IMPACT OF DISTRACTIONS ON MOBILE DEVICE USABILITY & ADOPTION

CONTEXTUAL USABILITY:

THE IMPACT OF DISTRACTIONS AND EXPECTATIONS ON PERFORMANCE, SATISFACTION, AND ADOPTION OF MOBILE DEVICES FOR WIRELESS DATA SERVICES

By

CONSTANTINOS COURSARIS, B.ENG, M.B.A.

A Thesis

Submitted to the School of Graduate Studies

In Partial Fulfillment of the Requirements

for the Degree

Doctor of Philosophy

McMaster University

Doctor of Philosophy (2006) McMaster University

(Business Administration) Hamilton, Ontario

TITLE: Contextual Usability: The Impact of Distractions and Expectations on Performance, Satisfaction, and Adoption of Mobile Devices for Wireless Data Services

AUTHOR: Constantinos K. Coursaris, B.Eng. (Carleton)

M.B.A. (McMaster)

SUPERVISOR: Dr. K. Hassanein & Dr. M. Head

NUMBER OF PAGES: xi, 197

ABSTRACT

Mobile devices are becoming increasingly popular, having already reached over 1.5 billion mobile subscribers. Such mobile devices propose increasing value to consumers found in "anytime/anywhere" connectivity, communication, and data services. Although progress has been made in terms of technological innovations, many mobile applications remain difficult to use, lack flexibility and robustness. Some key usability challenges facing m-Business (mobile business) applications include limited screen size and quality, limited input methods and navigation difficulties. Additionally, the mobile user has to share his or her attention between the application and the surrounding environment (e.g. visual and/or auditory stimuli). Furthermore, the user's state and personal characteristics (e.g. age, motion) may be key factors in their ability to use a mobile device. This context (i.e. technology, task, environment, and user characteristics) of use may have a significant impact on the usability of such devices.

This dissertation aims to support the claim that context impacts the usability of mobile devices, and in doing so it attempts to answer the following research questions: 1) What is the impact of distractions on the usability of mobile devices for wireless data services? 2) Does Expectancy-Disconfirmation Theory (EDT) help to explain the user's evaluative process of usability with a mobile device for wireless data services? 3) What is the impact of usability on consumers' behavioural intention towards using mobile devices for wireless data services? 4) Which factors become relevant when studying usability within a context of use?

Based on existing research, a framework is proposed and a research model is developed. An empirical study is performed to validate the model and answer the research questions outlined. To simulate a real-world setting where the context of using a mobile device is constantly changing, a laboratory experiment (2x2 factorial design) was performed involving 93 subjects. Distractions were simulated in this study in the form of either user motion or environmental noise (i.e. background auditory and visual stimuli). A structural equation modelling analysis confirmed the impacts of distractions on perceived usability (i.e. efficiency and effectiveness) of, and in turn the users' behavioural intention to use, a mobile device for wireless data services. The applicability of the Expectancy-Disconfirmation Theory (EDT) in explaining a mobile user's evaluative process of usability was also explored and was found moderately significant in explaining the impact of user expectations of perceived performance.

This dissertation contributes to theory and enhances our understanding of usable mobile devices for wireless data services. Implications for practice with respect to both interface design for mobile devices and marketing of mobile services are also presented.

ACKNOWLEDGEMENTS

This dissertation would not have been possible without the support of many people that deserve, at the very least, to be acknowledged. While in some cases it may not be feasible to list all of their names, I would like to at least recognize the group or organization they belong to.

To start, I would like to express my sincere gratitude to my two co-supervisors, Dr. Khaled Hassanein and Dr. Milena Head. With so much to be grateful for, their relentless attention to detail, passion for IS research, and on-going availability stand out in my mind. These were the critical factors in the highly informative guidance they have shown me and a successful collaboration we have shared over the last four years.

In addition to my supervisors, I was fortunate to have two exceptional dissertation committee members, Dr. Yufei Yuan and Dr. Nick Bontis. I am thankful to both of them because their support and guidance has been ongoing throughout my graduate studies.

Often the unsung heroes in a great story, the support staff at the DeGroote School of Business have been there for me every step along the way, starting with my MBA and continuing through to the end of my doctoral studies. Carolyn Colwell, Sandra Stephens, and Maddy Burley tried their hardest so that I would have an enjoyable and fruitful academic journey at the DeGroote School of Business.

It goes without saying that the knowledge and expertise acquired during my doctoral studies was consequent of exceptional instruction by the DeGroote faculty from all Departments. Furthermore, I would like to extend my sincere gratitude to the remaining members of our Management Science / Systems Department, and in particular Dr. Brian Detlor, who has been particularly supportive throughout my doctoral studies.

The last four years at McMaster University would not have been as memorable had it not been for the friendships made. Great researchers, but more importantly, great friends offered feedback and support all along the way. Dr. Gokul Bhandari, Ofir Turel, and Micheal Bliemel, it has been a pleasure working with you. I would also like to thank my closest friends not attending McMaster, who have always been there for me.

This dissertation and my doctoral studies would not have been feasible without the support of three external organizations: first, the Natural Sciences and Engineering Research Council of Canada (NSERC), who provided financial support for conducting this experiment; second, the Government of Ontario for the Ontario Graduate Scholarship I received two years in a row; and third, Research In Motion Inc. for their contribution of the BlackBerry 7250 mobile device that was used in this dissertation.

This dissertation was my springboard into my current position at Michigan State University. I thank the Faculty at the Department of Telecommunication, Information Studies, and Media, and the Administration of the College of Communication Arts and Sciences, for sharing my enthusiasm in Human Computer Interaction research, for inviting me to join them, and for accommodating my needs in this project.

Of course, I would not have achieved any academic or professional success had it not been for my loving parents, their support, and encouragement over the years. Who and what I am and what I will be is because of you – thank you for everything!

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

Figures

Tables

CHAPTER 1

RESEARCH MOTIVATION AND OBJECTIVES

1.1. Research Motivation and Importance of Topic

Mobile devices are becoming increasingly popular, having already reached over 1.5 billion mobile subscribers (OneUpWeb 2005). As consumers' technology fears and adoption costs are reduced, mobile devices are approaching "mainstream" status around the developed world. Such mobile devices propose increasing value to consumers found in "anytime / anywhere" connectivity, communication, and data services. Although progress has been made in terms of technological innovations, many mobile applications remain difficult to use, lack flexibility and robustness.

Key usability challenges facing m-Business (mobile business) applications include a technology system's interface attributes such as limited screen size and quality, limited input methods, and navigation difficulties. Additionally, the mobile user has to share his or her attention between the task (application) and the surrounding environment. Furthermore, the user's personal characteristics (e.g. age, culture) may be key factors in their ability to use a mobile device. This context of use (i.e. environment, user, task, and technology) may also have a significant impact on the usability of such devices. Consequently it is warranted to complement traditional usability studies that focus primarily on the interface with "contextual usability" research.

In the past, usability studies have typically examined technology attributes (e.g. website interface) in isolation. While that approach allows for a potentially more rigorous methodology by controlling for variables not of direct interest, conclusions reached may be of moderate significance, as any interaction effects present in the real world are not accounted for in the study. Research in the Information Systems (IS) field has been questioned / criticized in the past for focusing excessively on rigor and not paying enough attention to real world application and value (Lindroth et al. 2001).

The above considerations (i.e. growing demand for mobile devices, non-contextual view of past usability studies, need for more practical and closer to real-world research in IS) suggest directions for valuable future research in IS. The growing demand for mobile devices (CWTA 2006) is a result of both reduced technology adoption costs (e.g. free or low purchase price of mobile devices, affordable service plans) and increased mobility of individuals due to several factors (e.g. mobility of occupation) (Palen et al. 2001). These reasons combined generate increasing value for consumers found in anytime / anywhere connectivity and communication (Coursaris and Hassanein 2002). At the other end of the market continuum, as a result of increased competition network carriers have seen their Average Revenue Per User (ARPU) decrease in the last few years; one example that is representative of carriers around the world can be found here in Canada where the ARPU dropped by 30 percent during the period 1995-2001 (CWTA 2006). To combat this trend, carriers are promoting wireless data services that will help boost the ARPU and will contribute to alleviating some of the burden of deploying new infrastructure and the cost of third generation (3G) license fees. One such wireless data service is found in text messaging. According to the Canadian Wireless Telecommunications Association (CWTA), text messaging has reached a demand level of more than 2.7 million text messages per day in November 2004, a growth rate greater than 100% since the previous year and one greater than 400% over a three year period. The growth of text messaging (and other more recently introduced wireless data services) over this three year period resulted in an estimated growth of six percent for the ARPU. This demand suggests that text messaging is a widely used application for either of the following two purposes (Coursaris and Hassanein 2002):

- Consumers use it for communication with others
- Business users exploit it for either:
  - professional communication among employees, or
  - commercial use with their clients (text messaging offers an additional marketing channel)

Text messaging is just one application made possible by mobile devices. Organizers, e-mail, and the wireless Web are additional capabilities available on several mobile devices that could stimulate revenue growth for wireless network carriers. Therefore, investigating the usability of wireless data services made available on mobile devices is of significance to many parties. In particular, a study that can facilitate an enhanced understanding of the usability of text messaging, organizers such as the calendar and address book applications, and wireless Web browsing can translate into improved communication among consumers, businesses, and between the two. Such an understanding could also result in additional revenue for wireless service providers, network carriers, and indirectly for other m-Business market players (e.g. software developers), as it allows them to offer products and/or services that meet consumer expectations.

Extending from the earlier argument made regarding the non-contextual view of past usability studies, two recommendations can be made. First, there is tremendous opportunity to concurrently investigate multiple factors that may impact usability, while paying sufficient attention to methodological rigor. Secondly, past studies measured users' perceptions of technology usability mainly by means of a post-experiment questionnaire. This approach does not consider the evaluative process that a user employs in arriving at the conclusion of whether a device is usable or not. This phenomenon is consequent of usability studies being rooted in the 1970's when they were being performed almost exclusively from a technology-centred perspective. Hence, an opportunity arises for groundbreaking research in the ever-growing body of literature for usability as it relates to m-Business.

Considering the above discussion and addressing the need for laboratory research in IS to reasonably approximate the real-world setting (Galliers and Land 1987), the next section presents the objectives of this research that explores the antecedents and consequents of contextual usability for mobile devices.

1.2. Research Objectives

Usability has been a focus of discussion in both academic and trade press publications (Venkatesh et al. 2004). Although a plethora of usability issues are worthy of attention, we will focus on two dimensions that have been largely neglected — the context of use (Chan et al. 2002) and the user's affective process of evaluating usability (Dillon 2003). The research is therefore guided by the following questions:

1. What is the impact of distractions on the usability of mobile devices for wireless data services?
2. Does Expectancy-Disconfirmation Theory (EDT) help to explain a user's evaluative process of usability with a mobile device for wireless data services?
3. What is the impact of usability on consumers' behavioural intention towards using mobile devices for wireless data services?
4. What factors become relevant when studying usability within a context of use (i.e. situational specific attributes)?

This dissertation begins with a literature review of relevant topics to this research. A theoretical framework of contextual usability is then proposed along with the studied research model. The dissertation continues with a discussion on the research methodology. This study's data results will then be presented and analyzed. A discussion revolving around the above research questions will follow, along with a summary of research contributions to both theory and application, closing with the limitations of this study and suggestions for future research.

CHAPTER 2
MOBILE USABILITY & TECHNOLOGY ADOPTION: A LITERATURE REVIEW

With the objective of providing the background to this research, this chapter begins by introducing m-Business, where the discussion leads to the identification of usability as a key consumer concern in adopting mobile devices. Following a review of usability models and studies, a definition and measurement approach for usability is provided and the focus of this study is highlighted. The need to consider the context of use in usability studies is discussed next, supported by an overview of representative past studies that followed along this contextual path. This chapter concludes with an overview of representative past studies and a discussion on technology adoption theory and its relevance to m-Business.

2.1 Mobile Business

While there have been advancements in various fields over the last decade, progress made in information technology (IT) has arguably received the most attention (Lyytinen & Yoo 2002b). The dawn of electronic business (e-Business), highlighted by the launch of the World Wide Web in 1994 (W3C 2003), has led to the development of novel technological solutions that attempt to empower consumers and professionals alike. Enhanced computing devices, powerful middleware and applications, and sophisticated networks are being utilized in optimizing communication and transactional capabilities.

The vast majority of e-Business applications developed so far have concentrated on wired infrastructure, stationary devices and non-mobile users (Varshney & Vetter 2002a). The next evolutionary stage of e-Business is found in mobile business (m-Business) (Coursaris and Hassanein 2002).

Mobile business (m-Business) can be defined as electronic transactions enabled at least in part by mobile technology that may target businesses and consumers alike (Coursaris and Hassanein 2002). There are several mobile technologies that support m-Business. These are typically grouped as devices and networks (White 2005). Mobile devices range from small radio frequency identification (RFID) and global positioning system (GPS) chips to barcode scanners and wirelessly-enabled handheld personal computers. Mobile networks range from Bluetooth and RFID readers to mobile telecommunications networks and GPS. These mobile technologies are being used by organizations to help address their needs while offering opportunities for flexibility and customization. Wireless and mobile networks have experienced significant growth over the last few years, bringing the user base to an estimated one billion (Thai 2003). This user base is made up predominantly of cellular phone subscribers, but also of consumers using pagers, Personal Digital Assistants (PDAs), and laptops/notebooks. Growth in this arena is expected to continue and it is expected that over 40% of all business-to-consumer (B2C) e-Business will be initiated from the next generation of mobile devices, called smart phones. Smart phones are supported by Wireless Application Protocol (WAP), the most popular wireless communications standard currently (Varshney & Vettter 2002a). Despite this prediction, there are many who feel that mobile and wireless networks are still "a cure looking for a disease" (McGinity 1999). While there is demand by consumers for more mobile services, the mobile industry is still looking for its "killer application" (or "killer app") (Lehrer 2004), which some feel is the wireless Web. The question remains however, if customers want or even need the wireless Web and if it will ever replace the wired Web. Furthermore, a number of consumer concerns have been identified with m-Business, including (Coursaris and Hassanein 2002; Coursaris et al. 2003):

- *Usability:* Consumers are concerned with their ability to utilize the functions of mobile technology, e.g. mobile devices; Poor usability is a barrier to mobile technology acceptance
    - (Wong and Hiew 2005; Hassan and Li 2005; Turel and Yuan 2004; Gebauer and Shaw 2004; Varshney 2002; Chan et al. 2002; Kim et al. 2002; Buchanan et al. 2001; Lu and Yeung 1998; Nielsen 1993; Shakel 1991),

- *Security:* Consumers are concerned with the safety of their private electronic data (exchanged or stored)
    - (Misra and Wickamasinghe 2004; Walter 2004; Karnouskos 2004; Page et al. 2004; Camponovo and Cerutti 2004; Bussard et al. 2004; Ameri 2004; Gururajan 2002; Bask 2001; Dayal 1999; Hoffman 1999; Ovans 1999; Kalakota and Whinston 1996; Pesonen 1999),

- *Privacy:* Users are concerned that the mobile infrastructure may place their personal information at risk in terms of when and to what extent this information about them is used and/or communicated to others
    - (Karnouskos 2004; Harris 2001; Paul 2001; Branscum 2000; Green 1998; Agranoff 1993),

- *Trust:* Consumers are concerned with relying on other actors within the mobile information society
    - (Bussard et al. 2004; Karnouskos 2004; Shankar et al. 2002; McKnight et al. 2002; Belanger et al. 2002; Papadopoulou et al. 2001; Lee and Turban 2001; Light 2001; Jones et al. 2000),

- *Content:* Consumers are concerned with the lack of availability of mobile information and/or mobile services of interest
    - (Pousttchi et al 2004; Karnouskos, 2004; Yang 2002; Sohn 2000; Kaynama 2000; Jarvenpaa and Todd 1997; Rice 1997),

- *Cost:* Consumers are concerned with the total financial investment required to connect and/or access desired/needed mobile content
  - (Zmijewska 2004; Jain and Kannan 2002; Gallaugher et al. 2001; Krishnan et al. 1999; Bakos and Brynjolfsson 1999; McGinity 1999; Noll 1997; Noll 1997; Stidham 1992; Westland 1992; Mendelson and Whang 1990; Dewan and Mendelson 1990; Mendelson 1985),

- *Network dependability:* Consumers are concerned with the reliability, availability, and survivability of wireless networks
  - (Maitland et al. 2005; Islam and Fayad 2003; Malloy et al. 2002; Lyytinen & Yoo 2002a).

In spite of the above concerns, m-Business is already being used in diverse areas, ranging from entertainment to education, and from research to procurement (Malladi & Agrawal 2002). Wireless and mobile networks are increasingly becoming the preferred solutions, due to their flexibility and scalability (Varshney & Vetter 2000a). In the end, a number of factors (such as the issues described above) will come into play in determining the rate of m-Business adoption, and these factors will be addressed by those active participants in the m-Business market. As mentioned earlier, usability is one such factor and it is the focal point of this research.

2.2 Usability

2.2.1 History of Usability

Usability studies have their roots as early as the 1970's with the work of "software psychology". "Software psychology dealt with the utility of a behavioural approach to understanding software design, programming, and the use of interactive systems, and to motivate and guide system developers to consider the characteristics of human beings" (Carroll 1997). Evolving into the analysis of user requirements and the conception of Graphical User Interfaces (GUI), "usability engineering" had arrived (Nielsen 1993; Hermann & Heidmann 2002). A later stage that would form a subset of usability engineering concentrated on interfaces and came to be known as "information visualization" (Hornbaek 2003). The most recent exploration in the field of usability was coined "new usability" (Thomas & Macredie 2002) or "ubiquitous usability" (Hassanein and Head 2003). "Ubiquitous usability" is concerned with the context in which new products and services are being used. Varied situational contexts will result in emerging usability factors, making traditional approaches to usability evaluation inappropriate. This evolution of HCI is summarized in Table 2.1. The corresponding research has produced extensive resources in the form of usability guidelines and standards for various domains. Still, research in the two areas of context of use and a user's evaluative process of usability has been limited (Venkatesh et al. 2004). The importance of these two areas emerges from their importance in yielding a reasonable analysis during a usability study (Thimbleby et al. 2001; Maguire 2001). For example, a usability study would be of limited value if it were not to account for the following factors (Hassanein and Head 2003):

- User (e.g. prior relevant/computing experience, age, education, culture, state of motion)
- Environment (e.g. lighting, noise – music, speech, white noise)
- Task (e.g. complexity, interactivity)
- Technology (e.g. network dependability, interface design – input/output modes, size, weight, actual device vs. emulator)

The takeaway is that usability experiments need to consider the product and the user and be designed subject to those considerations in addition to the environmental and task-related constraints. This view is shared by several researchers whose work will be described in more detail in Section 3.2 (i.e. Usability and Context of Use).

TABLE 2.1: EVOLUTION OF HUMAN-COMPUTER INTERACTION (HCI)

Domain	Year	Distinctive Characteristics	Sample References
Software engineering	1970's	Goal: To establish the utility of a behavioural approach to understanding software design, programming, and the use of interactive systems, and to motivate and guide system developers to consider the characteristics of human beings. Two methodological axioms: 1) Assume the validity of the "waterfall" model (i.e. top-down decomposition and discretely sequenced stages with well-specified hand-offs) 2) Assume two central roles for psychology within this context: a) to produce a general description of human beings interacting with systems and software, a description which could be synthesized as a guideline for developers, and b) to verify directly the usability of systems and software as they were developed	Carroll 1997; Schneiderman 1980; Royce 1970
Usability Engineering	1980's	Analysis of user requirements and the conception of Graphical User Interfaces (GUI) Three key notions: 1) Propose iterative development is managed according to explicit and measurable objectives (i.e. "usability specifications") 2) Broaden the empirical scope of design through various techniques for user participation (e.g. *participatory*, *contextual*, *ethnographically informed*) 3) Cost effectiveness of prototyping tools (e.g. *by demonstration*), usability evaluation (*GOMS*, *impact analysis)*, user testing (e.g. *inspections*, *checklist*, *script-oriented*), analytical techniques (e.g. *GOMS* extended, *claims analysis*, *cognitive walkthrough*)	Hermann & Heidmann 2002; Carroll 1997; Kieras 1988; John 1990; Carroll & Rosson 1991; Polson, et al. 1992; Nielsen & Mack 1994; Good et al. 1986; Bentley et al. 1992
Information Visualization	1990's	Subset of usability engineering concentrated on interfaces	Hornback 2003
Ubiquitous / New usability	2000's	Concerned with the context in which new products and services are being used. Varied contexts will result in emerging usability factors, making traditional approaches to usability evaluation inappropriate.	Schmidt et al. 1999; Thomas & Macredie 2002; Tarasewich 2003; Hassanein & Head 2003; Yuan & Zheng 2005; Sarker &Wells 2003

2.2.2 Mobile Usability Studies Reviewed

The four variables discussed earlier comprising *context of use* (i.e. user, task, environment, and technology) will be used to classify any previous research that relates to the usability assessment of mobile applications and/or mobile devices. The benefit of using these variables for the literature review is found in both the structure it provides for the discussion to follow, as well as to help highlight any areas that are lacking investigation. The direction for this dissertation will thus emerge from this discussion. Table 2.2 classifies some selected studies reviewed here according to their author(s) and the four context-related variables (i.e. user, environment, task, and technology), which assists readers in scanning previous research according to any particular factor impacting usability. The following taxonomy is defined for the purpose of reviewing these studies:

- User:
    - Prior Computing/Relevant Experience: Novice (inexperienced) vs. Expert (experienced)
    - Motion: Static (immobile or stationary) vs. Mobile (in motion or moving)
- Environment: (either interruptions or distractions)
    - Experiment type: Lab (controlled) vs. Field (field study)
    - Stimuli (either interruptions or distractions – assumed none unless specified in study)
        - Auditory: Quiet (absence of noise/music) vs. Noisy (presence of noise/music)
        - Visual: Absence vs. Presence of visual cues (e.g. movement by other people)
        - Collocation: Alone vs. Public (being in near proximity of others)

- Task:
    - Type: Open (user defines outcome) vs. Closed (pre-defined outcome or goal)
    - Task descriptions (open/unstructured)
- Technology:
    - Device type: Emulator vs. Mobile device (e.g. mobile phone, PDA)
    - Interface - Input mode: Keyboard vs. Mouse vs. Speech vs. Character recognition (Jot and Graffiti)

TABLE 2.2: PREVIOUS EMPIRICAL USABILITY STUDIES RELATED TO M-BUSINESS

SAMPLE STUDIES	USER (subjects)	ENV'T	TASK	TECHNOLOGY
Nicholson et al. (2005)	Novice & Expert (36), Static & Mobile	Lab & Field	Closed, Text messaging	PDA Compaq iPAQ Nokia 3310 and 5510
Kaikkonen et al. (2005)	Novice (40) Static & Mobile	Lab & Field	Closed, 10 tasks, SMS & MobileWire	Nokia Series 40 mobile phone
Unknown (2005)	Novice, Experts (893)	Field	Closed, Examination	PDAs and HaPerT software
Juola & Voegele (2004)	Novice (48), Static	Lab	Closed, 2 sets of 3 tasks Talk or transfer data: picture, calendar entry, contact.	4 Bluetooth devices: Mobile phone, Headset, PC, Handheld
Kjeldskov & Stage (2004)	Expert ~70 Novice ~30 Static	Lab	Closed, 12 mobile tasks	PDA (not specified)
Chen & Vertegaal (2004)	Undefined (6), Static & Mobile	Lab, Interruption	Closed, Interruptions by calls, IM, email	Nokia mobile phone (undefined model, augmented)
Andon (2004)	Novice (9), Static & Mobile	Field	Open, typical work shift	Viewsonic Viewpad 1000 Windows 2000 Pro Health Sytems EmStat V7.9

Table 2.2 continued: Previous empirical usability studies related to m-Business

SAMPLE STUDIES	USER (subjects)	ENV'T	TASK	TECHNOLOGY
Kallinen (2004)	Novice, Casual, Experts (30), Static	Field, Noise, Co-location	Closed, Read a story on a PDA	Casio Cassiopeia Pocket PC Internet Explorer
Kaasinen (2003)	Novice, Casual, Experts (55), Static & Mobile	Lab & Field	Closed, 5 scenarios for using a GPS system	Benefon Esc! phone, WAP Pointers: Sonera, Bensa, Opas, Fakta, Garmin GPS 12 device, Magellan GPS module
Nagata & TNO (2003)	Expert (8), Static	Lab, Quiet & Interrupted	Closed, web browse & buy Interruptions by calls & IMs	Compaq iPAQ h3800 PocketPC vs. desktop
Fang et al. (2003)	Expert (24), Static	Lab	Closed, Web browsing	Compaq iPAQ 3760 PocketPC Pocket IE vs. SearchMobil
Rodden et al. (2003)	Expert ~70 Novice ~30 Static	Lab	Closed, 12 mobile tasks	PDA (not specified)
Karambelas et al. (2003)	Expert (8), Static	Lab	Closed, various tasks	PDA – overall
Hornbaek et al. (2003)	Novice (32), Static	Lab	Closed, 2 maps, locate objects	Display: computer screen
Wigdor & Balakrishnan (2003)	Novice (10), Static	Lab	Closed, 640 text phrases	Motorola i95cl mobile phone (augmented: Analog Devices ADXL202EB-232 accelerometer) Input: Multitap vs. TiltText
Palen & Salzman (2002)	Novice (19), Static & Mobile	Field	Open Phone calls, explore functionalities of device	Mobile phone

Table 2.2 continued: Previous empirical usability studies related to m-Business

SAMPLE STUDIES	USER (subjects)	ENV'T	TASK	TECHNOLOGY
Butts & Cockburn (2002)	3 Novice, 3 Casual 2 Experts (8), Static	Lab	Closed, 3 text input modes	PC Emulator
Waterson et al. (2002)	Novice (10), Static	Lab (5), Field (5)	Closed, find info on website; clickstream data vs. observation	PDA - Handspring Visor Edge w/ OmniSky wireless model
Buyukkokten et al. (2002)	Expert (N/A), Static	Lab	Closed, find information a page	Display: Emulator, various info presentation formats
Oquist & Goldstein (2002)	Novice (16), Static	Lab	Closed, subjects read a long & a short text	Display: PDA (Compaq iPAQ 3630 Pocket PC)
Watters et al. (2002)	Novice (84), Static	Lab	Closed, 2 table lookup tasks	Display: emulator
Sears & Arora (2002)	Novice (31), Static	Lab	Closed, 6 tasks, various	Input (Jot vs. Graffiti)
Bruijn et al. (2002)	Novice (30), Static	Lab	Closed, Web search	Nokia 6210 Rapid Serial Visual Presentation vs. WAP
Chittaro & Dal Cin (2002)	Novice (40) Static	Lab	Closed, Search & selection task	Nokia 7110 Navigation among cards vs. Single-choice lists
Kim et al. (2002)	Advanced (37) Static	Field	Open, mobile Internet tasks	Mobile Internet Phone
Suhm et al. (2001)	Novice (15) Static	Lab	Closed, 4 tests, speech recognition and error correction	Input: keyboard & mouse, speech, gesture, handwriting; Output: computer screen
Dyson and Haselgrove (2001)	Novice (36) Static	Lab	Closed, articles read, 6 types of questions	Display: computer screen, line length & reading speed

Table 2.2 continued: Previous empirical usability studies related to m-Business

SAMPLE STUDIES	USER (subjects)	ENV'T	TASK	TECHNOLOGY
MacKenzie et al. (2001)	Undefined (20) Static	Lab	Closed, 20 text phrases	Mandrake's GNU/Linux V7.2 Input: PC Concepts KB-5640 numeric keypad Output: 19" colour monitor LetterWise vs. Multitap
James & Reischel (2001)	Novice, Experts (20) Static	Lab	Closed, Text entry	Mobile phone Nokia 3210 Text entry: multi-tap vs. T9
Lindroth et al. (2000)	Range (12) Static	Lab, Field	Closed, 3 tasks, enter data	PDA (Palm V)
Pascoe et al. (2000)	Novice (1) Static	Field	Open, several tasks	3Com PalmPilot
Piolat et al. (1997)	Novice (54) Static	Lab	Closed, 2 tasks – reading	Display: emulator, page reading Scrolling vs. Page-by-page
Benbasat and Todd (1993)	Novice (48) Static	Lab	Closed, 1 task (sending a memo)	Display: emulator, & mouse: Direct manipulation vs. menus & Icons vs. text
Trumbly et al. (1993)	Novice (16), Expert (16) Static	Lab	Game simulation	Interface: dialogue style (menu/command), colour (yes/no), values (default/no), error messages (short/lengthy), help (automatic/upon request)

2.2.3 Findings from Mobile Usability Studies and Dissertation Research Direction

From the literature review on usability several observations can be made that can then serve as the basis for the research direction of this dissertation. First, most studies so far have involved subjects who were not in motion and almost all of them have been conducted in a controlled setting (i.e. laboratory), even though most of them investigated mobile applications and/or devices. Although naturalistic studies are very difficult to perform successfully, it is surprising to find that only a few scholars (Nicholson et al. 2005; Chen & Vertegaal 2004; Lindroth et al. 2001; Pascoe et al. 2000; Kim et al. 2002) have investigated a setting where the user is mobile. Hence, this research will attempt to involve both mobile (e.g. walking) and static (e.g. seated) participants.

