings. No differences were found between the advisor and automated versions of the telephone service, however younger customers did score these versions lower than their older counterparts ($p = 0.004$).

To test if there was a relationship between the quality metric and the age of the participants, correlation analysis was carried out. Pearson's correlation coefficient (r) was used to analyse the mean combined score for both SMS versions, and mean combined score for both telephone banking versions. The analysis showed there was no significant relationship for SMS banking quality and age ($r = -0.138$, $p = 0.144$). But there was a weak significant relationship between age and quality scores for telephone banking ($r = 0.230$, $p = 0.014$). This was a positive relationship, with quality rating increasing with age of the participant.

Commenting on their ratings, customers frequently mentioned the speed and efficiency of the SMS service options drove their ratings (22%). Contrarily, some 15% raised security concerns about the SMS services as the reason for their rating.

Customers also mentioned the following about the SMS service options: ease, convenience, written confirmation, modern, like using, cheap, keeping and referring back to the information, short with less information to give upfront, discrete. Some

other concerns were also mentioned about keeping messages, whether using a registered phone was secure enough, whether texting was secure, that the services were impersonal, OK for balance but not generally, or would depend on the situation.

Reasons for preferring the IVR automated telephone service included familiarity (5%) and security (3%) as well as being considered quick, easy, private (although without confirmation in writing), or just preferring phoning personally. Some negative issues raised included that the calls were too long for balance requests, were "old and slow", still "needs improvement" or were too complicated.

Finally, preferences for the advisor service were due to the preference for talking to someone (8%), security (7%) or familiarity (2%). Advisor calls were seen as easy, likable, private, personal and error tolerant. Some negative remarks included that they were "old fashioned" and asked "too many questions" (2%).

Reducing the Quality Rating data into a rank order results in the data presented in Table 5.9.

Rank	N	%
SMS best	47	40.5
Telephone best	54	46.6
Equal Rank	14	12.1
TOTAL	114	98.3*

Table 5.9: Rank Order for Designs (General)

* Two participants did not rank both services due to having no experience of the SMS version during the experiment.

Splitting the rankings by SMS message format (Balance-only or Balance and mini-statement) and telephone mode (advisor or automated) the responses compared to the automated service are shown in Table 5.10; comparisons to advisor services in Table 5.11.

None of the differences were significant (Chi-squared tests), however the trend indicates that the longer SMS messages were slightly more likely to be preferred compared to the automated service. The advisor services were generally preferred to SMS messages, again however it was the longer SMS with recent transaction information which was more equally preferred compared to the advisor version.

	Balance-only SMS	**Balance + Mini-statement SMS**
SMS best	11 (39.3%)	16 (53.3%)
Automated best	14 (50.0%)	9 (30.0%)
Equal Rank	3 (10.7%)	5 (16.7%)
TOTAL	28	30

Table 5.10: Rank Order for SMS Variants vs. Automated service

	Balance-only SMS	**Balance + Mini-statement SMS**
SMS best	8 (29.6%)	12 (41.4%)
Advisor best	16 (59.3%)	14(48.3%)
Equal Rank	3 (11.1%)	3 (10.3%)
TOTAL	27	29

Table 5.11: Rank Order for SMS Variants vs. Advisor service

5.3.4. Qualitative Analysis

Participants were asked what they *liked* about the SMS balance-only version, 88% of comments were positive about the experience of using the SMS service to get their balance (only). Other participants either did not mention any particular likes, or made neutral remarks.

Most participant comments (54%) were about the speed of the service, being described as mainly “very quick” or “quick”. Another 38% referred to the ease of use, simplicity or how straightforward the service was. Some (27%) mentioned that the process could be used anytime, anywhere, whilst away or “on the move”. Other comments included that the service was perfect or impressive (5%); had a quick response (4%); was convenient (4%); private (27%); helpful (2%); avoids eavesdropping (2%); uncomplicated (2%) and informative (2%). However, some mentioned security concerns – and these were added to the dislike comments.

When it came to *dislikes*, some 57% of the comments included some negative aspects of the service. The most frequent concern was the security of the SMS service (45%). A total of 22 participants mentioned their security concerns, whilst another 2 mentioned privacy issues and one simply declaring they “wouldn't use [mobile] for transactions.”

Other dislikes were mentioned just once, these included that the service was impersonal, fiddly, it was a chore, that the return text would have to be deleted, that

it was very public for personal banking information, or that they would have to have a good memory for text commands.

The most often mentioned improvement concerned either making the service more secure by using some password or similar system: “Some kind of password security or code/PIN No”, “To be reassured it was secure, to ask for security details”, “Anyone could get hold of your phone and the details would be in your sent messages”, “Introduce 2 stage security system with a letter form the bank with password, but this negates the simplicity of it.”

Other participants simply questioned the security, wanting more information: “as long as its secure then its fine”; “More information needed about the security behind SMS”; more information needed on security behind text service”. Another suggestion for improvements was the addition of available balance information:

Participants were asked what they *liked* about the SMS balance plus mini-statement version, some 88% of the comments were positive about the experience of using the SMS service to get their balance (with a mini-statement). Other participants either did not mention any particular likes, or made neutral remarks.

Again, similarly to the balance-only text message, most comments (55%) described the service as “very quick” or “quick”. Another 40% referred to the ease of use, simplicity or how straightforward the service was. Some (14%) mentioned that the

process was handy or convenient, and another 14% specifically mentioned liking the recent transaction list along with the balance.

Other participant comments included that the service was private or discrete enough for use in public areas (12%), gave good and sufficient information (10%), had a quick response (9%); offered a written record (9%).

When it came to *dislikes*, some 64% of the comments made by participants included some negative aspects of the service. The most frequent concern was the security of the SMS service (31%). A total of 22 participants mentioned their security concerns.

Participants (12%) mentioned a dislike of texting, or preference for phoning (2%). Some participants thought using mobile phones was fiddly or cumbersome (7%). Other dislikes mentioned once only were: concern about remembering the codes, the cost of texts or the fact that the service options were very limited if you could only ask for balance information.

The most often mentioned improvement concerned either making the service more secure by using some password or similar system – the comments were very similar to those offered by the Balance-only SMS group: “Should have more security checks, maybe security number”, “Would be good to know more about the security”, “Use a pin number.”

Participants were asked what they *liked* about the automated telephone banking version, a total of 85% of the comments were positive about the experience of using the automated service to get their balance. Others either did not mention any particular likes, or made neutral remarks.

Most comments (31%) were about the ease of use, simplicity or how straightforward the service was. Also, often mentioned was how quick or efficient the service was to use (29%). Many (27%) mentioned that the process felt secure, private and safe. Other comments included that is was familiar (20%) being the current service many participants chose to use; some liked the voice and thought the voice clear (7%); thought it was direct (2%) and understandable (3%)

With participants *dislikes*, 37% of the comments included some negative aspects of the service. The most frequent concern was the how long the call took, including the length of the security process (12%), a few found the service confusing or frustrating (5%), or that the voice recognition did not work (5%). Other aspects mentioned just once or twice included that some didn't like speaking, preferring key-in; that the process seemed slow; they had no written record or that they didn't like the computerised voice; or preferred speaking to advisors in person. The most often mentioned improvement concerned making the call quicker: "Speed it up", "Could make the process shorter", "Make the process quicker."

Participants were asked what they *liked* about the advisor version of telephone banking, a total of 86% of the comments were positive about the experience of going

through to an advisor to get their balance. Others either did not mention any particular likes, or made neutral remarks.

Most participant comments (49%) were about preferring the personal touch of speaking to an advisor, finding the voice and the human aspect appealing. Also, regularly mentioned was how friendly and helpful the advisors were (25%). Some 23% mentioned their confidence in the security of the advisor process and a further 11% mentioned the service being polite. Other comments included that is was easy to use (7%) clear (5%); reliable (5%); quick (4%) and flexible (2%).

With participant *dislikes*, 42% of the comments included some negative aspects of the service. The most frequent concern was the how long security process took or how many questions they were asked (19%), a few found the service inefficient, too long a call for a straightforward Balance-on-demand (12%). A couple of people mentioned that remembering recent transactions was a problem for them (4%). One participant disliked voice recognition in the first part of the call. Other aspects mentioned just once included that the service was dated, complicated, inconvenient and couldn't be used in public places. The most often mentioned improvement concerned shortening the process: "Shorten security", "Shorten security questions", "Less verification questions"

A total of 96% of the participants agreed that the information provided by text messages was clear. This was not affected by whether or not they received the balance only or the balance and mini-statement text message.

Participants were asked if there was enough information in the text messages, a few made comments: 6 participants from the Balance-only format group suggested that they needed more information than was present in the text message – several suggested the current balance and the available balance should be included (none asked for recent transactions). For the Balance and Mini-statement group, no participants asked for more information. Instead, 8 participants thought this balance and mini-statement was a good enough amount of information, while only 2 participants from the balance-only group made a similar comment. 6 participants complained about the balance and mini-statement format that it had "too much" information, that a balance would be sufficient, or needed to be presented in bold next to the supplementary material. Finally, two participants mentioned that the Balance-only text message also contained too much information, that the actual balance was all that was required.

Participants were asked on average how many times a month customers might use such a service (imagining that it cost no more than the usual charge for sending a text message), the range went from zero to 30 times a month. The responses offered are summarised in Table 5.12.

Potential Usage Frequency	
Mean frequency	2.68
N (%) selecting less than mean frequency	69 (59%)
N (%) selecting more than mean frequency	45 (39%)
Modal frequency	0
N (%) selecting modal frequency (zero)	39 (34%)
Median frequency	1.5
N (%) selecting less than median frequency	54 (47%)
N (%) selecting the median frequency (1.5 times / month)	4 (3%)
N (%) selecting more than median frequency	56 (48%)

Table 5.12: Potential Usage Frequency

A total of 19 (16%) participants mentioned that they may not need the SMS Balance-on-demand service due to using other methods like Internet banking, ATM's and phone calls to get the same information, feeling they are familiar and work well for their purposes.

Some participants thought cost concerns might affect how frequently they used the service (12, 10%) mentioning that they would use it "if it were free", "depending on cost". Similarly, some mentioned that other services were free: "I can use the Internet for free and this service would cost", others would use it "if I couldn't get to a bank and it is cheaper than a call".

Participants would use the service weekly or monthly to track and check their bills/direct debits/salary payments and generally make sure they were not going overdrawn, or whilst shopping to ensure they had enough money to make a purchase

(16, 14%); as a back-up method if they couldn't use their preferred contact point or had an urgent problem with a lost card or fraud concern (3, 3%); convenience might influence frequency of use (2, 2%) for example when away or travelling (3, 3%). One customer already used text alerts and thought this was sufficient for them.

Other issues mentioned in relation to how frequently the service would be used concerned: security issues which might prevent them using the service (8, 7%).

Participants were asked about circumstances where they might use SMS requests instead of telephone banking. Table 5.13 shows a summary of the types of response and the frequencies. A minority of participants could not think of any circumstances where they would use SMS in preference to telephone banking (9, 7.8% of cohort).

SMS instead of Telephone Banking	N	% of responses
Away / out / abroad / travelling	48	35%
Check balance prior to purchasing / shopping card problems	18	13%
In public /noisy places	17	12%
For speed / in a hurry	15	11%
No phone / PhoneBank / Internet option available	13	9%
Due to privacy / no speaking / no eavesdropping	11	8%
Emergency / last resort only	3	2%
No ATM / Bank nearby	2	1%
To check payments / pay day / DDs / cheques etc.	2	1%
Anytime	2	1%
Security concerns / if secure would use	2	1%
Other: short enquiries, after Banks close, at home, at work, if lost voice.	6	4%
TOTAL	139*	

Table 5.13: Circumstances where SMS may be used instead of Telephone Banking

* Some participants gave more than one category of response, resulting in this total.

Table 5.14 shows the mean scores (0-30 scale) of how likely participants thought they would be to use various service options to get their current account balance, where zero represented very unlikely to very likely at 30 points; scores above 15 indicate generally positive likelihood of usage whereas scores below 15 indicate that participants would be less likely to use these channels.

Service Option	Mean
Advisor	13.47
IVR	17.46
SMS	17.39
Internet Banking	21.42
ATM	21.62
Branch	11.61

Table 5.14: Usage Intentions – Balance requests via service channels

Comparing the rank of the proposed SMS service offerings to Advisor and IVR telephone options, the data is presented in Table 5.15.

Comparison of Ranks	PhoneBank Version	
	Advisor	Automated
SMS < Telephone	42 (36%)	51 (44%)
SMS = Telephone	2 (2%)	6 (5%)
SMS > Telephone	72 (62%)	59 (51%)

Table 5.15: SMS Usage Intentions c.f. Advisor and Automated

More participants would be likely to use the SMS service for balance enquiry than they would to use the advisor version of telephone banking, $p = .011$ (Chi-squared test on those ranking the two differently). However, for the automated version, there was no difference in the two groups ranking either SMS or automated above the other option.

The proportions favouring SMS to other telephone banking options did not depend on the SMS format experienced in the experiment.

As a final question, participants were asked what types of services or information they would like to be able to request using an SMS message. Responses were given on a 7-point scale, coded so that "definitely wouldn't use" scored 1 and "definitely would use" scored 7 points; 4 representing the neutral point on the scale. Table 5.16 shows the 16 service offerings in descending order of popularity. The table shows the mean score for each option and the percentage of participants who indicated that they "would definitely use" each option (e.g. score of 7-points on the 7-point scale).

SMS-on-demand: Service Options	Mean	% *would* definitely use
Request a balance	5.63	53%
Request mini-statement (6 transactions)	4.93	34%
Find nearest Lloyds TSB branch	4.70	27%
Request call back from bank	4.63	29%
Request full paper statement	4.03	21%
Request new cheque book	3.68	19%
Transfer money between own accounts	3.61	17%
Block a debit or credit card	3.60	25%
Change overdraft limit	3.33	16%
Standing order cancellation	3.30	17%
Request brochure, e.g. insurance, mortgage	3.25	11%
Stop a cheque	3.20	13%
Pay a bill	3.18	11%
Request insurance quote	2.90	9%
Request new PIN for bankcard	2.89	14%
Transfer money to third party accounts	2.83	10%

Table 5.16: Participant Requirements for SMS-on-demand services

Requesting a balance using SMS was the most popular option, scoring well over 5-points on the 7-point scale, with some 53% of participants saying they would definitely use this service and only 7% thinking they would not use it.

Mini-statements also were a popular option, scoring nearly 5-points on the scale, with some 34% definitely planning to use this service option; only 12% thinking they would not use it.

Branch location and call-back requests were also scored positively in terms of customer requirements but with only just over 25% of customers definitely planning to use such options, and less than 20% saying they would not use them.

Other options were less favoured, such as: *block credit/debit cards*, *changing overdraft limit*, *cancelling standing orders*, *stopping cheques*, *requesting insurance quotes*, *requesting new PIN for Bankcards*, or *making third party transfers*, with more than a third of customers saying they would not use an SMS request service for these types of banking tasks.

Participants were asked whether they would use an SMS request service such as the Balance-on-demand service they experienced, participants were offered four options: to use it if it cost no more than their standard network charge for sending SMS messages; use it and were willing to pay a small charge (to the Bank as well as the network charge); wouldn't use it or don't know. The results are shown in Table 5.17. Very few were willing to pay, the majority would use the service if it cost the standard charge they pay to send SMS messages, this was a significant bias, $p < .001$.

	N (%)
Use if cost standard SMS charge	59 (55%)
Would pay extra cost to Bank	8 (7%)
Wouldn't use	31 (29%)
Don't know	9 (8%)

Table 5.17: Use/Cost Issues

5.4. Discussion

It was predicted that there would be differences in the effectiveness of the two banking channels, and this proved to be correct with the telephone banking channel

having a significantly higher first time completion rate, though completion rates overall (three attempts) were high for both channels: SMS 91.5% and Telephone 94.1%. It should also be noted that these participants were recruited as telephone banking customers, and in this context the completion rates for the SMS service are comparable. It was also predicted that there would be a difference in satisfaction and again, overall, the telephone channel was rated significantly higher on the usability questionnaire and for the quality rating metric. It had lower scores for *confusion*, *flustered*, *stress*, *concentration* and *improvement needed*, higher scores for *knew what to do next*, *would use again*, *reliability*, *friendly*, *polite* and *confidence in security*. Scores for both channels overall and by version were above the neutral point indicating good usability. Converting the quality rating into a rank order preference did not though indicate a clear preference between the telephone or SMS channels.

Regarding the efficiency of the services, a large percentage of the positive comments regarding the SMS channel and versions focussed on the speed of the service, while for the advisor telephone service only a small percentage (4%) said that the speed of the service was something they liked about it, while 42% of the negative comments centred on frustration with how long the security process took to complete for an enquiry as simple as a balance request. This was also a concern for the automated service, but to a lesser extent. Comments made by the participants when making the quality comparison of the services seem to indicate that one of the main reasons driving their rating of the SMS services was speed and efficiency. Interestingly, there

was also no significant effect for the item on efficiency in the usability attitude questionnaire.

That the telephone channel scored higher for usability for a balance request may not be too surprising. The telephone is a familiar and mature technology in the banking sector, while the use of SMS is relatively new in banking. Reviewing the qualitative data it indicates that the most frequent concern over an SMS banking service was perceived security and this may be the main obstacle to overcome before SMS is considered as usable as a telephone banking service. Security has been shown to be a key factor in customer acceptance (O'Gorman, 2003; Schultz et al., 2001) in electronic banking, and particularly in mobile banking (Brown et al., 2003; Laurn and Lin, 2005). Though some have argued that security concerns (Laukkanen, 2007; Laukkanen and Lauronen, 2005; Suoranta, 2003) are not a reason for low adoption of mobile banking. Laukkanen (2007) argues that it is a feeling of uncertainty about making errors due to the size of the screen and slowness of access rather than the actual data security of the system, which is related more to the usability of mobile phones. This study suggests with SMS banking it is concerns with the actual security of the medium (data security) that is a major factor to the participants. Weir et al., (2009a) showed that customers were driven by their attitudes towards usability and convenience rather than security. SMS banking is more convenient than telephone banking, but it seems a combination of usability problems and security concerns are major factors in the low adoption of SMS banking. More research should be carried out to investigate the relationship between these two factors in the context of SMS banking.

A concept tied to security is trust, and participants will trust a telephone banking service more due to previous experience with it, and because it is an established banking channel. Trust (Kim et al. 2008; Lee and Chung, 2009) is an important factor in usage intentions of mobile banking. Lee and Chung (2009) argue that trust is the variable that most effects customer satisfaction with mobile banking, and that interface design quality is not as important a factor in building trust compared to information quality and system quality. As discussed previously, interface design quality for SMS banking is beyond the vendors control and is dependent on the model of mobile phone the customer owns. Interface design quality is obviously an aspect of usability, but most users will be experts in writing text messages on their own mobile phones. In this study where a generic mobile phone was used, the process of using SMS banking to get an account balance was found to score lower in usability compared to telephone banking. The participants in this study seemed to be relatively happy with the information quality of the SMS service but not the security, which is considered to be an aspect of system quality (Lee and Chung, 2009; Delone and Mclean, 1999). This study only looked at first time use of the SMS banking service for requesting a balance request, the participants had used either automated or advisor telephone services before, so with more use of the SMS service it could be argued that trust in the service would increase, which would impact on satisfaction. This may lead to SMS banking comparing more favourably towards a telephone service. This could be investigated in a more longitudinal study investigating user behaviour over an extended period.

The usage intentions of participants were also examined and it was found that more participants would be likely to use the SMS service for a balance enquiry than the advisor version of telephone banking, but they would still be more likely to use Internet banking or an ATM over SMS banking for this task.

The comparison of information tested on the SMS banking service (Balance only vs. Balance + Mini-statement) did not reveal any main significant effects for the usability satisfaction questionnaire or the quality measurement, though the Balance only message was less *confusing* and *stressful*. This is probably due to the visual affect of the extra information on the mobile phones small screen making it harder for the participant to immediately work out what is relevant to them. There is no strong evidence to pick between either version of the service, but it seems the extra information included in the mini-statement would not be missed by the customers if they requested their balance. This was a between-participants variable, so further work could be carried out whereby participants are shown both versions and allowed to state a preference.

Only one type of form factor was used in this experiment for the mobile device, along with the commonest text input method for mobile phones. Performance with SMS banking could vary slightly with different types of handheld devices, and with different text input methods. This should be explored in further research.

As for the results presented in Chapter 4, differences between the age groups were discovered. The older age group had lower first time completion rates with SMS

banking, and gave the traditional telephone banking much higher scores for satisfaction and quality compared to the SMS service. As discussed in Chapter 4 older users can suffer from visual, cognitive and motor impairments to varying degrees (Christopher, 1999; Jagacinski et al., 1995; Krampe and Ericsson, 1996; Kurniawan et al., 2006a; Walker et al., 1997), and have been shown to have problems with mobile phones (Ornella and Stephanie, 2006). It is argued these factors would impact on the usability of an SMS banking service for older users. This is due to the fact that mobile phones are not generally designed with the older user in mind. The difference in results here can also be attributed to these usability problems, and a negative relationship was discovered between overall usability attitude score for the SMS service and age.

The SMS usage patterns and experience levels are also a probable factor in the differences in scores, similarly to chapter 4. The older age group in general were found to send SMS less frequently than the younger (see Table 5.3), which follows the findings in chapter 4. The difference here to Chapter 4 was that the SMS banking service was being compared to a telephone banking service. There was a positive relationship between age and quality rating of telephone banking found in the experiment, and it is plausible that older users are more positive towards the traditional telephone banking, as it has been argued that older users are less positive in general to mobile phones (Ling, 2007). Older people are active users of the landline telephone, and older people also value personal contact highly (Blythe et al, 2005). It is also argued that older people do perceive a mobile phone's main function as being a communication device (Kurniawan, 2008), but rather as a safety

line. The study reported in this chapter looked only at balance requests via SMS banking compared to telephone banking. If further types of transactions were compared, such as funds transfers, it is arguable that the differences between age groups would be even more pronounced.

5.5. The SMS *Push* Confirmation Experiment

5.5.1. Experiment Design

To enable the participants to compare the different methods of funds transfer confirmation they were given simple background scenarios and tasks, where they had to undertake funds transfers between accounts held by the bank. For the telephone banking service a 'mirror' version of the actual automated telephone banking service employed by the bank was used. It was enhanced with the new SMS confirmation capability, and participants received messages on a Sony Ericsson K800i handset with a 'standard' keypad.

The experiment had three within-participants conditions: (1) no confirmation, (2) confirmation by updated balance and (3) confirmation by SMS. There was also a between-participants variable, with half of the participants in condition 3 (SMS) ringing the no confirmation version of the service (1) and half ringing the confirmation by updated balance version (2). Participants experienced all three within-participants conditions in a repeated-measures design.

The usability of each version of the service would be measured by the following independent variables: responses to individual items in a usability attitude questionnaire and quality rating with deduced preferences. For this experiment it was not deemed appropriate to time the participants' performance as a measure of efficiency, or compare completion rates for effectiveness as the process of the funds transfer would be the same.

The two between-participants factors were age (two groups) and gender. The experiment was designed to have the order of experience of the independent variables balanced across the participants of differing ages and genders. This was to ensure that each group would experience all possible orders equal amounts of time.