Second, user types have been predominantly novices. This may have been an appropriate choice for the studies in Table 2.2. However, since this research focuses on contextual usability, a better choice would be to control for user characteristics by soliciting participants from the entire spectrum of relevant experience. By doing so any effects associated with the learnability of the chosen device will be minimized. This will also enhance the generalizability of the results especially when attempting to study a user's intention to adopt a mobile device.

Third, interface types have been predominantly emulators. This approach introduces uncertainty in the results, since the outcome is derived from inappropriate settings, such as form factor, processor power and overall system performance, as well as misleading user-technology interaction and consequently misleading user experiences. For these reasons, this research used an actual mobile device that is representative of the current devices available in the market, as well as one that offers a suitable mode of input (e.g. built-on QWERTY keyboard, stylus-based data entry) for the aforementioned objectives.

Finally, the task types have been typically closed, a setting that allows for controlled behaviour in the experiment and objective data. Closed tasks are also recommended for this research. Such tasks could include authoring text messages, entering contact information, and browsing the wireless Web.

Overall, while a significant volume of literature exists in the domain of mobile usability given its young age, only 41 percent of mobile usability papers are empirical in nature (Kjeldskov and Graham 2003). Furthermore, most of these empirical studies (71 percent) have been carried out in a laboratory typically not considering the expected contextual use of these mobile interfaces. Hence a significant opportunity exists for future research in the area of contextual (mobile) usability.

2.3 Mobile Technology Adoption

Early adopters were dissatisfied with the purchase of mobile devices due to a poor usability experience. This poor usability was primarily due to confusing navigation and site structure designs, screen size limitations, and difficult to use input modes (Buchanan et al. 2001). The responsibilities of m-Business market players in satisfying consumers, including their usability expectations, were highlighted in the value network proposed by Coursaris and Hassanein (2002). This conceptual model identified the mobile consumer as the focal point for transactions between them and Content providers, Service providers, Network operators, Technology vendors, and Application developers. These market players are responsible from the beginning to the end of a consumer's typical buying cycle leading through to adoption of a product or service. Next, we cite a relevant theory that will help explain this expected adoption behaviour of m-consumers.

Rogers (1983) proposed that adoption of a new product by a consumer is consequent of a five-stage process; these stages are: awareness, interest, evaluation, trial, and adoption. Building on this, we sought to identify the catalysts for potential consumers as they move between each phase of this five-stage adoption process of wireless product(s) and/or service(s).

The process begins by information becoming available about a new product. This activity brings a consumer to the first phase of the process, namely awareness. Following third-party influence (business or personal) in the form of suggestions, the consumer becomes interested in learning more about the product. In doing so the consumer contrasts the newly found information against his/her own needs (or vice versa) and decides whether the product should be further evaluated. Assuming there is a qualified product, the consumer then weighs any perceived concerns (potentially against the level of needs being met or simply on an individual tolerance threshold, e.g. cost constraints), and decides whether to try out the new product. Once the product is tried, the experience is contrasted against any expectations formed a priori and a satisfactory assessment could result in adopting the product or service. Depending on the outcome of this experience, s/he decides to adopt the new wireless product(s) and/or service(s). Finally, insight or lessons learned from any of the stages in the adoption process may act as a stimulus for any of the earlier phases. For example, novel consumer needs may give rise to new solutions that would be pushed by industry at the information and/or suggestion phase. Figure 2.1 illustrates the catalytic adoption process described above. In this figure, dashed arrows connect the various catalysts representing the possibility that they can act as stimuli for any other phase in the consumer's buying behaviour for future products or services.

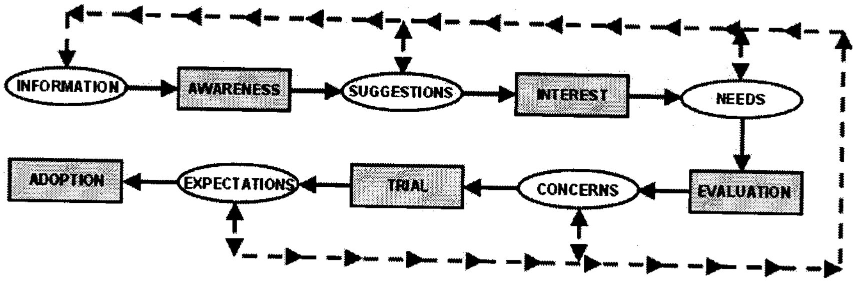


FIGURE 2.1: A CONSUMER'S CATALYTIC ADOPTION PROCESS OF A PRODUCT/SERVICE
(Extended from Rogers, 1983)

While including the catalysts between the five stages of the typical adoption process may seem trivial, it is important to do so as they in turn identify the party/parties that is/are most influential in each of the five stages. Starting with information and suggestions, they will typically be presented by any of the five market players described earlier in this section (other than the consumer). For example, if the product is text messaging, then any of the Network Carriers, Technology Vendors, and/or Service Providers could offer information and suggestions directly to the consumer. Then it would be up to the consumer to assess his/her personal need for text messaging ("needs"), contrast the concerns pertaining to this service ("concerns"), and decide whether or not to try it out. These steps are predominantly intrinsic and therefore the consumer has the most impact over them. If the experience lives up to the expectation that was formed from both external and internal stimuli ("expectations"), then the consumer would most likely adopt the product/service. Thus, it would appear that the five catalysts are in fact five variables that may impact directly or indirectly the adoption of m-Business. This claim needs to be validated empirically and the model itself can serve as the starting point of a more elaborate research plan.

Previous research has examined the impact of various factors on mobile technology adoption. A selection of representative studies in this area is provided in Table 2.3. Although it is not meant to be an exhaustive list, the cited studies are a fair representation of research in this area. It is interesting to note the variety of constructs and relationships considered in these mobile technology adoption studies. It is also noted that none of these studies have considered the impact of expectations on mobile technology adoption. As outlined in the foregoing analysis and in Figure 2.1, expectations appear to have an important role that needs to be examined. It is also interesting to observe that most of these studies have been carried out by means of questionnaires. A methodology that combines the measurement of both the perceived and objective scores for factors relevant to mobile technology adoption could enhance our understanding of this area.

TABLE 2.3: PREVIOUS RESEARCH STUDIES ON THE ADOPTION OF MOBILE TECHNOLOGY

SAMPLE STUDIES	METHOD (Sample size)	FOCUS	PATHS[1]
Mao et al. (2005)	Survey (273)	U.S. vs. Turkey: Mobile phone services	U.S. SAMPLE: P. Usefulness → BI P. Ease Of Use x BI Price x BI Accessibility x BI TURKISH SAMPLE: P. Usefulness → BI P. Ease Of Use → BI Price → BI Accessibility → BI
Pagani (2005)	Interview (56), Focus Groups (24), Phone Survey (1000)	U.S. vs. Italy: 3G Mobile Multimedia Services	P. Usefulness → BI P. Ease Of Use → BI Price → BI Speed → BI
Fang et al. (2003)	Survey (101)	Handheld devices for m-Commerce	P. Usefulness → BI P. Security → BI P. Playfulness → BI P. Ease Of Use → BI
Juola & Voegele (2004)	Survey, observation and device data recoding (48)	Various Bluetooth devices	Satisfaction → BI (statistics not reported)
Khalifa & Cheng (2002)	Survey (202)	Mobile Commerce	Subjective Norms → BI Attitude → BI P. Behavioural Control→BI Exposure→BI (indirectly)
Duda et al. (2000)	Survey, Observation, Interview (36)	WAP Services	Utility → A Usability → A System in- & output→A Feeling of control → A Speed → A

[1] Legend:
→ = supported; x = not supported
P. = Perceived; BI = Behavioual Intention to use; A = Acceptance

Building on the foregoing discussion and the body of literature identified in Table 2.3, this dissertation explores the role of usability in the relationship between user expectations and adoption of mobile technology while also studying the impact of distractions on mobile usability. The next chapter will review usability theory as it relates to the context of use, adoption, expectations, and cognition.

CHAPTER 3

RESEARCH MODEL

This chapter presents the theoretical basis for the research framework and model that will be described in later sections of this dissertation. The starting point is an examination of relevant usability theory from which the measurement dimensions of usability will be extracted, as well as the relationship of usability to technology adoption theory will be explored. Then we provide a theoretical background for context of use (or *context* in brief). Next we review relevant theory to the expectations construct, which was deemed to be relevant to usability in Section 2.3 (i.e. Mobile Technology Adoption). An understanding of usability, context, expectations, and technology acceptance will serve as the foundation for developing the research framework and model. Lastly, the discussion on cognition, attention, and performance will highlight the importance of studying distractions and their impact on usability. This research direction is in line with the opportunities for future studies that emerged following the literature review earlier in Chapter 2.

3.1 Usability and Technology Adoption

There have been several approaches to measuring usability put forth by scholars. One of these approaches was proposed by Nielsen (1993), where usability was measured as the learnability, efficiency, memorability, less errors, and satisfaction involved in a user's interaction with a technology. Rubins (1994) proposes similar usability dimensions, including learnability, effectiveness, usefulness, and attitude. A third measurement approach is put forth by the International Organization for Standardization (ISO). According to standard ISO 9241 (1998), usability is defined as, and measured in terms of, efficiency, effectiveness, and satisfaction (Bevan 2001). Lastly, Quesenbery (2003) defined usability in terms of five dimensions: efficiency, effectiveness, engagement (or satisfaction), error tolerance, and ease of learning.

For this study, usability will be measured following the ISO definition of usability:

- *Efficiency*: the level of resource consumed in performing tasks,
- *Effectiveness*: the ability of users to complete tasks using the technology, and the quality of output of those tasks,
- *Satisfaction*: users' subjective satisfaction to using the technology.

The ISO definition of usability was chosen for this study in part because it is the international standard (ISO-9241 1998) of measuring usability. The use of this standard allows for consistency with other studies in the measurement of efficiency, effectiveness, and satisfaction (Brereton 2004). Several mobile usability studies that have used these three constructs or a subset of them include the work of the following researchers: Unknown 2005; Kjeldskov and Stage 2004; Juola and Voegele 2004; Kallinen 2004; Wigdor and Balakrishnan 2003; Rodden et al. 2003; Nagata and TNO 2003; Butts and Cockburn 2002; Buyukkokten et al. 2002; Chittaro and Dal Cin 2002; Bruijn et al. 2002; Lindroth et al. 2000.

Further to this definition, Frokjaer et al. (2000) tested these three constructs for correlation in reference to usability. The results show that the three constructs should be considered discriminant, unless domain specific studies suggest otherwise, and that all three should be included in usability testing. Watters (et al. 2003) further argued that efficiency and effectiveness can be grouped under the concept of performance, which in turn impacts user satisfaction. Following suit to Frokjaer's suggestion, the impact of these factors on satisfaction will be tested in this study by focusing on mobile devices for wireless data services.

Since usability was argued to impact the growth of m-Business (i.e. poor usability hinders adoption), this research will test the impact of usability on the adoption of mobile devices.

The fit of the aforementioned definition of usability is considered next with respect to the Theory of Reasoned Action (TRA) (Fishbein and Ajzen 1975) and the Technology Acceptance Model (TAM) (Davis et al. 1989). According to TRA, a consumer's "behaviour is determined by his/her behavioural intention, and behavioural intention is determined by both the person's attitude and subjective norm concerning the behaviour in question", in this case being the use of mobile devices. Furthermore (in TRA), attitude is determined by the consumer's beliefs about consequences of performing the behaviour multiplied by the evaluation of those consequences. Usability is one such belief that may directly or indirectly impact a user's attitude towards using mobile devices (Chiu et al. 2004). Therefore it can be argued that usability impacts attitude, which in turn determines behavioural intention. Similar to TRA, TAM also argues that actual use of a computer system (e.g. mobile device) is impacted by the user's behavioural intention to use the technology. As subsequent studies have shown (Agarwal et al. 2000, Venkatesh 2000, Venkatesh 1999), there is a strong positive relationship between attitude towards use, behavioural intentions towards use, and actual use of a technology. Therefore, measuring only the consumer's behavioural intention towards using a mobile device for wireless data services may suffice in predicting actual usage of a technology (or adoption of mobile devices). It should be pointed out that although our definition of usability is inline with TAM regarding their respective explanatory power of mobile technology adoption, the two theories (usability and TAM) are distinct and no comparison between them can be made.

3.2 Usability and Context of use

The work of several scholars (Bevan and Macleod 1994; Thomas and Macredie 2002; Shami et al. 2005) attempting to identify additional variables that may impact usability and subsequently adoption, led to the conceptual emergence of *context of use* (herein referred to as context) as it relates to usability, also referred to as contextual usability. Schmidt et al. (1999) provide one of the earlier models for context, where contextual factors are hierarchically organized under two top level factors, namely human factors and physical environment. Each of the two categories is then further divided into three areas. First, human factors include the user (e.g. personal properties such as age and education), the social environment (e.g. co-location of others), and the task (e.g. spontaneous, planned, goals set). Second, the physical environment includes the location (e.g. absolute position, relative position), the infrastructure (e.g. resources for communication), and physical conditions (e.g. noise, light).

A similar taxonomy was used in the model proposed by Tarasewich (2003). The author groups all six sub-dimensions defined above by Schmidt et al. (1999) into four main areas, namely the environment, the participants, the activities, and the interactions occurring in this contextual setting. The difference between the two taxonomies is the grouping of all three physical environment sub-dimensions in one category (i.e. environment). Additional differences found in the Tarasewich model include the clear distinction given to interactions (e.g. group dynamics, time-of-day, participant-environment relationships), which also includes co-location of others (instead of the latter being a characteristic of both the user's social environment and the physical environment's location in the Schmidt et al. model). A similar taxonomy to the one proposed by Tarasewich was suggested by Hassanein and Head (2003). In their framework a similar set of four dimensions, namely user, task, interface, and environment, is supplemented with additional detail through a corresponding set of types and limitations for each dimension. More recently, Yuan and Zheng (2005) presented another taxonomy in a framework that proposes mobile workers, mobile tasks, mobile context, and mobile technology are fundamental aspects of mobile work. In their framework, context is described as the working place and work temporal structure in which mobile tasks are carried out by mobile workers. Another framework tailored to mobile interfaces was proposed by Lee and Benbasat (2003). Their work highlights three elements specific to the mobile setting, namely spatiality, temporality, and contextuality.

The latter refers to the circumstances surrounding the use of tasks on mobile devices, such as the degree of interactivity between the user and the interface.

The results of an exploratory project by Sarker and Wells (2003) gave rise to a more detailed breakdown of the type of factors that impact the usability and adoption of mobile devices, while context was a single construct in their overall model. Context referred to the economic and/or social factors, as well as to the critical mass of subscribers and the availability of services. The remaining variables in this model were the individual (e.g. age, culture), the technology (interface characteristics, network capabilities), the communication/task (e.g. number of interacting participants, goals), and the modality of mobility (e.g. extent and type, such as visiting, wandering). This categorization also presents slight variations to the previous models, two of which are: (i) technology is suggested to extend beyond the interface and include the network, (ii) modality of mobility is considered in isolation and not as part of the user's characteristics.

While there are other usability and/or adoption models that attempt to capture the essence of context, the three models described above provide sufficient theoretical foundation from which to build the proposed research framework described in later sections of this dissertation.

3.3 Expectations

As was discussed earlier, expectations may impact usability and adoption. In support of this argument a prominent theory is referenced from the field of consumer behaviour known as the Expectancy – Disconfirmation Theory (EDT) (for a review see Oliver 1997). Shown in Figure 3.1, EDT explains the process by which a consumer evaluates a product or service (note: "product" will be used in this discussion, as this study focuses on mobile devices for wireless data services). To start, a consumer forms expectations regarding a product as a result of any prior experience with the same or similar products, as well as the messages received through commercial marketing (e.g. advertisements) and/or opinions expressed by other consumers. Next, use of the product will facilitate the consumer's perception of the product's performance. Initial expectations and perceived performance combined result in a qualitative assessment, whereby unmet expectations would yield negative disconfirmation (or emotional dissatisfaction). In the event that expectations are met, confirmation occurs, and finally if expectations are surpassed, a consumer experiences positive disconfirmation and consequently satisfaction. Of interest is the case of expectancy confirmation, where even though it is a positive scenario, consumers do not typically experience strong feelings of satisfaction. The latter occurs only when performance expectations are surpassed by the actual experience with the product. Building on the basic EDT model, additional linkages have been shown in past studies. First, testing the concept that a consumer may experience "self-fulfilling prophecy", the direct effect of expectations on perceived performance has been supported (Olchavsky and Miller 1982; Spreng et al. 1996; reviewed by Yi 1990). Also, performance has been shown to impact satisfaction directly (Churchill and Surprenant 1982; Tse and Wilton 1988), expectations may also have a direct effect on satisfaction (Oliver 1997; Van Ryzin 2004), while there is a strong effect between the latter and a consumer's behavioural intention to use a product (Adnerson and Sullivan 1993, Cronin and Taylor 1992, Gotlieb et al. 1994; Mittal et al. 1999; Taylor and Baker 1994).

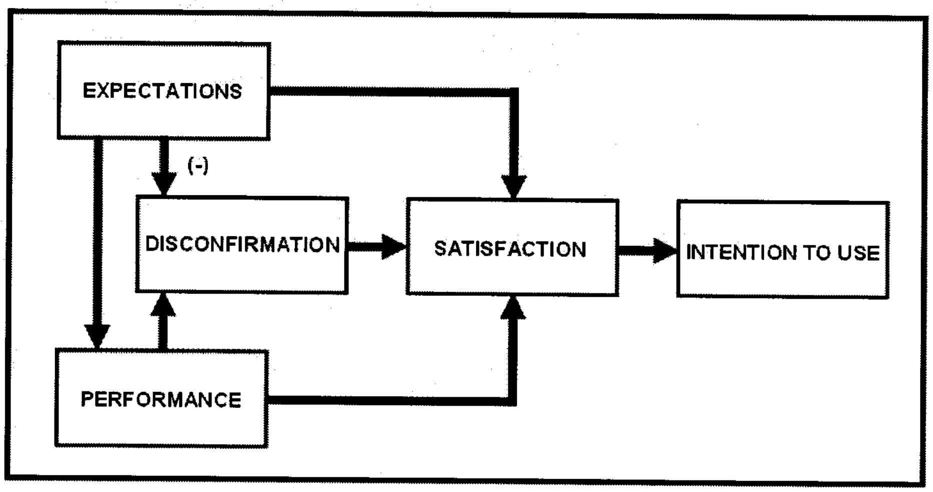


FIGURE 3.1: EXPECTANCY – DISCONFIRMATION THEORY (EDT)

EDT has been previously used in various disciplines, including Information Systems (IS) (Chiu et al. 2004; Bhattacherjee and Premkumar 2004; McKinney et al. 2002; Spreng and Chiu 2002). The evaluative process of satisfaction outlined by EDT provides a model for understanding and measuring end-user satisfaction with IS performance. In the domain of usability, EDT can be used to measure a consumer's feelings about device/system attributes relative to each of the two performance sub-dimensions, namely efficiency and effectiveness (Watters et al. 2003). This decomposition of the performance construct in EDT would essentially create a dyadic satisfaction formation process for the consumer, one for efficiency and one for effectiveness.

3.4 Cognition, Distraction, and Performance

In the context of usability performance is defined as the efficiency and effectiveness associated with performing tasks on a particular device. It then follows that any factors which may impact performance are relevant and should be accounted for in usability studies. Typically these studies have been performed in controlled laboratory settings where external variables (e.g. background noise), commonly referred to as "environmental noise", were absent (Kallinen 2004) in an attempt to uphold a rigorous methodology. "Environmental noise" refers to sounds and visual distractions that are present in the environment where the user is engaged in a particular activity (Wu and Newell 2003). By omitting environmental noise, however, such studies omitted factors that would typically be present in a real-world setting and are therefore limited in the external validity of their findings.

This limitation arises mainly from the observation that environmental noise negatively affects information processing and performance (Baker and Holding 1993; Rader 1993). Both short-term memory (also known as *working memory*) and attention span are subject to cognitive constraints (Baddeley 1986). Nicholson et al. (2005) describe cognitive load as "the total amount of mental activity imposed on the working memory at an instance in time." Any single distraction adds to the total cognitive stimuli (i.e. load) thereby reducing information processing efficiency and, by extension, performance (Nicholson et al. 2005; Eysenck 1984; Kahneman 1973).

Researchers need to pay close attention at the dyadic inverse relationship that exists between methodological rigour and relevance of findings (Lindroth et al. 2001).

Mason (1988) argues that both concepts, which he referred to as "tightness of control" and "richness of reality", are important dimensions in research. The selected trade-off point between the two dimensions should be determined by the researcher according to the study's objectives. Overall it can be argued that the more natural the experimental setting is the more relevant and applicable the results will be (Jarvinen 1999; Lindroth et al. 2001). Hence, there is a need to include "environmental noise" in usability studies when focusing on mobile devices, thereby increasing the study's real-world relevance.

Extensive literature focuses on sound and its impact on performance. A quiet environment has been shown to result in higher efficiency, while the presence of irrelevant sound lowers mental efficiency and performance due to the obligatory cognitive process of organizing unattended information (Banbury et al. 2001; Hughes and Jones 2003). Furthermore, there are two sound dimensions that have been explored with respect to the impact they have on performance. First, there is a distinction between noise and music. Both hinder performance (Stansfeld et al. 2000; Persson Waye et al. 2001), but music has been shown to have a negative impact on performance more than noise (Umemura 1992). A second dimension of sound is variability. Increased variability of background noise results in lower performance (Jones et al. 2000; Hughes and Jones 2001). Supporting the earlier point of each single distraction adding to the total cognitive stimuli, the combined exposure to noise and body vibration has been shown to increase the difficulty associated with performing a task more than that of each individual effect (Ljungberg et al. 2004).

Under the contextual usability model defined earlier, user characteristics have been said to be of relevance. Psychology studies on personality have found that there is interaction between a user's introversion/extraversion and the impact of background sound on any task (Furnham and Strbac 2001), as well as with the subject's task proficiency (Loeb 1981). The same study showed that task difficulty is important in considering the interaction effect of noise.

In summary, it is important to consider complex interactions including the effects of noise (auditory stimuli including music and meaningful speech) with task difficulty and task proficiency of the subject, personality of the subject, and the meaning and variability of noise. In addition, Ljungberg's et al. (2004) study supports the argument that the combination of a subject's motion (e.g. walking) and the presence of background noise (e.g. music, speech, white noise) would impact the subject's performance negatively, as they would have an additive effect on cognitive load. Thus, it can be inferred that the greater the level of distraction the more adverse its impact will be on performance.

Building on the theory described in this chapter thus far, a research framework for contextual usability will be proposed. This framework has a dual purpose. First, it serves as a roadmap for future research studies in the domain of usability. Second, this dissertation's research model will emerge from the contextual usability framework by identifying the selected set of variables investigated in our study.

3.5 Research Framework

From the earlier discussion on context (see Section 3.2) we posit that there are four main groups of variables that impact usability: *user*, *environment, task*, as well as *technology* characteristics. For the purpose of this research, the following descriptors are provided by assuming a scenario of mobile device usage. *User* refers to any individual's characteristics (e.g. age, gender) or state (e.g. walking, sitting) that is actively involved in a task involving a mobile device. *Environment* represents all external factors (beyond the user's direct control) that may have an impact on the user's satisfaction with the mobile device (e.g. room lighting). *Task* features include the nature of the activity (e.g. spontaneous, planned), its goals, and its corresponding complexity (measured by its consequent cognitive demand or task load). *Technology* accounts for all related items that could impact the user's experience with the mobile device. These include the device interface (e.g. input modes, such as keypads, speech-recognition), the device itself (e.g. hardware), and the supporting network (e.g. network dependability). While the characteristics shown in the research framework are adapted from past studies, this study will not engage in validating all of them, rather they are shown for completeness and are not meant to be exhaustive. The rationale behind the selection of the four main categories that define context is that each one contains unique characteristics, while interactions that may occur among any of them cannot be mapped since interactions are dynamic and constantly changing. For example, in one scenario there may be multiple users engaged in a single task, where interaction occurs among users and possibly with the environment (e.g. if users are in proximity of each other), and with the technology (e.g. if there are implications from concurrent use of applications). In a second scenario a single user may be involved in multitasking. Here, interaction occurs among tasks (e.g. increased complexity) and possibly with the technology (e.g. if the task involves use of the same infrastructure, thereby impacting the device and/or network). Therefore, interaction effects need to be considered as they occur and should not be grouped under any of the four main categories.

From the discussion on usability, it was argued that a comprehensive measurement of usability requires three constructs: efficiency, effectiveness, and satisfaction. This is shown in Figure 3.2, along with the grouping of efficiency and effectiveness under performance, which in turn impacts satisfaction (supported in the sections of literature review and theoretical foundations). Lastly, from the discussion on adoption, the single construct of behavioural intention to use a product is of most significance to this research.

Hence, a research framework (see Figure 3.2) was developed based on the literature review and the associated theoretical foundations described earlier for each of the major concepts, including context, usability, and technology adoption. As shown, these concepts are shown as dotted columns, in which various constructs are identified by solid rectangles. Furthermore, context was described as consisting of four groups of contextual factors: the user, task, technology, and environment. The properties of these contextual factors are encompassed by dashed rectangles. Such properties include sub-groups of potential constructs that are shown as underlined.

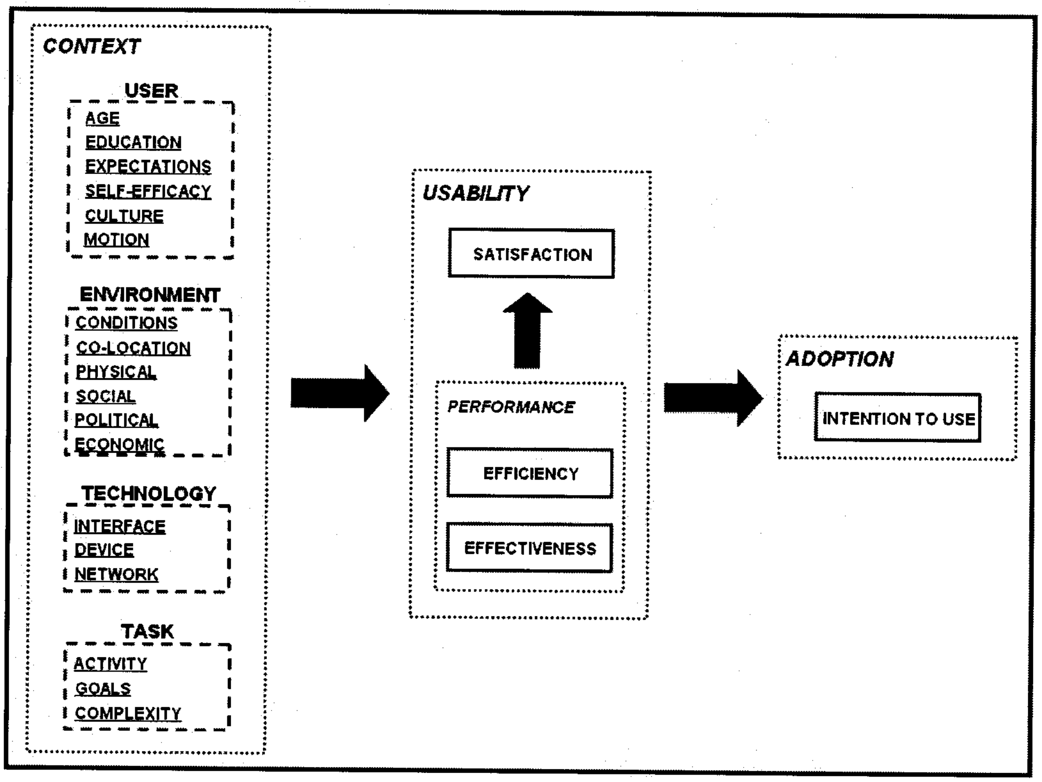


FIGURE 3.2: RESEARCH FRAMEWORK FOR CONTEXTUAL USABILITY

As can be seen from the illustrated framework, there are a number of potential antecedents to usability, while the most obvious consequent is that of behavioural intention to use a technology. Next, the research model for this dissertation will be developed building on the framework of Figure 3.2.

3.6 Research Model

Throughout the discussion in the literature review an increasing need for contextual usability studies became apparent. By leveraging the proposed research framework of Figure 3.2, a contextual usability study could be performed focusing on the user. This focus could either be in direct terms of user characteristics (e.g. age), or indirectly by exploring a user's interaction effects (e.g. interaction of user characteristics with the environment).

To this end, this study will control for individual characteristics with the exception of motion (i.e. seated vs. walking), while studying the impact of environmental characteristics (i.e. music and meaningful speech as background noise and collocation with others) acting as distractions to the user. The combined impact of these external factors on performance and in turn on satisfaction will be measured. Performance is decomposed into efficiency and effectiveness (supported above in Section *3.3. Expectations*). Hence, the study measures the impact of distractions on the perceived efficiency and the perceived effectiveness of the mobile device for wireless data services, which in turn impact the constructs of satisfaction with efficiency and effectiveness respectively. Finally, since satisfaction was shown to impact adoption (for a review see Oliver 1997), both satisfaction constructs (i.e. with respect to efficiency and effectiveness) are studied on their effect on the behavioural intention to use mobile devices for wireless data services. These relationships are captured in the proposed research model for this study shown in Figure 3.3, and the corresponding hypotheses are discussed in the next section.

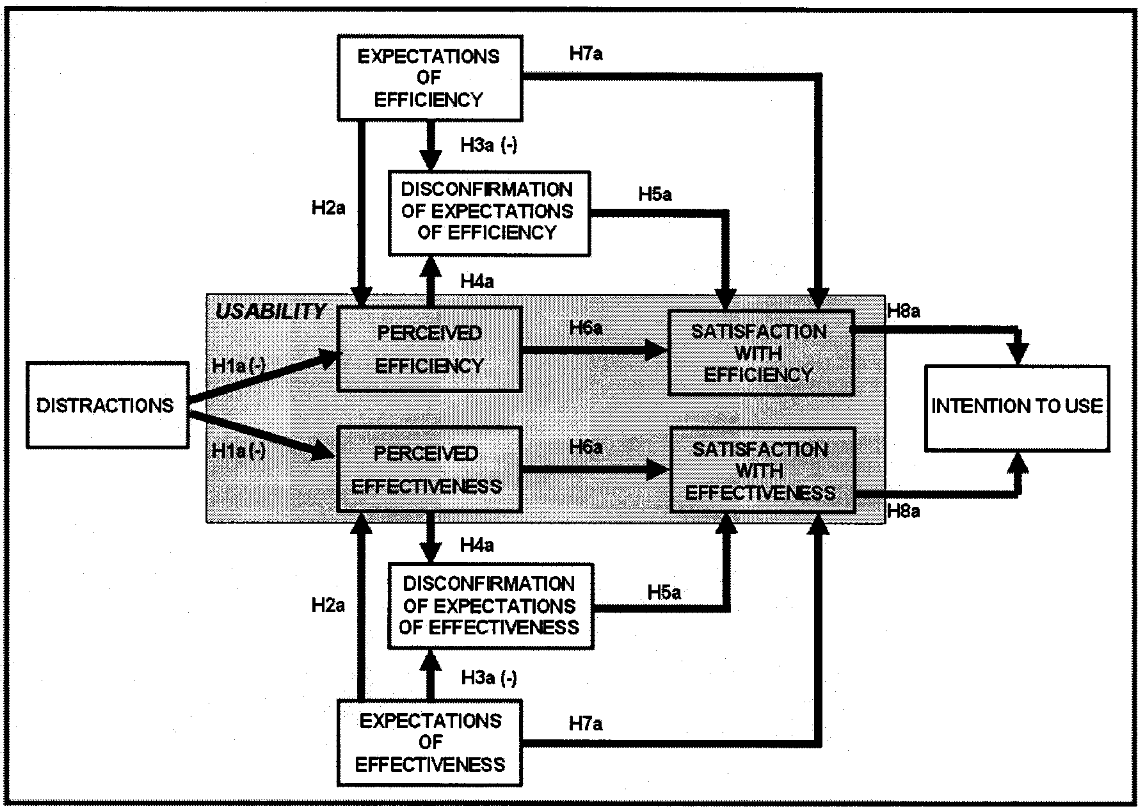


FIGURE 3.3: PROPOSED RESEARCH MODEL FOR CONTEXTUAL USABILITY STUDY

3.7 Research Hypotheses

Empirical research sets out to support either effects or theory application through its generalizability (Calder et al. 1981). Generalizability with respect to effects suggests that any effects demonstrated during a study would also be observed in the real world. Generalizability with respect to theory suggests that hypothesized theory is validated during the study. This research has a dual focus.

First, generalizability of theory application is hoped for the following: the proposed theory is that usability is not a unidimensional construct, nor is it a simple evaluation that a user arrives at instantly based on the experience with the product. Instead, it is argued that the evaluative process for usability is one that combines both a set of initial beliefs (i.e. expectations) and subsequent experience (i.e. perceived performance) with a product (i.e. mobile device). Generalizability here would support the applicability of EDT in usability studies.

Secondly, generalizability of effects application is intended for the following: it is hypothesized that distractions in the users' environment will impact their performance with a mobile device and their satisfaction with it. It is argued that distractions, such as music and co-location of others (serving as both auditory and visual stimuli) are of significance when a user is attempting to perform typical tasks on a typical mobile device. Generalizability here would support the claim of the cause-effect relationship between environmental distractions and performance with mobile devices.

We now turn our attention to the research model and the various hypotheses proposed (see Figure 3.3). Support for these hypotheses has already been provided in earlier sections of this chapter, but a recap of the key supporting literature for each of them is given below.

Starting with the exogenous construct, Distractions have been shown to have a negative effect on performance (Baker and Holding 1993; Rader 1993). Supported earlier (see Section *3.1 Usability and Technology Adoption*), performance is decomposed into efficiency and effectiveness. Hence, the following hypotheses are proposed:

H1a: Exposing users to higher levels of distractions will negatively influence their perceived efficiency of a mobile device for wireless data services.

H1b: Exposing users to higher levels of distractions will negatively influence their perceived effectiveness of a mobile device for wireless data services.

The case that consumers may experience "self-fulfilling prophecy" (i.e. what the individual expects of performance is eventually perceived), has been supported (Van Ryzin 2004; Oliver 1980, 1997; Spreng et al. 1996; Olchavsky and Miller 1982; reviewed by Yi 1990). Hence, there is a direct effect of expectations on perceived performance and the following hypotheses are proposed:

H2a: Higher user expectations of efficiency of a mobile device will positively influence the perceived efficiency of using the mobile device for wireless data services.

H2b: Higher user expectations of effectiveness of a mobile device will positively influence the perceived effectiveness of using the mobile device for wireless data services.

According to the Expectancy Disconfirmation Theory (EDT) (Oliver 1980; 1997), using a product will allow a consumer to form a perception about its performance. Initial expectations and perceived performance combined result in a qualitative assessment, whereby unmet expectations would yield negative disconfirmation (or emotional dissatisfaction). In the event that expectations are met, confirmation occurs, and finally if expectations are surpassed, a consumer experiences positive disconfirmation and consequently satisfaction. Therefore, and in accordance to EDT, disconfirmation is the gap between the initial expectations (pre-trial) and the realized performance (post-trial) from a product or service (Oliver 1980, 1997). Prior research has shown that there is a negative effect between expectations and disconfirmation, since the higher the initial set of expectations the more difficult it will be for them to be met by the experience with the product (Van Ryzin 2004; Oliver 1980, 1997). Hence, the following hypotheses are postulated:

H3a: Higher user expectations of efficiency of a mobile device will negatively influence the disconfirmation of user expectations of efficiency in using the mobile device for wireless data services.

H3b: Higher user expectations of effectiveness of a mobile device will negatively influence the disconfirmation of user expectations of effectiveness in using the mobile device for wireless data services.

On the other hand, performance has been shown to have a positive effect on disconfirmation, since the higher the performance, the more likely it will surpass expectations (Oliver 1980, 1997; Van Ryzin 2004). Hence, the following hypotheses are proposed:

H4a: Higher levels of perceived efficiency of a mobile device will positively influence the disconfirmation of user expectations of efficiency in using the mobile device for wireless data services.

H4b: Higher levels of perceived effectiveness of a mobile device will positively influence the disconfirmation of user expectations of effectiveness in using the mobile device for wireless data services.

Continuing along the remaining arguments of EDT, disconfirmation then has a direct effect on satisfaction (Oliver 1980, 1997; Van Ryzin 2004). Hence, the following hypotheses are proposed:

H5a: Higher levels of disconfirmation of user expectations of efficiency of a mobile device will lead to higher levels of user satisfaction with the efficiency of the mobile device for wireless data services.

H5b: Higher levels of disconfirmation of user expectations of effectiveness of a mobile device will lead to higher levels of user satisfaction with the effectiveness of the mobile device for wireless data services.

Using the product will allow a consumer to form a perception about its performance. Research has shown that perceived performance can have a direct effect on the consumer's satisfaction with the product (or service) (Churchill and Surprenant 1982; Tse and Wilton 1988). Hence, the following hypotheses are postulated:

H6a: Higher levels of perceived efficiency of a mobile device will lead to higher levels of user satisfaction with the efficiency of the mobile device for wireless data services.

H6b: Higher levels of perceived effectiveness of a mobile device will lead to higher levels of user satisfaction with the effectiveness of the mobile device for wireless data services.

In addition to performance and disconfirmation, expectations may also have a direct effect on satisfaction. This may occur when users are relatively unaware of the product's performance or have little psychological involvement during its use (Oliver 1997; Van Ryzin 2004). This dissonance reduction causes users to assimilate their satisfaction judgment to their initial expectations. For example, users' expectations of efficiency from the mobile device for wireless data services may be assimilated to a broader perspective on mobile devices. If this is the case, then a confirmation bias may result in the users' satisfaction judgment being formed accordingly to maintain their original, more general, assessment. Thus, the following hypotheses are proposed:

H7a: Higher user expectations of efficiency of a mobile device will lead to higher levels of user satisfaction with the efficiency of the mobile device for wireless data services.

H7b: Higher user expectations of effectiveness of a mobile device will lead to higher levels of user satisfaction with the effectiveness of the mobile device for wireless data services.

Lastly, there is a strong effect between satisfaction and a consumer's behavioural intention to use a product (Anderson and Sullivan 1993, Cronin and Taylor 1992, Gotlieb et al. 1994; Mittal et al. 1999; Taylor and Baker 1994). Hence, the following hypotheses are proposed:

H8a: Higher levels of user satisfaction with the efficiency of a mobile device will positively influence the user's intention to use it for wireless data services.

H8b: Higher levels of user satisfaction with the effectiveness of a mobile device will positively influence a the user's intention to use it for wireless data services.

Having stipulated the hypotheses pertaining to the proposed research model, the next chapter will discuss the methodology employed in carrying out the research study to validate these hypotheses.

CHAPTER 4

RESEARCH METHODOLOGY

This chapter describes the research methodology employed in this work. It begins with an overview of the experimental design and procedure. Details on the participants, tasks, instrument used in the study are provided next. This chapter will conclude with a description of the operationalization of both independent and dependent variables.

4.1 Experimental Design and Procedure

An empirical study was conducted to validate the proposed research model by testing our proposed hypotheses. The study was designed as a 2 x 2 factorial design (Factor 1: User motion; Factor 2: Auditory and visual stimuli in the environment). This design allowed for any differences found among the four groups of subjects to be attributed to the increased levels of distraction as a result of user motion and/or both visual and auditory cues in the environment. This approach was also used in usability studies by Watters (2003) and Chen and Vertegaal (2004), while a variant (2 x 3 factorial design) was employed by Kjeldskov and Stage (2004).

A total of 93 participants were recruited for the study, with a minimum of 20 subjects in each of the four treatments. This sample was well above the required total sample size of 40. The latter was determined based on the requirements of the methodology employed for data analysis, Partial Least Squares (PLS). More details on the selection of PLS and the sample size can be found in a later section (see Section *5.2*, *Data Analysis Technique*). The tasks involved in this study consisted of authoring text messages, entering data in both the calendar and address book applications, and browsing the wireless Web. Details on these tasks are described in a later section (see Section *4.3, Tasks*). These tasks were carried out on a Personal Digital Assistant (PDA) model chosen for its ability to accept user input by means of a built-in QWERTY keyboard: RIM's Blackberry 7250. This device is a fair representative of typical PDAs currently available in the market in terms of the supported functionality and general form factor. The built-in QWERTY keyboard was chosen as the input mode, because of its increasing popularity among new mobile devices released in the market, and its superior data entry performance over other input modes and convenience (e.g. carrying and attaching an external keyboard) (Soukoreff 2002).

Each subject participated in only one treatment group, and assignment of subjects to groups was fully randomized to control for confounding effects due to differences in subject characteristics. Every participant received $10 for their participation, which lasted between 30 and 45 minutes. Five minutes were allocated for the pre-test survey, ten minutes were used to train the subjects on the use of the device and the built-in QWERTY keyboard, twenty minutes were needed, on average, to run the experiment and have the subjects complete the required tasks, while another five minutes were allocated for the post-test survey. Participants were not allowed to interact with others during the experiment in an attempt to isolate the environmental conditions that were being tested. The experiment administrator did not offer any assistance throughout the experiment, again in an attempt to approximate the realism of the task.

All participants were physically separated by allowing for a minimum five-minute margin between sessions. Participants began by completing a pre-test survey. As will be described later (see Section *4.4, Instrument*), this survey gathered in part demographic data on the variables controlled in the study, including gender, education, and age. In addition, expectations were measured as they are exogenous constructs in this dissertation's proposed research model.

Then, the experimental procedure was described to the participants, following which a brief tutorial was given to them on how to use the mobile device for the study's tasks and finally an opportunity to carry out the practice runs for the various tasks. This experiment sequence is shown in Figure 4.1.

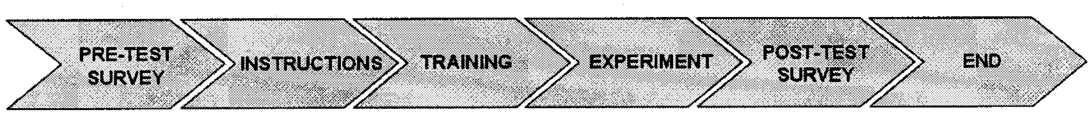


FIGURE 4.1: DISSERTATION EXPERIMENT SEQUENCE

Similar to Chen and Vertegaal (2004) and Nicholson et al. (2005), the four groups of subjects then conducted the experiment under varying cognitive loads. The tasks completed were the same in all experimental treatments, with only user motion and auditory/visual stimuli as changing parameters. To study the effect of user motion subjects being assigned to groups I and III were asked to complete the tasks while being seated. On the other hand, those subjects assigned to groups II and IV were asked to walk in a controlled environment (i.e. large room in a building), staying within the boundaries of an outlined path on the ground that is changing (i.e. non-linear), and walking at a steady pace. To study the effect of stimuli in the environment, subjects assigned to groups 1 and 2 were asked to undertake the tasks in the absence of any background noise (in this study music and speech) or visual stimuli (in this study the presence of and motion by other individuals). On the other hand, subjects assigned to groups 3 and 4 did so in the presence of such stimuli. Distractions in this study were either the isolated or combined effects of user motion or/with visual and auditory stimuli in the environment. Five individuals were hired to assist in this study and were asked to carry out specific roles during each observation: one actor walked alone on an assigned path (no marking were used to prevent the path from being visible by others), while two pairs of actors were asked to walk on their respectively assigned paths while carrying out conversations (scripts were assigned that contained varying topics, including talking about movies and relationships). Location of the actors' activities were distributed throughout the room, but at no time were any of the actors to come into contact with the test subject.

To increase the realism of the task, subjects needed to make their way through the menu to the appropriate screen (e.g. text message authoring screen) on their own.

The two levels for each of the two variables (motion and distraction) were set based on the mobile usability work by Maguire (2001), Chen and Vertegaal (2004), Kjeldskov and Stage (2004), and Nicholson et al. (2005). Aside from this work, there has been no other comparable study, to the best of our knowledge. Each of the two levels for these variables were expected to impact the efficiency and effectiveness of the mobile device and the wireless data services performed. Table 4.1 illustrates the experimental setup for this study as described above.

Following the experiment there was a post-test survey. This post-test survey was used to measure perceived efficiency and perceived effectiveness, disconfirmation of expectations from and users' satisfaction with technology efficiency and effectiveness, users' satisfaction with the respective performance dimensions, as well as the users' intention of future use of the mobile device and wireless data services.

To protect the anonymity of the participants, the written materials used by them were separated in three groups: i) the onymous consent form was separated from the remaining materials; ii) both the pre-test and post-test surveys were coded to link the two but in a manner not identifying the subject. Specifically, the code used was XY##, e.g. WD15, where:

- X was the group to which the subject was randomly assigned to, i.e. Seated (S) or Walking (W)
- Y was the group to which the subject was randomly assigned to, i.e. Quiet (Q) or Distracted (D) (note: these terms were used for convenience; Q involved the absence of both auditory and visual distractions, whereas D included the presence of both auditory and visual stimuli)
- # was assigned in sequence to help track the number of subjects in each treatment

TABLE 4.1: TREATMENT CONDITIONS (I.E. INDEPENDENT VARIABLES) FOR DISSERTATION STUDY

<table>
<tr><td colspan="2" rowspan="2">Experiment Independent Variables</td><td colspan="2">USER</td></tr>
<tr><td>Static</td><td>Mobile</td></tr>
<tr><td rowspan="2">ENVIRONMENT</td><td>Quiet & Alone</td><td>Group I</td><td>Group II</td></tr>
<tr><td>Noise & Co-location</td><td>Group III</td><td>Group IV</td></tr>
</table>

4.2 Participants

For the experiment a total of sample of 93 participants were recruited, having aimed for an equal representation of gender, all were native English speakers, and covered a broad range for age and education. These subjects were recruited from McMaster University and included students, staff, and faculty. The sample size was subject to both the relevant theory in the domain of usability, as well as the methodology used in analysing the data.

According to usability theory, Nielsen (1994) showed that four to five experimental subjects could find 80% of usability problems in any system, while Carroll (1997) suggested that an average of nine subjects in laboratory usability evaluations is enough for usability analysis.

The methodology that was used for data analysis is PLS. Support for this decision is provided in a later section (see Section *5.2 Data Analysis and Technique*). Based on the constraints imposed by the methodology, the required sample size is 10 times the number of items in the most complex formative construct (Chin and Newsted 1999). In this study, the most complex formative construct is that of Expectations of Efficiency with four items. Therefore, to use PLS for the analysis of this model a minimum of 40 participants was required.

To compensate for error (i.e. subjects' experimental sessions and/or questionnaires that cannot be used), we recruited twice the number of the required sample size. Thus, an additional 40 subjects were to be recruited for a total sample size of 80 (or 20 subjects per treatment) satisfying both the usability theory and methodology constraints. Given the random assignment of subjects in each treatment, the minimum of 20 subjects in any given treatment was reached after 93 subjects were recruited in total.

These participants were individuals who are students, faculty, or staff at McMaster University. To recruit these subjects an email invitation was sent to individuals through the university's aggregate e-mail directory. An announcement was posted as well on the university's website under an appropriate page (i.e. "Announcements"). Similar methodologies have been employed by other researchers, e.g. Pennington et al. (2004) and Chen et al. (2002). Subjects who wished to participate in the study first completed a consent form and a demographic questionnaire. The test subjects were randomly assigned to one of the four treatments. Furthermore, any prior experience with PDAs for the functionalities involved in this study (i.e. text messaging, directory/calendar entries, wireless Web browsing) was acceptable and documented, although no such subjects participated. The significant effect that user experience has on data entry performance has been shown by MacKenzie and Zhang (1999), where their subjects' performance improved by approximately 22 percent from the first to the fifth use of the mobile device (for the purpose of entering text), and continued to improve to an overall improvement of 36 percent after 50 attempts. To protect against the effects of such steep learning curves, participants were asked to perform training on text messaging (see *Section 4.3 Tasks*) prior to the measured test bringing all subjects within a comparable level of data entry speed.

4.3 Tasks

The experiment tasks involved the use of four applications: sending text messages, scheduling an appointment in the calendar, updating the address book, and browsing the Web on the mobile device. These four tasks create the most value, second only behind voice communication, for consumers (Jarvenpaa et al. 2004). It should be noted that all tasks were randomized within-application and between-applications, i.e. the order of the sentences that were entered during the text messaging task (or application), as well as the order of the four applications that were used were randomized across subjects. The four tasks are described below in detail:

Text messaging Task

Participants were given the set of sentences to be used for the text messages on cue cards. They were also given the following scenario (adapted from Pedersen 2002): "We now want you to focus on a text messaging service that is used to keep or get in contact with others. Examples of such service are sending text messages to friends and family. Using text messages to receive logos and ring tones is not relevant here."

The text messaging task involved participants entering ten different messages in total: six practice messages were sent during the training session and four messages sent during the actual experiment. These messages were taken from James and Reischel (2001) who compared two different input modes: T9 Text Input and Multi-tap. Under the Multi-tap approach, the language characters are divided onto eight keys found on the device. As a user hits a key the character is selected by circulating through the assigned character set to the particular key (e.g. one hit of the "2" button will call upon "A", two hits of the same will call "B" etc.). Under the T9 Text Input, a system developed by Tegic Communications, for a given word to be entered a user hits the various keys that correspond to the characters associated with that word, regardless of their respective assignment within the key. The built-in logic then assembles a word that is most frequently used given the possible character combinations. In the event that the word is not the one desired, the user has the option of browsing through all the remaining words that can be assembled. Participants in the James and Reischel (2001) study entered two sets of messages (five each) that were taken from either newspaper type content or from a chat session. The first message in each set served as a practice run. In our study, participants used all five newspaper messages and the first of the five chat messages as practice during the training session, while only the last four chat messages were sent during the experiment and considered in the post-hoc analysis. This approach accounted for novice users' steep learning curve experienced during the first few runs. In other words, the experience that the subjects would have from entering the messages during the actual experiment (after having practiced with the first six messages) is closer to the real-world experience that new users of wireless data services would have when they consider such use (e.g. a consumer, alone or with the help of a sales person, may try it out for a few sessions before making a purchase decision).

The six practice text messages were:

- this is the third year for the toy hall of fame awards
- the slinky and the jump rope are among five toys that will be designated Wednesday
- the bicycle and jacks will also be honoured
- those two playthings will formally be made part of the hall of fame this fall
- selections were chosen by a panel of educators
- hi joe how are you want to meet tonight

The four text messages sent during the actual experiment were:

- want to go to the movies with sue and me
- what show do you want to see
- we are meeting in front of the theater at eight
- let me know if we should wait

The order of presentation for the four messages was randomized across participants. Each participant attempted to enter all messages consecutively, while being exposed to the particular treatment conditions.

Calendar Task

Each participant was required to enter information (i.e. text) in the calendar of the mobile device and so was given the following instruction during the training session:

"We now want you to focus on a calendar service that is used to enter appointments and reminders for upcoming events. Examples of such services are entering meetings that occur on a weekly basis or a one-time visit to the dentist. For this task you are to enter an appointment in the calendar regarding a meeting (subject and location are not required) on December 4 (Sunday), 2005, from 9:30-18:00; the appointment will be repeated the following Sunday."

During the actual experiment, participants were asked to enter the following appointment (taken from Lindroth et al. 2001):

"Calendar - schedule task:

1. Start the device and open the application "Calendar"
2. Add an appointment Tuesday, November 21st at 15.00-17:30 with the text "test"
3. The appointment should be repeated every Tuesday for three weeks."

Address Book Update Task

Participants were asked to enter a person's contact information in the address book function of the mobile device. During the training, they were instructed to:

"We now want you to focus on an address book service that is used to manage contacts by storing relevant information. Examples of such service are entering first and last name, email, phone numbers, mailing addresses, etc. For this task you are to enter the following information in the corresponding fields (instructions are outlined below):

- Name: Jean Renault
- Title: Professor
- Work: 301 9696224
- E-mail: Jean_Renault@hotmail.com
- City: Paris"

During the actual experiment, participants were asked to perform the following task (taken from Lindroth et al. 2001):

"Address book – enter contact information task:

1. Start the device and open the application "Address Book"
2. Add a new address:
    - Name: Ib Rene
    - Title: Correspondent
    - Work: 98 12345
    - E-mail: Ib_rene_cairo@hotmail.com
    - City: Aalborg"

Web Browsing Task

Participants used the wireless Web with the goal of finding certain specific information. Sellen et al. (2002) identified several categories of Web searching activity that were broadly classified either as "finding" or "information gathering". Previous usability studies (Buyukkokten et al. 2001; deBruijn et al. 2002; Jones et al. 1999) used "finding" tasks, because they offer more control than open-ended "information gathering" ones. Following suit, a "finding task" was used in this experiment where participants looked for something specific. Each subject began by training on the mobile device Web browser by answering an initial practice question.

Starting from an index page, the subject independently followed links in pursuit of obtaining the answer to the given question. To avoid the impact of any external variables, such as network connection and signal, all web pages relevant to the search were stored (cached) on the mobile device. A script was given to the participants at the start of the task to put them into the right context during the actual experiment. The script read (adapted from Rodden et al. 2003):

"We now want you to focus on a wireless Internet service that may be used to locate information, make purchases, and engage in entertaining activities such as downloading ring tones and playing online games. For this task you are to access the wireless Internet portal (i.e. gateway to a wide range of information) available on the mobile device and navigate through the various links available to you as you attempt to find the answer to the following question:

- What is the first Air Canada flight out from Toronto to Montreal on August 3?"

Then, the subject attempted to find the information needed to answer the above question. During the actual experiment, subjects were given the following question to answer:

"What is the high temperature going to be in Hamilton tomorrow?"

If participants were unable to find the correct answer in two minutes during the actual experiment, they were asked to stop and the task was considered a failed one.

In addition to the arguments provided on the previous page regarding the choice to include a Web "finding task", the work of Kjeldskov et al. (2005) adds further support. Their work highlights the strong likelihood that users are using these Web-based applications to find information that is "closely related to their physical location and objects in their immediate surroundings" (Cheverst et al. 2000; Chincholle et al. 2002; Schmidt-Benz et al. 2002; Reid 2002; Umlauft et al. 2003). In addition, users are often making use of these Web-based applications while they are moving from one physical location to another (Kjeldskov et al. 2005). Thus, having participants search to find specific information (e.g. routes, timetables) on the wireless Web is a simulation of realistic and current applications.

4.4 Instrument

The questionnaire used for data collection contains scales that measure the various constructs shown in the proposed research model. These scales are shown in Appendix A. All scales were adapted from prior studies, which had established their reliability and validity. References to these studies are provided in the section of operationalization of variables that follows later (i.e. Sections *4.5.* and *4.6.*). The only scales that were modified after reviewing relevant literature were the first-order factors defined for efficiency and effectiveness. These changes and the rationale in performing them are described later in Section 4.6.1. (Perceived Measurement of Dependent

Variables) when the two concepts are operationalized. Otherwise, and in accordance to the advice of Fishbein and Ajzen (1975), Davis (1989), and Mathieson (1991), all instrument items were adapted to the use of the mobile device rather than to general IS use. When the questionnaire was conducted items within the same construct group were randomized to prevent systemic response bias. Effort was made to ensure consistency of scales. As such, all scales were anchored Highly Disagree (1), Moderately Disagree (2), Somewhat Disagree (3), Neutral – Neither Disagree Nor Agree (4), Somewhat Agree (5), Moderately Agree (6), Highly Agree (7), except for the items measuring Disconfirmation and Satisfaction. The scale anchors for these constructs are defined further below, under Perceived Measurement of Dependent Variables. Support of this instrument will be provided in this research through validity tests described later (see Section *5.6 Psychometric Evaluation of Measures*).

Before we proceed with a discussion on the operationalization of both dependent and independent variables, we define the variables that were controlled in this experiment.

Control Variables

Several dimensions were controlled for in this study, as there has been research to support their potential impact on usability and/or adoption. These control variables are: age, gender, and education. Control variables were measured as single-items respectively, as recommended by Gefen and Straub (1997), and taken from Khalifa and Cheng (2002).

4.5 Operationalization of Independent Variables

The experiment involved two variables, user motion and auditory and visual stimuli in the environment, each having two levels. The importance of these variables has been articulated by Maguire (2001) and Chen and Vertegaal (2004), where the objective is to create scenarios where distractions, and by extension cognitive demands, are increased during the subjects' use of a mobile device. User motion was explored as a dichotomous variable, where participants were either experiencing no motion (i.e. seated) or motion in the form of walking a variable path at a constant pace. The presence of varying motion was expected to produce a large effect size. This effect size was expected to be greater than one caused by motion in a straight path or at a constant pace. The latter may have become an automated function performed by the participant thus requiring very little cognitive effort and act as a minimal distraction (Kjeldskov and Stage 2004). By varying both the path and pace the user was distracted at a higher level than that of the control state (i.e. seated).

Similarly, auditory and visual stimuli in the environment were either absent or present. When present, distraction-causing cues were expected to reduce the participants' cognitive capacity to direct their attention to the task at hand. Distraction was caused by introducing auditory (i.e. background music and meaningful speech) and visual cues (i.e. presence of other people in proximity to the participant, some of whom are also walking). These distraction-causing cues were expected to produce a large effect due to the additive effect of the stimuli in the environment on the participants' cognitive load. A moderate effect could have been caused by introducing any of these distraction cues in isolation.

While the two variables (e.g. motion and both visual and auditory stimuli) are mutually exclusive, they share an underlying factor, cognitive load, and both act as distractions to the user in performing their primary assigned tasks (i.e. tasks of the experiment). Cognitive load (or workload) has been defined as either "the objective workload imposed by the task … or the subjective rating of the operator with regard to the demands of the task. In most cognitive workload theories workload refers to the information processing capacity of the operator" (MEGATAQ 1999). In this experiment both user motion and stimuli in the environment affected the participants' information processing capability. To capture the participants' cognitive workload a popular self-reported measure was used that also served as a manipulation check for the four treatments, namely the NASA Task Load Index (TLX) (Hart and Staveland 1988). After performing the experiment, participants were asked to respond to six 20-point scales that measure: Mental demand, Physical demand, Temporal demand, Effort, Performance, and Frustration (see Table 4.2). In previous studies that used the TLX, scales had been removed by the researchers if it was felt appropriate for the study (e.g. Bushey et al. 1999). Table 4.2 offers the Distractions questions that were used in this study.