5.5.2. Telephone Banking Service

The existing automated telephone banking dialogue confirms a transaction with:

"Thanks, your money has been transferred".

The existing service was enhanced with the new a new IVR capability, for the confirmation of new balance, the following prompt was used:

"The balance of your current account is now…."

For the confirmation versions that involve an SMS text message two new prompts were recorded:

"The balance of your current account is now….and we will send you a confirmation of this transaction by text message."

And

"We will send you a confirmation of this transaction by text message."

Each participant experienced one of these two prompts in a balanced experiment design. The inclusion of these new IVR prompts independently offered data on participant responses to both the addition of a spoken updated balance in the IVR and the addition of the SMS text message notification.

An example SMS funds transfer confirmation is shown in Figure 5.3.

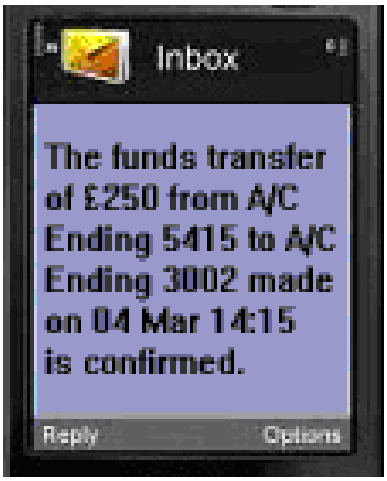


Figure 5.3: Example of SMS Transaction Confirmation

5.5.3. Scenario and tasks

Participants were told to imagine that they were customer of the bank and complete a task like the example below:

You want to transfer some money from your Lloyds TSB current account to your friend, Paul Williams's account. Call Lloyds TSB and use the automated service to transfer £300 to Paul Williams's account.

The participants were given all of the account details and information they would need to complete the tasks.

5.5.4 Usability Attitude Questionnaire

The design of the usability questionnaire used for this experiment followed standard practice (Likert, 1932) by using an equal number of negative and positive statements presented in a randomised order. The questionnaire used a 7-point Likert format that ranged from "Strongly Agree" (1) to "Strongly Disagree" (7). Following reversal of the polarity of positive questionnaire statements, a score of 7 consistently indicates a strongly positive attitude and 1 a strongly negative attitude. The questionnaire used consisted of 22 statements that address a range of usability attributes pertaining to human-computer interaction (HCI): *cognitive attributes* (level of concentration and degree of confusion), *the fluency and transparency of the service* (knowledge about what is expected, ease of use, degree of complication), and *quality attributes*

(efficiency of service, amount of improvement service is felt to require, reliability of service). The 22 questionnaire items are shown in Table 5.18.

1	I thought this service was too complicated
2	When using this service I didn't always know what I was expected to do
3	I thought the service was efficient
4	I would be happy to use this service again
5	I found this service confusing
6	The service was friendly
7	I felt under stress using this service
8	I thought the service was polite
9	I enjoyed using this service
10	I found this service frustrating to use
11	I felt flustered using this service
12	The service was easy to use
13	Using the service took a lot of concentration
14	I did not feel in control using this service
15	I feel this service needs a lot of improvement
16	I felt confident in the security of this service
17	I felt that the service was reliable
18	The service was too fast for me
19	I liked the voice
20	I would prefer to talk to a human being
21	I thought the voice was very clear
22	I felt confident my transaction was completed successfully using this service

Table 5.18: Usability Questionnaire Items

The questionnaire consists of the same core statements used in the previous experiments, with a few additions (Q18-22) specific to measuring the usability of automated telephone services, and one regarding confidence in completion of transaction. These changes were agreed upon by a panel of experts. The panel consisted of usability experts from the University of Edinburgh who all had extensive experience in questionnaire design.

All participants completed the usability attitude questionnaire following exposure to each of the three experimental conditions. The questionnaire was presented to the participants on a laptop computer and the order of questions was randomised for each participant. The mean scores for these usability attributes enabled direct comparisons of participants' attitudes towards the three different message formats.
Overall usability was determined by determining the mean for all the usability questions by all participants. Individual statements were also analysed separately to identify any specific usability issues that arose from the hands-on usage sessions.

5.5.5. Quality Metric and Exit Interviews

After experiencing both services, and completing the usability questionnaires, participants made a quality rating, recorded as a value on a 30-point linear scale labelled "poor" at the 0 end and "excellent" at 30. This quality rating involved evaluating both services against each other on the linear scale, and was also recorded to indicate an explicit preference, or no preference. The quality rating is a subjective satisfaction measurement, but unlike the usability questionnaire a participant is specifically asked to compare both services against each other which can be a useful result to compare with the questionnaire scores. If there is a discrepancy between the two, then the qualitative data from the one-to-one structured interviews can be consulted for an explanation.

In a semi-structured exit interview participants were given an opportunity to comment on their experiences and give their opinions on a range of related issues in closed type questions.

5.5.6. Procedure

The participant was given a task sheet and scenario with the instructions on for the first randomly assigned version of the service, which the researcher also read out aloud to them. The participant had already been introduced to the mobile phone in the first part of the experiment. They were allowed three attempts to complete the task. If they failed three times the researcher stepped in and helped them to complete the task. The participant then completed the attitude questionnaire. The researcher then gave the participant the next scenario and instruction sheet and the procedure was repeated until the participant had experience all three version of the service.

At the end of the session the participant completed a one-on-one 'exit' interview and were also asked to rate the overall quality of the three different version on the 30-point linear scale. The experiment was concluded by completion of a demographic questionnaire to collect factual data about the participant and their banking and mobile phone usage.

5.6. The SMS *Push* Confirmation Experiment: Results

5.6.1 Usability Attitude

The scores for the 22 usability attributes were averaged to obtain an overall usability score for each version. The mean overall usability scores with standard deviations are shown in Table 5.19.

Version	Mean (SD)
Current (N=110)	5.35 (0.43)
Current + Balance Update (N=110)	5.36 (0.44)
Current + SMS Confirmation (N=56)	5.45 (0.36)
Current + Balance Update + SMS Confirmation (N=52)	5.42 (0.45)

Table 5.19: Usability Scores for Funds Transfer Processes

The SMS confirmation scored the highest mean usability score, but this was not a statistically significant effect. A repeated measures ANOVA was performed on the means using age, gender and type (SMS or SMS + balance update) as between-participants factors. There was no main effect (d.f. = 1.96, $F = 0.746$, $p = 0.475$) or any significant interactions due to the main between-participant factors.

T-tests were used to analyse the data for the 22 individual attributes. Comparing the SMS confirmation to the current version of the service showed three attributes were statistically significantly higher ($p < 0.05$) for the SMS version, *confident completed*, *voice clarity* and *efficiency*. Comparing the SMS confirmation message with updated

balance to the current version using t-tests found that this SMS version scored lower for the attribute *friendly* ($p = 0.022$). Comparing the updated balance prompt version to the current version using t-tests found only one attribute to be significantly higher for the updated balance version, *ease of use* ($p = 0.033$).

5.6.2. Quality and Preference

To collect a quality rating for each version of the funds transfer process participants were asked to rate them comparatively on a 30-point scale. Table 5.20 shows the mean and standard deviation of the quality ratings awarded to the alternative version of the funds transfer process.

Version	Mean (SD)
Current (N=109)	17.34 (7.363)
Current + Updated Balance (N=109)	21.77 (5.231)
Current + SMS Confirmation (N=51)	22.27 (7.955)
Current + Updated Balance + SMS Confirmation (N=58)	23.04 (6.321)

Table 5.20: Quality Ratings for the Alternative Funds Transfer Processes

There was no significant difference between the two SMS versions (univariate ANOVA). T-tests were performed and found that the IVR with an SMS confirmation message after the funds transfer scored significantly higher than the current IVR design, $p = 0.014$. There was no statistically significant difference between the version with the updated balance and either SMS version. The updated balance

version scored significantly higher than the current version, $p < 0.001$ (paired samples t-test). There were no age or gender effects.

Positive comments regarding the current IVR service comprised some 41% of all remarks made by participants about this version, that it was “quick and short”, “concise, didn’t ask for balance and didn’t get one so no problem”, “quick and easy and you can chose another service rather than it giving you your balance right away”, “did exactly what I wanted it to do”. Another 48% of comments about the current service were negative: “no balance”, “too short and not enough information”, “too basic”.

Participants almost all gave positive comments about the updated balance prompt, with 87% of remarks positive about this version. The negative remarks included: “balance added but not needed, went in one ear and out the other.”, “[in real life] would have hung up before hearing end of message for text or balance”, “too much info”, “didn’t get a text”.

Similarly positive remarks were generally given regarding the SMS confirmation (without updated balance prompt), with 83% of comments positive. Those who had negative comments included: “didn’t need to wait for text”, “text is unnecessary”, “too much info”.

For participants with both the updated balance prompt and the SMS, 77% gave comments positive – but those who gave negative comments thought it was

"longwinded", "too much info" and "too long", some "would not want transaction by text message".

Reducing the quality rating into a rank order results in the data are presented in Table 5.21.

Version	Current	Current + Balance Upadate	Current + SMS	Current + Balance + SMS
Best	14 (13%)	35 (32%)	34 (67%)	34 (63%)
2nd Best	25 (23%)	60 (55%)	4 (8%)	17 (29%)
3rd Best	70 (64%)	14 (13%)	13 (12%)	7 (12%)
Total	109	109	51	58

Table 5.21: Rank Order for Designs (Split between three experiences)

Several participants ranked options equally and one participant did not complete this question, resulting in higher than 110 responses. Equal ratings were all awarded to the Balance Update and whichever SMS confirmation option had been experienced. The ranks for the SMS confirmation options involved smaller number of participants due to the between-participants comparison. There was very little differences between the proportions ranked best for the two alternative implementations of the SMS confirmation version. More people rated the version with both the Balance Update prompt and SMS confirmation second than those who just had the SMS confirmation.

Table 5.22 shows the split between the three options each participant experienced, showing a clear preference for the SMS confirmation options (both with and without the updated balance prompt).

Ranked Best overall	N	%
Current service	14	12.8
Current + Balance Update	35	32.1
Either SMS Confirmation option	68	62.4

Table 5.22: Rank Order for Designs (Split between three experiences)

5.6.3. Qualitative Analysis

Participants were asked about the SMS confirmation after the call and (67%) were positive about the idea commenting that it was a "good idea", "quite helpful" or "excellent". Most mentioned either that they liked the message because it provided extra confirmation and reassurance that the transfer had taken place, others talked about liking the written record which could be referred back to rather than the verbal prompt.

18 (18%) participants had negative views about the text message confirmation. Mostly, these were security concerns: "Didn't like it, if someone else has my phone, they'll know how much money I have", "I would not want that detail on my phone", "Good, but, some privacy issues if someone looks at your phone."

Others participant thought the text message was unnecessary, "Alright, but, didn't see the need", "Not necessary, could get while on the phone", "Not fussed about getting text", "Pointless", "Too cumbersome".

Asked about the Balance Update prompt, most participants (75%) were positive about the addition of this information commenting that it was a "good idea", "useful", "informative" and a "good and helpful feature". Some 16 (15%) also specifically mentioned being reassured that the transaction had gone through as planned.

Some 7 (7%) of the participants thought this prompt was unnecessary, commenting that it "feels irrelevant if you didn't want to know it", "unnecessary…slowed things down", "not of necessity to me, only towards the end of the month when lots of transactions happening", "I would not have held on the phone to hear that", "not necessary - if you wish to know balance you can ask". Several others disliked the idea for various other reasons.

A total of 99% of the participants agreed that the information provided by text messages was clear. Only one participant thought the information needed to be made clearer, others commented: "Security by only giving part of account number but enough to know it was from your account" and that although it was alright, it "could be a bit clearer". When asked if there was enough information in the text messages, most (93%) agreed.

Asked whether the SMS confirmation should be automatic, or whether they would like the choice to get the confirmation message, 66 (62%) participants wanted to be offered the choice. Another 38 (36%) wanted it to be automatically sent via the

automated telephone service. This was a bias in favour of the choice, $p = .008$ (Binomial test).

When asked if they wanted to have the option of entering the phone number they'd like to receive the SMS confirmation, 52 (50%) participants wanted to be able to do this. They commented that they may have multiple phones, or have changed, lost or had their phone stolen and therefore needed to be able to enter the correct details and ensure the texts arrived to the correct phone: "so you can chose which phone it goes to", "because you do not want your information going to any other phone", "you might have changed your number or your phone might be broken", "in case you change phone number" or "if the banks records are not up to date, change phone number". Some also commented that this option would be *preferable*, *for ease* and *helpful*.

However, nearly as many participants commented that they did not think that this would be a good idea, mainly their concerns were about security, about repetition and lengthening the phone call, that it would make the process *too complicated*, *difficult,* "it would slow it down, a hassle".

Offering the choice of sending the SMS message and inputting the mobile number brought up some additional security issues, such as how secure it would be to input the phone number at the same time as performing the transaction, and how long this would take, how prone to errors the process might be. The option of pre-registering a phone number and alternative numbers might be worth investigating. In addition

there is scope for customers to individually indicate their preferences as to whether texts would be automatic or offer a choice of confirmation process. These options show that there is scope for further research on how such services might be implemented and used securely in practice, and that generally, customers needed information on the security of using these services with their mobile phones and text messaging.

Participants were asked on average how many times a week would be acceptable to receive an SMS transaction confirmation, only 7 participants (7%) were happy to get these text messages as required, “After every transaction”, “Depends how many transactions go through”, “Depends how many transactions you have made”, “As many as required” and “No problem with any quantity”, “I would like a text message confirming any transaction on my account regardless of how many per week”.

Other participants gave values or ranges of options. The overall range went from zero to 14 times a week. The responses offered are summarised in Table 5.23
Only 9 (9%) thought they would use the service less frequently than once a week, all except one didn’t think they would use the service at all, with one person commenting “I only access my account once per month.” The most frequently selected option was once per week, with 40% of participants choosing this option.

Potential Usage Frequency	
Mean frequency	2.5
N (%) selecting less than mean frequency	69 (66%)
N (%) selecting more than mean frequency	27 (26%)
Modal frequency	1.0
Median frequency	1.0
N (%) selecting modal/median frequency (1/week)	42 (40%)
N (%) selecting less than modal/median frequency	9 (9%)
N (%) selecting more than modal/median frequency	49 (47%)

Table 5.23: Potential Frequency of Use per Week

Participants who selected once a week or slightly more frequently were concerned that they might get confirmation of any and all their banking activity, including some (13, 12%) selecting twice a week and some of those selecting up to 7 times per week.

Other participants wanted the texts to ensure they "keep close contact with my account", "after every transaction". Some customers selecting between three and fourteen times per week were more willing to receive multiple messages about transactions.

Table 5.24 shows how participants would prefer to receive transaction confirmations for various services (the confirmation channels offered were by SMS text message, letter, phone call or no confirmation).

Most participants liked confirmation to be via SMS or letter. Letter confirmation was slightly more popular overall, and more preferred for *change of address confirmation, complaints, PIN changes, application progress updates* and *setting up new arrangements*.

SMS confirmation was the preferred option for *funds transfers, change of phone number confirmation, large credits or one-off debits, overseas transactions and debit card purchases over the Internet.*

Confirmation Preference	SMS	Letter	Phone	None
After you make a funds transfer	74 (67%)	11 (10%)	6 (5%)	*20 (18%)*
After you contact Lloyds TSB to tell them about a change of address.	26 (23%)	80 (71%)	3 (3%)	3 (3%)
After you contact Lloyds TSB to tell them about a change of phone number.	49 (43%)	44 (38%)	15 (13%)	7 (6%)
After you have raised a complaint with Lloyds TSB and wish to receive progress updates	13 (11%)	71 (62%)	30 (26%)	1 (1%)
After you have contacted Lloyds TSB to request a change of PIN	27 (24%)	71 (64%)	8 (7%)	5 (5%)
After you have filled in a form such as a loan or mortgage application and wish to receive progress updates	18 (16%)	62 (55%)	31 (27%)	2 (2%)
When a large payment has been credited into your account.	52 (46%)	35 (31%)	14 (12%)	13 (11%)
When a large one-off payment has been debited out of your account	57 (50%)	28 (25%)	22 (19%)	7 (6%)
Setting-up a new payment arrangement, e.g. Standing Order or Direct Debit	33 (29%)	52 (46%)	17 (15%)	10 (9%)
When there has been an overseas transaction on your account	64 (56%)	24 (21%)	21 (18%)	5 (4%)
After a debit card transaction over the Internet	62 (55%)	18 (16%)	6 (5%)	*26 (23%)*
TOTALS	*475*	*496*	173	99

Table 5.24: Circumstances where SMS may be used instead of Telephone banking

* Some participants gave more than one category of response, resulting in different totals for each question. Percentages are based on total responses for that service.

5.7. Discussion

It was predicted that there would be differences in usability attitudes towards the three different versions of the telephone banking service, but this did not prove to be the case. Based on the attitude questionnaire results overall, usability was considered to be the same whether the participant received an SMS confirmation after a funds

transfer or not. Overall, SMS confirmations did increase scores for the attributes *confidence completed*, *voice clarity* and *efficiency* compared to the existing version (no confirmation). This was not the case for the updated balance version. That is not the end of the story though, as when measured with the quality metric the existing version (no confirmation) was rated significantly lower than the SMS confirmation versions and updated balance version. There were no differences between the SMS confirmation versions and the updated balance version of the service, but reducing the quality ratings to rank order preferences indicated a preference for the SMS versions of the service. It is not just usability that is important to users. The participants might have judged the usability of the different versions of the telephone banking service to be the same, but they still preferred a version that confirms the transaction to one that does not. This could point to the participants perceiving this version to be more useful.

From analysing the qualitative comments made by the participants it is clear that some form of confirmation is considered to be a good idea, whether it is an SMS confirmation or an updated balance. A large majority of participants were positive about an SMS confirmation and the confidence and reassurance it gave, but this was similar picture for an updated balance in the call. A small minority of participants expressed concerns about security regarding SMS confirmations, similarly to concerns over the use of SMS for balance requests in part one of the experiment. The scope for CRM and the ability to send SMS confirmations after other types of banking transactions can be seen as one of this channel's strengths. This was addressed by a set of questions in the exit interview. Most people would like

transaction confirmation to be via SMS or letter, with letter confirmation slightly more popular overall. SMS confirmation seemed to be most popular for monetary transactions on an account, were instant confirmation is required. Letter confirmation might be slightly more popular, but examining the data more closely it seems that most people still prefer this traditional communication for more complicated CRM issues, such as complaints, change of address, mortgages etc. A recent survey found that that customers still prefer traditional print mail compared to email communication from the bank (InfoPrint, 2007), and more would swap back to print if it had less of an environmental impact. Liljander et al. (2007) found that customers for an airline had a 'take it or leave' it response to mobile CRM. The results here seem to suggest that for transactions involving payments SMS feedback is desirable. This is likely to be because of the speed and convenience of SMS, for example the customer receives a confirmation after an overseas transaction on their account.

The comments made by the participants also suggested issues to be resolved regarding SMS communication. Only a tiny percentage of participants thought that they should receive SMS confirmations after every transaction because it could otherwise become an annoyance, e.g. after every ATM or debit card use, which could number several per day. As discussed above, it seems that participants would require SMS confirmation only after certain transactions. The majority of participants also thought SMS confirmation should be optional.

Interestingly, no interactions for age were found, unlike in part one of this experiment and in Chapter 4. This may suggest usability issues only arise when older

participants have to interact with the mobile handset more and instigate the transaction in a *Pull* service. For a *Push* service instigated by the bank the older participants do not have to interact so much by writing and sending an SMS, they just receive the SMS. It has been argued that some older users of mobile phones can panic when they receive an SMS because they find it intimidating to retrieve the message (Kurniawan, 2008), but this was not indicated in the usability attitude data. Also generally, confirmation like this from the bank may appeal to older participants which would make them more positive towards the use of SMS in this context. The use of SMS in this context is also more similar to SMS use in general, whereas SMS balance requests, or SMS transfer requests are a more novel use of SMS. Security may also seem less of a concern in a *Push* SMS from the bank.

5.8. Discussion of Chapter 5

This chapter details the results from a two part experiment to examine integrating an SMS channel into a multichannel banking environment as a 'next call avoidance' strategy. Both *Push* (automatic) and *Pull* (customer initiated) functionality was considered. The first part compared a *Pull* SMS banking service to the existing telephone banking service for balance requests. The second part investigated integrating SMS into the telephone banking channel by way of funds transfer confirmation messages: a *Push* service. The aim of the first part of the experiment was to compare the usability of an SMS banking service against an existing service in the form of telephone banking. The aim of the second part of the experiment was to investigated if an SMS confirmation after completing a funds transfer using telephone banking would improve the usability of this channel. The third aim of this

chapter was to gather data from participants as to the acceptability of integrating an SMS channel into the bank's existing services, and to investigate if potential customers would use it.

Based on a usability attitude questionnaire it was found that the telephone banking service was considered to be more usable for balance requests compared to the SMS banking service, but these results should not be too surprising as participants were experiencing the SMS service for the first time. The telephone service was also judged to be higher for quality compared to the SMS service. Converting the quality rating into a rank order preference did not indicate a clear preference between the telephone or SMS channels though. It was also found that more participants would be likely to use the SMS service for a balance enquiry than the advisor version of telephone banking, but they would still be more likely to use Internet banking or an ATM over SMS banking for this task. Practically, using SMS banking to request an account balance will generally be a much faster process (taking into account reception and coverage issues), and there were many positive comments regarding the speed of the SMS service, and many positive comments in general, but the most frequent concern about the SMS service was its security. It was argued that concern about security and trust in the SMS service played a role in its lower ratings, and that this is a significant factor in the low adoption of SMS banking. With extended usage these concerns may diminish. If a large potential customer base is targeted for the uptake of *Pull* SMS banking services and mobile banking in general, rather than just the technologically savvy or younger customers, then customer concerns about trust and security should be addressed by banks.

Differences in scores between the age groups were also discovered with the older age group being less positive to the SMS service compared to the younger group. These results follow the pattern found in Chapter 4 and add more evidence to the argument that an SMS channel will be used less by older customers due to usability issues (perceived and real).

In terms of usability, there was no difference between a telephone service that sends you an SMS balance confirmation after a funds transfer and one that did not. Each was as easy to use as the other, which was to be expected. The metric that measured the difference between the versions of the service better was the quality rating metric, which showed that the service without any type of confirmation was rated much lower than the versions of the service that did confirm the transaction in some manner. A large majority of participants were positive about an SMS confirmation and the confidence and reassurance it gave, but this was similar picture for a version that gave an updated balance confirmation during the actual call. These findings suggest that using SMS for CRM and as a medium to confirm banking transactions has possibilities and would be welcomed by customers. It was discovered that the type of transactions that participants believed would suit an SMS confirmation were monetary transactions on an account, where confirmation would be beneficial, for instance a debit card payment over the internet. In this case the SMS confirmation would help against fraudulent transactions.

There were no differences in scores found between the age groups for part two of the study. This may suggest usability issues with SMS banking only arise when older

participants have to interact with the mobile handset more and instigate the transaction in a *Pull* service.