TABLE 4.2: THE DISTRACTIONS SCALE

Scale items. 20-point Semantic Differential scale (Low/High) Instructions: Please answer all questions based on your experience with using the mobile device for wireless data services.	
Distractions Scale (TLX)	
Item Identifier	**Question**
TLX1	How much mental and perceptual activity was required (e.g., thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy/simple/forgiving (i.e. LOW) or demanding/complex/exacting (i.e. HIGH)
TLX2	How much physical activity was required (e.g., pushing, pulling, turning, controlling, activating, etc.)? Was the task easy/slow/slack/restful (i.e. LOW) or demanding/brisk/strenuous/laborious (i.e. HIGH)?
TLX3	How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred? Was the pace slow and leisurely (i.e. LOW) or rapid and frantic (i.e. HIGH)?
TLX4	How hard did you have to work (mentally and physically) to accomplish your level of performance? (LOW/HIGH)
TLX5	How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals? (GOOD/POOR)
TLX6	How insecure, discouraged, irritated, stressed and annoyed (i.e. LOW) versus secure, gratified, content, relaxed and complacent (i.e. HIGH) did you feel during the task?

The construct for Expectations of Efficiency was measured as a second-order factor using the first-order factors of effort, learnability, and time. These first-order factors were generated by reviewing literature relevant to "the level of resource consumed in performing tasks." As such, the resources considered were training time, effort, and task completion time (Frokjaer et al. 2000). Training time required is associated with the learnability of the technology (MacKenzie and Zhang 1999), i.e. how easy it is to learn to operate it. Therefore, the item used for training time is "learning to use the mobile device was easy", an item adapted from Gefen (2003). Effort is associated with ease of use. Therefore, the two items used (mobile device is user friendly; mobile device is easy to use) were adapted from McHaney et al. (2002) scales used for the Ease of Use construct in their model. Finally, task completion time has typically been an objective measure (it is also used in this study as one). By modifying an item from Butts and Cockburn (2001) we include "using the mobile device was fast" which helps us obtain a self-reported perception of the task-completion-time dimension. Table 4.3 offers the Expectations of Efficiency questions that were used in this study.

TABLE 4.3: THE EXPECTATIONS OF EFFICIENCY SCALE

Scale items. 7-point Likert scale (Strongly Disagree/Strongly Agree) **Instructions:** For each statement below, please circle the number to the right that best matches your opinion on each statement.	
Expectations of Efficiency Scale (ExpEffi)	
Item Identifier	**Question**
ExpEffi1	I expect that learning how to use a mobile device for wireless data services will be easy.
ExpEffi2	I expect that using a mobile device for wireless data services will be easy.
ExpEffi3	I expect that using a mobile device for wireless data services will be fast.
ExpEffi4	I expect that a mobile device will be user friendly for wireless data services.

Similar to the construct of Expectations of Efficiency, the second-order construct for Expectations of Effectiveness was measured using the first-order dimension of successful task completion. This first order factor was generated by reviewing literature relevant to "the user's ability to complete tasks using the technology" (ISO 9241). This ability to complete tasks is typically measured objectively by the percentage of tasks completed among the entire set of tasks required to complete the entire experiment (Chittaro & Dal Cin 2002). A perceived item would gauge the subjective evaluation of how successful a user was in completing a set of tasks. Hence, the item used is "able to complete all tasks successfully." Single-item constructs can be represented in structural equation modeling (SEM) such as the partial least squares (PLS) data analysis technique used in this dissertation (Gefen et al. 2000). Table 4.4 offers the Expectations of Effectiveness Scale.

TABLE 4.4: THE EXPECTATIONS OF EFFECTIVENESS SCALE

Scale items. 7-point Likert scale (Strongly Disagree/Strongly Agree) **Instructions:** For the statement below, please circle the number to the right that best matches your opinion.	
Expectations of Effectiveness Scale (ExpEffe)	
Item Identifier	**Question**
ExpEffe	I expect that I will be able to complete all wireless data services on a mobile device successfully.

4.6 Operationalization of Dependent Variables

The dependent variables of the proposed research model (see Figure 3.3), were primarily investigated by the self-reported measures included in a questionnaire. To supplement this set of data, we also gathered objective (device) data that were intended to enrich our post-hoc analysis. The two dimensions of collected data, perceived and objective are described next.

4.6.1 Perceived Measurement of Dependent Variables

It was pointed out earlier in this paper (see Section *3.3. Expectations*) that performance was decomposed into efficiency and effectiveness. The construct for Perceived Efficiency was measured as a second-order factor using the first-order factors for training time, effort, and task completion time. These factors are the same as those used in Expectations of Efficiency (see Section *4.5. Operationalization of Independent*

Variables). Similarly, the second-order construct for Perceived Effectiveness was measured using the first-order dimensions of successful task completion. These two scales are shown in Tables 4.5 and 4.6 respectively.

TABLE 4.5: THE PERCEIVED EFFICIENCY SCALE

Scale items. 7-point Likert scale (Strongly Disagree/Strongly Agree) **Instructions:** For each statement below, please circle the number to the right that best matches your opinion on each statement.	
Perceived Efficiency Scale (PerEffi)	
Item Identifier	**Question**
PerEffi1	Learning how to use the mobile device for wireless data services was easy.
PerEffi2	Using the mobile device for wireless data services was fast.
PerEffi3	The mobile device was user friendly for wireless data services.
PerEffi4	Using the mobile device for wireless data services was easy.

TABLE 4.6: THE PERCEIVED EFFECTIVENESS SCALE

Scale items. 7-point Likert scale (Strongly Disagree/Strongly Agree) **Instructions:** For the statement below, please circle the number to the right that best matches your opinion .	
Perceived Effectiveness Scale (PerEffe)	
Item Identifier	**Question**
PerEffe	I was able to complete all wireless data services on the mobile device successfully.

Disconfirmation of expectations can be measured in two ways: perceived (or subjective) and inferred (or subtractive). Perceived disconfirmation represents a subjective assessment of the difference between the perceived performance and expectation, in other words the degree to which the perceived performance was better or worse than initially expected (Tse and Wilton 1988; Chiu et al. 2004; Swan and Trawick 1981). Inferred disconfirmation, on the other hand, refers to the algebraic difference between the perceived performance and initial expectation (Tse and Wilton 1988; Chiu et al. 2004; Swan and Trawick 1981). Most studies measuring satisfaction use subjective disconfirmation because it considers disconfirmation as a distinct construct thereby overcoming any confounding issues consequent of an overspecified satisfaction model (Bhattacherjee 2001; Churchill and Surprenant 1982; McKinney et al. 2002; Oliver 1980; Spreng et al. 1996; Wirtz and Bateson 1999; Chiu et al. 2004). Similarly, we used perceived disconfirmation in this study and the scale was adapted from Oliver (1980) and Spreng et al. (1996). The first and second-order factors for Disconfirmation of Expectations of Efficiency were adapted from the factors measuring *Expectations of* and *Perceived Efficiency* as shown in Table 4.7. Similarly, the first and second-order factors for Disconfirmation of Expectations of Effectiveness were adapted from the factors measuring *Expectations of* and *Perceived Effectiveness* as shown in Table 4.8.

TABLE 4.7: THE DISCONFIRMATION OF EXPECTATIONS OF EFFICIENCY SCALE

Scale items. 7-point Likert scale (Much Worse Than I Expected/About What I Expected/Much Better Than Expected) **Instructions:** For each statement below, please circle the number to the right that best matches your opinion on each statement starting as: *"Compared to my expectations of efficiency, the"* (note: *efficiency* refers to the level of resources consumed in performing a task, e.g. time and effort).	
Disconfirmation of Expectations of Efficiency Scale (DisEffi)	
Item Identifier	**Question**
DisEffi1	Learnability (that is, the degree to which it was easy to learn how to use) the mobile device for wireless data services was…
DisEffi2	The time required to use the mobile device for wireless data services was…
DisEffi3	The user friendliness of the mobile device for wireless data services was…
DisEffi4	Ease of use of the mobile device for wireless data services was…

TABLE 4.8: THE DISCONFIRMATION OF EXPECTATIONS OF EFFECTIVENESS SCALE

Scale items. 7-point Likert scale (Much Worse Than I Expected/About What I Expected/Much Better Than Expected) **Instructions:** For the statement below, please circle the number to the right that best matches your opinion starting as: *"Compared to my expectations of effectiveness, the..."* (*effectiveness* refers to the ability of users to complete tasks using the system).	
Disconfirmation of Expectations of Effectiveness Scale (DisEffe)	
Item Identifier	**Question**
DisEffe	Ability to complete all wireless data services on the mobile device successfully was ...

Satisfaction, defined as a summary evaluation of the entire technology-use experience, was measured according to the scales developed by Spreng et al. (1996) that included four variables using cognitive and affective components to describe satisfaction (i.e. satisfaction, pleasure, delight, and contentment). Therefore, the same scales were used for *Satisfaction with Efficiency* and with for *Satisfaction with Effectiveness* but the participants were clearly instructed in the question to respond according to the dimension being measured. These scales are shown in Tables 4.9 and 4.10 respectively.

TABLE 4.9: THE SATISFACTION WITH EFFICIENCY SCALE

Scale items. 7-point Semantic Differential Scale **Instructions:** For each statement below, please circle the number to the right that best matches your opinion on each statement starting as: *"Thinking about my experience with the efficiency of this device for wireless data services, I feel ..."* (note: *efficiency* refers to the level of resources consumed in performing a task, e.g. time and effort)	
Satisfaction with Efficiency Scale (SatEffi)	
Item Identifier	**Question**
SatEffi1	Terrible (1)..Delighted (7)
SatEffi2	Very displeased (1)...................................Very pleased (7)
SatEffi3	Very dissatisfied (1)..................................Very satisfied (7)
SatEffi4	Frustrated (1)..Contented (7)

TABLE 4.10: THE SATISFACTION WITH EFFECTIVENESS SCALE

Scale items. 7-point Semantic Differential Scale **Instructions:** Please circle the number to the right that best matches your opinion on each statement starting as: *"Thinking about my experience with the effectiveness of this device for wireless data services, I feel..."* (*effectiveness* refers to the ability of users to complete tasks using the system)	
Satisfaction with Effectiveness Scale (SatEffe)	
Item Identifier	**Question**
SatEffe1	Terrible (1)...Delighted (7)
SatEffe2	Very displeased (1)....................................Very pleased (7)
SatEffe3	Very dissatisfied (1).................................Very satisfied (7)
SatEffe4	Frustrated (1)..Contented (7)

Behavioural Intention is defined as the strength of one's intention to perform a specified behaviour. The scales used in this study were adapted by questions from Venkatesh and Davis (2000) and are shown in Table 4.11.

TABLE 4.11: THE BEHAVIOURAL INTENTION TO USE SCALE

Scale items. 7-point Likely Scale (Strongly Disagree/Strongly Agree) **Instructions:** For each statement below, please circle the number to the right that best matches your opinion on each statement.	
Behavioural Intention to Use Scale (BI)	
Item Identifier	**Question**
BI1	Given that I had access to the mobile device, I predict that I would use wireless data services in the near future.
BI2	Assuming I had access to the mobile device, I intend to use wireless data services in the near future.

4.6.2 Objective Measurement of Dependent Variables

Objective measurement of the dependent variables stemmed from data collected during the experiment. Recording of data began as soon as the first key was hit and ended as soon as the last letter was typed. This was done by an application that was installed on the mobile device and was running in the background tracking all user keystroke activity. These data were used to determine the objective performance, i.e. both efficiency and effectiveness, of the mobile device for wireless data services. Observations were made regarding the relationships between the two data sets (i.e. objective and perceived) that were discussed in our post-hoc analysis (see Section 5.10). However, it was the data collected from the perceived measures alone that were used to support the proposed research model and hypotheses implicitly.

There are two commonly used models in measuring the objective efficiency of a technology: keystrokes and time (Jue 2004; Chittaro and Dal Cin 2002). Under the keystroke model, each keystroke represents an action performed by the user. Therefore, objective efficiency is defined as the average number of keystrokes made during a set of iterations for a particular task, which is calculated by taking the ratio of the total keystrokes performed for all iterations of a task to the total number of iterations performed for the same task. In the time model we are interested in the average time required to perform a task. This value was determined by the ratio of the total time required to perform a set of iterations for a particular task to the total number of iterations for the same task (note: task refers to a single application/service used one or more times).

In addition, a hybrid of the two models was used by Dunlop and Crossan (2000) and Gentner et al. (1983). In their studies the total time for each text message was be calculated by taking the total number of characters, dividing them by the average number of characters per word (i.e. 5) and by the time it took for the task to be completed in seconds and then multiplied by 60 to get a count of *words per minute* (*wpm*). This method was also used in our study for the text messaging task, while the two models of keystroke and time were used for all tasks.

Objective measurement of effectiveness was based on the task completion rate (i.e. what percentage of the tasks did the user complete successfully) and how accurate the completed task was. The latter could be measured through the *total error rate* (Soukoreff and MacKenzie 2004). The formula to determine this variable given below (Soukoreff and MacKenzie 2004):

- Total Error Rate = (INF+IBF+F)/(C+INF+IBF+F)*100% **Equation 4.1**
- Where:
    - Correct keystrokes (C): alphanumeric keystrokes that are not errors,
    - Incorrect and Not Fixed (INF) keystrokes: errors that go unnoticed and appear in the transcribed text,
    - Incorrect but Fixed (IBF) keystrokes: erroneous keystrokes in the input stream that are later corrected, and,
    - Fixes (F): the keystrokes that perform the corrections (i.e. delete, backspace, cursor movement)

Instead of the error rate, for the analysis performed in this study the accuracy rate was computed as 1-Total Error Rate.

CHAPTER 5

DATA ANALYSIS AND RESULTS

In this chapter the analysis and results of the experiment data are presented. First, a section on the pilot study is provided, followed by a discussion of the administration of the questionnaire. Next, we describe the statistical tests used to test for non-response, temporal, and common method biases, followed by the participants' demographics. The psychometric evaluation of measures follows through a variety of statistical tests conducted including those for reliability as well as construct convergent and discriminant validities. The final sections present the results from hypothesis testing, a discussion on effect sizes, the control variables, and lastly a post-hoc analysis involving the objective data collected.

5.1 Pilot Study

A pilot test involving three participants was performed having three objectives:

a) Check for any problems regarding the interpretation of the items contained in the measurement instrument. Pilot test participants were given the opportunity to offer qualitative feedback at the end of the experiment following the completion of the post-test questionnaire. Rewording of items or provision of explanations and/or definitions could be needed to the measurement instrument, depending to the feedback obtained.

Upon conclusion of the pilot study a few modifications were made in the questionnaire with respect to the wording of scale items.

b) Practice the manipulation of distractions. By observing the participants' reactions, any fine tuning could be done with respect to the ground path the participants had to follow, the volume of the background music playing, or the actors' scripts being used.

Upon conclusion of the pilot study the manipulation of distractions appeared to be well carried out (i.e. the pilot study ran smoothly).

c) Check that sufficient time is allocated for subjects to complete the entire session.

Upon conclusion of the pilot study, it was evident that a 45 minute window would be necessary to accommodate each participant, instead of the 30 minutes originally allotted. In addition, for three out of the four tasks (i.e. text messaging, address book update, and wireless web search) performed by the participants the task completion time ranged between 30 and ninety seconds, while the calendar task was completed within one to two minutes.

Given the above objectives, three subjects were sufficient for the pilot study, because no statistical analysis was involved. Instead, direct observations were made of the three subjects. Each participant was assigned to one of the treatments except for the control treatment (i.e. Group I, seated and no stimuli in the environment). No novel observations were likely to be made regarding the logistics of the experiment in the control treatment, due to the absence of any manipulations under this condition.

In summary, the pilot study was valuable in providing us with the opportunity to run the experiment and identify and problems relating to questionnaire interpretation and logistics.

5.2 Data Analysis Techniques

Data analysis was performed using Partial Least Squares (PLS) on the "PLS Graph Version 3.00 Build 1126" application. PLS allowed us to specify the construct relationships between one another (structural model), as well as with their underlying items (measurement model). Thus, data analysis provided support for both how well the items measured each construct, and how well the hypothesized relationships between constructs supported the theory. PLS features two additional advantages over other methodologies. First, it is expected that the data to be collected will be non-normal, particularly due to the individual thresholds that participants may display with respect to distraction tolerance; consequently, non-normality of data rules out LISREL as a viable methodology (Chin and Gopal 1995). Second, since PLS allows for multiple measures for each construct, paths among constructs would be more accurate estimates than those obtained through multiple regression. The latter would display downward bias in these estimates due to measurement error (Chin and Gopal 1995; Khalifa and Liu 2002). In addition, PLS offers researchers with the advantage of a holistic view when establishing measures to solve research problems. While confirmatory factor analysis (CFA) could utilize LISREL (Bollen, 1989, Jöreskog and Sörbom, 1993) for its numerous strengths, PLS is better suited for exploratory research PLS and can also be used in CFA (Fornell and Bookstein, 1982, Hulland, 1999). Lastly, another strength of PLS was that it required small to medium sample sizes (Chin 1998; Compeau and Higgins 1998).

Following Kaplan's (2000) recommendation on structural equation modeling analysis, the theoretically grounded proposed research model shown in Figure 3.3 was modified and its parameters re-estimated. This iterative process was performed until the goodness of the model was optimized. The model's goodness of fit was assessed by indices such as the constructs' variance explained and the significance of the model's path coefficients.

In addition to the PLS analysis, independent sample two-tail tests were performed to analyse the objective measures (device data) for efficiency and effectiveness during a post-hoc analysis (see Section 5.10).

5.3 Survey Administration

The entire set of data was collected during the period of August 1 to August 11th, 2005. The pilot study occurred on August 1 and a consistent set of data from the actual experiment was collected between August 5 and 11. The data set from the experiment included 93 cases. Of the 93 questionnaires collected, 87 were fully completed, while the remaining 6 had at least one required response field left unanswered. No pattern was evident regarding the omitted items. Furthermore, given the relatively low sample size required and the very short time frame during which respondents could participate, reporting a response rate would be misleading.

5.4 Participant Demographics

Typical users of mobile devices span the entire spectrum of demographics, including variables of gender, age, and education. By soliciting participants from the entire School of Business (not only students), all of the three variables have been controlled for. For the 93 participants involved in this study, analyses follow in terms of gender, age, and education. It should be noted that participants were screened for proficiency of the English language.

5.4.1 Gender

The sample consisted of 37 males (40%) and 56 females (60%). To test for a significant difference in terms of gender in the sample, a z-test was used. This non-parametric method is best suited given the enumerative data involved in the form of a frequency. With a one-tail p-value of 0.226, the sample is significantly dominated by females.

5.4.2 Age

Study participants were asked to indicate their age. Figure 5.1 reports their age in five-year bins. The sample is heavily represented by participants with ages between 18 and 30 (75% of participants), while the 21-25 age bin was the largest participant group. The mean and mode for ages reported were 28 and 20 respectively.

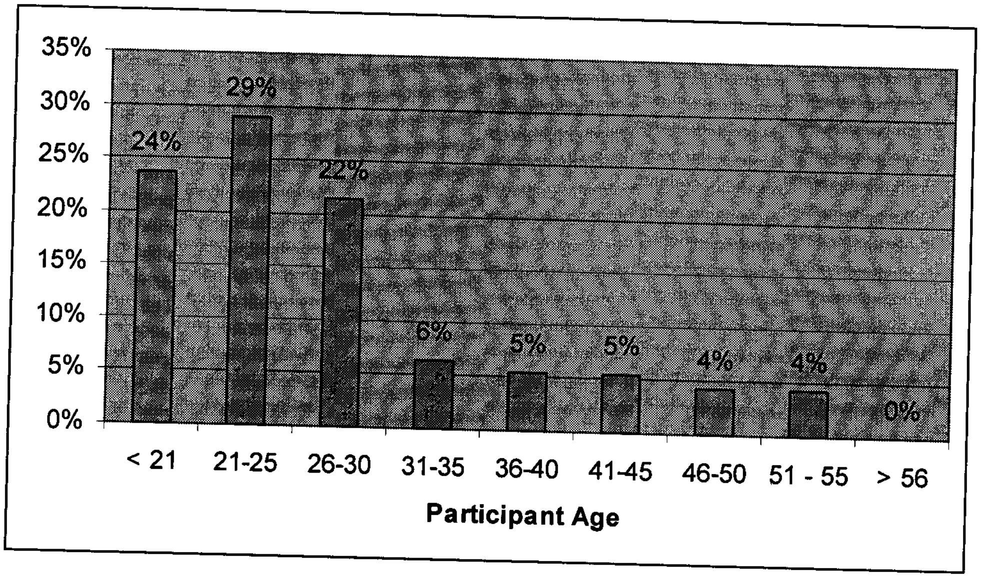


FIGURE 5.1: DISTRIBUTION OF PARTICIPANTS BY AGE

5.4.3 Education

Study participants were asked to indicate their highest level of education completed. Figure 5.2 reports this according to the share of the total sample. The sample is educated with 33% of participants currently pursuing or having completed graduate studies, 15% having completed undergraduate studies, 48% having completed some college/university level studies, and only 3% with high school education only.

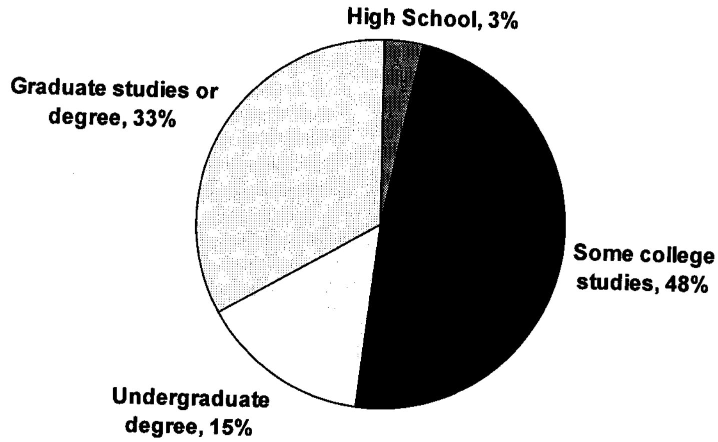


FIGURE 5.2: DISTRIBUTION OF PARTICIPANTS BY EDUCATION

5.5 Accounting for Biases

The use of survey methodology in this dissertation raises issues of bias. Two important considerations that may affect the integrity of the data include the impact of non-response and temporal biases. These biases will be explored through the test of Wilk's Lambda. A multivariate test statistic, the Wilk's Lambda test, indicates whether data group means are different. Ranging from 0 to 1, a value of 0 suggests means are different and 1 suggests that means are the same.

5.5.1 Temporal Bias

Although no follow up invitation was extended to solicit participants, and the entire study period lasted seven days, we tested for the existence of temporal bias by measuring the split-half reliability (McNemar 1969). The data were split into two groups:

first group included the first half of the data collected, while the second group contained the second half of the data collected. With respect to expectations (including disconfirmations of expectations) of both efficiency and effectiveness, usability (i.e. perceived efficiency, perceived effectiveness, satisfaction with efficiency and satisfaction with effectiveness), and adoption (i.e. intention to use), no difference among the data pertaining to two collection phases was found (Wilk's Lambda = 0.659; significance level = 0.732). Also, no statistically significant difference was found in participant-specific traits, namely age, gender, education, and distraction level (Wilk's Lambda = 0.992; significance level = 0.882), between the two halves of the data collected. Thus, no temporal bias exists in the data collected.

5.5.2 Non-response Bias

Study data was then tested for significant difference between those collected from completed questionnaires and those from incomplete questionnaires (Moore, 2000). Missing data can influence the interpretation of the results, and the strength of the confidence in the interpretations (Mohadjer et al. 1994). With respect to expectations (including disconfirmations of expectations) of both efficiency and effectiveness, usability (i.e. perceived efficiency, perceived effectiveness, satisfaction with efficiency and satisfaction with effectiveness), and adoption (i.e. intention to use), no difference among the data pertaining to the two sets of questionnaires (i.e. complete and incomplete) was found (Wilk's Lambda = 0.754; significance level = 0.978). Also, no statistically significant difference was found in participant-specific traits (Wilk's Lambda = 0.946; significance level = .180). Thus, no non-response bias was present for the data.

5.5.3 Common Method Bias

A potential hazard with using survey methodology is common method bias. This may occur when independent and dependent variables are provided by the same source. There is an even higher risk when participants respond to items that measure both independent and dependent variables within the same survey instrument (Bagozzi et al. 1991; Campbell and Friske 1967; Podsakoff et al. 1984; Podsakoff et al. 2003). To help alleviate some of this risk, participant trait information was collected and controlled for. However, to statistically test for common method bias, the data was rearranged (i.e., paired) so that every participant would provide responses to either the independent or dependent variables only. This way, no single participant would be providing responses to items tapping into both independent and dependent variables. A within-treatment random assignment of binary numbers was used to pair data sets (independent and dependent ones). This resulted in a sample of 46 cases, which met the PLS analysis required sample size threshold of 40 (i.e. ten times the number of items of the most complex construct).

The correlation of factor scores were then compared to see if a significant difference existed between the two data sets (i.e. full and half sample). The results in Table 5.1 show (through visual inspection) that there is minimal difference between correlations of factor scores using the total data set and the correlation of factor scores when participant data is paired. Thus, common method bias was not present in this study.

TABLE 5.1: COMMON METHOD BIAS

CONSTRUCT	CORRELATION WITH BI (n = 93) [1]	CORRELATION WITH BI (n = 46) [2]	ABSOLUTE DIFFERENCE
TLX	- 0.448	- 0.504	0.056
ExpEffi	0.335	0.360	0.025
ExpEffe	0.368	0.425	0.057
PerEffi	0.564	0.590	0.036
DisEffi	0.428	0.430	0.002
SatEffi	0.569	0.578	0.009
PerEffe	0.332	0.239	0.093
DisEffe	0.260	0.286	0.026
SatEffe	0.520	0.560	0.040

[1] Correlation of factor scores between exogenous variables (independent and dependent variables: TLX, ExpEffi, ExpEffe, PerEffi, DisEffi, SatEffi, PerEffe, DisEffe, SatEffe) with the right-most endogenous variable (BI) using total data.
[2] Correlation of factor scores using paired data.

5.6 Psychometric Evaluation of Measures

The evaluation of the measures and their corresponding constructs was conducted through various tests, starting with content validity. Content validity refers to the extent that instrument items are representative and comprehensive in measuring the various constructs in the proposed research model and it is assessed by inspecting the process taken to generate these items. Following recommendations by Cronbach (1971) and Kerlinger (1964), content validity stems from generating items from a universal pool. The instrument used in this study measures constructs consisting of items that have been taken from studies exhibiting strong content validity. Slight modifications from the items' original form were made to ensure applicability of the constructs' measures to the domain of study (i.e. usability of mobile devices for wireless data services). These modifications to, and the sources of, the items were described in Section 5.10, *Instrument*. Hence, content validity was satisfied since scales adapted were previously validated.

The factor loadings for the total set of items are summarized in Table 5.2. Shimp and Sharma (1987), Carmines and Zeller (1979), and Hulland (1999) suggest removing items with loading values less than 0.7 to ensure construct validity. Adherence to this procedure required the modification of only one scale (Distraction, measured by TLX) through the removal of two items: TLX5 and TLX6. While upon initial inspection it would appear that TLX2 requires removal as well, once item statistics were recalculated, the loading for TLX2 increased above the threshold. In addition, it should be pointed out that item PerEffi1 was not removed for the following three reasons: i) Its loading of 0.686 is just under the 0.7 threshold; ii) the item measures an important theoretical dimension of efficiency, namely learnability; iii) the same item when adapted in the context of Expectations of Efficiency (i.e. ExpEffi1) and Disconfirmation of expectations of Efficiency (DisEffi1) loaded highly at 0.900 and 0.780 respectively – removing PerEffi1 would also require the removal of those two items. After the removal of the non-valid items, each item was re-validated by testing its item-to-total correlation measure, where all items had higher measures than the 0.35 threshold suggested by Saxe and Weitz (1982). All subsequent analysis is based on the reduced 28-item measurement scale.

TABLE 5.2: ITEM STATISTICS

Item	Mean	Std. Dev	Loading	Error	Item-Total Correlations
TLX1	9.424	4.868	0.820	0.328	0.736
TLX2	6.174	4.439	0.662	0.562	0.526
TLX3	8.217	4.781	0.784	0.386	0.712
TLX4	8.511	4.889	0.860	0.261	0.824
TLX5*	14.903	4.150	-0.623	0.612	-0.279
TLX6*	12.409	5.458	-0.360	0.870	-0.208
ExpEffi1	5.132	1.343	0.898	0.193	0.717
ExpEffi2	5.198	1.310	0.912	0.169	0.781
ExpEffi3	5.231	1.309	0.796	0.367	0.695
ExpEffi4	5.077	1.408	0.710	0.496	0.652
ExpEffe	5.258	1.436	1.000		
PerEffi1	5.859	1.210	0.686	0.529	0.553
PerEffi2	5.457	1.500	0.907	0.178	0.800
PerEffi3	5.261	1.474	0.898	0.194	0.773
PerEffi4	5.109	1.544	0.768	0.411	0.619
DisEffi1	5.172	1.265	0.780	0.392	0.633
DisEffi2	4.774	1.461	0.860	0.261	0.719
DisEffi3	4.839	1.424	0.887	0.213	0.769
DisEffi4	4.570	1.611	0.809	0.345	0.669
SatEffi1	5.118	1.436	0.894	0.202	0.800
SatEffi2	5.172	1.316	0.931	0.134	0.863
SatEffi3	5.065	1.309	0.878	0.229	0.795
SatEffi4	4.849	1.713	0.886	0.215	0.791
PerEffe	5.581	1.690	1.000		
DisEffe	5.054	1.432	1.000		
SatEffe1	5.380	1.248	0.943	0.110	0.898
SatEffe2	5.337	1.369	0.959	0.081	0.922
SatEffe3	5.283	1.225	0.931	0.133	0.883
SatEffe4	5.228	1.446	0.925	0.144	0.862
BI1	5.451	1.648	0.976	0.048	0.904
BI2	5.451	1.688	0.976	0.048	0.904

Note: * denotes items removed from the subsequent analysis

To evaluate the discriminant validity of measures, one compares the loading of an item with its associated factor (i.e. construct) to its cross-loadings. The resulting matrix of loadings and cross-loadings used to test discriminant validity is shown in Table 5.3. For example, item TLX1 has a loading of 0.872 with its associated factor TLX. The value of 0.872 is higher than all of the remaining cross-loadings in that row. All remaining items had higher loadings with their corresponding factors in comparison to their cross-loadings. The results of this test show that there is confidence in the discriminant validity of the remaining measures and their corresponding constructs.

TABLE 5.3: MATRIX OF LOADINGS AND CROSS-LOADINGS

ITEM	TLX	ExpEffi	ExpEffe	PerEffi	DisEffi	SatEffi	PerEffe	DisEffe	SatEffe	BI
TLX1	**0.872**	-0.079	-0.057	-0.463	-0.264	-0.411	-0.281	-0.107	-0.289	-0.146
TLX2	**0.724**	-0.034	-0.044	-0.443	-0.374	-0.396	-0.199	-0.179	-0.212	-0.105
TLX3	**0.823**	-0.027	-0.026	-0.374	-0.255	-0.394	-0.157	-0.010	-0.253	-0.069
TLX4	**0.918**	-0.120	-0.021	-0.468	-0.309	-0.471	-0.270	-0.097	-0.291	-0.173
ExpEffi1	-0.087	**0.902**	0.747	0.249	0.088	0.212	0.212	0.083	0.109	0.138
ExpEffi2	-0.049	**0.912**	0.620	0.169	0.002	0.116	0.126	0.075	0.073	0.096
ExpEffi3	-0.099	**0.798**	0.499	0.260	0.151	0.255	0.086	0.083	0.050	0.051
ExpEffi4	-0.026	**0.710**	0.334	0.049	-0.125	0.015	-0.084	-0.036	-0.152	-0.035
ExpEffe	-0.045	0.688	**1.000**	0.366	0.078	0.263	0.224	0.056	0.124	0.099
PerEffi1	-0.236	0.293	0.522	**0.681**	0.327	0.411	0.319	0.202	0.277	0.171
PerEffi2	-0.522	0.248	0.383	**0.907**	0.505	0.658	0.420	0.258	0.428	0.228
PerEffi3	-0.480	0.188	0.254	**0.901**	0.623	0.680	0.360	0.298	0.416	0.166
PerEffi4	-0.439	0.083	0.122	**0.769**	0.492	0.569	0.254	0.246	0.383	0.179
DisEffi1	-0.224	0.054	0.059	0.405	**0.779**	0.531	0.334	0.214	0.238	-0.022
DisEffi2	-0.283	0.074	0.101	0.569	**0.860**	0.541	0.366	0.293	0.342	0.115
DisEffi3	-0.383	0.061	0.042	0.593	**0.887**	0.579	0.234	0.310	0.322	0.131
DisEffi4	-0.293	0.017	0.058	0.439	**0.809**	0.540	0.205	0.196	0.264	0.070
SatEffi1	-0.446	0.120	0.175	0.645	0.597	**0.893**	0.321	0.247	0.497	0.145
SatEffi2	-0.488	0.181	0.235	0.694	0.641	**0.930**	0.307	0.279	0.540	0.169
SatEffi3	-0.421	0.186	0.193	0.583	0.545	**0.878**	0.297	0.175	0.365	0.125
SatEffi4	-0.445	0.232	0.334	0.659	0.565	**0.886**	0.410	0.258	0.431	0.195
PerEffe	-0.275	0.138	0.224	0.414	0.340	0.373	**1.000**	0.425	0.422	0.226
DisEffe	-0.221	0.143	0.137	0.501	0.595	0.491	0.453	**1.000**	0.578	0.235
SatEffe1	-0.413	0.119	0.185	0.653	0.611	0.732	0.424	0.535	**0.943**	0.237
SatEffe2	-0.453	0.105	0.230	0.682	0.632	0.800	0.403	0.562	**0.959**	0.259
SatEffe3	-0.510	0.128	0.202	0.664	0.597	0.815	0.397	0.511	**0.931**	0.195
SatEffe4	-0.504	0.093	0.262	0.726	0.637	0.827	0.454	0.562	**0.925**	0.254
BI1	-0.421	0.307	0.349	0.546	0.396	0.545	0.336	0.227	0.487	**0.976**
BI2	-0.452	0.346	0.369	0.554	0.439	0.566	0.312	0.281	0.528	**0.976**

The internal consistency of constructs was tested by using the Fornell and Larcker (1981) measure (see Equation 5.1 where λ is the loading of each item). Results of tests for convergent validity (Bagozzi, 1981), discriminant validity (Bagozzi, 1981; Fornell and Larcker, 1981), as well as values for construct means and Cronbach's alpha can be found in Table 5.4.

$$\frac{\left(\sum \lambda_{yi}\right)^2}{\left(\sum \lambda_{yi}\right)^2 + \sum Var(\varepsilon_i)} \qquad \textbf{Equation 5.1}$$

All constructs had adequate reliability (Carmines and Zeller, 1979) and internal consistency well above the 0.7 threshold as prescribed by Nunnally (1978). Cronbach's alpha values and internal consistency values were as follows: TLX (0.85|0.86), ExpEffi (0.86|0.90), ExpEffe (1|1), PerEffi (0.84|0.89), DisEffi (0.85|0.90), SatEffi (0.92|0.94), PerEffe (1|1), DisEffe (1|1), SatEffe (0.96|0.97), and BI (0.95|0.98). In terms of convergent validity, Fornell and Larcker (1981) support an average variance extracted above 50%. The average variance extracted for each construct was as follows: TLX (70.4%), ExpEffi (69.4%), ExpEffe (100%), PerEffi (67.2%), DisEffi (69.7%), SatEffi (80.5%), PerEffe (100%), DisEffe (100%), SatEffe (88.3%), and BI (95.2%).

In addition to the above tests, convergent validity was tested by through an additional approach. By inspecting the t-test scores for the item loadings (Anderson and Gerbing 1988; Hatcher 1994), it became apparent that all t-values were significant at the 0.001 level. Thus, additional support for the constructs' convergent validity is provided.

In addition to the matrix of loadings and cross-loadings described in Table 5.5, discriminant validity can also be tested through the approach proposed by Fornell and Larcker (1981) as well as Hulland (1999). Fornell and Larcker (1981) suggest that the shared variance (see Equation 5.2 where λ is the loading of each item) between any two constructs should be less than the square root of the variance extracted by either of the individual constructs. In other words, values along the diagonal of the correlation matrix in Table 5.5 must be greater than the corresponding values in each row or column. Since this is the case, discriminant validity can be safely assumed.

$$\frac{\sum \lambda^2_{yi}}{\sum \lambda^2_{yi} + \sum Var(\varepsilon_i)}$$ **Equation 5.2**

TABLE 5.4: CONSTRUCT STATISTICS

Statistics	TLX	ExpEffi	ExpEffe	PerEffi	DisEffi	SatEffi	PerEffe	DisEffe	SatEffe	BI
Arithmetic Means[1] (all items)	9.325	5.159	5.258	5.421	4.839	5.051	5.581	5.054	5.307	5.451
Arithmetic Means[2] (used items)	8.082	5.159	5.258	5.421	4.839	5.051	5.581	5.054	5.307	5.451
Cronbach's[3] α Reliability	0.852	0.864	N/A*	0.844	0.854	0.919	N/A*	N/A*	0.956	0.950
Internal [4] Consistency	0.864	0.900	1.000	0.890	0.902	0.943	1.000	1.000	0.968	0.976
AVE (Convergent Validity[5])	0.704	0.694	1.000	0.672	0.697	0.805	1.000	1.000	0.883	0.952

TABLE 5.5: CORRELATION MATRIX AND DISCRIMINANT VALIDITY ASSESSMENT

Items	TLX	ExpEffi	ExpEffe	PerEffi	DisEffi	SatEffi	PerEffe	DisEffe	SatEffe	BI
TLX	**0.839**[6]									
ExpEffi	-0.084[7]	**0.833**								
ExpEffe	-0.045	0.688	**1.000**							
PerEffi	-0.526	0.239	0.366	**0.820**						
DisEffi	-0.358	0.063	0.078	0.607	**0.835**					
SatEffi	-0.503	0.201	0.263	0.721	0.655	**0.897**				
PerEffe	-0.275	0.138	0.224	0.414	0.340	0.373	**1.000**			
DisEffe	-0.122	0.078	0.056	0.309	0.307	0.269	0.425	**1.000**		
SatEffe	-0.315	0.060	0.124	0.465	0.352	0.513	0.422	0.797	**0.940**	
BI	-0.154	0.097	0.099	0.226	0.093	0.178	0.226	0.551	0.561	**0.976**

* N/A = Not Available for single-item constructs

Arithmetic mean of all items in each construct. Likert-type items are scaled from 1 to 7, except for TLX which is from 1 to 20.

2 Arithmetic mean of items used once items with loadings < 0.70 have been removed.

3 Cronbach's alpha (1951) using all measures above the 0.70 threshold as per Nunnally (1978).

4 Fornell and Larcker (1981) measure of internal consistency greater than 0.70 threshold. See Equation 4.1.

5 Fornell and Larcker (1981) measure of convergent validity greater than 0.50 threshold for the Average Variance Extracted (AVE). See Equation 4.2.

6 Fornell and Larcker (1981) measure of discriminant validity which is the square root of the average variance extracted compared to the construct correlations. Bold values are supposed to be greater than those in corresponding rows and columns.

7 Off-diagonal values are correlations. All correlation values are significant at 0.01 level (2-tailed).

CONSTRUCT DEFINITIONS

TLX – The acronym Task Load Index was used to represent Distractions, because the TLX scale was used to measure the latter.
ExpEffi – Expectations of Efficiency,
ExpEffe – Expectations of Effectiveness,
PerEffi – Perceived Efficiency,
DisEffi – Disconfirmation of expectations of Efficiency,
SatEffi – Satisfaction with Efficiency,
PerEffe – Perceived Effectiveness,
DisEffe – Disconfirmation of expectations of Effectiveness,
SatEffe– Satisfaction with Effectiveness,
BI – Behavioural Intention to Use

CONTROL VARIABLE MEANS

Gender = 60% females
Age = 28 years (mode: 20)
Education = Some college/university studies (48%)

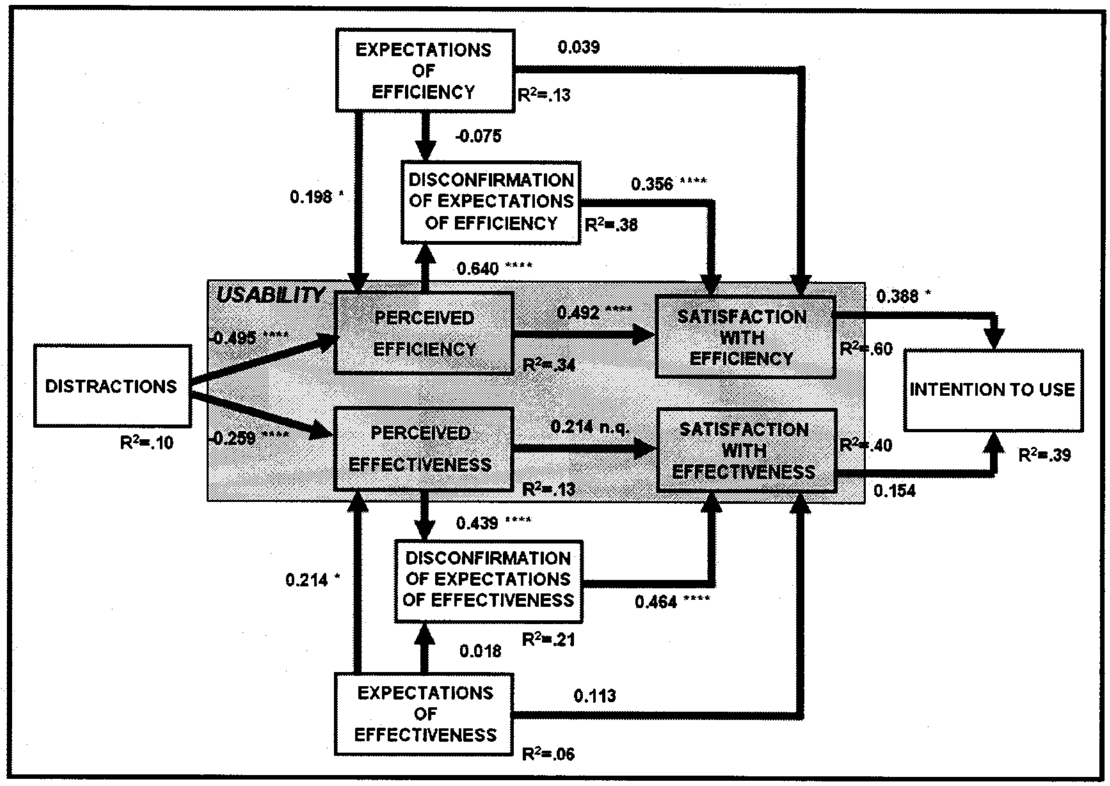
Upon review of the above statistics pertaining to the measurement model, where it appears that all constructs are reliable and display high convergent and discriminant validity, it is safe to assume that a valid data set was collected during this study.

5.7 Validation of Hypotheses

Having established the validity of our measurement instrument used for the data collection, this section presents the data analysis, with its findings seeking to support the proposed hypotheses of this dissertation.

First, a multiple linear regression equation was used to test the significance of the control variables. Three control variables were inserted into the regression that modelled the hypotheses. These included: i) GENDER = gender of participants; ii) AGE = age of participants in years, and iii) EDUCATION = participants' highest level of education completed (options included: high school; some college; undergraduate degree; graduate degree or studies). The three control variables are characterized by greater oscillations when series values are bigger in magnitude than when they are smaller. For this reason, a natural logarithmic transformation was calculated to make variation constant across levels of the series. Since the results of the regression showed that the control variables were insignificant, they were removed from the rest of the analysis. Analysis is performed by testing relationships for moderate significance at the 0.1 level and for strong significance at, or lower than, the 0.5 level. Relationships that fall short, but by a small margin, will be described as "not quite significant" (denoted as "n.q."; may also be found in literature as "marginally significant", Lowry 1999). This technique will allow us to investigate the particular relationship further, instead of disregarding it immediately as not-significant.

The proposed structural model shown in Figure 5.3 was tested by Jackknifing in PLS. This resampling procedure assesses the significance of PLS parameter estimates (Chin 1998, 2001). Jackknifing is just one of several PLS techniques that may be used in evaluation a research model. For example, Bootstrapping is another common PLS approach, but in general, estimations by either one approach should converge (Chin 1998). From the original 16 hypotheses, 10 were supported, two were not quite significant, and four were not supported. Table 5.6 presents the validation of these hypotheses in more detail. The discussion that follows will discuss what the proposed structural model has demonstrated through the majority of its linkages supported. The hypotheses discussed are organized by topic area: Cognition and Performance; Expectancy-Disconfirmation Theory; Usability; and Adoption.



n.q. not quite significant
* significant at 0.1 level
** significant at 0.05 level
*** significant at 0.01 level
**** significant at 0.001 level

FIGURE 5.3: THE PROPOSED STRUCTURAL MODEL

TABLE 5.6: HYPOTHESES VALIDATION

Hypotheses	From	To	Beta	t-Value	p-Value	Sig	Status
H1a	TLX	PerEffi	-0.495	4.230	< 0.001	****	supported
H1b	TLX	PerEffe	-0.259	2.444	< 0.025	**	supported
H2a	ExpEffi	PerEffi	0.198	1.823	< 0.1	*	supported
H2b	ExpEffe	PerEffe	0.214	1.787	< 0.1	*	supported
H3a	ExpEffi	DisEffi	-0.075	0.970	*0.332*		not supported
H3b	ExpEffe	DisEffe	0.018	0.339	*0.735*		not supported
H4a	PerEffi	DisEffi	0.640	9.674	< 0.001	****	supported
H4b	PerEffe	DisEffe	0.439	3.089	< 0.001	****	supported
H5a	DisEffi	SatEffi	0.356	3.520	< 0.001	****	supported
H5b	DisEffe	SatEffe	0.464	4.650	< 0.001	****	supported
H6a	PerEffi	SatEffi	0.492	4.463	< 0.001	****	supported
H6b	PerEffe	SatEffe	0.214	1.493	< 0.2	n.q.	not supported
H7a	ExpEffi	SatEffi	0.039	0.045	*0.964*		not supported
H7b	ExpEffe	SatEffe	0.113	1.229	*0.219*	n.q.	not supported
H8a	SatEffi	BI	0.388	1.945	< 0.1	*	supported
H8b	SatEffe	BI	0.154	0.942	*0.346*		not supported

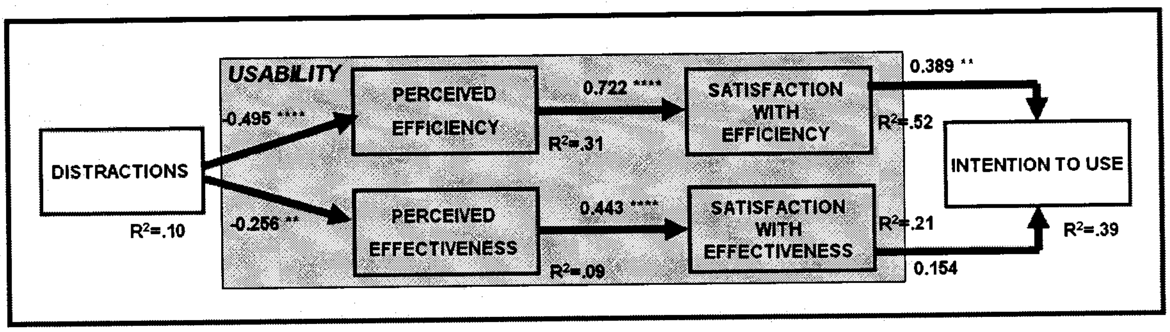
First, it was theorized that incremental cognitive load consequent of auditory/visual/motor stimuli present would detract from the users' performance with the mobile devices for wireless data services. There was strong statistical support for the corresponding hypotheses, H1a and H1b.

Second, the marketing theory of Expectations-Disconfirmation was being applied and tested for its relevance in IS. Overall, most hypotheses were supported, but with five out of 12 links not-supported, a closer examination is required. Upon inspection it becomes evident that there are two types of links that did not demonstrate statistical significance: i) Expectations of performance to Disconfirmation of expectations of performance (i.e. of both efficiency and effectiveness, H3a and H3B, respectively), and ii) Expectations of Performance to Satisfaction with Performance (i.e. of both efficiency and effectiveness, H7a and H7b). This consistency across both performance dimensions mandates the following two actions: first, the removal of those links from the proposed structural model and its re-analysis, and second a discussion on the inapplicability of the entire EDT theory in this study; this discussion will also be provided in the next chapter, "Discussion". On the other hand, the links between Expectations of performance to Perceived performance (H2a and H2b), Perceived performance to Disconfirmation of expectations of performance (H4a and H4b), Disconfirmation of expectations of performance to Satisfaction with performance (H5a and H5b), and Perceived performance to Satisfaction with performance (H6a and H6b) demonstrated statistical significance.

Third, on the topic of usability, both performance and satisfaction were decomposed from the originally single-construct definition into two components of Perceived Efficiency and Perceived Effectiveness. As mentioned above, these theoretical adaptations obtained strong statistical support for Perceived Efficiency (H6a), but not quite significant for Perceived Effectiveness (H6b).

Lastly, adoption of mobile devices was explored by measuring the behavioural intention of users upon obtaining hands-on experience with wireless data services. From the two hypotheses proposed, only Satisfaction with Efficiency (H8a) was shown to be statistically significant in impacting the aforementioned behavioural intention. The other link between Satisfaction with Effectiveness and Behavioural Intention (H8b) will be re-examined after the model has been re-run with the reduced set of hypotheses.

Following Kaplan's (2000) recommendation on structural equation modeling analysis, the theoretically grounded research model will next be modified and its parameters re-estimated in pursuit of an optimized model fit. Accordingly, the model was re-run in PLS in its basic form by removing the Expectations and Disconfirmation of expectations constructs (see Figure 5.4). This was done to test the strength of the fundamental relationships, namely those of usability, with respect to distractions and adoption. The PLS results are provided in Table 5.7, where the last column identifies the change in the t-values. By inspection, further, we observe that there were four important changes: i) both of the performance-related usability paths showed strong significance (i.e. the path from perceived effectiveness to satisfaction with effectiveness changed from non-significant to showing strong significance); ii) there was greater confidence for the effect of distractions on perceived efficiency than on perceived effectiveness; iii) the path between Satisfaction with Efficiency and Intention to Use changed from showing moderate to showing strong significance, the path of Satisfaction with Effectiveness to Intention To Use remained non-significant; and iv) there was a decrease of the variance explained among the usability-related constructs.



** significant at 0.05 level
*** significant at 0.01 level
**** significant at 0.001 level

FIGURE 5.4. THE BASIC STRUCTURAL MODEL (I.E. EDT WAS NOT INCLUDED)

TABLE 5.7. HYPOTHESES VALIDATION FOR BASIC STRUCTURAL MODEL

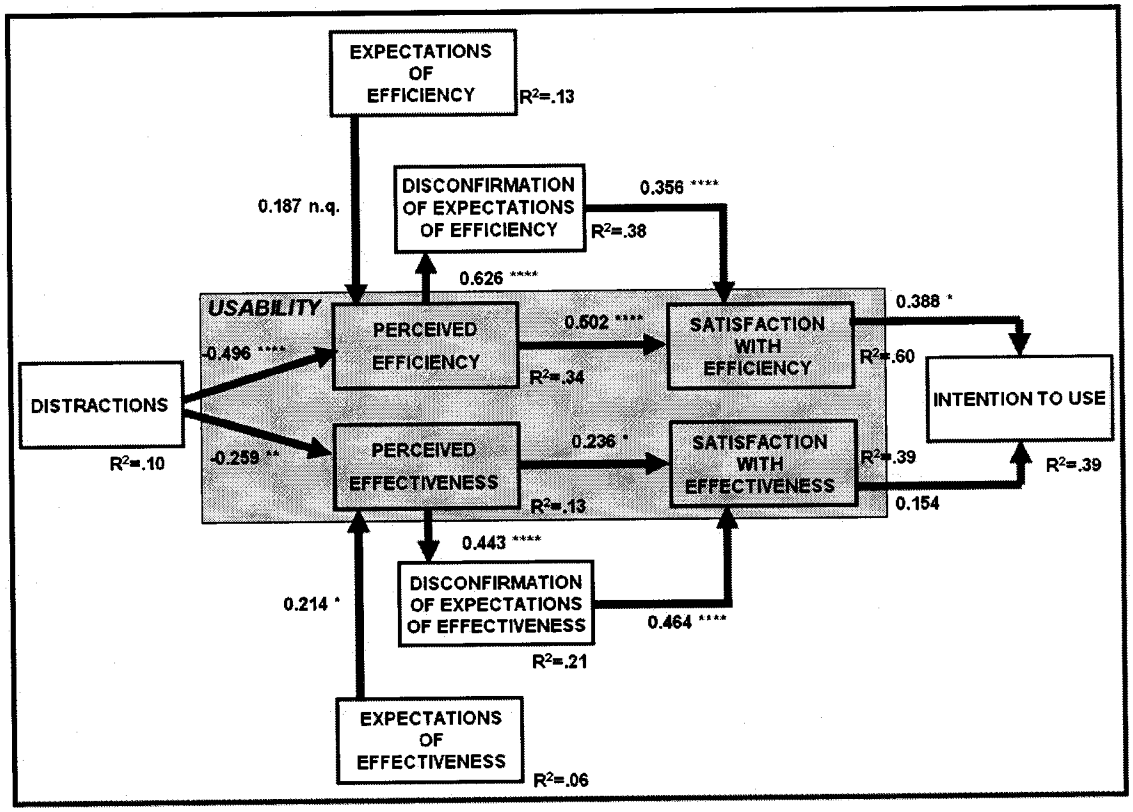
Hypotheses	From	To	Beta	t-Value	p-Value	Sig	Status	t-Val. Change[1]
H1a	TLX	PerEffi	-0.495	4.519	< 0.001	****	Supported	0.290
H1b	TLX	PerEffe	-0.256	2.058	< 0.05	**	Supported	-0.386
H6a	PerEffi	SatEffi	0.722	9.754	< 0.001	****	Supported	5.291
H6b	PerEffe	SatEffe	0.443	4.855	< 0.001	****	Supported	3.362
H8a	SatEffi	BI	0.389	2.175	< 0.05	**	Supported	0.230
H8b	SatEffe	BI	0.154	1.072	*0.346*		not supported	0.013

[1]Change in the t-value relative to the proposed structural model tested, shown in Figure 5.3

The basic model appears to have an improved fit in terms of the significance of the model's path coefficients (5 out of 6 are now significant, compared to only 10 out of 16 previously).

Again, and in line with Kaplan's (2000) recommendation, the model was re-run in PLS after having incorporated the EDT-related constructs, but also having removed the paths that were not supported in the first analysis earlier above. Specifically, the two paths from Expectations to Disconfirmation of expectations and the two paths from

Expectations to Satisfaction were removed (see Figure 5.5). This iteration was performed to assess whether the role of EDT in mobile usability is significant, but perhaps in a modified (compared to the full EDT model) way. The results are provided in Table 5.8, where the last column identifies the change in the t-values. Relative to the complete application of EDT, we observe that there were three important changes: i) the path from Expectations of Efficiency to Perceived Efficiency (H2a) dropped to not being quite significant, from being moderately significant at the 0.1 level before; ii) the path from Perceived Effectiveness to Satisfaction with Effectiveness (H6b) became moderately significant at the 0.1 level, from being not quite significant; iii) the path from Disconfirmation of expectations of Effectiveness to Satisfaction with Effectiveness (H5b) became less supported by showing moderate significance at the 0.1 level, from being significant at the 0.001 level before. Overall, it would appear that applying EDT and having removed the non-significant links (from the first run, i.e. the proposed structural model of Figure 5.3 earlier) adds value as many of the relationships were significant. This is further supported by the observation that the basic structural model of Figure 5.4 tested above did not change (i.e. five of the six hypotheses were supported). It should be pointed out that there were no significant changes in the variances explained for any of the constructs between the proposed (i.e. Figure 5.3) and the basic (i.e. Figure 5.4) version of the structural model.



n.q. not quite significant
* significant at 0.1 level
** significant at 0.05 level
*** significant at 0.01 level
**** significant at 0.001 level

FIGURE 5.5: THE REVISED STRUCTURAL MODEL (RE-RUN WITH SUPPORTED HYPOTHESES ONLY)

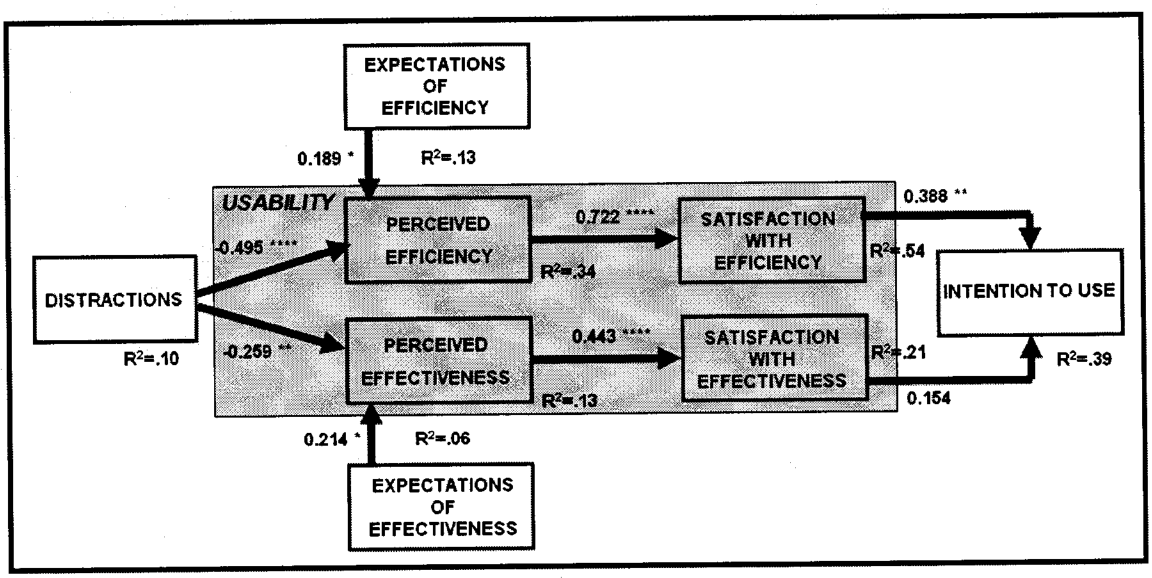
TABLE 5.8: HYPOTHESES VALIDATION FOR REVISED STRUCTURAL MODEL

Hypotheses	From	To	Beta	t-Value	p-Value	Sig	Status	t-Val. Change[1]
H1a	TLX	PerEffi	-0.496	3.990	< 0.001	****	Supported	-0.240
H1b	TLX	PerEffe	-0.259	2.444	< 0.025	**	Supported	0.000
H2a	ExpEffi	PerEffi	0.187	1.492	< 0.2	n.q.	Supported	-0.331
H2b	ExpEffe	PerEffe	0.214	1.787	< 0.1	*	Supported	0.000
H4a	PerEffi	DisEffi	0.626	9.012	< 0.001	****	Supported	-0.662
H4b	PerEffe	DisEffe	0.443	3.582	< 0.001	****	Supported	0.493
H5a	DisEffi	SatEffi	0.352	4.054	< 0.001	****	Supported	0.535
H5b	DisEffe	SatEffe	0.466	1.810	< 0.1	*	Supported	-2.840
H6a	PerEffi	SatEffi	0.502	5.254	< 0.001	****	Supported	0.790
H6b	PerEffe	SatEffe	0.236	1.923	< 0.1	*	Supported	0.430
H8a	SatEffi	BI	0.388	4.262	< 0.001	****	Supported	2.317
H8b	SatEffe	BI	0.154	1.001	*0.346*		not supported	0.059

[1]Change in the t-value relative to the proposed structural model tested shown in Figure 5.3

To further test the model, two variations were analysed: first, a simplified version removing the Disconfirmation of expectations constructs but maintaining the Expectations ones; second, a saturated model where all possible links were included. The discussion continues with the simplified version first.