In both parts of the study there were many positive comments regarding the use of SMS in a banking context, but there were also a few issues for further research that may prevent widespread adoption by customers. Only a tiny percentage of participants thought that they should receive SMS confirmations after every transaction because it could become an annoyance, e.g. after every ATM or debit card use, which could number several per day. Just over half of the participants would use an SMS service to request a balance on their account if it was introduced. Very few though, would be willing to pay, the majority would only use the service if it cost the standard charge they pay to send SMS messages. The mean potential usage frequency was 2.68 times per month for a balance request using the SMS service. This figure may potentially increase with extended use, and is likely to be higher in sub groupings of customers who are required to check their balance more frequently e.g. small business owner.

Chapter 6

6. The Usability of SMS as a 2-factor Authentication Process for Internet Banking Transactions

6.1. Introduction

This chapter discusses the results of an empirical investigation into the usability of SMS for 2-factor transaction authentication as a one-time passcode (OTP) generator. In the context of the authentication of an Internet Banking funds transfer, SMS is compared to a stand alone OTP generator device, a card reader device and the existing password method of authentication. The aim of the experiment was to compare the usability of SMS OTP generation to the existing password method and the alternative, competing devices for this type of online 2-factor authentication.

6.1.1 Background

Banks have been urged to introduce 2-factor techniques for authentication by the Financial Services Authority (Computer Fraud and Security, 2006) and Federal Financial Institutions Examination Council (FFIEC, 2005), and to replace or augment the single factor password method. The low security of knowledge based password verification has caused fraud problems (Sinclair and Smith, 2005). The password strategy of customer authentication is vulnerable to phishing and keystroke capture. There has been an obvious increase in these types of attacks by fraudsters to coincide with the increase in Internet banking usage (Hole et al., 2006). A major problem with passwords is that people tend to use the same password for many

different services and applications, so when one is compromised the security threat may not just apply to one website or service. For banks the loss of passwords can result in lost business and increased costs in customer support (Brown et al., 2004). It is these factors that are influencing banks to adopt 2-factor authentication.

2-factor devices, including OTP tokens and card readers, have been trialled and rolled out to customers by several UK Banks (BBC News, 2007b; Mathieson, 2007; Moneyextra.com, 2008), though there has been some argument as to how secure the devices actually are (Drimer et al., 2009; Flinders, 2009). For the financial sector, and from an ecommerce point of view, technology is the driver in producing secure solutions to customer authentication. But it is important that usability issues are not overlooked in the pursuit of this goal, as is the danger (Tognazzini, 2005). This usability-security trade off, where convenience, quality and usability are sacrificed for security, has been shown in recent research (Weir et al., 2009a). This research found that customers are more concerned with convenience and usability than security. It is convenience that is thought to be one of the main drivers for customers to bank online (Jayawardhena and Foley, 2000; Tan and Teo, 2000). The importance of online security to customers should not be overlooked, but it is something customers tolerate to a certain extent (Weir et al., 2009b). Security processes must have high usability for customers to accept them (Piazzalunga et al., 2005). Customers will also avoid security processes, or find ways around them if possible (Besnard and Arief, 2004; De Witt and Kuljis, 2006). This compromise of security, which is no use to banks or customers, is an argument for more usability research in this area.

It is important to find out how customers perceive 2-factor authentication devices and methods, and how usable they are. While there has been some previous research into the usability of such devices (Weir et al., 2009a; Weir el al., 2009b), their usability for Internet banking has not been fully investigated. There is much published research about security and trust in eCommerce and eBanking services, but there are few studies that directly examine authentication systems and how they relate to customer perceptions of usability, security and convenience. Nilsson et al., (2005) examined a 'security box', similar in function to the OTP device used in this study, and compared it with a single fixed password. They found that the security box device was considered to be more trustworthy and secure than the single password.

6.1.2. Mobile Phone SMS Authentication

Banks in Australia and Asia (Yeo, 2006) have already introduced SMS 2-factor authentication and there are companies offering SMS authentication solutions such as MobileKey[10]. Supporters (AlZomai et al., 2008) of SMS authentication argue that its main advantage is it uses an 'out of band' medium, the mobile phone network, to send the user their authorization code, which is separate to the Internet. This is different to the way various banks in the UK employ the functionality of card reader devices. A code is displayed on the web page of the Internet banking site and the user must enter this into their card reader device to proceed. The security of SMS authentication is based on the argument that it is difficult to attack the mobile phone

[10] Mobilekey: <http://www.visualtron.com/products_mobilekey.htm>

network (Josang, AlZomai, Suriadi, 2007). The other benefit of an SMS authentication solution for generating OTPs is that the system can be used with generic mobile phones, smartphones and PDA's, and the vast majority of customers will already own a mobile phone. Therefore it eliminates the requirement and costs for banks to rollout devices to their customers. Another advantage of using mobile phones is people are quick to notice if they are lost or stolen compared with bank cards (Brunswick, 2009). However, SMS authentication relies upon customers registering their mobile phone number with the bank, and more significantly it also relies on third parties: mobile operators. Network coverage is also a concern.

A system for SMS authentication has been successfully researched and tested (Aloul et al., 2009), but usability research into the use of mobile phones and SMS as a 2-factor authentication method is in its infancy. It has been argued that weaknesses in the SMS method can be improved by increased usability (AlZomai et al., 2008). A recent study (Weir et al., 2009b) is relevant to the research conducted here. It describes the usability evaluation of SMS authentication, an OTP device and a password method for logging in on Internet banking. Based on usability attitude measurements they found that overall, the SMS method scored higher than the password method, but that more experienced users of Internet banking scored the password method higher. Also, in general, the password was considered the most secure and convenient method by the participants.

6.1.3. Authentication Basics

There are essentially three components to a generic authentication model:

know – some security token which is secret to the customer (e.g. password or PIN)

have – some device assigned uniquely to an individual (e.g. ATM card or a code generating token)

are – some intrinsic properties of an individual (e.g. fingerprint or palm vein pattern)

For example, ATMs employ 2-factor authentication but with a relatively low security for the know component, a 4-digit PIN number (know) combined with the physical ATM card (have).

6.1.4. Hypotheses

The experiment reported here investigated which authentication method for Internet Banking transactions participants would find the most usable, and which they would prefer to use.

H1: The alternative authentication methods will be different in terms of efficiency, based on speed, with the password proving to be the most efficient.

H2: The alternative authentication methods will be different in terms of effectiveness, based on completion rates, with the password proving to be the most effective.

H3: The alternative authentication methods will be different in terms of participant attitude, based on the results of a usability questionnaire, and password will perform the best.

H4: The alternative authentication methods will differ in terms of a 0-30 point quality rating scale.

H5: The alternative authentication methods will be judged differently in terms of a 0-30 point security rating scale and the password will perform the worst.

6.2. Method

6.2.1. Experiment Design

To enable the participants to compare the four different authentication methods participants were given simple background scenarios, and three fund transfer tasks to carry out with each device. Three different working versions of a 'mirror' Internet banking site with added functionality for each authentication device were created. The three versions of Internet banking were fully functioning versions of the Case Bank's real service.

The experiment had three experimental conditions: (1) SMS, (2) Card Reader, (3) OTP device and (4) Password (existing service). Participants experienced all four conditions in a repeated measures design. The usability of each version of the service would be measured by the following dependent variables: timings for *efficiency*,

completion rates for *effectiveness*, responses to individual items in a usability attitude questionnaire and quality rating with deduced preferences for *satisfaction*.

A few caveats should be noted with the *effectiveness* variable. There was no ability to control and validate the OTP codes coming from the card reader and OTP device. This meant the validity of the OTP codes from the OTP device and card reader could not be checked by the prototype web pages used here. The interface would accept any 6 digit access code for SMS and the OTP device, and any 8 digit code for the card reader. At login the password was validated on the web page and errors were treated realistically with the expected error messages. At the funds transfer stage the interface accepted any combination of letters as long as there was the correct number for that persona's password. This kept the conditions on similar footings and validity, but in this respect the effectiveness data has limitations. To add to realism though, the access codes could not be used more than once.

A security metric was also included to measure the participants' perception of security for each device directly. The two between-participants factors were Internet usage (low/high) and mBanker status (50% of customers would be signed up for the banks limited SMS banking service). The experiment was designed to have the order of experience for the four conditions balanced across the participants for the two between-participant variables. Due to no-shows and incomplete data sets, the final recruitment cohort was roughly balanced for these factors.

6.2.2. Participants

A cohort of 88 participants was recruited in Edinburgh. There were 33 female participants and 55 male participants (Table 6.1). The age range of the participants was 20 to 76 and the mean age was 44.

	18 – 45	**46 and over**	**Total**
Females	18 (20.5%)	15 (17.0%)	33 (37.5%)
Males	31 (35.2%)	24 (27.3%)	55 (62.5%)
Total	49 (55.7%)	39 (44.3%)	88 (100%)

Table 6.1: Participant demographics

Participants were all registered eBankers for the Case Bank's Internet banking website and had logged on to the service within the last 3 months. They were asked how frequently they logged on to Internet banking per month (during recruitment) and were grouped into low and high usage based on whether they used Internet banking more or less than 16 times per month. Some 48 participants were in the low use category (15 times a month or less) and 40 participants were high frequency users (16 times a month or more). The full range of usage frequencies is shown in Table 6.2.

Use IB	**N**
Daily	40
2-3 times / week	27
Weekly	12
2-3 times / month	7
Monthly	2
Less frequently	0
Total	88

Table 6.2: Frequency of Use of Internet Banking

Some 69 (78%) of participants used Internet Banking for balance enquiries, 19 (22%) checked their statements, 62 (70%) reported performing transfers online, 46 (52%) performed bill payments, 20 (23%) used the standing order functions, 23 (26%) managed their direct debits and 14 (16%) performed other tasks including: paying employees and other wages, paying credit cards, cheques and cardnet services. Internet bankers had been using the service for between 3 months to 12 years, with the mean time of use being just over 3 years, the median and modal usage time being 3 years.

All 88 participants in this cohort used a mobile phone. 39 (44%) of these had signed up for Text alerts, which is the bank's limited *Push* SMS banking service. At the time of this research the Text alert service sends registered customers weekly account balances and high/low balance alerts. Including mBanker status as a variable would allow a comparison of the data from users of SMS banking with non-users.

6.2.3 Authentication Methods and Devices

SMS Authentication

The mobile phone used in this experiment was a Sony Ericsson K-700 model. The mobile phone method for authentication makes use of server-based security using SMS text messages to send customers time-stamped access codes. Delivery of a one-time access code only takes a few seconds to the mobile after the user has entered transaction details. The mobile phone number must first be registered with the bank; however, this step was not included in the experiment sessions.

Figure 6.1 shows an illustration of the mobile phone used in the experiment sessions. Participants were all given the same phone for reasons of experimental control.

Figure 6.1: The Sony Ericsson K-700 as a Hardware Authentication Token

Figure 6.2 shows an illustration of the text message received with the access code, as seen by participants in the experiment session. The access codes sent in the experiment were randomly generated.

Figure 6.2: An SMS Sends the Access Code to the Mobile

As participants were not all familiar with mobile phone use, or this particular model, some assistance was given to those who needed it, in locating and opening the text message.

The instructions for the SMS passcode were always the same; no instructions for using the phone were included on the Internet banking pages. A screenshot of the confirmation page for the SMS passcode is shown in Figure 6.3.

Figure 6.3: Instructions – Text Message Version, All Uses

Existing Password Authentication

The Case Bank's Internet banking authentication process is currently based on single-factor two-step authentication process: customers are identified with a User ID

that is generated by the Bank. The authentication information consists of two secret tokens: customer-inspired Password and Memorable Information – the *know* model. Authentication is applied in a two-step process, firstly a full password is entered.

Secondly the interface requests three characters (selected at random) from the Memorable Information to be entered. The interface will not accept keystrokes, but instead forces mouse activation and selection in three separate drop-down menus.

This process is combined with transaction authentication required to securely confirm banking transactions using full disclosure of the password. This process is conducted on the confirmation screen, after transaction details have been input on a previous screen, see Figure 6.4.

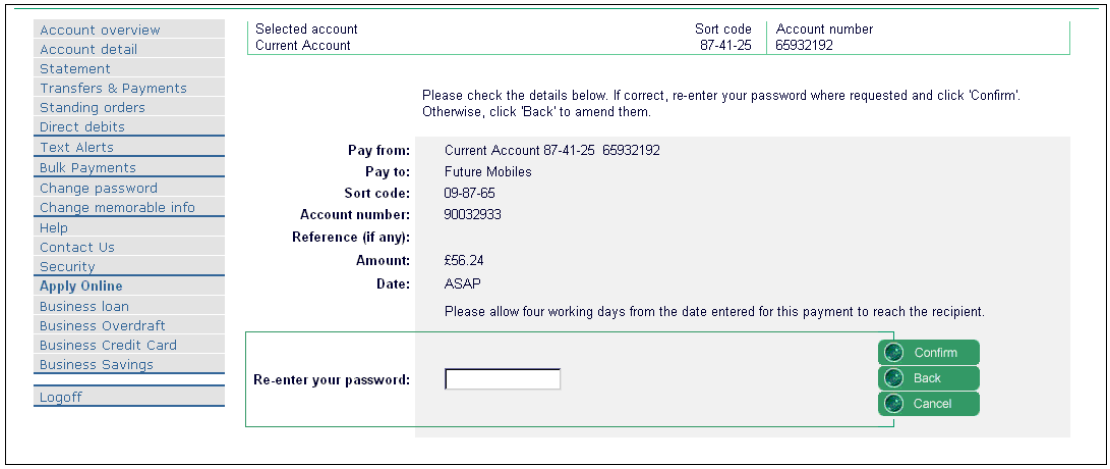


Figure 6.4: Transaction Confirmation Screen with Password Entry

The APACS Card Reader

The APACS[11] card reader, VASCO DIGIPASS 830 was used in the experiment, see Figure 6.5. The device offers three modes of operation. In the Identify mode the card reader functions as an OTP generator, therefore only the Identify mode was used in the comparison with alternative devices and methods for the experiment.

The procedure for the Identify mode involved successfully entering a valid bankcard and PIN, whereby the device offers an 8-digit code to be read by the customer and keyed in to a field on the Internet banking page.

Figure 6.5: VASCO DIGIPASS 830 APACS Card Reader

On first use of the device, the screens for payment confirmation were adapted to display a picture of the device, the Identify mode button and full instructions, see Figure 6.6.

[11] APACS. Remote card authentication. Available from: <http://www.apacs.org.uk/payments_industry/new_technology2.html>

On subsequent tasks using this device, the instructions read, above the passcode prompt:

"Use the Identify function of your card reader to obtain a new Passcode."

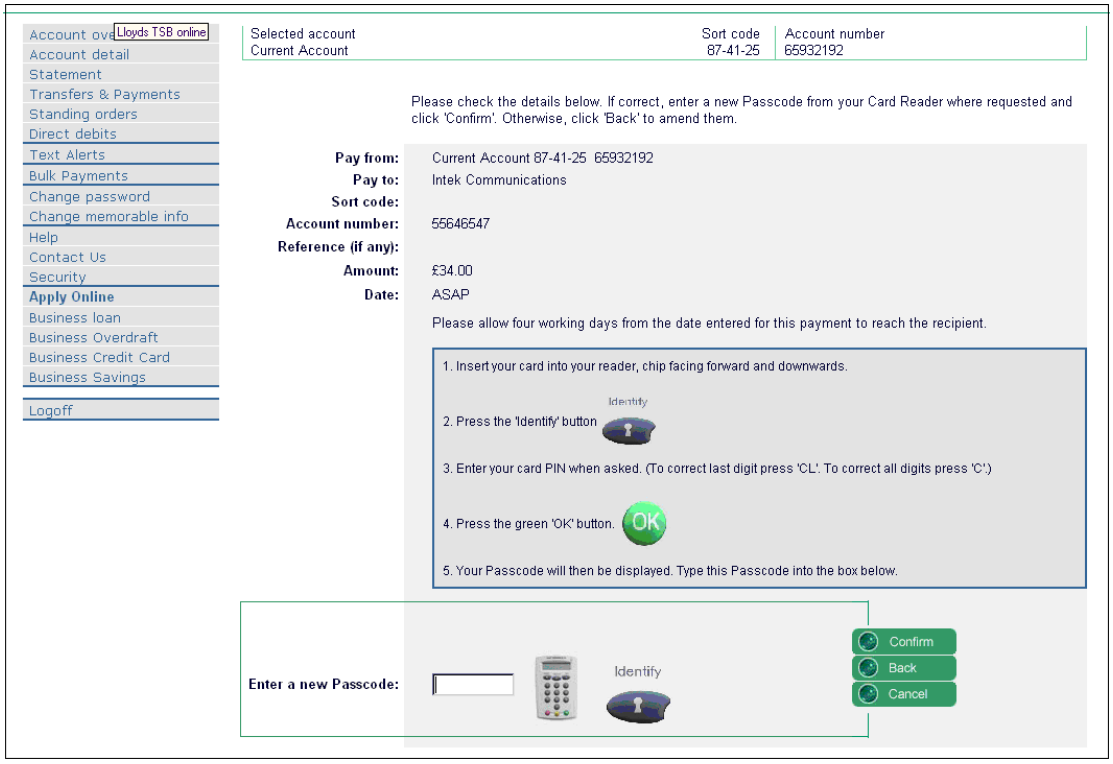


Figure 6.6: Payment Confirmation Screen for Identify Mode – First Use

OTP Device

The OTP device was the VASCO GO3[12]. It features an (up to) 8-character LCD display and a single push button. An on-board real time clock provides time synchronous encryption. One push on the button and a unique one-time access code is shown on the display to be entered at the passcode prompt.

[12] VASCO. Digipass GO 3 OTP device. Available from: <http://www.vasco.com/products/digipass/digipass_go_range/digipass_go3.aspx>

The key for the generation of an access code on a customer's OTP device is linked via their account to the same key for code generation on the internal banking server. The authenticating server uses the same algorithm as the device, using the time and serial code of each customer's device to verify their inputted access code. The code is updated every 30-60 seconds. Typically the server accepts the only the last three codes. No PIN entry is required.

Figure 6.7 shows an illustration of the OTP device as it was presented to participants in the experiment sessions.

Figure 6.7: The OTP Device as a Hardware Authentication Token

The instructions (open as default for the first use, and available from the help link for subsequent tasks) for the OTP device are shown in Figure 6.8.

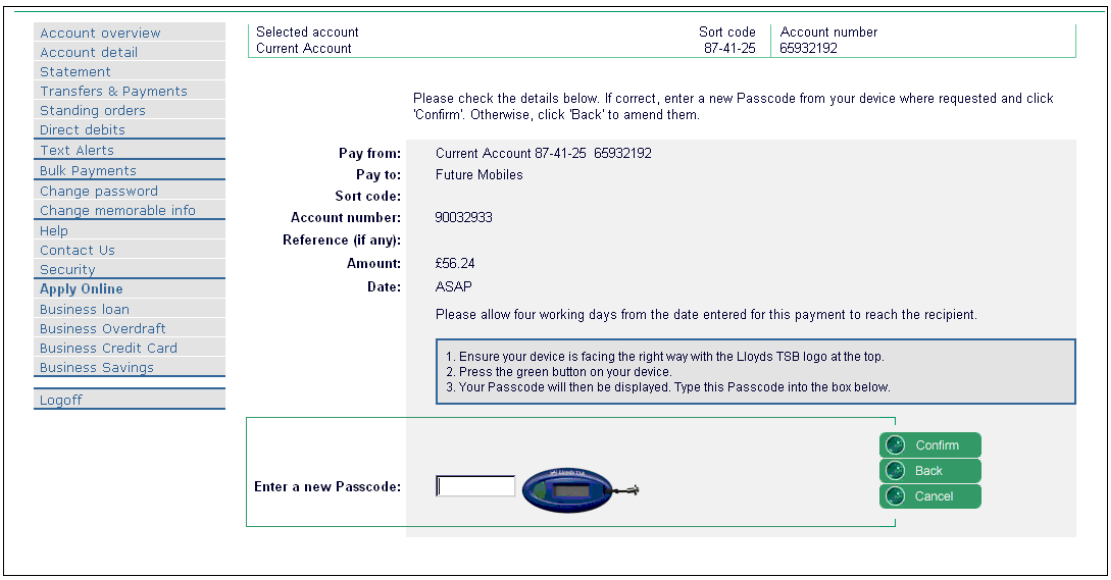


Figure 6.8: Instructions and Passcode Entry Field for the OTP Device – First Use

On subsequent tasks the instructions read, above the passcode prompt:

"Press the green button on your device to obtain a new Passcode."

6.2.4 Scenario and tasks

Participants were told to imagine that they were customer of the bank and were asked in each scenario to log on to Internet banking and perform three payment transactions. Task times were recorded for each task confirmation step (including the authentication step) automatically by server logs.

Example task:

You have been asked to use Internet banking to complete the following tasks.

Make a payment of £**185.50** to **Jones Stationary**

Make a payment of £**209.99** to **Loomes Publishing**

Make a payment of £**305.22** to **Star Utilities**

The participants were given all of the account details and information they would need to complete the tasks.

6.2.5 Usability Attitude Questionnaire

The design of the usability questionnaire used for this experiment followed standard practice (Likert, 1932) by using an equal number of negative and positive statements presented in a randomised order. The questionnaire used a 7-point Likert format that ranged from "Strongly Agree" (1) to "Strongly Disagree" (7). Following reversal of the polarity of positive questionnaire statements, a score of 7 consistently indicates a strongly positive attitude and 1 a strongly negative attitude. The questionnaire used consisted of 24 statements that address a range of usability attributes pertaining to human-computer interaction (HCI): *cognitive attributes* (level of concentration and degree of confusion), *the fluency and transparency of the service* (knowledge about what is expected, ease of use, degree of complication), and *quality attributes* (efficiency of service, amount of improvement service is felt to require, reliability of service). The 24 questionnaire items are shown in Table 6.3.

1	I found this method confusing to use
2	I had to concentrate hard to use this method
3	I got flustered when using this method
4	I felt under stress while using this method
5	I found this method frustrating to use
6	I thought this method was too complicated
7	Using this method I always knew what I was expected to do
8	I felt in control while using this method
9	This method was easy to use
10	The instructions for completing this method were clear
11	I would be happy to use this method again
12	I felt this method was reliable
13	I thought this method was quick to complete
14	I feel that this method needs a lot of improvement
15	I found this method 'user-friendly'
16	I liked using this method
17	I did not enjoy using this method
18	This method was difficult to understand
19	This method did not match my expectations
20	I found this method trustworthy
21	I felt confident in the security of this method
22	I found this method convenient to use
23	I felt that this method was unhelpful
24	It was hard to know what details to key in with this method

Table 6.3: Usability Questionnaire Items

All participants ($N = 88$) completed the usability questionnaire following exposure to each of the three experimental conditions. The usability questionnaire was presented to the participants on a laptop computer and the order of questions was randomised for each participant. The mean scores for these usability attributes enabled direct comparisons of participants' attitudes towards the three different message formats. Overall usability was determined by determining the mean for all the usability questions by all participants. Individual statements were also analysed separately to identify any specific usability issues that arose from the hands-on usage sessions.

6.2.5. Quality and exit interviews

After experiencing both services, and completing the usability questionnaires, participants made a quality rating, recorded as a value on a 30-point linear scale labelled “poor” at the 0 end and “excellent” at 30. This quality rating involved evaluating all four authentication methods against each other on the linear scale, and was also recorded to indicate an explicit preference, or no preference. The quality rating is a subjective satisfaction measurement, but unlike the usability questionnaire a participant is specifically asked to compare devices against each other which can be a useful result to compare with the questionnaire scores. If there is a discrepancy between the two, then the qualitative data from the one-to-one structured interviews can be consulted for an explanation. In the exit interview participants were given an opportunity to comment on their experiences and give their opinions on a range of related issues.

6.2.6. Security Metric

The security metric used the same 30-point linear scale as the quality measure. This time the scale was labelled security and the participants evaluated each authentication method against the others on the scale. The security rating is a subjective measurement that can be a useful result to compare with the usability data. This type of measure has been validated and used successfully before in similar research (Weir et al., 2009a; 2009b).

6.2.7. Procedure

The researcher verbally introduced the first (randomised) authentication method/device and told the participant that onscreen instruction would be provided for them. For example:

"For this method you will also need this mobile phone to receive text messages providing unique and frequently changing ***Passcodes****. Instructions will be provided on the Internet Banking Website, ask me if you need help using the phone."*

The participant was given a task sheet with the instructions and account details on, which the researcher also read out aloud to them. They were allowed three attempts to complete each task. If they failed three times the researcher stepped in and helped them to complete the task. Once the participant had completed the three tasks the research asked them to complete the usability questionnaire. The next authentication method/device was then introduced and the procedure repeated until the participant had experienced all four methods.

At the end of the session the participant completed a one-on-one 'exit' interview and were also asked to rate the overall quality of the four different methods on the 30-point linear scale. The experiment was concluded by completion of a demographic questionnaire to collect factual data about the participant.

6.3. Results

6.3.1 Efficiency – Task Timings

The time each participant took on each task was logged. Table 6.4 shows the mean time in seconds for each authentication method on the three tasks, and the combined mean time for all three tasks. The data are also illustrated in Figure 6.9.

Method/Device	1st Task	2nd Task	3rd Task	Overall
Password	26.3 (7.7)	24.2 (8.1)	20.4 (7.5)	23.6 (7.5)
OTP Device	25.9 (5.0)	23.3 (4.9)	22.9 (4.9)	24.2 (4.8)
SMS	41.5 (6.9)	38.1 (8.3)	37.4 (8.1)	38.9 (7.4)
APACS	54.5 (8.5)	49.6 (9.8)	42.2 (11.1)	48.5 (9.0)

Table 6.4: Mean times (S.D) in seconds by task and Authentication Method

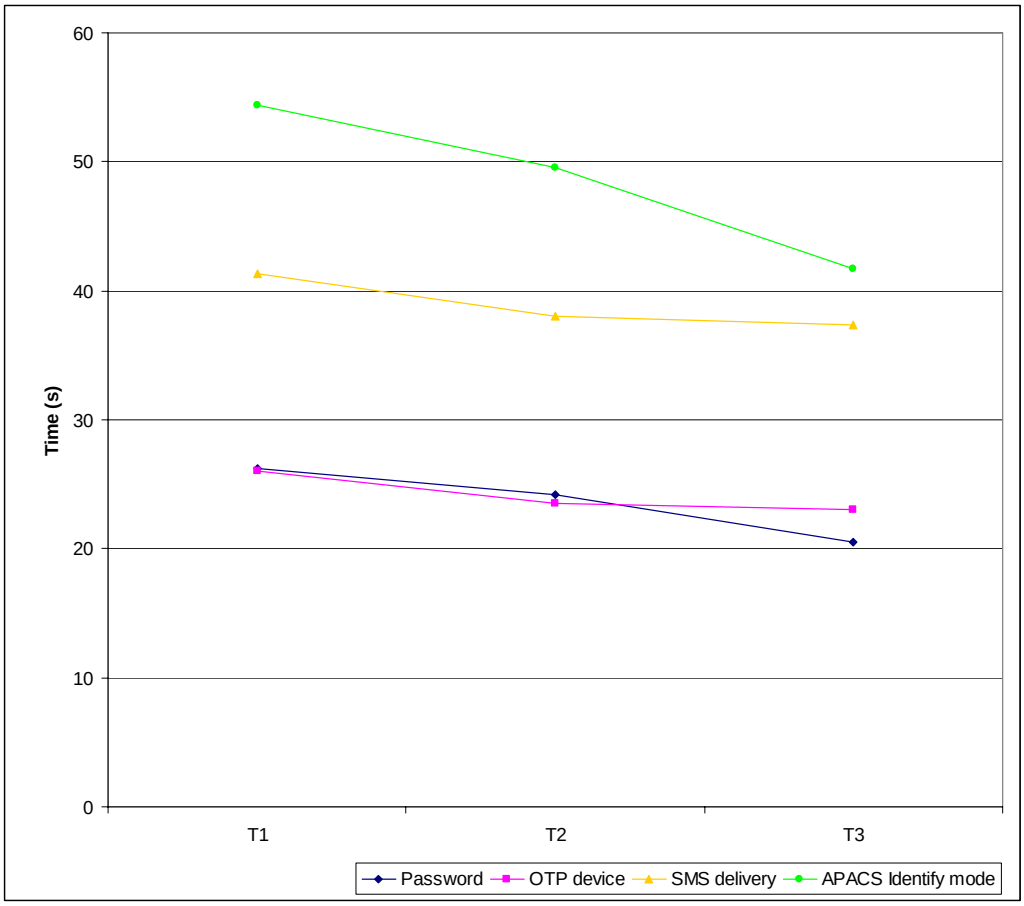


Figure 6.9: Mean Task Times for Authentication Methods at Tasks 1, 2 and 3.

A repeated-measures Analysis of Variance (ANOVA) was carried out on the timing data taking the mean time for each task and each authentication method as the within-participants variables and Internet banking (IB) usage frequency (low or high) and mBanking status as the between-participants variables. There were no significant interactions for usage frequency or mBanking status (user of Text Alerts or not) as the between-participants variables. Deviation from Sphericity was taken into account by applying the Greenhouse-Geisser adjustment in the analysis. There was a very highly significant difference between the four methods (d.f. = 3, $F = 239.06$, $p < 0.001$). There was also a very highly significant difference between the three tasks (d.f. = 1.81, $F = 289.79$, $p < 0.001$), again with no significant interactions between

tasks due to usage frequency or mBanking usage. Finally, there was a highly significant interaction between time, method and task (d.f. = 3.46, $F = 42.32$, $p < 0.001$), with no interactions with usage frequency or mBanking status. There were no between-participant effects due to IB usage frequency or mBanking usage.

Pair-wise tests were performed using Fisher's Least significant difference (LSD) method and these showed that there were significant differences between all the methods except between the OTP device and password (both equally fastest over all three tasks). These results are summarised in Table 6.5.

		Mean Difference	**Sig. (*p*)**
Password	OTP Device	-0.389	0.704
Password	SMS	-15.359	<0.001
Password	APACS	-25.143	<0.001
OTP Device	SMS	-14.969	<0.001
OTP Device	APACS	-24.754	<0.001
SMS	APACS	-9.784	<0.001

Table 6.5: Pair-Wise Comparisons of Mean Task Times for Each Method

There were also significant differences for each task with mean usage speeding up with each task in sequence ($p < 0.001$ for all comparisons). Finally, for each method, each task was also quicker to complete. For the SMS method, the overall reduction was 4.1s and this was significant for both comparisons ($p < 0.001$ for first to second and $p = .001$ for second to third - analysis of the different times between tasks per method using paired samples t-tests). The card reader achieved the highest reduction in speed between the first and third tasks performed (12.3s), which is an indicator of an increased complexity in early use. The differences in task times in each case were

significant ($p < 0.001$). The password (currently used by the cohort) decreased by 5.9s over the same three tasks, the reduction between tasks 1, 2 and 3 were all significantly different, $p < 0.001$. However, participants were using persona details and passwords rather than their own memorised ones, which may affect the magnitude of these timings in reality. The OTP method reduced overall by 3.0s over the three uses, again the first to second task was highly significantly different ($p < 0.001$) and for the second to third ($p = 0.001$).

The final result from the paired samples t-tests of the different methods at each task indicated that initially there was no difference in time for the OTP and password methods, similarly there was no difference in timing between them after task 2, however, after task 3 the password was significantly quicker to complete than the OTP device, $p = 0.013$. All other methods were significantly different for each task confirmation process following the main effect.

Differences in scores between the age groups were found in Chapter 4 and Chapter 5, so a final analysis was conducted including age group as a between-participants factor. The design was not pre-balanced for age as this detail of the cohort was not know at recruitment, but there was a reasonable split with 55.7% in the 18-45 group and 44.3% in the 46 and over group. A repeated measures analysis of variance (ANOVA) was performed on the mean total times, with method as the within-participants factor, and age group as the between-participant variable. The analysis did not show any significant interaction between method and age (d.f. = 3, $F = 67.69$, $p = 0.284$). There was also no between-participants effect for age. The analysis was

repeated for the individual times on first usage. Again, this did not reveal any significant interaction for age (d.f. = 3, $F = 53.29$, $p = 0.380$).

To find out if there was a relationship between age and time taken correlation analysis was performed. The analysis did not reveal any significant relationship between them overall, or by individual task.

6.3.2. Effectiveness

The percentage of participants to complete their first task with each method without help were calculated using the Adjusted Wald Method (Sauro and Lewis, 2005; Lewis and Sauro, 2006). These are shown in Table 6.6.

Method	Mean (%)	Interval (%)
SMS	89%	81-95
APACS	87%	79-93
OTP	91%	84-96
Password	94%	89-99

Table 6.6: Successful first time completion rates

A repeated measures analysis of variance (ANOVA) was performed on the overall successful first time completion rates, with method as the within-participants factor, and Internet Banking (IB) usage frequency (low or high) and mBanking status (user of Text Alerts or not) as the between-participant variables. The analysis showed no significant difference between methods (d.f. =2.90, $F = 1.78$ $p = 0.153$), and no within-participants effects. A repeated measures analysis of variance (ANOVA) was

also performed on the completion rates, with method as the within-participants factor, and age as the between-participant factor. There was no significant interaction between method and age (d.f. =2.90, $F = 2.40$ $p = 0.071$), but there was a between-participant effect for age ($p < 0.001$), with the older participants having lower completion rates in general.

The main confusion with the card reader was whether or not to input a space in passcode (as shown on the card-reader device screen) – although the system accepted both formats. A main error was placing the card in the reader the wrong way round which was perhaps surprising given the ubiquity of Chip & PIN payment services. Some participants also got to the point of typing in the passcode, but then didn't know to press the confirm button. Some participants were unsure as to whether or not to leave the card in after a transaction was completed and a few were also unsure as to how to activate the device - "*How do I turn it on/off*". Generally, the token was well understood, although one participant aimed the device at the screen thinking it behaved like a remote control.

6.3.3. Satisfaction - Usability Attitude Questionnaire

The scores for the 24 usability attributes were averaged to obtain an overall usability score for each method of authentication. The mean usability scores for the four methods were as follows: Password scored highest with a mean score of 5.89 (S.D. = 0.74), second was the OTP device with 5.55 (S.D. = 1.06), third came SMS with 5.11 (S.D. = 1.21) and the APACS had the lowest score with 4.68 (S.D. = 1.22).

A repeated-measures Analysis of Variance (ANOVA) was performed on the overall mean usability scores with method as the within-participant factor and usage frequency (low or high) and mBanking status as the between-participant factors. There was a very highly significant difference between the four levels of the within-participant variable (d.f. = 3, $F = 27.21$, $p < 0.001$). There were no significant interactions between the methods with mBanking or usage frequency. There were also no between-participant effects.

Pair-wise tests were performed using the LSD method and these showed that there were significant differences between each method. The password was significantly more usable than the OTP device ($p = 0.021$), the SMS ($p < 0.001$) and the APACS card reader ($p < 0.001$). The OTP device was also significantly more usable than the SMS ($p = 0.001$) and the APACS ($p < 0.001$). Finally, SMS was significantly more usable than APACS ($p = 0.004$). These results are summarised in Table 6.7.

		Mean Difference	**Sig. (*p*)**
Password	OTP Device	0.334	0.021
Password	SMS	0.772	<0.001
Password	APACS	1.214	<0.001
OTP Device	SMS	0.438	0.001
OTP Device	APACS	0.880	<0.001
SMS	APACS	0.442	0.004

Table 6.7: Pair-Wise Comparisons of Mean Usability Score for Each Method

A repeated measures analysis of variance (ANOVA) was also performed on the overall mean scores with method as the within-participants factor, and age as the

between-participant factor. There was a significant interaction between method and age (d.f. = 2.94, $F = 3.73$ $p = 0.012$), with the older group giving SMS a much lower score, but the OTP a higher score compared to the younger group. There was no between-participant effect for age.

To check for a relationship between age and attitude to usability of each method, correlation analysis was performed. The analysis revealed a weak but significant relationship with SMS authentication ($r = -0.257$, $p = 0.016$). This was a negative relationship, which meant attitude to usability decreased with age. There were no significant correlations between age and the remaining three methods: password ($r = -0.078$, $p = 0.470$), OTP ($r = 0.081$, $p = 0.451$) and APACS ($r = -0.013$, $p = 0.903$).

Repeated Measures ANOVAs were run on the individual attributes in the same way as the overall mean scores. The main results are shown in Table 6.8 along with the results of pair-wise comparisons (LSD). There were again no interactions or between-participant effects due to mBanking usage status. There were also no effects due to the frequency of use of Internet banking.

The password had significantly ($p < 0.05$) higher scores compared to SMS for 21 of the 24 usability attributes. For 12 of the attributes the OTP device scored significantly higher ($p < 0.05$) compared with SMS, but for the other half of the attributes the OTP and SMS were considered equivalently usable. In comparisons between SMS and the APACS device, SMS scored significantly ($p < 0.05$) higher for 14 of the 24 attributes. The APACS device scored significantly ($p < 0.05$) higher

than SMS for *reliability* and *confidence in security*. The SMS method is considered the least reliable of the four methods, and evokes the least confidence in the security of the process. The password scored significantly higher than the OTP device in terms of usability for 11 attributes, and APACS for 23 out of 24. The OTP device scored significantly ($p < 0.05$) higher compared with APACS for 21 of the 24 attributes.

Attribute	Significance of Main Effect	Main Pair wise comparisons[1]
Confusion	<0.001[#]	P > O = S > ID
Concentration	<0.001	P > O > S > ID
Flustered	<0.001	P > O = S > ID
Stressful	<0.001	P > O = S > ID
Frustration	<0.001	P > O > S > ID
Complication	<0.001[#]	P > O > S > ID
Knew what to do	<0.001[#]	P > O = S > ID
In control	<0.001	P > O = S = ID
Easy to use	<0.001[#]	P = O > S > ID
Instructions clear	<0.001[#]	P = O = S > ID
Would use again	<0.001	P > O > S = ID
Reliable	0.001[#]	P = O = ID > S
Quick	<0.001	P = O > S > ID
Needs improvement	0.001	P = O > S = ID
User-friendly	<0.001	P = O > S > ID
Liked using	<0.001	P >= O > S = ID
Enjoyment	<0.001	P = O > S = ID
Difficult to understand	<0.001[#]	P > O = S > ID
Matched expectations	<0.001	P >= O = S = ID
Trustworthy	-[#]	-
Confident in security	0.012[#]	P = O = S < ID
Convenient	<0.001	P > O > S > ID
Unhelpful	<0.001	P = O > S = ID
Key-in details	<0.001[#]	P >= O = S > ID

Table 6.8: Individual Usability Scores and Overall Comparisons for each Method

Interactions of 3 or more levels are not included in the tables.
P = Password; O = OTP device; S = SMS; ID = APACS Identify mode;
[1] Only the overall comparison between all four methods is shown here
[#] In these cases Mauchly's test was significant, therefore the Greenhouse-Geisser estimates of sphericity were applied to correct the degrees of freedom in calculating the results.

6.3.4. Satisfaction – Quality and Preference

To collect a quality rating for each of the authentication methods participants were asked to rate the four alternative authentication methods comparatively on a 30-point scale. The mean rating scores for the four methods were as follows: the OTP device scored highest with 20.02 (S.D. = 7.79), followed by Password with 19.00 (S.D. = 8.25), SMS scored lower with 15.05 (S.D. = 9.63), and APACS with 14.13 (S.D. = 8.54).

A repeated-measures Analysis of Variance (ANOVA) was performed on the means with method as the within-participant factor and usage frequency (low or high) and mBanking as the between-participants factors. There was a very highly significant difference between the four levels of the within-subject variable (d.f. = 3, $F = 9.00$, $p < 0.001$) but no significant interactions with usage frequency or mBanking status. There were no between-participant effects.

In terms of the main effect on the quality ratings due to method, inclusion of pair-wise comparisons in the repeated-measures ANOVA are summarised in Table 6.9.

		Mean Difference	**Sig. (*p*)**
Password	OTP Device	-0.833	0.542
Password	SMS	4.234	0.010
Password	APACS	5.119	<0.001
OTP Device	SMS	5.066	0.001
OTP Device	APACS	5.951	<0.001
SMS	APACS	0.885	0.551

Table 6.9: Pair-Wise Comparisons of Quality Ratings for Each Method

The password and OTP Device scored significantly higher in terms of overall quality compared with the SMS and APACS ($p \leq 0.010$ for all comparisons). There was no significant difference in the relative quality of the password and OTP device, or the SMS and APACS methods.

A repeated measures analysis of variance (ANOVA) was also performed on the overall mean scores with method as the within-participants factor, and age as the between-participant factor. There was a significant interaction between method and age (d.f. =2.78, $F = 3.71$ $p = 0.014$), with the older group giving SMS a much lower score, but OTP a higher score compared to the younger group. There were no between-participant effects for age.

To check for a relationship between age and quality rating of each method, correlation analysis was performed. The analyses revealed a weak but significant relationship with SMS authentication ($r = -0.293$, $p = 0.006$). This was a negative relationship, which meant quality rating decreased with age. There were no significant correlations between age and the remaining three methods: password ($r = -0.002$, $p = 0.988$), OTP ($r = 0.199$, $p = 0.063$) and APACS ($r = -0.088$, $p = 0.417$).

The quality scores were also reduced to a rank order of preference, see Table 6.10.

Method	Password	OTP Device	SMS method	APACS Identify
Best	30 (34.1%)	32 (36.4%)	18 (20.5%)	11 (12.5%)
2nd Best	19 (21.6%)	31 (35.2%)	17 (19.3%)	20 (22.7%)
3rd Best	23 (26.1%)	15 (17.0%)	26 (29.5%)	25 (28.4%)
Worst	9 (18.2%)	10 (11.4%)	27 (30.7%)	32 (36.4%)
Total	88	88	88	88

Table 6.10: Rank Order of Preference for the Alternative Authentication Methods

It can be seen from Table 6.10 that the OTP device and the Password received similar top ratings by the highest amount of participants, with SMS and APACS receiving considerably less top ratings.

6.3.5 Security Measurement Analysis

A security rating of the four authentication methods was then taken on the same scale as the quality rating. This allowed participants perceptions of security towards the alternative methods to be compared. The mean rating scores for the four methods were as follows: the APACS device scored highest with 22.00 (S.D. = 6.67), followed by the OTP device with 19.64 (S.D. = 7.20), SMS scored lower with 15.72 (S.D. = 8.28), and password with 13.28 (S.D. = 8.70).

A repeated-measures Analysis of Variance (ANOVA) was performed with the security rating data taking the mean rating for each method as the within-participant factors and usage frequency (low or high) and mBanking as the between-participants factors. Deviation from Sphericity was taken in account by applying the Greenhouse-

Geisser adjustment in the analysis. There was a very highly significant difference between the four levels of the within-participant factor (d.f. = 2.67, $F = 21.95$, $p < 0.001$) but no interactions with mBanking or usage, and no between participant effects.

Pair-wise tests (LSD) showed that there was no significant difference between the password and the SMS method, but that the APACS card reader was judged to be very significantly higher in terms of overall security than the alternate methods. The OTP device was also rated significantly more secure than the SMS and the password. These results are summarised in Table 6.11.

		Mean Difference	**Sig. (*p*)**
Password	OTP Device	-6.183	<0.001
Password	SMS	-2.058	0.135
Password	APACS	-8.469	<0.001
OTP Device	SMS	4.125	0.001
OTP Device	APACS	-2.286	0.025
SMS	APACS	-6.411	<0.001

Table 6.11: Pair-Wise Comparisons of Security Ratings for Each Method

A repeated measures analysis of variance (ANOVA) was also performed on the overall mean scores with method as the within-participants factor, and age as the between-participant factor. Deviation from Sphericity was taken into account by applying the Greenhouse-Geisser adjustment in the analysis. There was a significant interaction between method and age (d.f. =2.70, $F = 3.12$ $p = 0.031$), with the older group giving SMS a lower score and the password a higher score. The older participants also gave the OTP the highest score rather than the APACS. There was no between-participant effect for age.

To check for a relationship between age and perceived security, correlation analysis was carried out. There was a weak, but significant ($r = 0.234$, $p < 0.024$), positive correlation for the password method.

To check for a relationship between usability attitude and security, correlation analysis was conducted. This analysis showed that correlations were strong and positive for SMS ($r = 0.424$, $p < 0.001$), OTP ($r = 0.315$, $p < 0.003$) and APACS ($r = 0.311$, $p < 0.003$), but not for the password ($r = 0.151$, $p < 0.160$). As attitude to usability increased the perception of security increased. These results indicate that attitude was influence by perceptions relating to security for the three novel authentication methods.

6.3.6 Qualitative Analysis

The exit interview gathered comments and preferences on a wide range of issues related to the different methods for transaction authentication. Participants were asked what they liked, disliked and could suggest as improvements about each of the four different authentication methods they had used in the experiment.

The SMS message delivered to a mobile phone was liked for its familiarity and the fact that many people usually have their phones with them: “Using the mobile is a good idea as I have it with me all the time”, “Mobile method was easy and good”, “You always have it with you”, “It was simple, felt secure.” A primary concern for this method was the timeliness of the receipt of text messages: “Can’t be bothered

waiting for texts & what if I'm using the phone at the time?" "Mobile, I work in a basement with no reception, unfeasible", "You might be out of range", "The text could take a while to arrive." Suggested improvements for this method centred on the font size of the text message, with most indicating a larger font would be better (in the experiment the font chosen was the default for the model of mobile phone – font size would typically be customised by the customer on their handset.)

The Password method was praised for its simplicity and ease of use: "Password worked well", "I can create it and not rely on any technology", "Password is the best and easiest to use, most convenient", "Preferred the password - more accustomed to it". There were however some concerns over the password security in comparison to the other methods (all using an additional piece of hardware): "Didn't feel secure. Anyone could get it", "Concerned about the key-loggers for passwords", "Password has already been entered, so is useless." In terms of improving the password method, one participant suggested that "an additional question could be used".

The OTP device was praised for its ease of use and portability: "Dead easy to use - put on key-ring", "By far the simplest and most straightforward", "Quite neat", "Small and simple, keep track of it, more secure." Participants were concerned about the requirement to carry the device about with them (similarly for the Card Reader): "Could lose it", "Feels like a throwaway toy", "Too easy to use and lose." Some suggestions for improving included: "Add a password to enter alongside", "Could be larger for display purposes", "Have a button to clear, replaceable battery with secure chip", "It needs a proper attachment, maybe include a fingerprint reader."