First, the model was re-tested by removing the Disconfirmation of expectations constructs. This was done in part following Kaplan's (2000) recommendation for iterative data analyses in pursuit of the desired model fit, but also because Disconfirmation of expectations has received mixed support in past studies. After analysing this simplified model, shown in Figure 5.6, in PLS four changes arise in the proposed links (see Table 5.9): i) the Expectations of Efficiency to Perceived Efficiency link (H2a) changed from being not quite significant to being moderately significant at the 0.1 level; ii) the Perceived Effectiveness to Satisfaction with Effectiveness link (H6b) changed from being not quite significant to showing strong significance at the 0.001 level; iii) the Satisfaction of Efficiency to Behavioural Intention (H8a) changed from being moderately significant at the 0.1 level to having strong significance at the 0.05 level. It should be noted that the variance explained in the simplified model changed for two constructs: i) Satisfaction with Efficiency increased from 0.60 to 0.64; ii) Satisfaction with Effectiveness decreased from 0.39 to 0.21. Overall, the simplified model has a higher predictive power than the previous two versions in terms of the ratio between supported and proposed hypotheses (i.e. improved from having moderate or strong support for 10 out of 16 hypotheses to having such support for 7 out of 8 hypotheses).



* significant at 0.1 level
** significant at 0.05 level
*** significant at 0.01 level
**** significant at 0.001 level

FIGURE 5.6: THE SIMPLIFIED STRUCTURAL MODEL

TABLE 5.9: HYPOTHESES VALIDATION FOR SIMPLIFIED STRUCTURAL MODEL

Hypotheses	From	To	Beta	t-Value	p-Value	Sig	Status	t-Val. Change
H1a	TLX	PerEffi	-0.496	4.459	< 0.001	****	Supported	0.470
H1b	TLX	PerEffe	-0.259	2.444	< 0.025	**	Supported	0.000
H2a	ExpEffi	PerEffi	0.189	1.813	< 0.1	*	Supported	0.322
H2b	ExpEffe	PerEffe	0.214	1.787	< 0.1	*	Supported	0.000
H6a	PerEffi	SatEffi	0.722	10.142	< 0.001	****	Supported	4.889
H6b	PerEffe	SatEffe	0.443	4.855	< 0.001	****	Supported	2.932
H8a	SatEffi	BI	0.389	2.175	< 0.05	**	Supported	0.252
H8b	SatEffe	BI	0.154	1.072	*0.346*		not supported	0.071

Lastly, the saturated model was tested (not shown here due to excessive links). The need to explore all linkages arises from the novel application of this model in its entirety in IS and that the wireless data services studied present a growing area of research lacking an extensive theoretical base. The saturated model contains a total of 31 relationships, of which 15 are new compared to previously presented models. These new links are described in Table 5.10 and the previously hypothesized and validated links in Table 5.11. Paths that were shown to be non-significant (i.e. p-value <0.1) were not reported ("n.r.").

TABLE 5.10: HYPOTHESES VALIDATION FOR SATURATED MODEL – NEW RELATIONSHIPS

From	To	Beta	t-Value	p-Value	Sig	Status
TLX	ExpEffi	0.013	0.085	n.r		Rejected
TLX	DisEffi	-0.081	0.595	n.r		Rejected
TLX	SatEffi	-0.153	0.851	n.r		Rejected
TLX	ExpEffe	0.013	0.287	n.r		Rejected
TLX	DisEffe	-0.096	0.480	n.r		Rejected
TLX	SatEffe	-0.125	1.711	< 0.1	*	Supported
TLX	BI	-0.160	1.005	n.r		Rejected
ExpEffi	BI	0.158	1.177	n.r		Rejected
ExpEffe	BI	0.078	0.149	n.r		Rejected
PerEffi	PerEffe	0.341	2.107	< 0.05	**	Supported
PerEffi	BI	0.140	0.816	n.r		Rejected
PerEffe	BI	0.096	1.054	n.r		Rejected
DisEffi	BI	0.128	1.005	n.r		Rejected
DisEffe	BI	-0.157	1.313	n.q.		Rejected
SatEffi	SatEffe	0.055	7.1558	< 0.001	****	Supported

TABLE 5.11: HYPOTHESES VALIDATION FOR SATURATED MODEL – HYPOTHESIZED RELATIONSHIPS

			Non-Saturated Model			Saturated Model			
Hypotheses	From	To	Beta	t-Value	Status	Beta	t-Value	Δ Beta	Status
H1a	TLX	PerEffi	-0.495	4.230	supported	-0.491	4.139	-0.004	Supported
H1b	TLX	PerEffe	-0.259	2.444	supported	-0.089	0.793	-0.170	*not supported*
H2a	ExpEffi	PerEffi	0.198	1.823	supported	0.202	2.027	0.004	Supported
H2b	ExpEffe	PerEffe	0.214	1.787	supported	0.101	0.611	-0.113	*not supported*
H4a	PerEffi	DisEffi	0.640	9.674	supported	0.598	6.139	-0.042	Supported
H4b	PerEffe	DisEffe	0.439	3.089	supported	0.415	2.780	-0.024	Supported
H5a	DisEffi	SatEffi	0.356	3.520	supported	0.346	3.308	-0.010	Supported
H5b	DisEffe	SatEffe	0.464	4.650	supported	0.189	2.110	-0.275	Supported
H6a	PerEffi	SatEffi	0.492	4.463	supported	0.421	2.751	-0.071	Supported
H6b	PerEffe	SatEffe	0.214	1.493	not supported	0.091	0.692	-0.123	not supported
H8	SatEffi	BI	0.388	1.945	supported	0.219	0.823	-0.169	*not supported*
H8b	SatEffe	BI	0.154	0.942	not supported	0.055	0.180	-0.099	not supported

Based on the above analysis, three observations can be made: i) Out of 15 newly proposed hypotheses only three show significance, from which only two show strong significance; ii) The two paths with strong significance are both directed from the usability-related constructs of Efficiency to those of Effectiveness (i.e. from Perceived Efficiency to Perceived Effectiveness and from Satisfaction with Efficiency to Satisfaction with Effectiveness); iii) there is moderate significance for the path between Distractions and Satisfaction with Effectiveness. From these observations two conclusions can be drawn. First, there is an apparent relationship between Efficiency and Effectiveness, whereby the former impacts the latter. In other words, the more efficient users perceive a mobile device to be for wireless data services, the more effective (and arguably useful) they perceive that device to be for the intended purpose. Second, the negative (and weak) relationship between Distractions and Satisfaction with Effectiveness is likely due to chance, given the lack of theoretical support. However, one possible explanation would be that users associate dampened effectiveness consequent to distractions not to those stimuli but to the device itself.

In the saturated model, three of the previously validated hypothesized relationships lost their statistical significance, while another two demonstrated a lower level of significance (only one path demonstrated slightly improved significance). Having three paths rejected in this saturated model does not mean that they are not valid. Given the high number of additional paths introduced in the saturated model, PLS attempted to assign factor scores with the end goal of optimizing the variance explained by all the constructs and proposed links. In this process, part of the variance explained for the theoretically supported constructs and paths was redistributed to optimize for total variance. Therefore, when the path from Satisfaction with Efficiency to Satisfaction with Effectiveness was introduced a portion of the significance shifted from the previously single link between the former and Behavioural Intention. Similarly, when the Distraction to Satisfaction with Effectiveness was introduced, the link between the former and Perceived Effectiveness was negatively impacted.

Shifting the discussion back to the proposed structural model, the variance explained is reviewed next. Overall, the model demonstrated high explanatory power. The R-square of the Behavioural Intention construct was 0.39, or 39% of the variance in user intentions to adopt mobile devices for wireless data services. It should be noted that the R-square values for Distractions, Expectations of Efficiency, and Expectations of Effectiveness were relatively small (i.e. 0.10, 0.13, and 0.06 respectively). These values do not necessarily pose a threat to the model's validity. Particularly in behavioural science research low R-square values are common and often the amount of actual association between constructs is higher than the variance accounted for by R-square (Cohen 1988, p. 532-535). Low R-square values have also been reported in many technology adoption studies (Davis et al. 1989; Davis 1993; Chau and Hu 2002; Moon and Kim 2001). An additional explanation for the low R-square values in this model may be that these constructs were associated with only a single construct (TLX). Relative to multi-relationship models, these single or few-relationship associations often provide low R-square values (Nunnaly 1978). In the event that additional non-correlating constructs are introduced as antecedents to these constructs with low R-square values (i.e. ExpEffi and ExpEffe), the scores would increase considerably.

In summary, the research model was analysed in five variations:

i) The proposed model that included all of the EDT constructs and paths;

ii) A basic distractions-usability-adoption model (see Figure 5.4) that served as a benchmark for assessing the core relationships in the absence of EDT;

iii) A revised model (see Figure 5.5) that included only the supported paths from the proposed model (see Figure 5.3);

iv) A simplified model (see Figure 5.6) that included Expectations but excluded Disconfirmations of expectations; and

v) A saturated model (Figure not shown) to test for new relationships not in the originally proposed model (see Figure 5.3).

From the findings described above, the fourth variation (i.e. including Expectations but excluding Disconfirmations of expectations) appears to offer the best fit for the following three reasons. First, this simplified model shown in Figure 5.6 did not reduce the observed significance between distractions, usability, and adoption. Second, this simplified model was selected because it supported the highest ratio in terms of significant to non-significant paths (seven of eight) relative to the other possible models. Finally, the selected simplified model of Figure 5.6 did not lose considerable variance explained (i.e. six and eight percent respectively) for the two satisfaction constructs, as a result of removing the two Disconfirmation of expectations constructs and the corresponding paths.

5.8 Effect Size

One way to assess the quality and predictive power of a model is to perform an analysis of R-square values of the most endogenous construct, here Behavioural Intention. Since the interpretation of these values is like that of linear regression, Chin's (1998) formula can be applied, where the effect size "f" is obtained as shown in equation 5.3:

$$f^2 = \frac{R^2_{\text{included}} - R^2_{\text{excluded}}}{1 - R^2_{\text{included}}} \qquad \textbf{Equation 5.3}$$

Interpreting the formula, the effect size of each independent construct is directly related on the R-square extracted when the given construct is included in the model and negatively affected by the R-square extracted when this construct is removed from the model. Adhering to this formula, the R-square and corresponding f value were determined in sequence for all of the antecedents of Behavioural Intention. The results, based on the selected simplified model shown in Figure 5.6, are shown in Table 12 and the following values may be used as reference points: 0.02 (small), 0.15 (medium), and 0.35 (large) (Cohen 1988). By inspection, we observe that Satisfaction with Efficiency had a medium-to-large effect, while we proved that the impact of Satisfaction with Effectiveness on Behavioural Intention is not significant.

TABLE 5.12: EFFECT SIZES ON BEHAVIOURAL INTENTION

$R^2_{included} = 0.393$	SatEffi	SatEffe
$R^2_{excluded}$	0.352	0.386
F^2	0.27	0.05
Effect	medium-large	Small

A similar approach was utilized in analysing the effect of antecedents on Satisfaction with both Efficiency and Effectiveness. Table 5.13 shows that the effects of Perceived Efficiency and Disconfirmation of expectations of Efficiency on Satisfaction with Efficiency are significant (i.e. large and medium respectively). Table 5.14 shows that the effects of Perceived Effectiveness and Disconfirmation of expectations of Effectiveness on Satisfaction with Effectiveness are significant (i.e. low and medium respectively). A plausible explanation for the difference in the effect size of perceptions on satisfaction (between those for efficiency and effectiveness) could be due to the reduced variance explained for effectiveness consistently across all related constructs. It is foreseeable that if the measurement of effectiveness improved through an enhanced scale that would capture additional dimensions, then the effect would also increase. It is also interesting to see that expectations of both efficiency and effectiveness had a low effect on satisfaction with each of them respectively.

TABLE 5.13: EFFECT SIZES ON SATISFACTION WITH EFFICIENCY

$R^2_{included} = 0.604$	ExpEffi	PerEffi	DisEffi
$R^2_{excluded}$	0.602	0.468	0.526
F^2	0.01	0.34	0.20
Effect	very low	large	Medium

TABLE 5.14: EFFECT SIZES ON SATISFACTION WITH EFFECTIVENESS

$R^2_{included} = 0.397$	ExpEffe	PerEffe	DisEffe
$R^2_{excluded}$	0.385	0.362	0.227
F^2	0.02	0.06	0.28
Effect	low	low	medium

Lastly, examining the effect of Distractions (TLX) on perceived performance, Table 5.15 shows that the effects of the former are significant in both cases (medium to large for Perceived Efficiency; small to medium for Perceived Effectiveness). Again, in the event that Effectiveness is captured better, it is likely that the effect of Distractions will also increase.

TABLE 5.15: EFFECT OF DISTRACTIONS ON PERCEIVED PERFORMANCE (EFFICIENCY & EFFECTIVENESS)

$R^2_{included} = 0.305$	TLX on PerEffi	$R^2_{included} = 0.133$	TLX on PerEffe
$R^2_{excluded}$	0.09	$R^2_{excluded}$	0.072
f^2	0.31	F^2	0.07
Effect	medium-large	**Effect**	small-medium

5.9 Control Variables

Study participants were asked to provide information regarding their gender, age, and education. Earlier, in Sections 5.4.1 through 5.4.3, it was reported that 60% were female, the mean and mode for ages reported were 28 and 20 respectively, and lastly 97% had at least some college/university level education (half of who had at least completed these and/or higher level studies). This personal information can be used as control variables: i) Gender (0 – male, 1 – female); ii) Age (absolute values, range: 18-55); iii) Education (0 – high school, 1 – some college/university studies, 2 – undergraduate degree, 3 – graduate studies/degree). These control variables were analysed by running the model excluding them (uncontrolled), including them one at a time, and lastly including them all at the same time (controlled) in PLS. The results are shown in Table 5.16. Chin (1998, p.316) suggests that any changes R-square values may indicate substantial impact of an independent construct on dependents one(s). By inspection, we observe the following:

i) Education has the least influence over the dependent constructs, with R-square values increasing slightly for Perceived Efficiency (PerEffi) and Satisfaction with Effectiveness (SatEffe), and moderately for Expectations of Effectiveness (ExpEffe) and Behavioural Intention (BI);

ii) Age had little impact over Perceived Effectiveness (PerEffe), and Behavioural Intention (BI), and considerable impact on Distractions (TLX), and Expectations of Efficiency (ExpEffi);

iii) Gender had little impact over Expectations of Efficiency (ExpEffi), Perceived Efficiency (PerEffi), and Satisfaction with Effectiveness (SatEffe), and considerable impact over Distraction (TLX), Expectations of Effectiveness (ExpEffe), and Behavioural Intention (BI). Overall, the fully controlled model improves the R-square values for all dependent constructs except for Satisfaction with Efficiency (SatEffi), which remained unchanged.

TABLE 5.16: IMPACT OF CONTROL VARIABLES ON R-SQUARE VALUES OF DEPENDENT CONSTRUCTS

	TLX	ExpEffi	ExpEffe	PerEffi	PerEffe	SatEffi	SatEffe	BI
Controlled Model	0.10	0.13	0.06	0.34	0.13	0.52	0.21	0.39
Gender	0.05	0.01	0.04	0.34	0.12	0.52	0.21	0.36
Age	0.04	0.10	0.00	0.32	0.13	0.52	0.20	0.34
Education	0.00	0.00	0.03	0.33	0.12	0.52	0.21	0.36
Uncontrolled Model	0.00	0.00	0.00	0.32	0.12	0.52	0.20	0.33

Next, the path coefficients (Beta) and significance levels (t/p-values) between the control variables and the dependent constructs were reviewed and presented in Table 5.17. Strong beta coefficients and corresponding t-values indicated that women are affected more by distractions than men, while the negative beta coefficients between Gender and ExpEffe suggest that women had higher Expectations of Effectiveness from the mobile devices for wireless data services than men did. Age had strong negative path coefficient and corresponding t-value for Expectations of Efficiency. This would suggest that older participants have higher expectations from mobile devices for the intended wireless data services than younger ones.

TABLE 5.17: PATH SIGNIFICANCE BETWEEN CONTROL VARIABLES AND DEPENDENT CONSTRUCTS

		TLX	ExpEffi	ExpEffe	PerEffi	SatEffi	PerEffe	SatEffe	BI
Gender	Beta	0.237	-0.109	-0.175	-0.113	0.041	0.059	-0.096	-0.128
	t-Value	2.591	0.885	1.983	0.8944	0.770	0.543	1.347	1.213
	p-Value <	0.05	n.r	0.05	n.r	n.r	n.r	n.r	n.r.
	Validation	sign.	n.s.	sign.	n.s	n.s.	n.s.	n.s.	n.s.
Age	Beta	0.216	-0.353	-0.099	0.038	-0.065	-0.096	0.038	-0.144
	t-Value	1.632	3.150	0.900	0.164	0.389	0.751	0.322	0.775
	p-Value <	n.r.	0.01	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.
	Validation	n.s.	sign.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
Education	Beta	0.026	0.088	0.137	0.066	0.019	0.035	0.048	0.177
	t-Value	0.343	0.452	1.224	0.291	0.627	0.294	0.864	1.239
	p-Value <	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r
	Validation	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.

5.10 Post-Hoc Analysis

In addition to the self-reported data collected through the use of the questionnaire, objective (i.e. device) performance data were collected through a software application that was running in the background of the mobile device. This application collected the following data during each of the assigned tasks (i.e. e-mail, calendar, address book, wireless Web):

- Task completion time (seconds) – "*Time*"
- User keystroke activity (key presses / characters) – "*Keys*"
- Navigation wheel rolls (rolls) – "*Rolls*"
- Navigation wheel clicks (clicks) – "*Clicks*", and
- Words Per Minute (ratio of Task completion time to User keystroke activity) – *WPM*

Furthermore, the investigator noted for each wireless Web task its completion status by inspecting the accuracy of the retrieved information. These data were analysed to achieve the following three objectives:

1) Determine the impact of context of use on user performance with mobile devices for wireless data services, by varying the context (in this case user motion and environmental noise) and investigating for any significant differences between treatments. Specifically, the data were analysed to:

a) Assess the impact of user motion on the objective efficiency and objective effectiveness of mobile devices for wireless data services,

b) Assess the impact of environmental noise on the objective efficiency and objective effectiveness of mobile devices for wireless data services,

c) Assess the impact of the interaction effect of user motion and environmental noise on the objective efficiency and objective effectiveness of mobile devices for wireless data services

2) Determine if any relationships existed between self-reported and device data;

In the interest of space and the vast amount of data collected, only those of statistical significance will be reported here.

To address the first objective, we revisit the relevant theory on cognition and user performance. It has been shown that the higher the cognitive load resulting from distractions in the environment and/or user motion, the lower the performance will be (Ljungberg's et al. 2004; Banbury et al. 2001; Hughes and Jones 2003). This argument was tested for each of the four goals described under the first objective above. Efficiency was measured in terms of task completion time (Time) and the words per minute (WPM). Effectiveness was measured in terms of accuracy and successful task completion rates. In obtaining support for each goal, the methods used were ANOVA and Tukey's test for post-hoc pair-wise comparisons between the treatment pairs for each of the four wireless data services of this study. Thus:

a) *Treatment 1 (seated in isolation) was expected to demonstrate: i) faster times and higher WPM, and ii) higher accuracy and higher successful task completion rates, than treatment 2 (walking in isolation).*

There were no significant differences in any of the objective data analysed. Thus, the impact of motion was not significant for both the objective efficiency and the objective effectiveness of a mobile device for wireless data services.

b) *Treatment 1 (seated in isolation) was expected to demonstrate: i) faster times and higher WPM, and ii) higher accuracy and higher successful task completion rates, than treatment 3 (seated and exposed to auditory and visual stimuli in the environment).*

There were no significant differences in any of the objective data analysed. Thus, the impact of environmental noise (i.e. auditory and visual stimuli) was not significant for both the efficiency and the effectiveness of a mobile device for wireless data services.

c) *Treatment 1 (seated in isolation) was expected to demonstrate: i) faster times and higher WPM, and ii) higher accuracy and higher successful task completion rates, than treatment 4 (walking and exposed to auditory and visual stimuli in the environment).*

For this pair-wise comparison, analysis of variance results are reported in Table 5.18, which includes all variables that showed statistically significant differences between treatments (to a moderate significance at the 0.1 level or higher significance); these were:

i) SMS task completion time and WPM for each of the four messages sent

ii) SMS mean time and mean WPM for all four messages (see Figure 5.7)

The above variables were examined closer through Tukey's test. The means that are significantly different across treatments are reported in Table 5.19. Time and WPM are significantly different between groups 1 and 4 (except for the WPM of the third text message sent, *sms3wpm*; this difference was moderately significant with a p-value of .06), but Keys are not. This suggests that the combined effect of auditory/visual stimuli and user motion impacts the duration of the task. Text entry accuracy and task completion rates among treatments were not affected by the stimuli. One possible explanation for this could be that the users' increased effort for accurate key entry and resulted in the increased time for each key entry and the prolonged duration of the task.

TABLE 5.18: ANALYSIS OF VARIANCE FOR DEVICE DATA

ANOVA						
Dependent Variable		**Sum of Squares**	**Df**	**Mean Square**	**F**	**Sig.**
sms1time	Between Groups	2025.943	3	675.314	2.051	0.114
	Within Groups	23701.027	72	329.181		
	Total	25726.969	75			
sms1wpm[1]	Between Groups	164.258	3	54.753	3.451	0.021
	Within Groups	1142.215	72	15.864		
	Total	1306.473	75			
sms2time	Between Groups	479.717	3	159.906	2.119	0.105
	Within Groups	5432.823	72	75.456		
	Total	5912.540	75			
sms2roll	Between Groups	3.446	3	1.149	3.363	0.023
	Within Groups	24.593	72	0.342		
	Total	28.039	75			
sms2wpm	Between Groups	129.440	3	43.147	3.065	0.033
	Within Groups	1013.511	72	14.077		
	Total	1142.950	75			
sms3time	Between Groups	789.564	3	263.188	2.461	0.069
	Within Groups	7699.897	72	106.943		
	Total	8489.461	75			
sms3wpm	Between Groups	106.024	3	35.341	2.231	0.092
	Within Groups	1140.570	72	15.841		
	Total	1246.594	75			
sms4time	Between Groups	455.512	3	151.837	3.133	0.031
	Within Groups	3488.916	72	48.457		
	Total	3944.428	75			
sms4wpm	Between Groups	191.843	3	63.948	3.772	0.014
	Within Groups	1220.608	72	16.953		
	Total	1412.451	75			
smsavgt[2]	Between Groups	812.824	3	270.941	2.769	0.048
	Within Groups	7045.970	72	97.861		
	Total	7858.794	75			
Smsavgr	Between Groups	0.828	3	0.276	2.278	0.087
	Within Groups	8.727	72	0.121		
	Total	9.555	75			
smsavgw	Between Groups	124.445	3	41.482	3.154	0.030
	Within Groups	946.933	72	13.152		
	Total	1071.377	75			

Notes:

[1]First column reports the iteration and variable analysed for text messaging (i.e. sms):
- sms1time = Task completion time for the first text message sent;
- sms1wpm = Words Per Minute for the first text message sent;
- sms2time = Task completion time for the second text message sent;
- sms2roll = Number of thumbwheel rolls for the second text message sent;

- sms2wpm = Words Per Minute for the second text message sent;
- sms3time = Task completion time for the third text message sent;
- sms3wpm = Words Per Minute for the third text message sent;
- sms4time = Task completion time for the fourth text message sent;
- sms4wpm = Words Per Minute for the fourth text message sent;
- smsavgt = Average task completion time for all four text messages sent;
- smsavgr = Average number of thumbwheel rolls for all four text messages sent;
- smsavgw = Average Words Per Minute for all four text messages sent.

TABLE 5.19: ANALYSIS OF TUKEY'S TEST ON TREATMENT CONDITIONS

Tukey HSD	Treatment		Mean Difference				
Dependent Var.	(I)	(J)	(I-J)	Std. Err.	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Sms1wpm[1]	1	4	3.646	1.366	**0.045**	0.053	7.239
Sms1wpm	3	4	3.697	1.274	**0.025**	0.347	7.048
Sms2wpm	1	4	3.777	1.287	**0.023**	0.392	7.161
Sms3time	1	4	-9.420	3.547	**0.047**	-18.749	-0.091
Sms3wpm	1	4	3.493	1.365	*0.060*	-0.098	7.083
Sms4time	1	4	-6.707	2.388	**0.032**	-12.986	-0.427
Sms4wpm	1	4	4.631	1.412	**0.009**	0.916	8.345
smsavgt[2]	1	4	-9.271	3.393	**0.039**	-18.195	-0.346
Smsavgw	1	4	3.782	1.244	**0.017**	0.511	7.054

Legend:

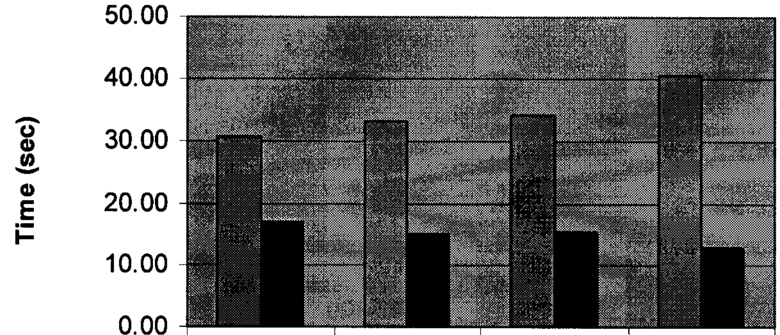
- Dependent Variables:

[1]First column reports task, treatment, and variable, e.g. sms1wpm refers to the users' WPM (Words Per Minute) during the first SMS task;

[2]"avg" refers to the average value of a variable among all four treatments (e.g. smsavgt is the average Time, t, for all four sms tasks; "r"=Rolls, "w"=WPM). For detailed description, refer to the legend for Table 5.18.

- Treatment (I) and (J): Tukey's test reported refers to the pair-wise comparison for the two treatments (numbers) shown

- Sig. (Significance): Variables reported in this table were limited to those with a t-value of .1 or less. T-values for variables with significant differences between treatments at a 0.05 level of significance or higher are shown in bold.



	1	2	3	4
Mean Completion Time	30.55	33.07	34.08	40.55
Words Per Minute (WPM)	17.07	15.23	15.39	12.97

Treatment

FIGURE 5.7: MEAN TIME AND MEAN WPM FOR ALL FOUR SMS MESSAGES – BY TREATMENT

For the sake of completion, mean completion times for the three tasks that did not experience a significant effect are shown in Figure 5.8 with respect to treatment condition. One explanation for the non-significance of the differences in means could be the large variance reported for the low qualified sample size. Qualified sample size refers to the total number of cases used in the data analysis excluding those of participants that did not follow instructions resulting in unusable data sets for the objective data analysis. As such, the number of cases used in this analysis (by means of ANOVA and Tukey's test for post-hoc pair-wise comparisons, and not by PLS) decreased from 93 to 51, and furthermore, the cases in each treatment ranged between nine and 16. Furthermore, another possible explanation for the lack of significance between differences of means may be that the low task complexity and involvement of the Web task in the study and the response lag of the webpage (consistent between subjects and treatments) visited provided users with sufficient cognitive allowance for the additive stimuli and thus, performance was not impacted significantly. The descriptive statistics for these three tasks are found in Appendix B.

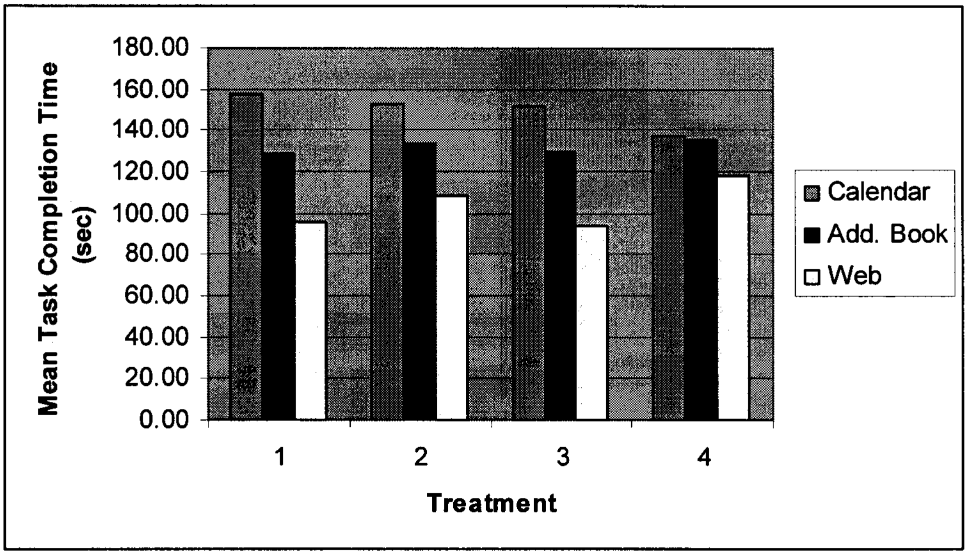


FIGURE 5.8: MEAN TIME FOR CALENDAR, ADDRESS BOOK, AND WEB TASKS – BY TREATMENT

With respect to the task completion success rate for the four tasks, only the Web task presented failed cases (17 out of 93 or 18.3%). To determine the binomial asymmetric confidence intervals of successful task completion, the incomplete beta function approach was used (Sauro 2005). The confidence intervals for a successful web task overlapped significantly across all four treatments (analysis of each treatment is shown in Tables 5.20-5.23 respectively), once again supporting that there are no significant differences between the four treatments in terms of the completion time for the web task. This finding was supported by the ANOVA on the self-reported item measuring perceived successful task completion on the device, which also showed no significant difference between treatments.

TABLE 5.20: WEB TASK SUCCESS RATE BINOMIAL CONFIDENCE INTERVALS – TREATMENT 1

Alpha	N	x	f-low	f-high	Lower CI	P	Upper CI	CI Width
0.05	17	14.0000	2.6872	5.065	57%	82%	96.2%	40%
Normal Approximation using t-distribution								
N	P	SE	T	CI	Low	High		
17	0.82	0.0932	2.1199	19.8%	62.2%	101.8%		
Normal Approximation using z-distribution								
N	P	SE	Z	CI	Low	High		
17	0.82	0.0932	1.9600	18.3%	63.7%	100.3%		

TABLE 5.21: WEB TASK SUCCESS RATE BINOMIAL CONFIDENCE INTERVALS – TREATMENT 2

Alpha	N	x	f-low	f-high	Lower CI	P	Upper CI	CI Width
0.05	19	18.0000	3.1668	39.472	74%	95%	99.9%	26%
Normal Approximation using t-distribution								
N	P	SE	T	CI	Low	High		
19	0.95	0.0500	2.1009	10.5%	84.5%	105.5%		
Normal Approximation using z-distribution								
N	P	SE	Z	CI	Low	High		
19	0.95	0.0500	1.9600	9.8%	85.2%	104.8%		

TABLE 5.22: WEB TASK SUCCESS RATE BINOMIAL CONFIDENCE INTERVALS – TREATMENT 3

Alpha	N	x	f-low	f-high	Lower CI	P	Upper CI	CI Width
0.05	23	17.0000	2.2784	2.925	52%	74%	89.8%	38%
Normal Approximation using t-distribution								
N	P	SE	T	CI	Low	High		
23	0.74	0.0915	2.0739	19.0%	55.0%	93.0%		
Normal Approximation using z-distribution								
N	P	SE	Z	CI	Low	High		
23	0.74	0.0915	1.9600	17.9%	56.1%	91.9%		

TABLE 5.23: WEB TASK SUCCESS RATE BINOMIAL CONFIDENCE INTERVALS – TREATMENT 4

Alpha	N	x	f-low	f-high	Lower CI	P	Upper CI	CI Width
0.05	17	11	2.528481551	3.019	38%	65%	85.8%	47%
Normal Approximation using t-distribution								
N	P	SE	T	CI	Low	High		
17	0.65	0.1157	2.1199	24.5%	40.5%	89.5%		
Normal Approximation using z-distribution								
N	P	SE	Z	CI	Low	High		
17	0.65	0.1157	1.9600	22.7%	42.3%	87.7%		

There were significant differences in terms of the task completion time and WPM of the text messaging task, but none for the other three tasks. Thus, the combined impact of environmental noise (i.e. auditory and visual stimuli) and user motion was significant on the objective efficiency but not on the objective effectiveness of a mobile device for text messaging. Also, this combined effect of environmental noise and user motion was not significant for both the objective efficiency and the objective effectiveness of the remaining three wireless data services, i.e. calendar update, address book entry, and Web search.

Addressing the second objective posed earlier (i.e. are there any relationships between self-reported and device data), a comparison was sought between the factor score for self-reported items and the corresponding objective performance measures thereby identifying any relationship between the two data sets. Learnability was the only dimension that mapped closely between a self-reported item and a device-based variable, because:

i) SMS was the only task that was carried out in multiple iterations (i.e. four messages were sent);

ii) Subjects' perception of learnability was measured directly by the PerEffi1 item; and

iii) Words-Per-Minute (WPM) could be used as the objective measure to determine if there was any performance improvement between the first and last SMS message sent.

As evident in Tables 5.24 through 5.26 there is a significant improvement in the users' task performance (measured through WPM) between the first and fourth text message sent, thus learnability occurred for most of the participants. By performing an

ANOVA on the PefEffi1 item we sought to determine if this phenomenon was realized (understood) by the users. The ANOVA shown in Table 5.27 reported no significant difference in terms of the perceived learnability between treatments. An aggregate PerEffi1 score of 5.83 +/- 1.27 would suggest that the users' moderate agreement in that learning how to use the device was easy was partly consequent of the 16.5 percent improvement in the SMS task completion time. Caution is warranted because users responded to this item, like all items in the measurement instrument, at the device level and not the individual task level. Therefore, perceived learnability reported is consequent of the users' overall experience with the mobile device, even though each of the remaining tasks was performed only once.

TABLE 5.24: PAIRED SAMPLES STATISTICS

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	SMSWPM1	13.5606	76	4.49710	.51585
	SMSWPM4	15.8042	76	4.53318	.51999

TABLE 5.25: PAIRED SAMPLES CORRELATIONS

		N	Correlation	Sig.
Pair 1	SMSWPM1 & SMSWPM4	76	.728	.000

TABLE 5.26: PAIRED SAMPLES TEST

		Paired Differences					T	Df	Sig. (2-tailed)
		Mean	Std. Dev'n	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	SMSWPM1 - SMSWPM4	-2.244	3.331	.382	-3.005	-1.483	-5.873	75	.000

TABLE 5.27: ANOVA FOR ITEM "LEARNING HOW TO USE THE MOBILE DEVICE FOR WIRELESS DATA SERVICES WAS EASY."

ANOVA

Source of Variation	*SS*	*Df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	8.725248	3	2.908416	1.019412	0.387996	2.706999
Within Groups	253.9199	89	2.853033			
Total	262.6452	92				

In summary, the impact of distractions caused by each discriminant distraction dimension (i.e. user motion or environmental noise) was shown to be non-significant on the objective usability of most tasks used in this experiment. However, the impact of distractions consequent of the interaction of distractions (in this study user motion and environmental noise) was shown to have a significant effect on the objective efficiency of mobile devices for text messaging. The prevalent non-significance of distractions on objective performance may be due to one of two reasons:

First, the simulated level of distractions in this experiment may have not surpassed the participants' threshold of cognitive load beyond which a degradation of performance would have occurred. One exception was in the case of text messaging, where it appears that the combined effect of the simulated distractions (i.e. fourth treatment) surpassed the individuals' threshold of distractions tolerance and had a negative impact on performance. This is in line with Harbluk and Noy (2002) who suggest that distractions should surpass a threshold of cognitive capacity to result in a degradation in performance.

Second, the lack of a significant effect of distractions on the performance with any of the other three tasks (other than text messaging) also suggests that task characteristics (e.g. task complexity) may be relevant factors in whether certain cognitive loads will result in a corresponding degradation of performance. Hence, the level of task complexity for the selected tasks combined with the simulated distraction levels in this study may have not been high enough to induce a high enough cognitive load for the user to result in a lower level of objective performance. Task complexity is an objective and discriminant attribute independent of user characteristics (Campbell 1988). Its positive impact on cognitive load and its consequent negative impact on performance have been supported in past studies (Nicholson et al. 2005; Elteren 2004; Sweller et al. 1998; Moreno and Bodenhausen 1999; Kahneman 1973).

Future research could study the effect that incremental levels of the same type of distraction would have on the objective (and perceived) usability of mobile devices for wireless data services. The impact of task complexity on the objective (and perceived) usability of mobile devices for wireless data services could also be studied.

In this chapter we validated the simplified structural model of Figure 5.6 based on the research model proposed in Chapter 3 (see Figure 3.3). In the next chapter, we address the research questions that guided this study and continue with a discussion on the implications to both theory and practice. We end by identifying the inherent limitations of this study, as well as suggestions for future research.

CHAPTER 6

DISCUSSION AND CONCLUSIONS

The previous chapter presented results that describe users' experience with the usability of mobile devices for wireless data services. In particular, the impact of distractions and expectations on usability, as well as the impact of usability on users' behavioural intention to use mobile devices for wireless data services was measured and analysed. The key learning from that discussion was that the context of use of a mobile device for wireless data services is a key consideration in both its perceived and objective usability.

This chapter aims at achieving four goals. First, the research questions posed at the beginning of this dissertation will be answered primarily by leveraging the findings that emerged from the data analysis and the validation of hypotheses. Second, contributions to both theory and practice are discussed. Third, limitations of this empirical investigation are outlined. Lastly, opportunities for future research are described.

6.1 Answers to Research Questions

6.1.1 Research Question 1

RQ1: *What is the impact of distractions on the usability of mobile devices for wireless data services?*

This research question relates to the suggestion that incremental cognitive load (acting as a distraction to the primary task at hand) consequent of auditory/visual/motor stimuli present would detract from the users' performance with mobile devices for wireless data services (Baker and Holding 1993; Rader 1993; Ljungberg's et al. 2004; Banbury et al. 2001; Hughes and Jones 2003). Accordingly the following hypotheses were proposed:

H1a: *Exposing users to higher levels of distractions will negatively influence their perceived efficiency of a mobile device for wireless data services.*

H1b: *Exposing users to higher levels of distractions will negatively influence their perceived effectiveness of a mobile device for wireless data services.*

There was strong statistical support for both hypotheses. With respect to H1a, the hypothesis was supported in its argument that distractions negatively influence a user's perception of efficiency of a mobile device for wireless data services ($\beta = -0.495$; p-value < 0.001). With respect to H1b, the hypothesis was supported in its argument that distractions negatively influence a user's perception of effectiveness of a mobile device for wireless data services ($\beta = -0.259$; p-value < 0.025).

Overall, the impact of distractions on performance was supported by a greater confidence interval for perceived efficiency ($t = 4.459$; p-value < 0.001) than for perceived effectiveness ($t = 2.444$; p-value < 0.025). Referring to the attributes defined for efficiency, there appears to be greater confidence in that increased levels of distractions will result in users perceiving wireless data services on mobile devices to be either less intuitive, more time consuming, and more demanding in learning how to use them, than more difficult to complete them. This observation was supported through by the analysis of the effect size. The effect of Distractions (TLX) on perceived performance was significant in both cases (0.31 or medium to large for Perceived Efficiency; 0.07 or small to medium for Perceived Effectiveness). In other words, distractions caused users to invest more time in completing a task while preserving a comparable degree of accuracy between treatments.

Beyond the users' perception, we were also interested in finding out how distractions impacted actual performance. Accuracy was measured for each task through device data and there was no significant difference between treatments. Successful task completion was measured both through both a self-reported item and through device data, and in both cases there was no significant difference between treatments. On the other hand, there was a difference in the time it took to complete a task, particularly for text messaging (the "heaviest" task in terms of data entry). Put simply, users appeared to react to the context of use (incremental level of distractions) by taking longer to complete a task, while ensuring a comparable rate of successful task of completion and accuracy between treatments.

Overall, participants in Group 1 (i.e. seated and in isolation) demonstrated faster task completion times for all tasks, than those in Group 2 (i.e. walking in isolation), Group 3 (i.e. seated and exposed to background auditory and visual stimuli), and Group 4 (walking and exposed to background auditory and visual stimuli). It should be noted that although there was a difference in the task completion time for all tasks (i.e. efficiency), users experienced a statistically significant difference only with text messaging. The task completion times for each group were 30.55s, 33.07s, 34.08s, and 40.44s respectively. Statistically significant difference was found for participants between the control group (Group 1, seated and no background stimuli) and those in Group 4, i.e. when the users were exposed to both types of cognitive load, user motion and auditory and visual stimuli in the environment. Thus, there was no significant effect on actual effectiveness neither between treatments nor between tasks. One possible explanation for this could be that the users' increased effort for accurate key entry resulted in the increased time for each key entry and the overall prolonged duration of the task.

6.1.2 Research Question 2

RQ2: *Does Expectancy-Disconfirmation Theory (EDT) help to explain a user's evaluative process of usability with a mobile device for wireless data services?*

This research question relates to the examination of the applicability of EDT in mobile usability studies. The relevant constructs here are: Expectations of Efficiency, Disconfirmation of expectations of Efficiency, Perceived Efficiency, and Satisfaction with Efficiency, as well as the same set of constructs but focusing on Effectiveness instead of Efficiency. We begin with the first dyad of Expectations-Performance, namely the relationships between Expectations of Efficiency and Perceived Efficiency (H2a), as well as the one for Expectations of Effectiveness and Perceived Effectiveness (H2b):

H2a: *Higher user expectations of efficiency of a mobile device will positively influence the perceived efficiency of using the mobile device for wireless data services.*

H2b: *Higher user expectations of effectiveness of a mobile device will positively influence the perceived effectiveness of using the mobile device for wireless data services.*

This dyad was found to be significant. It was theorized that incrementally higher expectations would result in a higher level of perceived performance. Stemming from a "self-fulfilling prophecy" phenomenon (Van Ryzin 2004; Oliver 1980, 1997; Spreng et al. 1996; Olchavsky and Miller 1982; reviewed by Yi 1990), hypotheses H2a and H2b expressed this with respect to efficiency and effectiveness respectively. The argument of H2a that higher user expectations of efficiency of a mobile device would positively influence the perceived efficiency of using the mobile device for wireless data services received moderate support ($\beta = 0.189$; p-value < 0.01). Similarly, moderate support ($\beta = 0.214$; p-value < 0.01) was received for the argument of H2b that higher user expectations of effectiveness of a mobile device would positively influence the perceived effectiveness of using the mobile device for wireless data services.

A couple of observations can be made here. First, since the relationships were shown to be moderately significant, this could suggest that there are other moderating factors that were not observed. Such factors could include self-efficacy and culture, two dimensions that were not explored in this research given the mixed research results of corresponding past studies. Also, another plausible explanation for this moderate level of significance is that perceived performance of innovative new products and services is less susceptible to the impact of expectations than perhaps more mainstream products and services. This may arise due to the absence of any normative factors (such as beliefs regarding wireless devices) contributing to the formation of lower expectations, thus creating a moderate level of expectations across the entire sample.

Next we discuss the set of relationships related to the expectations-disconfirmation dyad, namely the relationships between Expectations of Efficiency and Disconfirmation of expectations of Efficiency (H3a), as well as the one for Expectations of Effectiveness and Disconfirmation of expectations of Effectiveness (H3b).

Next we discuss the set of relationships related to disconfirmation. Although the selected simplified model excluded the disconfirmation constructs, we will take the time to discuss the corresponding relationships and any implications that arose from the findings. These relationships between constructs are: expectations to disconfirmation, perceived performance to disconfirmation, disconfirmation to satisfaction, and expectations to satisfaction. We begin with the former relationship, expectations to disconfirmation.

Prior research suggests expectations have a negative effect of on disconfirmation, since the higher the initial set of expectations the more difficult it will be for them to be met by the experience with the product (Van Ryzin 2004; Oliver 1980, 1997). Accordingly, the following two hypotheses were proposed:

H3a: *Higher user expectations of efficiency of a mobile device will negatively influence the disconfirmation of user expectations of efficiency in using the mobile device for wireless data services.*

H3b: *Higher user expectations of effectiveness of a mobile device will negatively influence the disconfirmation of user expectations of effectiveness in using the mobile device for wireless data services.*

The corresponding dyad of hypotheses (H3a and H3b) was not found to be significant. The support for H3a was (β = -0.075; p-value = 0.332) and for H3b was (β = 0.018; p-value = 0.735). Also, we observe that the path for H3a was in the correct direction (i.e. expectations negatively influence disconfirmation), whereas the path of H3b, although positive, it was near zero (and a negative beta path coefficient). As such, further investigation of these relationships is needed.

On the other hand, perceived performance was expected to have a positive effect on disconfirmation, since the higher the performance, the more likely it will surpass expectations (Oliver 1980, 1997; Van Ryzin 2004). Accordingly, the following hypotheses were proposed:

H4a: *Higher levels of perceived efficiency of a mobile device will positively influence the disconfirmation of user expectations of efficiency in using the mobile device for wireless data services.*

H4b: *Higher levels of perceived effectiveness of a mobile device will positively influence the disconfirmation of user expectations of effectiveness in using the mobile device for wireless data services.*

Both of these hypotheses received strong support. The claim that higher levels of perceived efficiency of a mobile device will positively influence the disconfirmation of user expectations of efficiency in using the mobile device for wireless data services was strongly supported ($\beta = 0.640$, p-value < 0.001). Also, the claim that higher levels of perceived effectiveness of a mobile device will positively influence the disconfirmation of user expectations of effectiveness in using the mobile device for wireless data services was also strongly supported ($\beta = 0.439$, p-value < 0.001).

In turn, disconfirmation was proposed to have a direct effect on satisfaction (Oliver 1980, 1997; Van Ryzin 2004). Accordingly, the following hypotheses were proposed:

H5a: *Higher levels of disconfirmation of user expectations of efficiency of a mobile device will lead to higher levels of user satisfaction with the efficiency of the mobile device for wireless data services.*

H5b: *Higher levels of disconfirmation of user expectations of effectiveness of a mobile device will lead to higher levels of user satisfaction with the effectiveness of the mobile device for wireless data services.*

This dyad of hypotheses received strong support. The argument of hypothesis H5a, that higher levels of disconfirmation of user expectations of efficiency of a mobile device will lead to higher levels of user satisfaction with the efficiency of the mobile device for wireless data services, received strong significance ($\beta = 4.054$; p-value < 0.001). Strong significance ($\beta = 1.810$; p-value < 0.001) was also received for the argument of hypothesis H5b that higher levels of disconfirmation of user expectations of effectiveness of a mobile device will lead to higher levels of user satisfaction with the effectiveness of the mobile device for wireless data services. Support for both hypotheses shows that users experiencing positive disconfirmation regarding the performance of a mobile device exhibited satisfaction with its performance.

In addition to performance and disconfirmation, expectations may also have a direct effect on satisfaction. This may occur when users are relatively unaware of the product's performance or have little psychological involvement during its use (Oliver 1997; Van Ryzin 2004). Accordingly, the following hypotheses were proposed:

H7a: *Higher user expectations of efficiency of a mobile device will lead to higher levels of user satisfaction with the efficiency of the mobile device for wireless data services.*

H7b: *Higher user expectations of effectiveness of a mobile device will lead to higher levels of user satisfaction with the effectiveness of the mobile device for wireless data services.*

This dyad of expectations-to-satisfaction hypotheses, H7a and H7b, however, was not supported. The two hypotheses received path and significance scores of (β = 0.039; p-value = 0.964) and (β = 0.113; p-value = 0.219) respectively. This suggests that the users' expectations from the efficiency and/or effectiveness of mobile devices regarding wireless data services varied. For some participants, such expectations were assimilated to a broader perspective on mobile devices, which would then result in a confirmation bias. For others, including novice users, the applicability of any existing mental models (Norman 1988) may have caused them to maintain a more narrow perspective of expectations from wireless data services. This omnipresent varied impact of heterogeneous expectations on satisfaction (Wellbery, 2005; Toyone et al. 2005) suggests it is an unreliable mechanism for influencing consumer satisfaction.

The last dyad relevant to EDT is also relevant to usability. User perceived efficiency and perceived effectiveness were expected to lead to satisfaction with the efficiency and satisfaction with effectiveness of a mobile device respectively (Churchill and Surprenant 1982; Tse and Wilton 1988). Accordingly, the following hypotheses were proposed:

H6a: *Higher levels of perceived efficiency of a mobile device will lead to higher levels of user satisfaction with the efficiency of the mobile device for wireless data services.*

H6b: *Higher levels of perceived effectiveness of a mobile device will lead to higher levels of user satisfaction with the effectiveness of the mobile device for wireless data services.*

This dyad received strong support. Specifically, the argument that higher levels of perceived efficiency of a mobile device will lead to higher levels of user satisfaction with the efficiency of the mobile device for wireless data services, received strong support ($\beta = 0.722$; p-value < 0.001) when the construct of disconfirmation of expectations of efficiency was not included in the analysis, and similarly strong support when the latter was included in the analysis ($\beta = 0.492$; p-value < 0.001). On the other hand, the argument that higher levels of perceived effectiveness of a mobile device will lead to higher levels of user satisfaction with the effectiveness of the mobile device for wireless data services, received strong support ($\beta = 0.443$; p-value < 0.001) when the construct of disconfirmation of expectations of efficiency was not included in the analysis, while it found to be not quite significant when the latter was included in the analysis ($\beta = 0.214$; p-value < 0.2).

In summary, with respect to EDT we observed that in the absence of the Disconfirmation construct, the path from Expectations of Efficiency to Perceived Efficiency and the path from Perceived Effectiveness to Satisfaction with Effectiveness gained statistical confidence in their claims. The mediating role of performance between expectations and satisfaction was further supported in this dissertation, this time in the context of a mobile device for wireless data services. On the other hand, the inclusion of the disconfirmation construct appeared problematic in that the core path from Perceived Effectiveness to Satisfaction with Effectiveness did not receive support, a finding that would contradict a core usability relationship. Hence, EDT appears not to be readily applicable in mobile usability studies. A variant of EDT, however, appears to offer high explanatory power in a user's evaluative process of the usability of a mobile device for wireless data services. This variant proposes that Satisfaction with a performance dimension (either efficiency or effectiveness) is influenced by the perceived assessment of that performance dimension, which in turn is influenced by the original set of Expectations of that performance dimension prior to trial (actual usage).

6.1.3 Research Question 3

RQ3: *What is the impact of usability on consumers' behavioural intention towards using mobile devices for wireless data services?*

This research question focuses on the most endogenous construct studied, behavioural intention (BI) towards using a mobile device for wireless data services. It was theorized that there is a strong effect between satisfaction and a consumer's behavioural intention to use a product (Anderson and Sullivan 1993, Cronin and Taylor 1992, Gotlieb et al. 1994; Mittal et al. 1999; Taylor and Baker 1994; Chiu et al. 2004). Accordingly, the following hypotheses were proposed:

H8a: *Higher levels of user satisfaction with the efficiency of a mobile device will positively influence the user's intention to use it for wireless data services.*

H8b: *Higher levels of user satisfaction with the effectiveness of a mobile device will positively influence the user's intention to use it for wireless data services.*

There was very strong statistical support for just one of these two hypotheses. The claim articulated by hypothesis H8a that higher levels of user satisfaction with the efficiency of a mobile device will positively influence the user's intention to use it for wireless data services received strong support ($\beta = 0.389$; $p\text{-value} < 0.001$). On the other hand, the claim by hypothesis H8b that higher levels of user satisfaction with the effectiveness of a mobile device will positively influence the user's intention to use it for wireless data services did not receive statistical support ($\beta = 0.154$; $p\text{-value} = 0.346$). These values were obtained after analysing the data relevant to the simplified model shown in Figure 5.6 selected from the discussion in chapter 5 (i.e. model excluded both Disconfirmation constructs, and hence all associated relationships were dropped). These findings were supported by an analysis of the constructs' respective effect size. The effect on BI by satisfaction with efficiency was medium to large (i.e. 0.27).

From the above findings, what becomes particularly interesting is that consumers appear to place more emphasis on efficiency than on effectiveness when it comes to personal (as opposed to professional) purchase decisions in this context. Drawing upon the dimensions measured for efficiency, we can conclude that satisfaction with learnability, ease of use, and task completion time are prevalent in a user's decision making process of a mobile device for wireless data services, while satisfaction with successful task completion seems to not hold true across a significant portion of the population.

6.1.4 Research Question 4

RQ4: *What factors become relevant when studying usability within a context of use (i.e. situational specific attributes)?*

This research began with an exploratory objective of understanding the importance of context of use in assessing the usability of mobile devices. By achieving this objective we effectively created a conceptual road map for mobile usability studies, such as the one performed for this dissertation. Following a literature review, a research framework was proposed (see Figure 3.1) that introduced a taxonomy for contextual factors that includes: the User, the Environment, the Technology, and the Task. For each of those four categories, a set of sub-groups was identified, each having the capability of introducing secondary level dimensions that could be operationalized and may impact usability.

In addition to identifying the factors relevant to mobile usability at a high level of abstraction through the proposed framework, we sought to assess the impact that a few of these variables have in the context of mobile device use for wireless data services.

The complexity of human behaviour has drawn significant interest in studying which user characteristics may be of relevance in mobile usability studies. While the list of user characteristics was not meant to be exhaustive, this dissertation identified a set of factors that are proposed to be of significance in mobile usability studies. These factors include the user's age, gender, education, self-efficacy, expectations, and physiological state. As we saw earlier, in our study we controlled for the first three, i.e. age, gender, and education. Next, we will discuss these three variables in terms of their impact on the usability of a mobile device for wireless data services.

Strong beta coefficients and corresponding t-values (see Table 5.17) further indicated that women were affected more by distractions than men during their use of a mobile device for wireless data services, a finding in line with the results of Bruni's work on the impact of instant messaging on task performance (2004). Similarly, moderate significance ($t = 1.632$; see Table 5.17) was found for age suggesting that older users are also more susceptible to distractions when using mobile devices for wireless data services. The following implications emerge from these findings: i) if women are distracted more than men, it is likely that applications that require more attention will not be as popular - such applications could include both professional and leisure applications; ii) similarly, older consumers are less likely to be satisfied with the use of an application that requires significant attention than younger users are. Hence, an opportunity exists for manufacturers to offer special accessories for mobile devices increasing their usability to these target market segments. These products will not only introduce a new stream of revenue, but also boost sales of the mobile devices directly, particularly for these target market segments.

Turning the discussion to the impact of the control variables on EDT, the following two observations were made. First, age had a considerable impact on Expectations of Efficiency, which was further qualified by the strong negative path coefficient and corresponding t-value for Expectations of Efficiency. This would suggest that older participants have higher expectations from mobile devices for the intended wireless data services than younger ones. Devices targeting an older market segment could incorporate appropriate form factors, such as larger buttons, bigger displays, paying attention to colours used in the interface, and many others that would support an efficient performance with it.

Second, gender had considerable impact over Expectations of Effectiveness. This was further qualified by the negative beta coefficient, which would suggest that women had higher expectations of Effectiveness than with the efficiency for wireless data services than men did. This can be linked to the gender differences found with respect to aptitude, and particular with respect to multitasking with small motor activity (Franken 2002). Accordingly, Noble et al. (2003) suggested that gender would influence considerations of usability, which was also evident in our study. Moving forward with this knowledge, mobile device manufacturers and carriers could avoid a common pitfall of game design: male-friendly design. Everything in videogames, from "the environment (the marketing, the merchandising, the image of the industry) to the system (the laughably phallic joystick, the original Xbox controllers which are too big for my hands, the color scheme of the Xbox) are male-friendly. The attitude seems to be, "Maybe some women play our games, but we don't really know, and frankly, we don't care."" (Pinckard 2002). Applying the finding from our study (i.e. gender impacts expectations of effectiveness) and the lesson from the gaming industry in the context of mobile devices, close attention to both the device design as well as the related promotions is warranted. Attention to such gender issues could result in a more positive experience from female consumers, which in turn will impact their intention towards using such mobile devices and wireless data services.

Finally, age, gender, and education did not have a significant impact on a user's behavioural intention to use a mobile device for wireless data services (see Table 5.17). This is an interesting observation, as it suggests that any individual, regardless of age, gender, and education, is just as likely to use a mobile device for wireless data services.