Participants regarded the Card reader method to be the most secure: “Most secure, most modern”, “Card reader felt more secure”, “Good security.” Some participants disliked the complication in the operation of this method compared to other OTP methods: “I would need the instructions every time”, “There were too many steps”, “It’s longer and more complicated.” The (comparatively) larger form factor was also a concern: “Quite big and chunky”, “Bulky”, “Too many numbers, too big”, “I wouldn’t want to carry this around.” No specific improvements to this method were expressed other than the suggestion of more explicit instructions online and the removal of the space embedded in the Passcode.

The participants were asked if they thought a password was secure enough to confirm an online transaction. The results can be seen in table 6.12.

Secure Enough?	N	%
Yes	48	55.0%
No	30	34.0%
Don’t Know	10	11.0%

Table 6.12: Are Passwords Secure Enough to Confirm Transactions?

Over half the cohort felt that passwords were indeed secure enough to confirm transactions.

The participants were asked if they though it should be made mandatory to use one of the 2-factor devices instead of a password. The results can be seen in Table 6.13.

Mandatory?	N	%
Yes	36	40.9%
No	41	46.6%
Don't Know	11	12.5%

Table 6.13: Mandatory device usage to confirm transactions?

There was divided opinion on this question with 46.6% of the opinion that it should not be mandatory.

The participants were then asked if customers should be able to choose which 2-factor device they wished to use. The results can be seen in Table 6.14.

Device choice?	N	%
Yes	69	78.4%
No	15	17.0%
Don't Know	4	4.5%

Table 6.14: Participant to choose device?

The majority of the cohort believed that the choice of device should reside with the customer.

The participants were then asked if this choice would effect how often they used Internet banking. The results can be seen in Table 6.15.

IB usage with device	N	%
More frequently	8	9.1%
Less frequently	4	4.5%
About the same	76	86.4%
Don't know	0	0.0%

Table 6.15: Device impact on IB usage frequency?

A clear majority of participants thought that a customer-selected device would render their Internet banking usage frequency unchanged.

The participants where asked what were the advantages of 2-factor devices. A significant number of participants highlighted added security as the main advantage over the current method: "If you can't get access to the device, you can't hack my account", "Different way to do security looks good for bank", "Random number good, people use similar passwords", "Introduce complication which would make fraud less likely."

The main disadvantage given was the impact on accessing accounts if the device was stolen: "Something else to carry around", "Accessibility", "Something to lose", "I would want to know how they work", "Might lose the device", "Not user friendly. Time consuming. Might forget to take the device with me", "If it's not with you - you won't be able to use your account", "An extra thing to carry, lose or break."

The participants were asked if they had used the APACS, the OTP device or SMS authentication before. The results can be seen in Table 6.16.

Previous use?	Card reader		Token		Text message	
	N	%	N	%	N	%
Yes	19	21.6%	15	17.0	7	8.0
No	68	77.3%	22	81.8	76	86.4
Don't Know	1	1.1%	1	1.1	5	5.7

Table 6.16: Previous experience with the three new authentication methods.

Over three-quarters of participants had no previous experience of using devices like the ones used in the experiment. More participants had previously used a card reader than either of the other two methods. The text message method was the least experienced. The card reader was mainly used for competitor online banking access. The token was mainly used for computer / work access. The text message method had been used for promotional alerts and other specialised services such as m-parking (using a mobile phone for payment at parking meters).

6.4. Discussion

The results from this experiment show that the existing password method of authentication for Internet banking funds transfers generally performed better for usability, based on efficiency and satisfaction, compared to a 2-factor method using SMS generated OTPs. Judged on effectiveness it was predicted that the password would have significantly higher completion rates compared to the SMS 2-factor method, but this did not prove to be the case. Completion rates were high for all four methods. As discussed in the methods sections the completion data has limitations and so should be treated with caution. It is rather, an indicator of the problems customers will have with the authentication methods. There were no differences

found in scores for any metrics between high users of Internet banking and low users, or between mBanker and non mBankers. Those participants that already used the case bank's SMS service did not have a more positive attitude towards the use of SMS for authentication of online transactions. A related study (Liljander et al., 2007) found that users of sophisticated mobile services and applications had a more positive attitude towards a proposed mobile ecommerce service. This was not found to be the case here.

The participants used each authentication method three times, and for the vast majority it was the first time they had used SMS (or the OTP device and APACS card reader) for authentication. In this context it is not surprising that the traditional password method performed so well. The superior usability of the password is consistent with the fact that it is used as the current method of Internet banking transaction authentication, and is therefore well used and understood by participants in this cohort. In fact, over half of the participants considered a password to be secure enough to confirm an online transaction.

By their inherent nature 2-factor authentication devices add on more steps to the process, so will perform worst for efficiency (based on time taken) against a password, and arguably for satisfaction. However, it should be noted that participants were using persona details and passwords rather than their own memorised ones, which may affect the magnitude of these timings in real life. The difference in timing would still exist though, and would probably be significant. SMS authentication is inherently slower, in first time use it took nearly twice as long to complete the task

compared with the password. It adds more steps to the authentication process, which was born out in the timing data. The timing data for the SMS condition was artificial in the respect that the time the SMS took to be received was controlled. In reality it might actually be slower because it will be dependent on the network and reception, and from participant comments it seems they did consider these real life issues.

In an experiment (Weir, et al., 2009b) where participants had to create and memorise their own password it was found that SMS authentication scored higher for usability. This was not the only difference though, as in that experiment the password was validated by the input of three randomly requested digit. Here, the password was fully disclosed when entered. This suggests that as the password process becomes more complicated it will score lower for usability, and the advantages of the SMS method will come to the fore. With the SMS method the user only had to type in a 6 digit number, but with the password they had to input three randomly requested digits using drop down menus. The context of the study (Weir, et al., 2009b) was also different though, as authentication was used at the logging in stage, rather than for payments. This factor might affect user perceptions of security. It should also be noted that in Weir et al., (2009b) the participant cohort contained experts and novices, and customers from a competing bank. Here all of the participants were users of the case bank's Internet banking service. Further research should be carried out to investigate these issues in the context of payment authentication in Internet Banking, as this is the most typical context of use in the UK.

Previous research (Weir at al., 2009a) has shown that customers are driven by attitudes towards usability and convenience rather than their perceptions of security. The performance of the SMS authentication method compared with the password seems to follow this argument. The password model of authentication offers ease of use, speed and convenience (Tognazzini, 2005; Zviran and Haga, 1999), but offers low security. Here the password method was the current procedure followed by the bank. The password is fully disclosed at the logging in stage. Authentication of payments requires the customer to fully re-type the same password again. If the password is already compromised then this method offers very little in the way of extra security. The SMS OTP generation offers a much higher 2-factor level of security against online fraud, but this did not seem to be considered an advantage by the participants. Digging deeper though, the results from the security measurement actually show that the participants did not think the SMS method was more secure than the password. Even though it was rated higher on this measure, there was no significant difference between its score compared with the password. This result follows those in Chapter 5 where it was found that the participants had concerns about the security of the SMS medium. This is a serious issue. It seems actual fears about the security of SMS delivery is hampering its perceived usability as an authentication method. The tasks in this experiment required relatively low level transactions (all under £500), it would be interesting to see if there would have been a difference in the results if the transactions were in £1000s rather than £100s. If participants were more convinced about the security of the SMS method the results may have been different. This should be further investigated.

To get a clearer picture of the usability and suitability of SMS OTP authentication its performance needs to be compared against the alternatives: the OTP device and the APACS card reader. Banks will continue to push 2-factor authentication onto their customers as a means of cutting their losses due to fraudulent transactions so it is important to know how the alternatives compare with SMS. In a straight comparison between SMS, OTP device and APACS the OTP device came out on top for satisfaction and efficiency metrics, but not for effectiveness. In fact, the OTP performed very well against the password for efficiency and quality where there was no difference between the two. With attitude to usability though, the password still had a higher score.

In a recent similar study (Weir et al., 2009b) SMS authentication was found to be as usable as an OTP device. As discussed above, in that experiment the context was authentication at the logging on stage of Internet banking. Here in the context of authenticating online payments an OTP device scored higher for usability compared to the SMS method. This different context might be a reason for the different findings. Weir et al., (2009b) did find very varied preferences for the SMS method and the participants were split between those who hated the idea, and those who liked the idea, but not enough to pick it as their favourite method. The results here seem to show that the participants were driven by attitudes towards usability and convenience, as in this case both the OTP device and SMS offered increased security with 2-factor authentication. The advantage SMS authentication has over an OTP device is the ubiquity of the mobile phone, and the majority of customers will have their mobile phone on them, which should make it more convenient. Customers

would have to carry the fob around with them, or keep it in a safe location where they can always find it (if banking at home). A few participants did comment that they liked this about the SMS option, but surprisingly the convenience of SMS did not show in the quantitative results when compared to the OTP device. This might point to the participants being concerned with other issues around SMS authentication, and one of these factors would be security, as discussed above. This was shown in the results from the security measure where the OTP device was considered to be a much more secure medium compared with SMS. When examining the usability questionnaire results there is no difference between the two methods for the attribute *confidence in security*, but the OTP device does score higher for *convenience* and *reliable.*

There were a number of comments from participants regarding the reliability of the SMS method. They were concerned with the length of time they would have to wait to receive the SMS OTP, and this might be an important factor in the SMS methods performance here, along with perceived problems with mobile reception. These types of concerns with SMS might be allayed with continued use, but if participants had a choice they would choose the OTP, and the SMS method would not get the opportunity to convince them of its merits.

The SMS method compared more favourably to the APACS method, and this is a very interesting result as the APACS device is being rolled out to customers by a large number of UK banks. In fact, the APACS device generally performed the worst out of all four authentication methods. Generally, participants saw the increased steps

in using the card reader device as analogous to better security, but also many mentioned this as being complicated and cumbersome. The SMS was more efficient than the APACS device based on task timings in this study, and it also had a higher score for attitude. Both devices had equal scores for quality. In reality, it appears that UK banks are putting their stall behind card reader devices and their increased functionality as the 2-factor method of choice. The findings in this study though, suggest that customers would find an SMS OTP method more usable compared to the APACS device. The one advantage the APACS method has is its extra functionality, with three modes of operation, and as such it is a multipurpose device.

The results from the security measure add more evidence to the argument that customers place usability and convenience over security when judging the device or method (O'Gorman, 2003; Schultz et al., 2001; Weir et al., 2009a), and the previous research describing the potential conflict between usability and security in authentication (Besnard and Arief, 2004; Furnell, 2005; Johnston et al., 2003; Nodder, 2005; Renaud, 2005). The APACS card reader was considered to offer a higher level of security compared to the two other 2-factor methods (SMS and OTP device) and the password. This shows that the participants know the APACS would be better for their security with payment authentication in Internet banking, but the majority would prefer to stick with the simple password because it is more usable and convenient. Weir et al., (2009b) also report this contradiction with user attitudes to security levels, and suggest that customers do not actually understand how OTP works. They argue that this may have affected their participants' ratings of security and this could be the case here as well. This factor should be further researched.

This study did not measure participants performance over extended repeated use, and as with the previous chapters, some of the questions raised here may be answered by a longitudinal study in the field. For each method the time taken to complete the tasks reduced from first to third, which showed that even with this relatively small amount of exposure the devices became quicker to use. Over time the perceptions of security and usability for each device might also change.

In Chapter 4 and Chapter 5 differences in how the age groups rated the SMS banking services were found. Differences were not as strong in this study, with no differences in the timing data, but there was a difference in scores in the usability attitude metric and the quality metric. The older group were less positive about the SMS method for both, which ties in with the results from previous chapters. Negative relationships were found between age and usability attitude and age and perceived quality for the SMS method. The older age group were though, more positive about the OTP device compared to the younger group. In the quality measurement the older group rated the OTP device top and the SMS method last, while with the younger group the scores for the Password, OTP device and SMS were closer together, with the APACS rated last. The older group's preference for the OTP device is probably due to its simplicity. There was also a positive correlation between age and security for the password. The older the participant the more positive they were about the security of the password method. The SMS and APACS methods are more complicated, and older users will find them more difficult to use, and also, as has been discussed, older participants are more ambivalent to text messaging in general. This was shown in

general by the fact that with all four methods the older participants had lower first time completion rates.

Chapter 7

7. Conclusions and Suggestions for Further Work

The research outlined in this thesis was undertaken to contribute knowledge to the emerging area of mobile banking. Its aim was to advance understanding by providing reasons for the highly unsatisfactory adoption of mobile banking services by customers worldwide. The financial services sector is investing a considerable sum into the provision of mobile banking services for their customers, but they are not seeing a return on their investment. A series of three usability engineering experiments were designed and carried out which would provide insights on how best to realise the practical application of SMS banking services. The findings from these studies will help improve the usability of mobile banking services. A large body of qualitative data was gathered to help understand what people think and feel about SMS banking.

This thesis proposed that there are three general functions of SMS in electronic banking: transactions, communication/CRM and security.

Transactions

Ordering a new cheque book or PIN number, requesting a mini statement, transferring money or making a payment, these are all types of banking transactions that could, and are, offered by an SMS service. These types of services would generally be *Pull* (customer initiated), but will also be *Push* (automatic) for confirmation.

Communication/CRM

SMS can be used as a one-to-one business to customer communication channel and offers massive potential for CRM. SMS can be used for marketing of the banks services and products, confirmation of transactions made by customer with the bank via another channel (e.g. Internet, telephone banking), confirmation of contact with the bank via another channel, confirmation of appointments, complaints etc. These types of service will generally be of the *Push* type.

Security

SMS can be used as method of adding 2-factor authentication to online transactions, and potentially telephone transactions. SMS can be used to generate one time passcodes (OTP). These types of service will be of the *Push* type.

7.1. Summary of Evidence

Three empirical usability evaluations were carried out to explore customers' perceptions and attitudes towards using these functions of SMS banking.

Chapter 4 detailed a usability comparison of three alternative message formats for an SMS banking service: abbreviations, numbers and free-form. Participants used all three formats to carry out three banking transactions in a repeated measures experiment. Many of the existing SMS banking services use abbreviations, for example for an account balance the users sends ACBAL, but no research has been conducted to find out if abbreviations is the most usable format. In choosing a

message input format banks must recognise that text messaging is still a relatively tricky thing to do due to the methods of text entry available on mobile devices, and the lack of a standard user interface, or even a standard layout of the keypad. The aim of the experiment was to find the most usable message input format to use in a *pull* SMS banking service. It was proposed that the choice of numbers as an input message format would be the simplest and most succinct. The problem with a numbers based system is the lack of mapping between the numbers and what function they represent, which would initially result in the service being difficult to learn to use. The second message format was the one that has been implemented by a number of banks: abbreviations. If was proposed that if the abbreviations are short and intuitive then this system could be almost as fast and easy to use as a numbers system. The mapping between the abbreviation and function should be more obvious than with numbers. The free-form message format has one inherent advantage: the user can write SMS requests in their own words. This sidesteps the potential usability concern of requiring the user to memorise meaningless numbers, or potentially non-intuitive abbreviations. It was proposed that the disadvantages of free-form are an increased cost of development and longer messages requiring more time and effort.

The results in Chapter 4 showed that numbers and abbreviations were the most beneficial to the user. The participants used each version three times over a short duration, and as it was the first time they had used each version they had access to an SMS banking card detailing the relevant abbreviation or number command to use. In this context, it had been predicted that numbers would be the most efficient version of the service to use, and it was perceived to be faster by the participants, but in

actuality, participants only performed faster in the cheque book task compared with abbreviations. It was also predicted that numbers would perform best for *satisfaction*, but again there were no significant difference between overall attitude scores for abbreviations and numbers, or the quality ratings.

True to predictions, free-form in general performed the worst for usability. It was proposed that the reasons for this can be deduced from its inherent limitations when applied to a mobile phone user interface. It should be noted that this did not impact on how successful participants were in completing the tasks. Indeed, in the case of a funds transfer task free-form actually had a significantly higher first time completion rate compared to numbers. It was argued that free-form performed the worst because it forced people to write text messages in grammatically correct English. In reality many people use abbreviations, or textish. Performance may improve over repeated use, and Free-form theoretically has high learnability and memorability, but the first impressions of a free-from service, as shown in the results from chapter four, may put users off.

It was argued that it is still possible that the numbers format is not as intuitive because of a lack of natural mapping between the commands and the services that they represent, and this factor could impact on its learnability and memorability. It may be that the nature of the study had been more favourable on the numbers format, and that in a longitudinal study numbers would not perform as well. This argument, of course, could also be applied to a service using abbreviations, but it seems plausible that abbreviations, if chosen well, might have higher learnability and

memorability. Regarding the results from chapter four it was recommend that users would find an SMS banking service that used either numbers or abbreviations to be equally usable, but it was suggested that abbreviations may prove to be the most usable format based on an assumption of higher learnability and memorability.

Chapter 5 detailed a two part empirical investigation into the integration of an SMS channel into a banks multichannel environment as a strategy for 'next call avoidance.'

The first experiment, the SMS *Push* service experiment, compared an SMS channel using the abbreviations message format to the existing telephone banking channel for balance requests. Two SMS versions: SMS balance-only and SMS balance with Mini-Statement, and two telephone banking services: Automated and Advisor, were compared by measuring their usability. Chapter 4 discovered that an SMS balance on-demand service was the one most likely to be used by customers. The capability to provide balance on-demand initiated by *Push* SMS message requests from customers would remove the need for the customer to call telephone banking for a balance, and would offer potential cost reductions and improved customer experience. The introduction of any new banking channel requires research to compare it with available channels in terms of usability and customer satisfaction, but there has been little research into comparisons of the usability of mobile banking and competing banking channels.

The results from this experiment showed that there were differences in the effectiveness of the two banking channels, with the telephone banking channel having a significantly higher first time completion rate, though completion rates overall were high for both channels. The participants were recruited as telephone banking customers, and in this context the completion rates for the SMS channels are positive findings. Overall, the telephone channel was rated higher for attitude to usability and for the quality. Scores for both channels overall and by version were above the neutral point indicating good usability. Converting the quality rating into a rank order preference did not though indicate a clear preference between the telephone or SMS channels. A large percentage of the positive comments regarding the SMS channel and versions focussed on the speed of the service. Comments made by the participants seem to indicate that one of the main reasons driving their rating of the SMS services was speed and efficiency.

It was argued that the reasons for the telephone channel scoring higher for usability for a balance request may not be too surprising. The telephone is a familiar and mature technology in the banking sector, while the use of SMS is relatively novel. The qualitative data indicated that the most frequent concern over SMS banking is perceived security and this may be the main obstacle to overcome before SMS is considered as usable as telephone banking. Security is a key factor in customer acceptance (O'Gorman, 2003; Schultz et al., 2001) in electronic banking, and particularly in mobile banking (Brown et al., 2003; Laurn and Lin, 2005). Though some argue that security concerns (Laukkanen, 2007; Laukkanen and Lauronen, 2005; Suoranta, 2003) are not a reason for low adoption of mobile banking. It was

argued that the results from this study suggest that with SMS banking it is concerns with the actual security of the medium, data security, that is a major factor to the participants.

A concept tied to security is trust, and it was argued that participants will trust a telephone banking service more due to previous experience with it, and because it is an established banking channel. It has been found that trust (Kim et al. 2007; Lee and Chung, 2009) is an important factor in usage intentions of mobile banking. Security is considered to be an aspect of system quality (Lee and Chung, 2009; Delone and Mclean, 1999), and system quality is an important factor in building trust, and trust in turn affects customer satisfaction. This study only looked at first time use of the SMS banking service for requesting a balance request, the participants had used either automated or advisor telephone services before. It was argued that with more use of the SMS channel trust in the service would increase, which would impact on satisfaction. This may lead to SMS banking comparing more favourably towards a telephone service. The usage intentions of participants were also examined and it was found that more participants would be likely to use the SMS service for a balance enquiry than the advisor version of telephone banking, but they would still be more likely to use Internet banking or an ATM over SMS banking for this task.

The second experiment detailed in Chapter 5, the SMS *Pull* service experiment, investigated integrating SMS into the existing telephone banking channel by way of funds transfer confirmation messages. It evaluated participants' attitudes to the communication/CRM function of SMS banking. If banks can furnish transaction

confirmation SMS alerts on transactions carried out using their channels and thereby remove the need for customers to call an advisor in order to enquire and double-check transaction details, this would also results in cost reductions for the bank, and a better user experience for the customer with additional convenience, reassurance and confidence. There has been limited research into mobile CRM, and the research that has been carried out has had mixed results, and with such little research so far there was a need for usability research on mobile CRM.

Chapter 5 showed that attitude to usability was considered to be the same whether the participant received an SMS confirmation after a funds transfer or not. The quality and preference results showed that the existing version of telephone banking (no confirmation) was rated lower than the SMS confirmation versions and updated balance version. There were no differences between the SMS confirmation versions and the updated balance version of the service, but reducing the quality ratings to rank order preferences indicated a preference for the SMS versions of the service. It was argued that the usability of the different versions might be judged to be the same, but the participants judged the utility of a version that confirms the transaction to be higher than one that does not. A large majority of participants were positive about an SMS confirmation and the confidence and reassurance it gave, but this was a similar picture for an updated balance in the call. A small minority of participants expressed concerns about security regarding SMS confirmations, similarly to concerns over the use of SMS for balance requests in part one of the experiment.

The findings from chapter 5 suggest that using SMS for CRM and as a medium to confirm banking transactions has possibilities and would be welcomed by customers. It was discovered that the type of transactions that participants believed would suit an SMS confirmation were monetary transactions on an account, where confirmation would be beneficial, for instance a debit card payment over the internet. In this case the SMS confirmation would help against fraudulent transactions.

Chapter 6 detailed the findings from the SMS security experiment, which evaluated the usability of the security function of SMS. The SMS OTP authentication method was compared against the existing password method, the OTP device method and the card reader method. Banks are moving to introduce 2-factor techniques for authentication to replace or augment the single factor password method. The password strategy of customer authentication is vulnerable to phishing and keystroke capture. SMS authentication's main advantages are suggested to be that it uses an 'out of band' medium, it is difficult to attack the mobile phone network, it is a system that can be used with generic mobile phones and the vast majority of customers will already own a mobile phone. Previous research has found that security processes must have high usability for customers to accept them (Piazzalunga et al., 2005), so it is important to find out how customers perceive 2 factor authentication devices and methods, and how usable they are. There has been some previous research into the usability of such devices but their usability for Internet Banking has not been fully investigated.

Chapter 6 revealed that the existing password method of authentication for Internet banking funds transfers generally performed better for usability compared to a 2-factor method using SMS generated OTPs. The participants used each method three times, and for the vast majority it was the first time they had used SMS (or the OTP device and APACS card reader) for authentication. It was argued that in this context it was not surprising that the traditional password method performed so well. It is used as the current method of Internet banking transaction authentication, and was therefore well used and understood by participants. Over half of the participants considered a password to be secure enough to confirm an online transaction.

The results in Chapter 6 contradict part of the findings from a related experiment (Weir et al., 2009b). It was argued that this is may be due to the different context of the experiments, which were authentication at logging in vs. authentication for funds transfers, the type of password process examined and the different user groups studied. It was also argued that it is possible that as the password process becomes more complicated it will score lower for usability, and the advantages of the SMS method will come to the fore. The SMS OTP generation offers a much higher 2-factor level of security against online fraud, but the results from Chapter 6 did not seem to indicate that this was considered an advantage by the participants. The results from the security measurement showed that participants did not think the SMS method was more secure than the password. This result followed those in Chapter 5 where it was found that the participants had concerns about the security of the SMS medium. It is argued that this is a serious issue, as it seems actual fears

about the security of SMS delivery are hampering its perceived usability as an authentication method.

Chapter 6 showed that in a straight comparison between SMS, OTP device and APACS card reader the OTP device comes out on top for usability. These results again contradict Weir et al., (2009b) were SMS was considered as usable as a similar OTP device. The convenience of SMS compared to the OTP device did not show in the quantitative results. It was argued that this pointed to the participants being concerned with other issues around SMS authentication, and one of these factors would be security, as discussed above. This was shown in the results from the security measure where the OTP device was considered to be a much more secure medium compared with SMS. Other factors that concerned the participants were the reliability of SMS OTPs and the length of time they would have to wait to receive them. It was argued that these types of concerns with SMS might be allayed with continued use, but that if participants had a choice they would choose the OTP, and the SMS method would not get the opportunity to convince them of its merits.

Chapter 6 showed that the SMS method compared more favourably to the APACS card reader, which is a very interesting result as card readers are being rolled out to customers by a large number of UK banks. It was found that the APACS device generally performed the worst out of all four authentication methods. It was argued that participants saw the increased steps in using the card reader device as analogous to better security, but also as being complicated and cumbersome. The findings in

Chapter 6 suggest that customers would find an SMS OTP method more usable compared to the APACS card reader device.

It was argued that the results in Chapter 6 add more evidence to the argument that customers place usability and convenience over security when judging the device or method (O'Gorman, 2003; Schultz et al., 2001; Weir et al., 2009a), and the previous research describing the potential conflict between usability and security in authentication (Besnard and Arief, 2004; Furnell, 2005; Johnston et al., 2003; Nodder, 2005; Renaud, 2005). The APACS card reader was found to offer a higher level of security compared to the two other 2-factor methods (SMS and OTP device) and the password. It was argued that this shows that the participants know the APACS card reader would be better for their security with payment authentication in Internet banking, but the majority would prefer to stick with the simple password because it is more usable and convenient. Weir et al., (2009b) also reported this contradiction with user attitudes to security levels, and suggest that customers do not actually understand how OTP works. This may affect people ratings of security with these devices.

Chapter 6 found no differences in scores for any metrics between high users of Internet banking and low users, or between mBanker and non-mBankers. Those participants that already used the Case Bank's SMS service did not have a more positive attitude towards the use of SMS for authentication of online transactions. A related study (Liljander et al., 2007) found that users of sophisticated mobile services

and applications had a more positive attitude towards a proposed mobile ecommerce service. This was not found to be the case in Chapter 6.

7.2. SMS Banking and Older Users

In all three of the experiments on SMS banking detailed in this work the results from the younger age group were compared with those of the older group. Older users suffer from visual, cognitive and motor impairments to varying degrees (Christopher, 1999; Jagaacinski et al., 1995; Krampe and Ericsson, 1996; Kurniawan et al., 2006a; Walker et al., 1997). Each of these impairments could impact on how usable an SMS banking service would be to older users, so it was of interest to investigate the data recorded here to see if this would indeed be the case. The term 'older users' does not just apply to the over 60s, or the retired, as a decline in visual acuity is noticeable by the mid-40s (Kurniawan et al., 2006a), and this decline would make reading the text on a mobile phone's small screen much more difficult. Older users can also suffer from a reduced ability to cope with repetitive fast movements. These are the types of movements required when entering text into a mobile phone. Older users can also have difficulty with finer movements and motor coordination, due to stiffening of joints and arthritis, which would cause difficulty when using the small keypads on mobile phones and this has been shown in recent research (Ornella and Stephanie, 2006). Chapter 4 showed a correlation between age and speed typing text messages, with time taken increasing with age. It also showed that older users had lower completion rates when writing text messages required for SMS banking transactions. Chapter 5 also found this result. In all three chapters older users'

attitude scores were lower than those of the younger age group, and the quality scores followed a similar pattern. Chapter 5 found a negative relationship between overall attitude to usability for an SMS balance request service and age. Chapter 6 found that the older group was less positive about the SMS authentication method for attitude and quality, and negative relationships were found between age and usability attitude and age and perceived quality for the SMS method. Interestingly, no interactions for age were found for a *Push* SMS funds transfer confirmation in Chapter 5. This may suggest usability issues only arise when older participants have to interact with the mobile handset more and instigate the transaction in a *Pull* service. A confirmation like this from the bank may appeal to older participants which would make them more positive towards the use of SMS in this context. The use of SMS in this context is also more similar to SMS use in general, whereas SMS balance requests, or SMS transfer requests are a more novel use of SMS.

SMS usage patterns and experience levels are also a probable factor in the differences in scores between age groups. The older age group in general were found to send SMS less frequently than the younger group. There is some debate in the HCI community as to how to explain the SMS usage patterns of older users (Ling, 2007), with one side arguing that it is down to usability problems, and the other arguing that there are more complicated sociological reasons, with the attitude of the older users playing a part in their low adoption of the technology. The SMS usage patterns in this study show that the younger age group in general send SMS more frequently than the older group (see Tables 4.2 and 5.3). It is probable that the difference in experience levels between the two groups has also contributed to the difference in

results. To answer the question as to whether lower usage levels, and the consequent lower experience levels, are due to the usability issues surrounding mobile devices, or are caused by sociological explanations, is beyond the scope of this thesis. It is clear though, that older users are more ambivalent towards SMS. It is plausible that older users are more positive towards traditional telephone banking, as it has been argued that older users are less positive in general to mobile phones (Ling, 2007). Older people are active users of the landline telephone, and older people also value personal contact highly (Blythe et al, 2005). The implications to the financial service sector for these findings are that they should think seriously about the target audience for SMS banking, and mobile banking in general, as older user will find the service more difficult to use and may not be convinced to try them in the first place. The marketing strategies for mobile banking might need to be altered to focus on a specific younger, more technology savvy customer base, and this might affect its cost/benefit ratio.

7.3. The Security Factor

The findings in the work reported here suggest that some customers are not convinced in the security of SMS as a banking channel. It can be argued that this is probably a very important factor in the low adoption of SMS banking by customers. This was most seriously shown in Chapter 6 where it seemed actual fears about the security of SMS delivery were hampering its perceived usability as an authentication method. SMS was not considered more secure than using a password. For a significant group of customers the perception is that SMS as a medium is insecure

for banking service. The implication is that banks must educate their customer base and reassure them of the security of SMS banking; otherwise it will remain a major factor in the low adoption of SMS banking.

7.4. Limitations

The limitations of the individual experiments have been discussed in the relevant chapters. In general, for each of the studies reported here the one off exposures to the designs and devices restricted the depth of the studies. They did not specifically test how easy the designs, devices and interfaces were to remember to use and learn, but did offer qualitative data from the participants to address these issues.

The major limitation was the experimental research method, and the argument over realism – in the tasks being attempted, the contexts of use of the systems under test, and the users' motivation and background knowledge. For instance, mobile devices are used on the move and in various locations so it could be argued that laboratory tests do not simulate the context where mobile devices would be used in the real world. For usability experiments in general the experimental situation only approximates reality, but real world behaviour is inferred from measurements and observations collected in the laboratory.

A further limitation is that the work carried out here was in the context of the UK banking industry. Cultures and people are different so there are questions in how

generalizable the findings are to the rest of the world. For instance, many countries are further ahead than the UK in the provision of mobile banking services.

7.5. Further Work

The potential for further work was discussed throughout this thesis and will be expanded upon here. In general, it would be of interest to explore the use of the functions of SMS banking in longitudinal studies to explore levels of expertise. It would be interesting to discover how extended use of SMS banking might influence customer trust in the service, and also customer attitude. Extended use might also result in an uptake in usage patterns.

Further research is needed into the usability of SMS payments systems for transactions, and how they compare to alternative mobile payment applications, such as Monilink. Attitudes to the security of these type of methods for bill, and third party payments will be an important factor in the adoption of any such system.

There is also the potential and need for further work regarding the relationship between usability and security with SMS banking. This thesis has found that customers fears regarding security are a major factor in the low adoption of SMS mobile banking. Potential future work might investigate an added level of security with SMS banking in the form of password, or PIN type authentication, and how this would impact on the usability and perceived security of the service. Several participants suggested this would be an improvement to the SMS banking service.

Research could investigate this extra level of security for transactions and for SMS 2-factor security applications. The trade-offs between any increased layers of security and the usability of the relatively straightforward process of sending a text message will need to be investigated in this context. Methods for increased security could involve downloading software clients to the mobile phone that offer another level of authentication. From within the 'closed' application the customer can send SMS and carry out banking tasks. Alternatively, for SMS transactions, the extra authentication step could be carried out using another channel, for instance an IVR, which is a solution employed by PayPal mobile. It would also be of interest to see if the education of customers regarding the security of SMS and the purpose of OTPs would affect their attitude in further studies.

The mobile phones used throughout this thesis were generic handsets. Further research could investigate the usability of SMS banking with other form factors and interaction styles such as touch screen devices like the iPhone. It will be important to discover how this would impact on the usability of SMS and mobile banking.

Finally, there is a need for further usability research comparing SMS banking and existing channels, such as Internet banking, on desktop and handheld devices.

7.6. Conclusions

Abbreviations would be the most usable message input format to use for SMS banking service transactions. Number and abbreviations performed to generally

equal levels on usability metrics, but abbreviations is recommended based on an assumption of higher learnability and memorability. SMS for balance transactions using abbreviations was found to be less usable when compared with the existing telephone banking channel. Attitude to the SMS channel was positive indicating its general usability was acceptable, but telephone banking would be preferred for usability and quality reasons. Using SMS for CRM and as a medium to confirm banking transactions has possibilities and would be welcomed by customers. It was discovered that the transactions that participants believed would suit an SMS confirmation were monetary transactions on an account, as confirmation would help against fraudulent transactions. SMS as security for authentication of Internet banking funds transfers was found to be lower in usability compared to the existing password method and a standalone OTP device. It was considered to me more usable than a card reader device, which is the device being pushed by several UK banks. For a significant group of customers the perception is that SMS as a medium is insecure for banking services. Older user will find SMS banking services less usable than younger users and are more ambivalent regarding SMS in general.

The empirical work presented in this research has contributed to the knowledge of mobile banking and will provide insights on the practical application of SMS banking and services. The data serve to support the thesis that usability is a significant factor in the low customer adoption of SMS and mobile banking services. Related to usability issues are customer concerns over the security of SMS as a banking channel.

Bibliography

Adams, A., Lunt, P., Cairns, P., 2008. A qualititative approach to HCI research. In: Cairns, P., Cox, A., (Ed.), Research Methods for Human-Computer Interaction. Cambridge, UK: Cambridge University Press, pp. 138–157.

Agarwal, R., Venkatesh, V., 2002. Assessing a firm's web presence: A heuristic evaluation procedure for the measurement of usability. Information Systems Research 13 (2), 168-186.

Aladwani, A., M., 2001. Online banking: a field study of drivers, development challenges, and expectations. International Journal of Information Management 21 (3), 213-225.

Aloul F., Zahidi, S., El-Hajj, W., 2009. Two factor authentication using mobile phones. IEEE/ACS International Conference on Computer Systems and Applications, 641-644.

AlZomai, M., Josang, A., McCullagh, A., Foo, E., 2008. Strengthening SMS-Based authentication through usability. International Symposium on Parallel and Distributed Processing with Applications, ISPA'08, 689-694.

Anderson, A.H., Newlands, A., Mullin, J., Fleming, A.M., Doherty-Sneddon, G., Van der Velden, J., 1996. Impact of video-mediated communication on simulated service encounters. Interacting with Computers 8, 2, 193-206.

Balakrishnan, V., Yeow, P., H., P., 2008. A study of the effect of thumb sizes on mobile phone texting satisfaction. Journal of Usability Studies 3 (3), 118-128.

Balakrishnan, V., Yeow, P., H., P., Ngo, D., C., L. 2005. An investigation on the ergonomic problems of using mobile phones to send SMS. In: P.D., & P.T., McCabe (Eds.) Contemporary Ergonomics 2005 (pp.195-199). United Kingdom: Taylor & Francis.

Bennett, J., L., 1984. Managing to meet usability requirements: Establishing and meeting software development goals. In: Bennett, J., L., Case, D., Sandelin, J., Smith, M., (Eds.) Visual Display Terminals: Usability Issues and Health Concerns, 161-184. Englewood Cliffs, NJ: Prentice-Hall.

Betsch, H., P., Schwieren, C., Gutig, R., 2001. I like it but I don't know why: a value-account approach to implicit attitude formation. Personality and Social Psychology Bulletin 27 (2), 242-253.

Black, N., J., Lockett, A., Ennew, C., Winklhofer, H., McKechnie, S., 2002. Modelling consumer choice of distribution channels: an illustration from financial services. International Journal of Bank Marketing 20 (4), 161-173.

Blythe, M., A., Monk, A., F., Doughty, K., 2005. Socially dependable design: The challenge of aging populations for HCI. Interacting With Computers 17, 672-689.

Besnard, D., Arief, B., 2004. Computer security impaired by legitimate users. Computers and Security 23 (3), 253-264.

Brooke, J., 1996. SUS: a "quick and dirty" usability scale. In P W Jordan, B Thomas, B A Weerdmeester & A L McClelland (eds.) Usability Evaluation in Industry. London: Taylor and Francis.

Brown, I., Cajee, Z., Davies, D., Stroebel, S., 2003. Cell phone banking: predictors of adoption in South Africa – an exploratory study. International Journal of Information Management, 23 (5), 381-94.

Brown, A. S., Bracken, E., Zoccoli, S., Douglas, K., 2004. Applied Cognitive Psychology, 18 (6), 641 - 651.

Brunswick, S., 2009. eCommerce fraud – time to act. Card Technology Today, 21, (1), 12-13.

Cains, P., Cox, A., L., 2008. Research methods for human-computer interaction. Cambridge University Press.

Carrington, V., 2005. Txting: the end of civilization (again)? Cambridge Journal of Education, 35, 2, 161-175.

Centeno, C., 2004. Adoption of internet services in the acceding and candidate countries, lessons from the internet banking case. Telematics and Informatics 21 (4), 293-315.

Chin, J., P., Diehl, V., A., Norman, K., L., 1988. Development of an instrument measuring user satisfaction of the human-computer interface. Proceedings of SIGCHI'88, ACM Press., New York, 213-218.

Computer Fraud and Security, 2006. Regulator warns banks to face up to online fraud. News article from Computer Fraud & Security, Volume 2006, Issue 2, February 2006, Page 4.

Corston, R., Coleman, A., 2003. A crash course in SPSS. WileyBlackwell.

Cronin, M., J., 1997. Banking and Finance on the Internet. John Wiley and Sons.

Culwin, F., Faulkner, C., 1997. The integration of usability engineering considerations within the C.S./S.E. curriculum. The proceedings of the 13th annual eastern small college computing conference, Stockton, New Jersey, 169-178.

Curran, K., Woods, D., O Riordan, B., 2006. Investigating text input methods for mobile phones. Telematics and Informatics 23, 1-21.

Davies, F., Mouthinho, L., Curry, B. 1996. ATM user attitudes: a neural network analysis. Marketing Intelligence and Planning 14 (2), 26-32.

Delone, W., H., McLean, E., R., 1992. Information systems success: The quest for the dependent variable. Information System Research 3 (1), 60-95.

DeWitt, A. J., Kuljis, J., 2006. Is usable security an oxymoron? Interactions, 13 (3), ACM Press, 41-44.

Drimer, S., Murdoch, S., J., Anderson, R., 2009 (in press). Optimised to fail: Card readers for online banking. Financial Cryptography and Data Security '09. Springer LNCS.

Dutton, R., T., Foster, J., C., Jack, M., A., Stentiford, F., W., M., 1993. Identifying usability attributes of automated telephone services, Proc. European Conference on Speech Communication and Technology (Eurospeech 93), 1335-1338.

Faulkner, X., Culwin, F., 2005. When fingers do the talking: a study of text messaging. Interacting with Computers 17, 167-185.

Faulkner, L., 2003. Beyond the five-user assumption: Benefits of increased sample sizes in usability testing. Behaviour Research Methods, Instruments and Computers 35 (3), 379-383.

Faulkner, X., 2000. Usability engineering. Palgrave Macmillan.

Fornell, C., and Wernerfelt, B. 1987. Defensive marketing strategy by customer complaint management: A theoretical analysis. Journal of Marketing Research, 24, 337-346.

Friedman, Z., Mukherji, S., Roeum, G., K., Ruchir, R. 2001. Data input into mobile phones: T9 or keypad? Student Online HCI Research Experiments. Retrieved June 10, 2006, from http://www.otal.umd.edu./SHORE2001/mobilePhone/index.html

Frokjaer, E., Hertzum, M., Hornbaek, K., 2000. Measuring usability: are effectiveness, efficiency and satisfaction really correlated? CHI Letters 2, 1, 345-352.

Furnell, S., 2005. Why users cannot use security. Computers and Security 24, 274 – 279.

Gerrard, P., Cunningham, J., B., 2003. The diffusion of internet banking among Singapore consumers. International Journal of Bank Marketing 21 (1), 16-28.

Grabner-Krauter, S., Kaluscha, E., A., 2003. Empirical research in on-line trust: a review and critical assessment. International Journal of Human Computers Studies 58 (6), 783-812.

Guttman, L., Suchman, E., A., 1967. Intensity and a zero point for attitude analysis. In. M., Fishbein (Ed.), Readings in Attitude theory and Measurement, New York, Wiley, pp. 267-276.

Hudson, W., 2002. The lost world of e-banking. ACM SIGCHI Bulletin – a supplement to interactions, 7-7.

Hole, K.J., Moen, V., Tjostheim, T., 2006. Case study: online banking security, IEEE Security & Privacy, 4 (2), 14-20.

Hyvarinen, T., Kaikkonen, A., Hiltunen, M., 2005. Placing links in mobile banking applications. Proceedings of the 7th International Conference on Human Computer Interaction with Mobile Devices and Services, 63-68.

International Organisation for Standardisation (ISO) 9241-11, 1998. Ergonomic requirements for office work with Visual Display Terminals (VDTs) – Part 11: Guidance on Usability.

International Organisation for Standardisation (ISO) 13407, 1999. Human Centred Design Process for Interactive Systems.

Jack, M., A., Foster, J., C., Stentiford, F., W., M., 1993. Usability analysis of intelligent dialogues for automated telephone services. Proc. Joint ESCA/NATO workshop on Applications of Speech Technology, 149-152.

Jagacinski, R., J., Liao, M., J., Fayyad, E., A., 1995. Generalized slowing in sinusoidal tracking in older adults. Psychology of Aging 9, 103-112.

James, C., L., Reischel, K., M., 2001. Text input for mobile devices: Comparing model prediction to actual performance. Proceedings of CHI'01 Human Factors in Computing Systems, ACM Press., New York, 365-371.

Jarvenpaa, S., Lang, K., R., Takeda, Y., Tuunainen, V., K., 2003. Mobile commerce at crossroads: An international focus group study of users of mobile handheld devices and services. Commun. ACM 46, (12), 41-44.

Jayawardhena, C., Foley, P., 2000. Changes in the Banking Sector - the case of Internet Banking in the UK. Internet Research: Electronic Networking, Applications and Policy 10 (1), 19-30.

Johnston, R., & Mehra, S. (2002). Best-practice complaint management. Academy of Management Journal 16 (4), 145-154.

Johnston, J., Eloff, J., H., P., Labuschagne, L., 2003. Security and human computer interfaces. Computers and Security 22 (8), 675-84.

Josang, A., AlZomai, M., Suriadi, S., 2007. Usability and privacy in identity management architectures. In: Brankovic, L., Steketee, C., (Ed.), Fifth Australasian

information security workshop (Privacy Enhancing Technologies) AISW'07, 68 CRPIT, Ballarat, Australia, ACS, 143-152.

Kaikkonen, A., Kallio, T., Kekalainen, A., Kankainen, A., Cankar, M., 2005. Usability testing of mobile application: A comparison between laboratory and field testing. Journal of Usability Studies, 1, 1, 4-16.

Karajaluoto, H., 2002. Selection criteria for a mode of bill payment: empirical investigation among Finnish bank customers. International Journal of Retail & Distribution Management, 30 (6), 331-9.

Karajaluoto, H., Mattila, M., Pento, T., 2002. Electronic banking in Finland: consumer beliefs and reactions to a new delivery channel. Journal of Financial Services Marketing 6 (4), 346-61.

Kasesniemi, EE-L., Rautiainen, P., 2002. Mobile culture of children and teenagers in Finland, in: Katz, J., E., Aarkhus, M., (Eds.), Perpetual Contact. CUP, Cambridge, 170-193.

Khodawindi, D., Pousttchi, K., Wiedemann, D., G., 2003. Akzeptanz mobiler Bezahlverfahren in Deutschland. In Pousttchi, K., Turowski, K., (Eds.) Mobile commerce – Anwendungen und perspektiven. Proceedings zum 3. Workshop mobile commerce, Augsburg, 42-57.

Kim, G., Shin, B., Lee, H., G., 2008. Understanding dynamics between initial trust and usage intention of mobile banking. Information Systems Journal. OnlineEarly Articles.

Kim, K., Prabhakar, B., 2000. Initial trust, perceived risk, and the adoption of internet banking. Proceedings of the 21st International Conference on Information Systems, 537-43.

Kim, J., Moon, J., Y., 1998. Designing emotional usability in customer interfaces trustworthiness of cyber-banking system interfaces. Interacting with Computers 10, 1-29.

Krampe, R., T., Ericsson, K., A., 1996. Maintaining excellence: deliberate practice and elite performance in young and older pianists. Journal Experimental Psychology General 125, 331-359.

Kuniavsky, M., 2003. Observing the user experience: A practitioner's guide to user research. San Francisco: Morgan Kaufmann.

Kurniawan, S., 2008. Older people and mobile phones: A multi-method investigation. Interntaionl Journal of Human-Computer Studies, 66, (12), 889-901.

Kurniawan, S., H., King, A., Evans, D., G., Blenkhorn, P., L., 2006a. Personalising web page presentation for older people. Interacting with Computers 18, 3, 457-477.

Kurniawan, S., Nugroho, Y., Mahmud, M., 2006b. A study of the use of mobile phones by older persons. Proceedings of CHI'06 Human Factors in Computing Systems, ACM Press., New York, 989-994.

LaLomia, M.J., Sidowski, J.B., 1990. Measurements of computer satisfaction, literacy, and aptitudes: A review. Int. J. Human-Computer Interaction 2, 3, 231-253.

Landauer, T. K., 1988. Research methods in human-computer interaction. In: Helenader, M. (Ed.), Handbook of Human-Computer Interaction. North-Holland, Amsterdam, pp. 905-928.

Laukkanen, T., Lauronen, J., 2005. Consumer value creation in mobile banking services. International Journal of Mobile Communications 3 (4), 325-328.

Laukkanen, T., 2007. Internet vs mobile banking: comparing customer value perceptions. Business Process Management Journal 13 (6), 788-797.

Lee, K., C., Chung, N., 2009. Understanding factors affecting trust in and satisfaction with mobile banking in Korea: A modified Delone and Mclean's model perspective. Interacting with Computers 21 (5-6), 385-392.

Lee, Y., Kim, J., 2002. From design features to financial performance: a comprehensive model of design principles for online stock trading sites. Journal of Electronic Commerce Research 3 (3), 128-143.

Lee, M.S.Y., McGoldrick, P.F., Keeling, K.A., Doherty, J., 2003). Using ZMET to explore barriers to the adoption of 3G mobile banking services. International Journal of Retail & Distribution Management 31 (6), 340-8.

Lee, Y., S., Hong, S., W., Smith-Jackson, T., L., Nussbaum, M., A., Tomioka, K., 2006. Systematic evaluation methodology for cell phone user interfaces. Interacting with Computers 18, 304-325.

Lewis, J., R., Sauro, J., 2006. When 100% really isn't 100%: Improving the accuracy of small-sample estimates of completion rates. Journal of Usability Studies Issue 3, 1, 136-150.

Lichtenstein, S., Williamson, K, 2006. Understanding Consumer Adoption Of Internet Banking: An Interpretive Study in the Australian Banking Context. Journal of Electronic Commerce Research 7 (2), 50-66.

Likert, R., 1932. A technique for the measurement of attitudes. Archives of Psychology 140.

Liljander, V., Polsa, P., Forsberg, K., 2007. Do mobile CRM services appeal to loyalty program customer. International Journal of e-Business Research 3 (2), 24-40.

Lindgaard, G., 2009. Early traces of usability as a science and as a profession. Interacting with Computers 21 (5-6), 350-352.

Ling, R., 2000. 'We will be reached': the use of mobile telephony among Norwegian youth. Information Technology and People 13, 2, 102-120.

Ling, R., 2004. The mobile connection – the cell phone's impact on society. Morgan Kaufmann, San Francisco.

Ling, R., 2007. SMS og hvordan elder blir utestengt. I Proitz, L., Luders, M and Rasmussen, T. Livet I og utenfor skjermene. Uniersitetetsforlaget, Oslo.

Liao, Z., Cheung, M., T., 2002. Internet based e-banking and consumer attitudes: an empirical study. Information and Management 39, 283-295.

Lockett, A., Littler, D., 1997. The adoption of direct banking services. Journal of Marketing Management 13 (8), 791-811.

Love, S., Dutton, R., T., Foster, J., C., Jack, M., A., Nairn, I., A., Vergeynst, N., A., Stentiford, F., W., M., 1992. Towards a usability measure for automated telephone services. Proc. Institute of Acoustics Speech and Hearing Workshop, 14 (6), 553-559.

Love, S., Dutton, R., T., Foster, J., C., Jack, M., A., Stentiford, F., W., M., 1994. Identifying salient usability attributes for automated telephone services. Proc. International Conference on Spoken Language Processing (ICSLP-94), 1307-1310.

Love, S., 1997. The Role of Individual Differences in Dialogue Engineering for Automated Telephone Services (PhD thesis, University of Edinburgh).

Luarn, P., Lin, H., H., 2005. Toward an understanding of the behavioral intention to use mobile banking. Computers in human behavior 21 (6), 873-91.

Mallat, N., Rossi, M., Tuunainen, V., K., 2004. Mobile banking services. Communications of the ACM 47 (5), 42-46.

Mayhew., D., J., 1999. The usability engineering lifecycle: A practitioner's handbook for user interface design. Morgan Kaufmann.

Mathieson, S., A., 2007. Barclays and Lloyds lead e-banking security. Infosecurity 4 (4) 8.

McGuffin, M., J., Balakrishnan, R., 2005. Fitts' Law and expanding targets: experimental studies and designs for user interfaces. ACM Transactions on computer-human interaction (TOCHI) 12 (4), 388-422.

Miller, R., B., 1971. Human ease of use criteria and their tradeoffs. IBM report TR 00.2185. Poughkeepsie, NY:IBM Corporation.

Moore, D., S., 2001. Statistics: Concepts and controversies. WH Freeman and Co, NY.

Nilsson, M., Adams, A., Herd, S. 2005. Building security and trust in online banking. In: CHI '05 Extended Abstracts on Human Factors in Computing Systems (Portland, OR, USA, April 02 - 07, 2005), CHI '05. ACM Press, New York, NY, pp. 1701-1704.

Neilsen, J., 1993. Usability Engineering. Cambridge, MA: Academic Press.

Nodder, C., 2005. Users and trust: a Microsoft case study. In: Cranor., Garfinkel. (Ed.), Security and Usability. O'Reilly, pp, 129-124.

O'Gorman, L., 2003. Comparing passwords, tokens and biometrics for user authentication. Proceedings of the IEEE 91 (12), 2021-40.

Ornella, P., Stephanie, B., 2006. Universal Designs for Mobile Phones: A Case Study. Computer Human Interaction 2006 (Work in Progress). Quebec, Canada.

Peevers, G., McInnes, F., Morton, H., Matthews, A., Jack, M., A., 2009a. The mediating effects of brand music and waiting time updates on customers' satisfaction with a telephone service when put on-hold. International Journal of Bank Marketing 27 (2 & 3), 202-217.

Peevers, G. and McInnes, F., 2009b (in press). Laboratory studies. In: Love, S. (Ed.), Handbook of Mobile Technology Research Methods. Nova, Hauppauge NY.

Peevers, G., Douglas, G., Jack, M. A., 2008. A usability comparison of three alternative message formats for an SMS banking service. International Journal of Human-Computer Studies 66, (2) 113-123.

Piazzalunga, U., Savaneschi, P., Coffetti, P., 2005. The usability of security devices. In: Cranor and Garfinkel (Ed.), Security and Usability, O'Reilly, pp. 221-242.

Pousttchi, K., Schurig, M., 2004. Assessment of today's mobile banking applications from the view of customer requirements. Proceedings of the 37th Hawaii International Conference on System Sciences. IEEE.

Preece, J., Rodgers, Y., Sharp, H., 2002. Interaction design: Beyond human-computer interaction New York, NY. Wiley.

Preece, J., Rodgers, Y., Sharp, H., Benyon, D., Holland, S., Carey, T., 1994. Human-computer interaction. Addison-Wesley.

Polatoglu, V., N., Ekin, S. 2001. An empirical investigation of the Turkish consumers' acceptance of internet banking services. International Journal of Bank Marketing 19 (4), 156-65.

Renaud, K., 2005. Evaluating authentication mechanisms. In: Cranor., Garfinkel. (Ed.), Security and usability. O'Reilly, pp, 103-128.

Richardson, A., 2005. Turning Transactions Into Relationships: Metrics for Usability in the Financial Services Sector, PhD thesis, University of Edinburgh.

Root, R.W., Draper, S., 1983. Questionnaires as a software evaluation tool. Proceedings of CHI 83. ACM, New York.

Rugimbana, R., Iversen, P. 1994. Perceived attributes of ATMs and their marketing implications. International Journal of Bank Marketing 12 (2), 30-35.

Ryu, Y., S., Smith-Jackson, T., L., 2006. Reliability and Validity of the Mobile Phone Usability Questionnaire (MPUQ). Journal of Usability Studies 2 (1), 39-53.

Sathye, M., 1999. Adoption of internet banking by Australian consumers: an empirical investigation. International journal of Bank Marketing 17 (7), 324-34.

Sauro, J., Lewis, J., R., (2005). Estimating completion rates from small samples using binomial confidence intervals: comparisons and recommendations. Proceedings of the Human Factors and Ergonomics Society Annual Meeting (HFES 2005) Orlando, FL.

Schultz, E., E., Proctor, R., W., Lien, M., C., Savendy, G., 2001. Usability and security: an appraisal of usability issues in information security methods. Computers and Security 20 (7), 620-34.

Shackel, B., 1991. Usability – Context, framework, definition, design and evaluation. In: Shackel, B., Richardson, S., J., (Ed.), Human Factors for Informatics Usability, 21-37. NY: Cambridge University Press.

Shackel, B., 1981. The concept of usability. Proceedings of IBM Software and Information Usability Symposium. IBM, NY, 1-30.

Shaw, M., Gardner, D., Thomas, H., 1997. Research opportunities in electronic commerce. Decision Support Systems 21, 27-33.

Shih, Y., Fang, K., 2004. The use of a decomposed theory of planned behaviour to study internet banking in Taiwan. Internet Research 14 (3), 213-223.

Sillence, E., Barber, C., 2004. Integrated digital communities: combining and competition. Interacting with Computers 16, 1, 93-113.

Sinclair, S., Smith, S.W., 2005. The TIPPI Point: Towards trustworthy interfaces, IEEE Security and Privacy, 68-71.

Sohail, M., S., Shanmugham, B., 2003. E-banking and Customer Preference in Malaysia: An Empirical Investigation. Information Sciences 150, 207-217.

Soriano, C., Raikundalia, G.K., Szajman, J., 2005. A usability study of short message service on middle-aged users. Proceedings of OZCHI 2005, Canberra, Australia.

Stamoulis., 2002. An approach and model for assessing the business value of e-banking channels: evaluation an communication. International Journal of Information Management 22 (4), 247-261.

Suh, B., Han., I., 2002. Effect of trust on customer acceptance of internet banking. Electronic Commerce Research and Applications 1 (3), 247-263.

Suoranta, M., 2003. Adoption of mobile banking in Finland. Phd, Jyva¨skyla¨, Finland, Jyva¨skyla¨ Studies in Business and Management 28.

Suoranta, M., Mattila, M., 2004. Mobile banking and consumer behaviour: new insights into the diffusion pattern. Journal of Financial Services Marketing 8 (4), 354–366.

Tan, M., Teo, T., S., H., 2000. Factors Influencing the Adoption of Internet Banking. Journal of the Association for Information Systems 1, 1-42.

Tiwari, R., Buse, S., 2007. The mobile commerce prospects: A strategic analysis of opportunities in the banking sector. Hamburg: Hamburg University Press.

Tiwari, R., Buse, S., Herstatt, C., 2007. Mobile services in banking sector. The role of innovative business solutions in generating competitive advantage. Proceedings of the 8th International Research Conference on Quality, Innovation and Knowledge Management.

Tognazzini, B., 2005. Design for usability. In: Cranor., Garfinkel. (Ed.), Security and Usability. O'Reilly, pp, 31-46.

Tullis, T., S., Stetson, J., N., 2004. A comparison of questionnaires for assessing website usability. Proceedings of UPA'04 .

Tullis, T., Albert, B., 2008. Measuring the user experience: Collecting, analyzing, and presenting usability metrics. Morgan Kaufmann.

Walker, N., Philbin, D., A., Fisk, A.D., 1997. Age-related differences in movement control: adjusting submovement structure to optimise performance. Journal of Gerontology: Psychological Sciences 52B, 1, 40-52.

Weir, C., S., Anderson, J., A., Jack, M., A., 2006. On the role of metaphor and language in design of third party payments in eBanking: usability and quality. International Journal of Human-Computer Studies 64 (8), 771-785.

Weir, C., S., McKay, I., Jack, M., A., 2007. Functionality and usability in design for eStatements in eBanking services. Interacting with Computers 19 (2), 241-256.

Weir, C., S., 2008. Investigating the Relationship between usability, preferences and usage intentions when banking online, PhD thesis, University of Edinburgh.

Weir., C., S., Douglas, G., Carruthers, M., Jack, M., 2009a. User perceptions of security, convenience and usability for ebanking authentication tokens. Computers and Security 28, 47-62.

Weir, C., Douglas, G., Richardson, T., Jack., M., 2009b. Usable Security: User Preferences for authentication methods in eBanking and the effects of experience. Interacting with Computers (In Press).

Weiss, S., 2002. Handheld Usability. John Wiley & Sons, LTD, West Sussex.

Whiteside, J., Bennett, J., Holzblatt, K., 1988. Usability engineering: Our experience and evolution. In: Helander, M., (Ed.), Handbook of human computer interaction, 91-817. New York: North-Holland.

Zviran, M., Haga, W., J., 1999. Password security: an empirical study. Journal of Management of Information Systems 15 (4), 161-85.

Ziefle, M., 2002. The influence of user expertise and phone complexity on performance, ease of use and learnability of different mobile phones. Behaviour and Information Technology, 21 (5), 303-311.

URLs

Ahmed, Z., 2004. Youth drives India's mobile phone revolution. BBC News, April. Available from: < http://news.bbc.co.uk/1/hi/business/3585257.stm >

BBC News, 2005a. Vodafone buys into India's Bharti, 28th October. Available from: < http://news.bbc.co.uk/1/hi/business/4384258.stm >

BBC News, 2005b. A back-to-basics mobile launched, 21st May. Available from: < http://news.bbc.co.uk/1/hi/technology/4566809.stm >

BBC News, 2006. Texting record for New Year's Eve, 4th January. Available from: < http://news.bbc.co.uk/1/hi/technology/4582060.stm >

BBC News, 2007a. Enfield's cash gift to the world, 27 June. Available from: < http://www.bbc.co.uk/london/content/articles/2007/06/26/cash_machine_feature.shtml >.

BBC News, 2007b. Banks issues net security device, 2 October. Available from: < http://news.bbc.co.uk/1/hi/business/7023743.stm >.

Christopher, P., 1999. Older adults – Special considerations for special people. Available from:
< http://www.gsu.ed/~mstswh/course/it7000/papers/newpage31.htm >

Eatoni Ergonomics Inc, 2006. Eatoni's LetterWise Predictive Text System. Available from: http://www.eatoni.com/wiki/index.php/LetterWise.

Federal Financial Institutions Examination Council FFIEC, 2005. Press release:
< http://www.ffiec.gov/pdf/authentication_guidance.pdf >.

Flinders, K., 2009. Security weakness found in online banking card readers. ComputerWeekly.com, 27 February. Available from: <
http://www.computerweekly.com/Articles/2009/02/27/235062/security-weakness-found-in-online-banking-card-readers.htm >.

Garson, G., D., 1998. Data levels and measurement. From Statnotes: Topics in Multivariate Analysis. Available from:
http://www2.chass.ncsu.edu/garson/pa765/datalevl.html#likert

InfoPrint Solutions Company. Marketers are still not targeting their customers effectively, press release. Available from: <
http://www.infoprintsolutionscompany.com/internet/wwsites.nsf/vwwebpublished/ai_pr030408_marketing_survey_xn >.

Monyextra.com, 2008. UK Banking Sector fire fighting response to online fraud, 28 August. Available from: < http://www.moneyextra.com/news/uk-banking-sector-fire-fighting-response-to-online-fraud.html >.

The Mobile Data Association, 2009. The Q4 2008 UK mobile report (press release), 13 February. Available from: < http://www.themda.org/mda-press-releases/the-q4-2008-uk-mobile-trends-report.php >.

Martin, O., 2010. Texting was never actually designed for the consumer market. The Guardian. Available from: <
http://www.guardian.co.uk/business/2010/jan/01/texting-never-designed-for-consumers >

Rumpa, G., 2005. Are you banking more on your mobile? Times of India, 6th February. Available from:
http://timesofindia.indiatimes.com/articleshow/1013140.cms

The Register 2007, Anam introduces SMS payments Available at:
http://www.theregister.co.uk/2007/06/18/hash_cash_sms_payments/

Ward, L., 2004. Texting 'is no bar to literacy'. Guardian Unlimited, December 23rd. Available from:
http://education.guardian.co.uk/schools/story/0,5500,1378951,00.html

Yeo, V., 2006. Asia's banks split on SMS authentication. ZDNet Asia, October. Available from: < http://www.zdnetasia.com/insight/security/0,39044829,61960225,00.htm >.

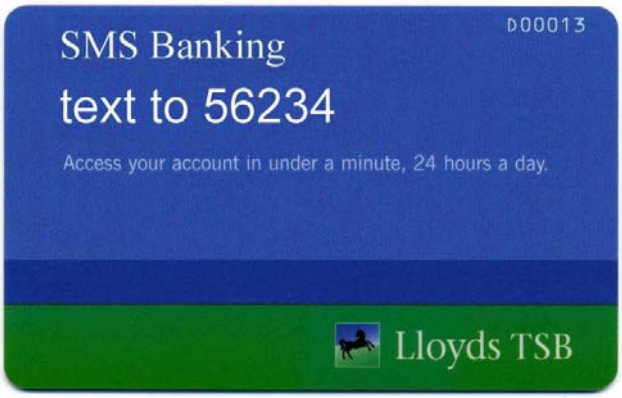
Appendix A: Materials from Chapter 4

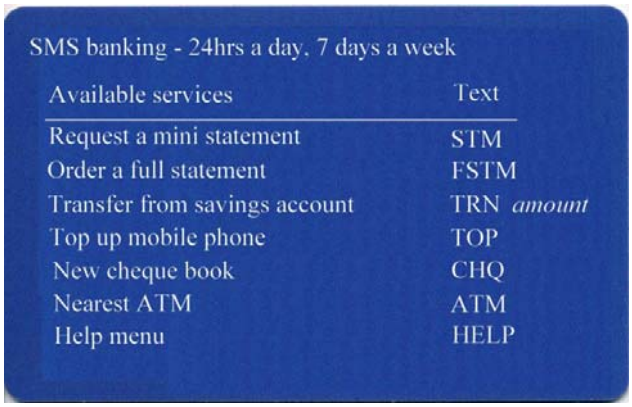
Task Sheets

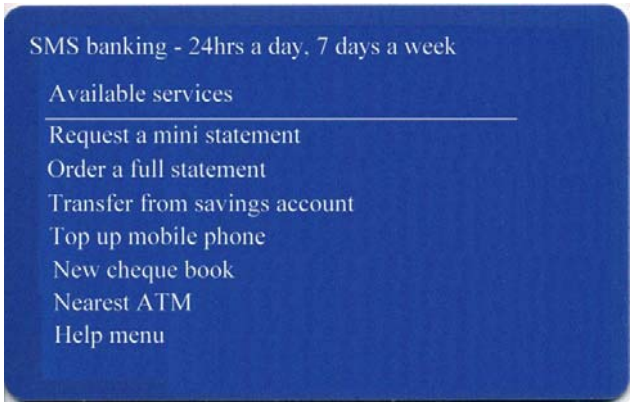
Each participant carried out the three tasks with each version of SMS banking

- Order a new cheque book using the SMS banking service
- Get a mini-statement using the SMS banking service
- Transfer £200 from your savings account to your current account using the SMS banking service

SMS Banking customer cards







SMS banking - 24hrs a day, 7 days a week

Available services	Text	
Request a mini-statement	1	
Order a full statement	2	
Transfer from savings account	3	*amount*
Top-up mobile phone	4	
New cheque book	5	
Nearest ATM	6	
Help menu	HELP	

Exit interview questionnaire

"Today you have experienced three different versions of SMS text message banking. I am now going to ask you some questions to find out what you thought of them."

1. Was there anything in particular that you liked about the version A of the SMS text messaging approach? ***Hand participant the SMS banking card for version A as a reminder***

……..
…….
……..
……..

2. Was there anything about this version that you disliked, or that could be improved?

……..
…….
……..
……..

3. Was there anything in particular that you liked about version B of the SMS text messaging approach? ***Hand participant the SMS banking card for version B as a reminder***

……..
…….
……..
……..

4. Was there anything about this version that you disliked, or that could be improved?

……..
…….
……..
……..

5. Was there anything in particular that you liked about version C of the SMS text messaging approach? ***Hand participant the SMS banking card for version C as a reminder***

……..
…….
……..
……..

6. Was there anything about this version that you disliked, or that could be improved?

……..
…….
……..
……..

7. "Please place the pointers representing the different versions of SMS banking on the scale. Where 0 = poor and 30 = excellent. I'd like you to order and rate them by preference. Feel free to adjust them after you have placed all three on the ruler."

Show participant the SMS banking cards for each version before they place them on the ruler

SMS version	*Ruler Measurement*
A	
B	
C	

Comments

……..
…….
……..
……..

8. I'm now going to ask you for your views about the text messages the Bank sent to you in our experiment.

Show screen shot of mini statement response

Do you think this message contains enough information?

❑ Yes
❑ No

Comments

……..
…….
……..
……..

Show screen shot of transfer confirmation message response

Do you think this message contains enough information?

❑ Yes
❑ No

Comments
……..
…….
……..
……..

Show screen shot of cheque book ordering confirmation message response

Do you think this message contains enough information?

❑ Yes
❑ No

Comments
……..
…….
……..
……..

9. If Lloyds TSB introduced an SMS text messaging banking service like the one you've said you prefer, would you use it?

❑ Use it, but only if it was free of charge
❑ Use it, and willing to pay a small charge
❑ Don't know

Comments – why? How much would you pay for the service?

……..
…….
……..
……..

10. I would now like to find out what kind of services you would like to see available with SMS banking.

Hand participant services questionnaire

For each of the services listed indicate on the scale how likely you would be to make use of it, if it was offered by Lloyds TSB. There is also a space at the bottom for you to write in any services that you think we have missed out.

11. I would now like to ask you your views on the abbreviations we used in version A of the SMS text messaging service. ***Show participant customer card for version A***

Hand participant sheet with list of abbreviations

As you can see on the sheet there is a list of the abbreviations used in the experiment. If you like an abbreviation tick the box next to it. If not, please write down what you think the abbreviation should be in the space on the right.

12. If this kind of SMS text messaging service was introduced by Lloyds TSB what do you think it should be called?

……..
……
……..
……..

13. And finally, do you have any other comments that you would like to add?

……..
……
……..

Q.10

For each of the following services please indicate how likely you would be to use it if it was available with SMS banking.

	Definitely use						**Definitely would not use**
Request a mini statement.	❑	❑	❑	❑	❑	❑	❑
Order a full statement.	❑	❑	❑	❑	❑	❑	❑
Transfer from savings account.	❑	❑	❑	❑	❑	❑	❑
Top-up mobile phone.	❑	❑	❑	❑	❑	❑	❑
Order a new check book.	❑	❑	❑	❑	❑	❑	❑
Find the nearest ATM.	❑	❑	❑	❑	❑	❑	❑
Pay bills.	❑	❑	❑	❑	❑	❑	❑
Stop cheque.	❑	❑	❑	❑	❑	❑	❑
List of recent debits from account.	❑	❑	❑	❑	❑	❑	❑
List of recent credits to account.	❑	❑	❑	❑	❑	❑	❑
Account balance.	❑	❑	❑	❑	❑	❑	❑
Transfer to savings account.	❑	❑	❑	❑	❑	❑	❑
Find the nearest Lloyds TSB branch.	❑	❑	❑	❑	❑	❑	❑

Other services you would like to see available to you via SMS banking

……

……

……

……

Q.11

Please read the list of abbreviations below. Tick the corresponding box if you like the abbreviation, or write down what you think it should be in the space provided.

	Like	**Suggestions**
STM (for a mini-statement)	❑	…….
FSTM (to order a full statement)	❑	…….
TRN (to transfer from savings account)	❑	…….
TOP (for a mobile phone top-up)	❑	…….
CHQ (to order a new cheque book)	❑	…….
ATM (to find the nearest ATM)	❑	…….

Now please read the list of services below and write down what you think a suitable abbreviation would be for it.

	Suggestions
Pay bills	…….
Stop cheques	…….
List of recent debits from account	…….
List of recent credits to account	…….
Account balance	…….
Transfers to savings account	…….
Find the nearest Lloyds TSB branch	…….

Demographic Questionnaire

1. Age: ……. **2. Gender:** ❑ female ❑ male

3. Occupation:.....
(if retired or unemployed – previous occupation)

4. How long have you banked with Lloyds TSB?……..

5. Do you use Lloyds TSB's Internet banking service?

1. ❑ Yes
2. ❑ No

comments……

……

6a. Do you use Lloyds TSB's Text alert service?

1. ❑ Yes **(go to question 7)**
2. ❑ No

6b. (if no) Do you receive text messages on your mobile phone?

1. ❑ Yes
2. ❑ No **(go to question 8)**

(**if no**) comments……..

……..

7. Do you receive Text alerts from other banks, businesses or services?

1. ❑ Yes
2. ❑ No

(**if yes**) details……..

…….

8a. Do you send text messages from you mobile phone?

1. ❑ Yes
2. ❑ No

(**if no**) comments……..

……..

8b. (if yes) How often do you send text messages?

1. ❑ A few times a day or more
2. ❑ Daily
3. ❑ A few times a week
4. ❑ Weekly
5. ❑ Monthly
6. ❑ less often

8c. (if yes) Do you use predictive text when you write text messages?

1. ❑ Yes
2. ❑ No
3. ❑ Sometimes

(**if no**) comments……..

……..

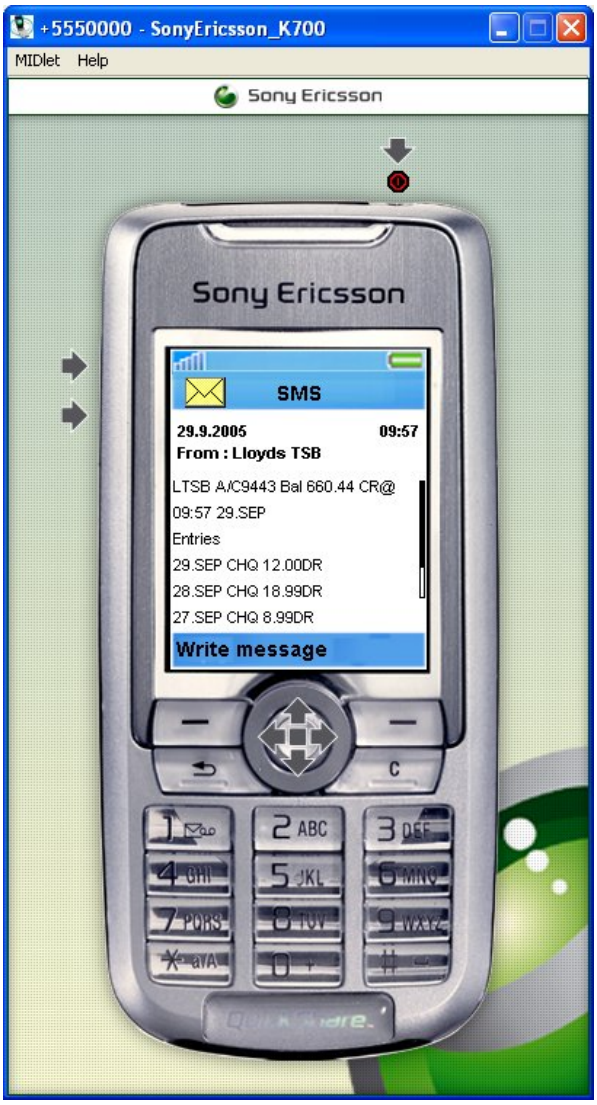
9. What type of mobile phone do you own? *If they can't remember ask them if they have it on them so they can show you*

……..

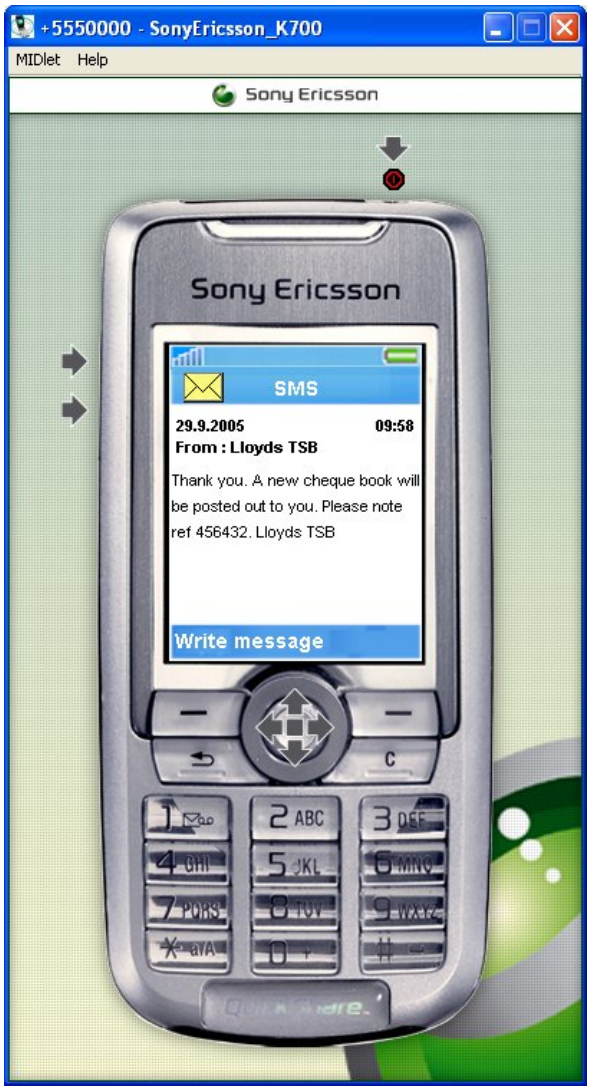
……..

Confirmation Messages

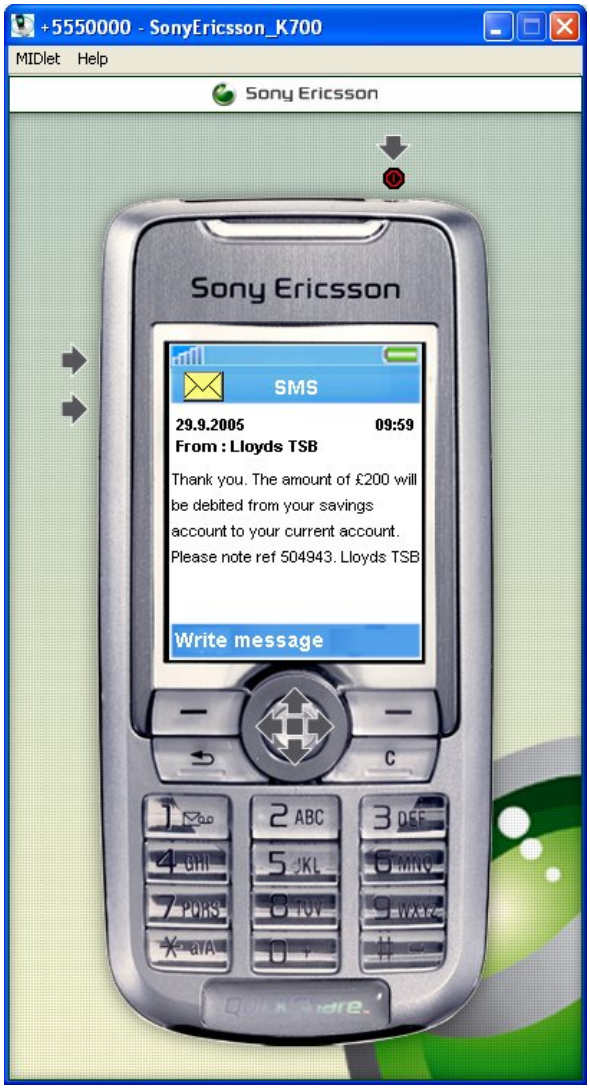
Mini statement



Cheque book


+5550000 - SonyEricsson_K700
MIDlet Help
Sony Ericsson
Sony Ericsson
SMS
29.9.2005 09:58
From : Lloyds TSB
Thank you. A new cheque book will be posted out to you. Please note ref 456432. Lloyds TSB
Write message

Transfer


+5550000 - SonyEricsson_K700
MIDlet Help
Sony Ericsson
Sony Ericsson
SMS
29.9.2005 09:59
From : Lloyds TSB
Thank you. The amount of £200 will be debited from your savings account to your current account. Please note ref 504943. Lloyds TSB
Write message

Appendix B: Materials from Chapter 5

Demographic Questionnaire

1. Age: ……. 2. Gender: □ male □ female

3. Occupation:.....
(if retired or unemployed – previous occupation)

4. Do you use any of Lloyds TSB's mobile phone banking services?
□ Yes □ No

If yes
Please describe which services you use
…….
…….

If no
Why do you not use any services?
…….
…….

5a. Do you send SMS text messages using your mobile phone?
□ Yes □ No

5b. (if yes) How often do you send SMS text messages?

1□ A few times a day or more
2□ Daily
3□ A few times a week
4□ Weekly
5□ Monthly
6 □ Less often

5c. (if yes) Do you use predictive text when you write SMS text messages?
□ Yes □ No □ Sometimes

6. Do you receive SMS text message alerts from other banks, businesses or services?
□ Yes □ No
(if yes)
details……..

Example Task Sheets from Part 1

Imagine you are Mrs L Brown

You want to get your balance on your Lloyds TSB current account.
Call Lloyds TSB and use the automated service to get your balance.

You will need to use the information below.

Security Number 846134

Imagine you are Mrs L Brown.

You want to get your balance on your Lloyds TSB current account.
Call Lloyds TSB and speak to an advisor to get your balance.

You will need to use the information below.

Recent transaction £59.99 at Next, yesterday

Your Lloyds TSB Branch Cameron Toll, Edinburgh

Date of Birth 21st August 1970

Imagine you are Mrs F MacKay.

You want to get your balance on your Lloyds TSB current account.

Text Bal 5390

to 61119 to request your current account balance using your registered mobile phone.

Example Task Sheets from Part 2

Imagine you are Mr H Evans.

You want to transfer some money from your Lloyds TSB current account to your friend, Paul Williams's account.

Call Lloyds TSB and use the automated service to transfer **£300** to Paul Williams's account.

You will need to use the information below.

Your details:

Security Number: **405992**

Paul Williams's details:

Account number **50944492**
Sort Code **87-13-09**

Amount to transfer: £300

Imagine you are Mrs S Jones.

You want to get transfer some money from your Lloyds TSB current account to your friend, Brenda White's account.

Call Lloyds TSB and use the automated service to transfer **£350** to Brenda White's account.

You will need to use the information below.

Your details:

Security Number: **197423**

Brenda White's details:

Account number **33958785**
Sort Code **87-43-90**

Amount to transfer: £350

Exit Interview for Part 1

"Today you've experienced two different ways of getting your current account balance….

The first one was when you phoned and spoke to an advisor and the second one was when you send a text message on your mobile and received a text message reply"

1. What did you like most about using the first version, for which you spoke to an advisor to get your balance?

.....

.....

.....

2. What did you dislike about using the first version, for which you spoke to an advisor to get your balance?

.....

.....

.....

3. Do you have any suggestions for improvements?

.....

.....

.....

4. What did you like most about using the second version which was the text message service to get your balance?

.....

.....

.....

5. What did you dislike about using the second version which was the text message service to get your balance?

.....

.....

.....

6. Do you have any suggestions for improvements?

.....

.....

.....

7. Please rate each method you have used along the ruler between Poor and Excellent using these magnets. Explain your ratings:

	Overall Rating (30cm scale)
1st Method (Advisor)	
2nd Method (SMS)	

Why?

.....

.....

8. Was the information clear in the text message about your balance?

❑ Yes ❑ No

Comments

.....

.....

9. Was there enough information in the text message?

❑ Yes ❑ No

Comments

.....

.....

10. If Lloyds TSB were to offer you the ability to request a balance by SMS text message like you experienced here today would you be more or less likely to recommend Lloyds TSB to a friend or colleague?

- ❑ More likely
- ❑ Less likely
- ❑ No difference
- ❑ Don't know

Comments – why?

.....

.....

11. If Lloyds TSB were to offer you the ability **to request a balance by SMS text message like you experienced here today and it would cost no more than the usual charge for sending a text message, on average how many times per month would you use it, if at all?**

[ ] times per month

Comments – why?

.....

.....

12. Under what circumstances do you imagine you would use an SMS text message request to get a balance rather than by telephoning Lloyds TSB?

.....

.....

→ ***Change labels on ruler to 'very unlikely', 'very likely'.***

13. Using the ruler provided, please indicate how likely you think you would be to use either PhoneBank, an SMS text request, Internet Banking, an ATM or the branch to get you current account balance.

	Overall Rating (30cm scale)
PhoneBank - advisor	
PhoneBank – automated service	
SMS	
Internet Banking	
ATM	
Branch	

Comments – why?

.....

.....

14. I would now like to find out what kind of services or information you would like to be able to request using an SMS text message to Lloyds TSB.

Hand participant services questionnaire A

For each of the services listed indicate on the scale how likely you would be to make use of it, if it was offered by Lloyds TSB. There is also a space at the bottom for you to write in any services or information that you think we have missed out.

Questionnaire A

For each of the following, please indicate how likely you would be to use it if you could request it via an SMS text message.

	Definitely would use						**Definitely would not use**
Request a balance.	❑	❑	❑	❑	❑	❑	❑
Request a mini-statement list of your last 6 transactions	❑	❑	❑	❑	❑	❑	❑
Request a new cheque book.	❑	❑	❑	❑	❑	❑	❑
Request a new PIN for your bank card.	❑	❑	❑	❑	❑	❑	❑
Pay a bill.	❑	❑	❑	❑	❑	❑	❑
Transfer money between your own accounts.	❑	❑	❑	❑	❑	❑	❑
Transfer money to someone else's account.	❑	❑	❑	❑	❑	❑	❑
Stop a cheque.	❑	❑	❑	❑	❑	❑	❑
Find the nearest Lloyds TSB branch.	❑	❑	❑	❑	❑	❑	❑
Block a debit or credit card	❑	❑	❑	❑	❑	❑	❑
Request a brochure, e.g. insurance, mortgage	❑	❑	❑	❑	❑	❑	❑
Request an insurance quote	❑	❑	❑	❑	❑	❑	❑
Request a Standing Order cancellation	❑	❑	❑	❑	❑	❑	❑
Request a change to your overdraft limit	❑	❑	❑	❑	❑	❑	❑
Request a full paper statement	❑	❑	❑	❑	❑	❑	❑
Request someone from the bank to call you back	❑	❑	❑	❑	❑	❑	❑

Other services you would like to see available to you via SMS request

……

……

……

Exit Interview for Part 2

"Today you've experienced three different ways of transferring money and in each case you phoned the automated service."

"In the first version you experienced, you heard this" - play message B, show laminate B

"In the second version you experienced, you heard this" - play message A, show laminate A

"And in the third version you experienced, you heard this" - play message C, show laminate C "and you received a text message"

1. Please rate each version you have used along the ruler between Poor and Excellent using these magnets. Explain your ratings:

	Overall Rating (30cm scale)
1st version (FEIVR+BAL)	
2nd version (FEIVR)	
3rd version (FEIVR+SMS)	

Why?

.....

.....

2. When completing the transfer on the phone, in the first version you were given an updated balance. What did you think of this feature?

.....

.....

3. When completing the transfer on the phone, in the third version you received an SMS text message confirmation. What did you think of this feature?

.....

.....

4. Was the information clear in the text message confirming your transfer?

❑ Yes ❑ No

Comments

.....

.....

5. Was there enough information in the text message confirming your transfer?

❑ Yes ❑ No

Comments

.....

.....

6. Assuming there was no cost to you, would you prefer the automated service to give you the choice of getting an SMS text message confirmation or should it be automatic?

❑ Choice ❑ Auto ❑ Don't know

Comments

.....

.....

7. Should the automated service give you the option of entering the phone number you'd like to receive the text message on?

❑ Yes ❑ No ❑ Don't know

Comments

.....

.....

8. If Lloyds TSB were to start using SMS text message transaction confirmation messages like you experienced here today would you be more or less likely to recommend Lloyds TSB to a friend or colleague?

❑ More likely
❑ Less likely
❑ No difference
❑ Don't know

Comments – why?

.....

.....

9. Thinking about the kind of services you would like to receive an SMS confirmation for.

Hand participant services questionnaire B

There are different ways you could receive a banking confirmation – by SMS text message, by letter or by phone call. For each of the services listed indicate on the scale how you would prefer to receive a confirmation if it was offered by Lloyds TSB. There is also a space at the bottom for you to write in any services or information that you think we have missed out.

10. If Lloyds TSB were to send you SMS text messages to inform you about transactions on your accounts, how many times a week do you think would be an acceptable number to receive?

☐ times per week

Comments – why?

.....

.....

11. Thinking back to the balance request you sent by SMS text message earlier, if Lloyds TSB were to offer an SMS text message request service would you use it?
[*Read options*]

❑ Use it, but only if it cost no more than standard charge for sending an SMS text message
❑ Use it and willing to pay a small charge
❑ Wouldn't use it
❑ Don't know

Comments

.....

.....

If willing to pay a charge
11b. What kind of charge would you prefer? [*Read options*]

❑ A small fee every time you send a request How much……..
❑ A monthly subscription cost How much……..
❑ Don't know
❑ Other……

Comments

.....

.....

12. Do you have any final comments?

....

....

..

Questionnaire B

For each of the following services please indicate how you would prefer to receive a confirmation.

	SMS text message	**Letter**	**Phone Call**	**None**
After you make a funds transfer	❑	❑	❑	❑
After you contact Lloyds TSB to tell them about a change of address.	❑	❑	❑	❑
After you contact Lloyds TSB to tell them about a change of phone number.	❑	❑	❑	❑
After you have raised a complaint with Lloyds TSB and wish to receive progress updates	❑	❑	❑	❑
After you have contacted Lloyds TSB to request a change of PIN	❑	❑	❑	❑
After you have filled in a form such as a loan or mortgage application and wish to receive progress updates	❑	❑	❑	❑
When a large payment has been credited into your account.	❑	❑	❑	❑
When a large one-off payment has been debited out of your account	❑	❑	❑	❑
Setting-up a new payment arrangement, e.g. Sanding Order or Direct Debit	❑	❑	❑	❑
When there has been an overseas transaction on your account	❑	❑	❑	❑
After a debit card transaction over the Internet	❑	❑	❑	❑

Other services for which you would like to receive a confirmation

.....

.....

Appendix C: Materials from Chapter 6

Demographics and Technographics

(Completed by researcher)

1. Date of birth ……. (dd/mm/yyyy)

2. Gender [ ] Female
[ ] Male

3a. What is your position in the company?.....

3b. How many employees in your company?.....

4. Do you use Lloyds TSB's Internet banking service? [ ] Yes [ ] No *(Pay and Exit)*

If Yes: **5a. How frequently do you use the Internet banking service for accessing your company account?**

[ ] Daily
[ ] 2-3 times per week
[ ] Weekly
[ ] 2-3 times per month
[ ] Monthly
[ ] Less frequently
[ ] Other

(specify).....

5b. What type of transactions do you use the Internet banking service for...? *(Read options, tick all that apply)*

[ ] Balances
[ ] Statements
[ ] Transfers between your accounts
[ ] Bill payments
[ ] Standing orders
[ ] Direct debits

[ ] Other

(specify).....

5c. How long have you been using the Internet Banking service?

.....

6. Do you own a mobile phone? [ ] Yes [ ] No

7. Do you use Lloyds TSB's Text Alert service? [ ] Yes [ ] No

(Text Alerts are SMS messages sent to your mobile phone)

8. Do you use Lloyds TSB's Internet Banking service on your mobile phone/PDA?

[ ] Yes [ ] No

Comments: e.g. why not?……

…..……..……

9. Do you use any other Internet Banking services on your mobile phone/PDA?

[ ] Yes [ ] No

Comments……

……

Experiment Tasks and Login Details

Participants' tasks were all bill payments, with details changed in each task sheet to help engage them with the different methods. Each task sheet was associated with different dummy persona details, and an appropriate Bankcard for the dummy company was issued with the Card Reader version.

Task Sheet A:

Business banking login details:

User ID:	659321921
Password:	monkeys
Memorable Info:	glasgow121

As a company employee, you have been asked to use Lloyds TSB Business Banking to complete the following tasks.

- Make a payment of £**342.32** to **Zam Zam International**
- Make a payment of £**199.99** to **Oliver Bonas**
- Make a payment of £**113.23** to **Intek Communications**.

Task Sheet B:

Business banking login details:

User ID:	871564219
Password:	giraffe
Memorable Info:	rush1192

As a company employee, you have been asked to use Lloyds TSB Business Banking to complete the following tasks.

- Make a payment of £**185.50** to **Jones Stationary**
- Make a payment of £**209.99** to **Loomes Publishing**
- Make a payment of £**305.22** to **Star Utilities**

Task Sheet C:

Business banking login details:

User ID:	446900323
Password:	embolden
Memorable Info:	crisps884

As a company employee, you have been asked to use Lloyds TSB Business Banking to complete the following tasks.

- Make a payment of £254.40 to **United Utilities plc**
- Make a payment of £189.99 to **Future Mobiles**
- Make a payment of £140.00 to **Budget Retail**

Task Sheet D:

Business banking login details:

User ID:	557948132
Password:	finesse
Memorable Info:	blazer544

As a company employee, you have been asked to use Lloyds TSB Business Banking to complete the following tasks.

- Make a payment of £402.02 to **Williams Van Hire**
- Make a payment of £154.40 to **Gourmet Catering**
- Make a payment of £307.56 to **AGM Consultancy**

Exit questionnaire

"Thinking about those four methods of confirming Internet banking transactions…"

1. What did you like most about the different methods?

.....

.....

.....

2. What did you dislike about the methods?

.....

.....

.....

3. Do you have any suggestions for improvements?

.....

.....

.....

4. Please rate each method you have used along the ruler between Poor and Excellent using these four magnets. Explain your ratings:

	Overall Rating (30cm scale)
Password	
OTP device	
Mobile phone - Text message	
APACS device	

Why?

.....

.....

.....

(Remove magnets and use laminated SECURITY 'tag' for the following rating)

5. Please rate the four methods anywhere from Poor to Excellent in terms of *SECURITY*

	Security Rating (30cm scale)
Password	
OTP device	
Mobile phone – Text message	
APACS device	

Why?.....

.....

.....

6. Do you think that inputting a password is secure enough to confirm a transaction in Lloyds TSB's Internet banking service?

[] Yes [] No [] Don't Know

Comment?.....

.....

7. Should the Bank make it mandatory to use one of these devices instead of a password for confirming transactions?

[] Yes [] No [] Don't Know

Comment?.....

.....

8. If the Bank decides to rollout this approach, should they offer customers a choice of device?

[] Yes [] No [] Don't Know

Comment?.....

.....

9. If the Bank gave you a choice of device, would you be inclined to use Internet Banking…(read options)

[] More frequently [] Less frequently [] About the same [] Don't Know

Comment?.....

.....

10a. What are the advantages of using these devices?

Comment?.....

.....

.....

10b. What are the disadvantages of using these devices?

Comment?.....

.....

.....

11. Have you used any of these devices as a security method before?

a.) Card Reader: [] Yes [] No [] Unsure

What

for/comments?.....

.....

.....

b.) Token: [] Yes [] No [] Unsure

What

for/comments?.....

.....

.....

c.) Text Message: [] Yes [] No [] Unsure

What

for/comments?.....

.....

.....

12. Do you have any other comments about what you've experienced today?

.....

.....