The second of the four dimensions defining the context of use is the environment. This dimension refers to elements beyond the user's control but may impact the usability of mobile devices. Such elements include lighting, temperature, and other items pertaining to the physical environment. In addition, the environment may refer to non-physical attributes, such as political, social, or economic conditions. In this study, the impact that the environment may have on the usability of mobile devices for wireless data services was studied by manipulating auditory and visual stimuli. Specifically, participants were asked to perform tasks either in the absence or presence of background music, meaningful speech, and other individuals walking in their visual horizon without actually interacting with them. These manipulations were present in two out of the four treatments (groups three and four). As seen in the earlier discussion on Data Analysis and Results in Chapter 5, these distractions had a significant effect on the perceived usability of subjects. Specifically, the effect of distractions on perceived performance was significant in both cases (0.31, a medium to large effect for Perceived Efficiency, and 0.07, a small to medium effect for Perceived Effectiveness). The simulated distractions, however, did not have a significant effect on the objective usability of the mobile device for wireless data services. This phenomenon will be explained in later sections of this dissertation, but in brief, it may have occurred due to a simulated level of distractions that was too low to cause any significant effect with actual performance, or it may be consequent of a low task complexity associated with the tasks designed in this study.

The third dimension defining the context of use is the task. Task related elements that may of relevance in usability include the corresponding goals, the degree of interactivity, or the level of complexity for the particular mobile interface. In this study, four different tasks were used, each with their own interface design, level of interactivity, and complexity. However, these were not intentionally manipulated for the purpose of understanding their impact on mobile usability. Instead the tasks were selected because of their popularity among current mobile users. Still, relevant findings show that the calendar, address book, and wireless Web tasks were resistant to the effects of the distractions simulated in this study. On the other hand, users exhibited a lower level of performance with the text messaging task when it was carried out in the presence of both types of distractions (user motion and environmental auditory and visual stimuli). While the focus of this study was not on task characteristics, these results make a strong case for a future study of task complexity and its impact on mobile usability.

The last of the four dimensions defining the context of use is the technology. This dimension extends beyond the interface to include the specifics of the overall system (e.g. elements relevant to performance, such as memory and battery) and the underlying technology infrastructure (e.g. network characteristics such as bandwidth, reliability, dependability, and availability) supporting its use. This experiment involved the use of a handheld mobile device, specifically the BlackBerry 7250. Wireless networks were not used in this study (websites were stored on the device), and the same device was used between groups. Therefore no findings arose in this study with respect to these technology elements. The only technology variant in this study was the application interface. Its implication on usability was described in the previous paragraph.

6.2 Contributions

6.2.1 Theoretical and Methodological Contributions

This section describes the contributions to theory and methodology of this dissertation. Such contributions will be presented in the context of the theoretical concepts involved in this study, namely usability, distractions, expectations, and adoption. Methodological contributions will be discussed last.

6.2.1.1 Usability

Efficiency and Effectiveness have in the past been measured as an aggregate construct, namely performance (Watters et al. 2003), which in turn was shown to impact satisfaction (Churchill and Surprenant 1982; Tse and Wilton 1988). The work of Frokjaer et al. (2000), however, suggested that efficiency and effectiveness are discriminant dimensions. In the event this is true, satisfaction should also be measured through discriminant scales. This follows from the literature on EDT that suggests performance impacts satisfaction. Hence, decomposition of the antecedent into two constructs could suggest the decomposition of the consequence into two constructs as well, maintaining the correspondence of items measured by each construct.

Post-hoc analysis offers additional support on the argument for the decomposition of satisfaction into two discriminant constructs, one for efficiency and one for effectiveness. First, by referring to the correlation matrix for the assessment of discriminant validity (see Table 5.5), the square root of the average variance extracted for Perceived Efficiency (0.820) was greater than its correlation with Perceived Effectiveness (0.414). Similarly, the square root of the average variance extracted for Perceived Effectiveness was greater than its correlation with Perceived Efficiency. Thus, according to the work of Fornell and Larcker (1981), there is strong support that the two constructs are discriminant. This finding provides further support for the independent measurement of the two dimensions in future usability studies. In similar fashion, we assess the discriminant nature of the two corresponding satisfaction constructs. The square root of the average variance extracted for Satisfaction with Efficiency (0.897) was greater than its correlation with Satisfaction with Effectiveness (0.513). This was also the case for Satisfaction with Effectiveness, for which the square root of the average variance extracted was greater than its correlation with Satisfaction with Efficiency. Hence, there is support that the two Satisfaction constructs are discriminant.

As an additional check to the inspection of the correlation matrix described above on the discriminant nature of the satisfaction constructs, we observed the research model's fit between the decomposed performance and a hypothesized single satisfaction construct measuring overall satisfaction for which data was collected during the study. The variance explained for satisfaction by the two decomposed constructs of efficiency and effectiveness was 54 and 21 percent respectively, whereas a single overall satisfaction construct only accounted for 59 percent.

Another set of contributions to theory emerges from the post-hoc analysis on the participant demographics for which we controlled, as well as the between-treatment device data analysed.

First, when comparing the device data between males and females, there were no significant differences in terms of actual performance.

Second, there is a significant improvement in the users' task performance (measured through WPM, a measure of keystroke activity relative to task completion time) between their first and last (fourth) text message sent. Thus, learnability generally occurred for the participants between treatments with respect to text entry (see Tables 5.24 through 5.27).

Third, there was no significant difference in terms of the perceived learnability between treatments.

Fourth, an aggregate score of 5.83 +/- 1.27 for the first item of the Perceived Efficiency construct (i.e. PerEffi1) that measures learnability suggests that the users' moderate agreement in that learning how to use the device was easy was partly consequent of the 16.5 percent improvement in the SMS task completion rate. Caution is warranted users responded to this item, like all items in the measurement instrument, at the device level and not the individual task level. Therefore, perceived learnability reported is also consequent of the users' experience with the remaining tasks even though each was performed only once.

6.2.1.2 Distractions

First, the measurement instrument incorporated NASA's TLX scale to measure the level of distraction experienced by the participants. This scale worked very well in measuring distractions, which also served as a manipulation check for the four treatment conditions. Referring to the item statistics in Table 5.2 (see *Chapter 5 Data Analysis and Results*), it becomes apparent that only the fifth and sixth item of the TLX scale, as expected, did not load. These items do not measure distractions, rather their consequences (i.e. performance and frustration level), thus they were removed from the subsequent data analysis presented in the previous chapter. Furthermore, the remaining four items had loadings ranging from 0.662 to 0.860, with only the second item being below 0.7. This is reasonable, because the second item measured physical load experienced by the participants. While it is reasonable to assume that physical load translates to a cognitive load for the participant, there are additional factors that would affect the latter that are not tapped into by this item.

As indicated in the post-hoc analysis, the effect of distractions caused by each discriminant distraction dimension (i.e. user motion or environmental noise) was shown to be non-significant on the objective usability of most tasks used in this experiment. However, the impact of distractions consequent of the interaction of distractions (in this study user motion and environmental noise) was shown to have a significant effect on the objective efficiency of mobile devices for text messaging. The prevalent non-significance of distractions on objective performance on other tasks may be due to one of two reasons:

First, the simulated level of distractions in this experiment may have not been high enough to cause a degradation of performance. This is in line with Harbluk and Noy (2002), who suggest that distractions should surpass a threshold of cognitive capacity to result in a degradation of performance. In the case of text messaging it appears that the combined effect of the simulated distractions (i.e. fourth treatment) surpassed the individuals' threshold of distractions tolerance and had a negative impact on performance. The relationship between distractions and performance has been supported by the work of several researchers who found that user motion, and auditory and visual stimuli (e.g. collocation) increase the user's cognitive load resulting in reduced task performance (Smith et al. 2005; Fockert et al. 2001; Pontari and Schlenker 2000; Moreno and Bodenhausen 1999; Eysenck 1984; Kahneman 1973).

Second, the lack of a significant effect of distractions on the performance with any of the other three tasks (other than text messaging) also suggests that the level of task complexity for the selected tasks in this study may have not been high enough to induce a high enough cognitive load for the user to result in a lower level of objective performance. Task complexity is an objective and discriminant attribute independent of user characteristics (Campbell 1988). Extensive literature offers support for the positive relationship between task complexity and cognitive load leading to a subsequent degradation of performance (Nicholson et al. 2005; Elteren 2004; Sweller et al. 1998; Moreno and Bodenhausen 1999; Kahneman 1973).

6.2.1.3 Expectations

The marketing theory of Expectations-Disconfirmation was applied and tested for its relevance in IS. Overall, most hypotheses were supported, but with five out of 12 links not receiving support, a closer examination was warranted. Upon inspection it became evident that there are two types of links that did not demonstrate statistical significance: i) Expectations to Disconfirmation, and ii) Expectations to Satisfaction. This was consistent across both performance dimensions, i.e. efficiency and effectiveness. In addition, and although some paths were supported, the overall fit of the model was worse and in particular for the core paths of usability. This is not a surprising result, as the applicability of EDT has received mixed support in other studies (Van Ryzin 2005). Hence, the applicability of EDT in its original form for mobile usability studies is cautioned. A variant of EDT where expectations is introduced as an antecedent to perceived performance independently is likely to yield results of stronger statistical support.

6.2.1.4 Methodology

One of the observations emerging from the literature review was that most past usability studies left much to be desired in terms of realism. This was a conscious decision by the researchers at the time in an attempt to ensure a rigorous methodology, and possibly due to limitations of resources. Our research attempted to approximate real-world scenarios for the tested product and services, while upholding a rigorous methodology. The manipulations set as part of the experiment setup worked well in introducing participants to varying levels of distractions though likely contextual usage scenarios (e.g. alone, walking, noisy environment, close proximity of other individuals). The assigned tasks represented the most commonly used applications on mobile devices. However, while the scope of text messaging, address book, and calendar tasks is limited (i.e. these applications enable a focused set of user activities), use of the wireless Web could be carried out in several different ways. The chosen approach was one of "Web finding" Sellen et al. (2002). This type of task is the most common approach in mobile usability studies (Buyukkokten et al. 2001; deBruijn et al. 2002; Jones et al. 1999), because it offers more control over the research than an open-ended "information gathering" task. The latter refers to tasks where users are asked to locate information with minimal direction. Adding to the rigor of "Web finding" tasks, the designed Web task was a fair representative of the cognitive demands associated with goal-oriented searches performed by mobile device users.

Furthermore, the measurement instrument worked well in serving as the manipulation check for distractions. In addition, the operationalization of EDT-related constructs that was borrowed from marketing theory also worked well in our study. Finally, the adopted scale for measuring the user's behavioural intention to use a mobile device for wireless data services was also validated.

Adding to the earlier points on realism, the designed manipulations worked well. Not only was this study closer to a real world scenario than most usability studies in the past, but also it provided insight and opportunities for future research in related areas. For example, the cognitive load of user motion appears to be higher than the combined effect of both visual and auditory stimuli in the environment. Also, text messaging was found to be significantly impacted only in the presence of aggregate distractions by the user's motion and the background music, conversations, and pacing of other individuals. Hence, our research design facilitated the exploration of important dimensions of contextual usability that would otherwise not have been attended to with research performed in great distance from real-world use of mobile devices and wireless data services.

6.2.2 Practical Contributions

This section describes practical contributions of this dissertation. Such contributions will be presented in the context of the practical implications emerging from this study with respect to usability, distractions, expectations, and adoption.

6.2.2.1 Usability

This dissertation studied usability by decomposing the performance construct into two dimensions of efficiency and effectiveness and measuring satisfaction through discriminant constructs for each of them. This approach received strong statistical support from a theoretical perspective. In addition, there are practical implications of this finding. By using a single satisfaction construct we lose valuable information in terms of which performance dimension(s) has a significant effect on the subjects' future intention to use a mobile device for wireless data services. To argue that satisfaction with the overall performance will affect a consumer's intention to use a mobile device seems intuitive, whereas a decomposed satisfaction dimension offers relevant insight to which performance dimension becomes critical in such decision scenarios.

How important trial with (i.e. opportunity to use) a mobile device is for consumers in forming their subjective opinions on the degree of satisfaction with it can be seen through the effect size analysis. The effect of Perceived Efficiency on Satisfaction with Efficiency was large (i.e. 0.34 – see Table 5.13). Reflecting on the second-order factors of Perceived Efficiency, this effect would suggest that satisfaction with the ease of use, learnability, and/or task completion time of a mobile device is significantly impacted by an individual's use and perceived performance of it. On the other hand, the effect of Perceived Effectiveness on the Satisfaction with Effectiveness was low (i.e. 0.06 – see Table 5.14). Implications from this finding include: i) the necessity of wireless carriers to provide consumers with the opportunity to try a mobile device; ii) the importance of device designers and manufacturers to focus on the efficiency of a mobile device when it is intended for consumers.

Furthermore, it was observed that web tasks presented no significant difference in terms of user performance between treatments, whereas distractions had a significant negative impact on text entry. This would suggest that Web- rather than text-driven tasks are a better fit for the varied situational use of a mobile device. For example, Web-based menu-driven selection would be significantly more efficient when compared to a text-based order entry form. This finding is transferable to any application that can be run on a mobile device.

Lastly, a relationship was found between perceived efficiency and perceived effectiveness, whereby the former impacts the latter. The observation that the more efficient users perceive a mobile device to be for wireless data services leads to a perception that the device is more effective (and arguably more useful) gives rise to important practical considerations. First, design features and the overall form factor of a mobile device should focus on supporting usability considerations, such as ease of use, learnability, and short task completion cycles/times. In doing so, such designs will inherently further boost a favourable impression on the effectiveness of the mobile device. Second, implementation of these design considerations will bring mobile devices closer to a state of ubiquitous and pervasive computing, which is likely to generate increased revenue across the m-commerce supply chain, from device manufacturers to service providers.

6.2.2.2 Distractions

We begin by revisiting observations relating to the impact of distractions on user performance with a mobile device for wireless data service. Participants in Group 4 (walking and exposed to visual and auditory stimuli in the background) experienced a greater level of performance degradation than those in any of the other three groups. This has significant implications on the expected contextual use of a technology. Since a typical setting in which consumers may use a mobile device for wireless data services while walking involves auditory and/or visual stimuli in the environment, it calls for specific requirements regarding the design of mobile interfaces (Lim and Benbasat 2000). Such requirements would address the changing levels of distractions typically enabled through adaptable and location-aware mobile systems (Gebauer et al 2004; Liang and Wei 2004; Rao and Minakakis 2003).

Next, we explore any implications for practice regarding user demographics and distractions. Strong beta coefficients and corresponding t-values further indicated that women were affected more by distractions than men (see Table 5.17), which is in agreement with the results of Bruni's (2004) work, which examines the impact of instant messaging on task performance. The following implications emerge from these findings: i) if women are less robust to distractions than men in the context of a mobile device for wireless data services, it is likely that such applications will not be as popular with women - such applications could include both professional and leisure applications; ii) similarly, older consumers are less likely to be satisfied with the use of an application that requires significant attention compared to younger users. Hence, manufacturers could benefit by offering special accessories for mobile devices increasing their usability to these target market segments. These products will not only introduce a new stream of revenue, but also boost sales of the mobile devices directly, particularly for these target market segments.

It was mentioned earlier that the three tasks other than text messaging (i.e. Calendar, Address Book, and wireless Web) presented no significant differences between treatments. This would suggest that such applications are more likely to be used regardless of the contextual use of a mobile device. Hence, manufacturers and wireless carriers could benefit by increasing the availability of such services. As seen by Kjeldskov et al (2005), the wireless Web in particular seems to offer an efficient form of interaction with a mobile device across contexts. Thus, we can expect to see a plethora of emerging web-based applications for mobile devices. Such applications would enable users to find information (e.g. routes, timetables) while they are moving from one physical location to another (Kjeldskov et al. 2005) that is "closely related to their physical location and objects in their immediate surroundings" (Cheverst et al. 2000; Chincholle et al. 2002; Schmidt-Benz et al. 2002; Reid 2002; Umlauft et al. 2003).

Finally, following from the negative relationship between Distractions and Satisfaction with Effectiveness found during the data analysis, manufacturers and developers should attempt to make these mobile devices as adaptable as possible to the conditions they are being used in, or otherwise they run the risk of having their product criticized/evaluated negatively perhaps inappropriately.

6.2.2.3 Expectations

Although the relationships between expectations of performance and perceived performance (i.e. efficiency and effectiveness) were shown to be moderately significant, this level of significance could suggest that there are other moderating factors that were not observed. Such factors could include self-efficacy and culture, two dimensions that were not explored in this research given the mixed research results of corresponding past studies. If this is the case, then industry would be best off manipulating those factors through customer relationship management (CRM) initiatives. Such techniques would leverage consumer profiling in shaping expectations and consequently perceptions of the company's products.

On the other hand, perceived performance was shown to have a positive effect on satisfaction with performance. Combining this observation with its consequence (i.e. positive effect on behavioural intention to use a mobile device) and applying them to industry highlights the importance of mobile device usability. If performance with a device is perceived to be high, then it is likely that the consumer will continue to use a product, thus growing the company's subscriber / client base and its revenue. Coupling this intuitive conclusion with the observed relatively more dominant importance of efficiency over effectiveness, manufacturers and service providers are given more information that could be leveraged in their production and marketing initiatives.

In addition, industry should focus on generating both realistic and favourable expectations. Such expectation formation will in turn positively impact consumers' perceived performance of mobile devices. As shown earlier, this is a highly supported causality path leading to the consumers' behavioural intention to use mobile devices for wireless data services.

The mediating role of perceived performance between expectations of performance and satisfaction with performance can be leveraged by industry in terms of developing tailored marketing campaigns. These campaigns would aim to shape expectations through carefully crafted advertising messages, and offering consumers the opportunity to try these wireless data services on a mobile device.

Turning the discussion to the impact of the control variables on EDT, the following three observations were made. First, education had a moderate effect on Expectations of Effectiveness. This is an expected result, if we assume that one passes more critical judgment the more evaluative tools they have, tools that they adopt during formal training and education.

Second, age had a considerable impact on Expectations of Efficiency, which was further qualified by the strong negative path coefficient and corresponding t-value for Expectations of Efficiency. This would suggest that older participants have higher expectations from mobile devices for the intended wireless data services than younger ones. Devices targeting an older market segment could incorporate appropriate form factors, such as larger buttons, bigger displays, paying attention to colours used in the interface, and many others that would support an efficient performance with it.

Third, gender had little impact over Expectations of Efficiency, Disconfirmation of Efficiency expectations, and considerable impact over Expectations of Effectiveness. The relationship between gender Expectations of Effectiveness was further qualified by the negative beta coefficients, which would suggest that women had higher expectations of Effectiveness than with the efficiency for wireless data services than men did. This observation may be linked to the gender differences found with respect to aptitude, and particular with respect to multitasking with small motor activity (Franken 2002). Accordingly, Noble et al. (2003) suggested that gender would influence considerations of usability, which was also evident in our study. Moving forward with this knowledge, mobile device manufacturers and carriers could avoid a common pitfall of game design: male-friendly design. Everything in videogames, from "the environment (the marketing, the merchandising, the image of the industry) to the system (the laughably phallic joystick, the original Xbox controllers which are too big for my hands, the color scheme of the Xbox) are male-friendly. The attitude seems to be, "Maybe some women play our games, but we don't really know, and frankly, we don't care."" (Pinckard 2002). Thus, close attention to both the mobile device design as well as the related promotions could result in a more positive experience from female consumers, which in turn will impact their intention towards using such mobile devices and wireless data services.

6.2.2.4 Adoption

Of particular interest is the finding that consumers are more likely to place emphasis on efficiency than on effectiveness when it comes to personal purchase decisions for a mobile device. Dimensions like learnability, ease of use, and time required to complete a task are prevalent in the decision making process of a mobile device for wireless data services, while successful task completion seems to be less relevant for a significant portion of the population.

In addition, the designed Web task used in our study was a fair representative of the cognitive demands associated with goal-oriented searches performed by mobile device users. A few examples of Web-related tasks requiring a similar level of cognitive demand include searching for directions, locating needed "last minute" (timely) information, downloading coupons that can be scanned at stores' registers (Taylor 2004), and even mobile blogging. M-blogging, as it is commonly known as, refers to tasks on mobile devices, where users navigate to particular websites to enter real-time information about their lives. This information often includes photos or videos taken by users on the go. Other Web applications include navigating to specific websites to download ring tones and wallpapers, to receive Web-based promotions in real time and based on the users location via text messages (through selectable Web links), or to perform financial transactions, such as balance updates or fund transfers in just a few clicks. Although it is uncommon for consumers in North America to perceive the value of such mobile tasks, mobile consumers in Europe and Asia have a different opinion (Taylor 2004). For example, with response rates reaching up to eight percent, companies can in turn leverage Web-based promotions delivered via text messages to boost their bottom line. What all of these examples have in common is that users navigated the Web with specific goals, and after a short series of clicks they were able to retrieve the needed product or service (e.g. information). The experience is so simple that it can occur even while walking, particularly for the 12-24 year old youth market (OneUpWeb 2005).

6.3 Limitations

To successfully solicit the sample size required to complete this study, the test subjects were university students, staff, and/or faculty. This constraint reduces the generalizability regarding individuals' intention to use mobile devices for wireless data services. However, controlled randomization of the subjects facilitated the collection of a sample that covers the entire spectrum in terms of age, gender, and education.

Furthermore, the tasks were simulated in a laboratory setting. As such, any sense of urgency (or other contextual emotional response) that a user might experience in a real-setting did not arise here, other than those triggered by mobility, the visual and auditory environment. While this is a limitation in terms of the realism of the study, it is a means of controlling for additional variables that could not be otherwise measured during the experiment. Consequently, the generalizability of our results is reduced regarding the emotional context of use, but maintains high internal validity. In addition, the experiment was carried out at a single point in time and is therefore limited in its strength to evaluate usability over a period of time.

According to the definition of usability used in this dissertation, context of use is a critical factor. Therefore, if the same input interfaces are to be tested for a different application other than the tasks tested, or if the tasks are to be completed by other input-interface types than the built-on QWERTY keyboard, then the findings from the study will need to be validated. The same limitation applies for testing the impact of the user's mobility and the visual environment on other input-interfaces and/or applications.

Finally, this research did not delve into such important issues as task complexity and task interactivity, as they were beyond its intended scope.

6.4 Future research

Given the interdisciplinary nature of this dissertation, several avenues may be explored for future research. The first way would be to conduct a study that collects both device- and self-reported data regarding the efficiency and effectiveness of mobile devices at the task level. This study would then compare these data sets so that conclusions could be drawn for each task (or application) independently. Thus, a closer attention between the measurement of perceived and device data is needed. A more focused investigation for performance-related attributes (e.g. time, learnability, error rate, success rate, etc.) is warranted. By conducting mobile usability studies that triangulate data between self-reported, observed, and device data, richer and closer to real-world analyses would be feasible.

Another avenue for future research would be to study the impact of expectations on satisfaction formation, subject to the level of performance and potentially also to the level of normative expectations (i.e. how a user expects the device should - not would - perform) (Schommer 2000).

Revisiting the findings of our study in terms of EDT, there were two types of links that did not demonstrate statistical significance: i) Expectations to Disconfirmation, and ii) Expectations to Satisfaction. Future research could aim to explain the varied impact of expectations of performance on disconfirmation of expectations of performance and on satisfaction with performance. Also, since both the direct path from performance to satisfaction, and the mediated path by disconfirmation, received significant support, usability research should explore the validity of each approach.

From the finding that gender influences the perceived effectiveness of a mobile device for wireless data services, future research could further explore the impact of gender on expectations of mobile devices. If women have lower initial expectations than men do, then the differences found here in terms of perceived effectiveness would be further supported.

Finally, a recommendation for future research, and based on the findings in this dissertation, future mobile usability studies should measure the performance dimensions of efficiency and effectiveness through discriminant constructs.

REFERENCES

Agarwal, R., Prasad, J. and Zanino, MC (1996). Training experiences and usage intentions: a field study of a graphical user interface. International Journal of Human-Computer Studies, 45, 215-241.

Agranoff, M.H. (1993). Controlling the Threat to Personal Privacy. Journal of Information Systems Management, Summer.

Ameri, A. (2004). The Five Pillars of Information Security. Risk Management, NY, July 2004, Vol. 51, 7, pg 48-49.

Anderson, J.L. and Maheswaran, D. (1997). The effect of cultural orientation on persuasion. Journal of Consumer Research, Vol. 24, December, pp. 315-28.

Anderson, R. J., Heath, C. C., Luff, P., and Moran, T. P. (1993). The social and the cognitive in human-computer interaction. International Journal of Man-Machine Studies, 38, 999-1016.

Andon, C. (2004). Usability analysis of wireless tablet computing in an academic emergency department. Capstone project, School of Medicine, Oregon Health & Science University.

Armstrong, L., Phillips, J.G., & Saling, L.L. (2000). Potential determinants of heavier Internet usage. International Journal of Human-Computer Studies, 53, 537-550.

Arriba, M. and Brugos, J. (2002). Development of Adaptive Web Sites with Usability and Accessibility Features. Proceedings from Adaptive Hypermedia (AH) 2002, LNCS 2347, 501-504.

Baber, C. and Mellor, B. (2001). Using critical path analysis to model multimodal human-computer interaction. International Journal of Human-Computer Studies, 54, 613-636.

Baddeley, A. (1986). Working Memory. London: Oxford, University Press.

Bagozzi, R., Yi, Y. & Philips, L. (1991). "Assessing Construct Validity in Organizational Research", Administrative Science Quarterly, 36, 421-458.

Baker, M.A. and Holding, D.H. (1993). The Effects of Noise and Speech on Cognitive Task Performance. Journal of General Psychology, 120 (July), 339-355.

Bakos, Y. and Brynjolfson, E. (1999). Bundling Information Goods: Pricing, Profits, and Efficiency. Management Science, 45(12), 1613-1650.

Banavar, G. and Bernstein, A. (2002). Software Infrastructure and Design Challenges for Ubiquitous Computing Applications. Communications of the ACM, 45(12), 92-96.

Banbury, S.P., Macken, W.J., Tremblay, S., Jones, and D.M. (2003). Auditory distraction and short term memory: phenomena and practical implications. Human Factors, Vol. 43, No. 1, pp. 12-29.

Bass, F., Gordon, K., Ferguson, T.L., and Githens, M.L. (2001). DIRECTV: Forecasting Diffusion of a New Technology Prior to Product Launch. Interfaces, 31:3, Part 2 of 2, May-June 2001, pp. S82-S93.

Belanger, F., Hiller, J., and Smith, W.J. (2002). Trustworthiness in electronic commerce: the role of privacy, security, and site attributes. Journal of Strategic Systems, Vol. 11, 245-270.

Benbasat, I. and Todd, P. (1993). An experimental investigation of interface design alternatives: icon vs. text and direct manipulation vs. menus. International Journal of Man-Machine Studies, 38, 369-402.

Bennett, J.L. (1986). Tools for building advanced user interfaces. IBM Systems Journal, 25, 3/4, 354-368.

Bentley, R., Hughes, J. A., Randall, D., Rodden, T., Sawyer, P., Shapiro, D., and Sommerville, I. (1992). Ethnographically-informed system design for air-traffic control. Proceedings of CSCW'92: Computer Supported Cooperative Work, Toronto, pp.123-129. New York: Association for Computing Machinery.

Bernstein, A. (2000). How can cooperative work tools support dynamic group processes? Bridging the specificity frontier. In Proceedings of the Computer Supported Cooperative Work Conference. ACM Press, 2000.

Bevan, N. (2001). International standards for HCI and usability. International Journal of Human-Computer Studies, 55, 533-552.

Bevan, N. and Macleod, M. (1994). Usability measurement in context. Behaviour and Information Technology, 13, 132-145.

Bhattacherjee, A. (2001). An empirical analysis of the antecedents of electronic commerce service continuance. Decision Support Systems, Vol. 32, 2, pg. 201-214.

Bhattacherjee, A. and Premkumar, G. (2004). Understanding Changes in Belief And Attitude Toward Information Technology Usage: A Theoretical Model and Longitudinal Test. MIS Quarterly, Vol. 28, 2, June 2004, pg. 229-254.

Bieber, M., Vitali, F., Ashman, H., Balasubramanian, V. and Oinas-Kukkonen, H. (1997). Fourth generation hypermedia: some missing links for the World Wide Web. International Journal of Human-Computer Studies, 47, 31-65.

Blandford, A. and Rugg, G. (2002). A case study on integrating contextual information with analytical usability evaluation. International Journal of Human-Computer Studies, 57, 75-99.

Borgman, C. (1986). The user's mental model of an information retrieval system: an experiment on a prototype on-line retrieval catalog. International Journal of Man-Machine Studies, 24(1), 44-47.

Branscum, D., (2000). Guarding on-line privacy. Newsweek, 135(23), 77-78.

Brereton, E. (2005). Don't neglect usability in the total cost of ownership. Communications of the ACM, 47 (7), July 7, 10-11.

Brewster, S. (2002). Overcoming the Lack of Screen Space on Mobile Computers. Personal and Ubiquitous Computing, 6, 188-205.

Brooke, J. (1996). SUS: A Quick and Dirty Usability Scale. In P. W. Jordan, B. Thomas, B. A. Weerdmeester and I. L. McClelland (eds). Usability Evaluation in Industry, 189-194. London, UK, Taylor & Francis.

Bruijn, O. D., Spence, R., Chong, M. Y. (2002). Rsvp browser: Web browsing on small screen devices. Personal and Ubiquitous Computing, 6(4), 245-252.

Bruni, S.(2004). The role of Instant Messaging on Task Performance and Level of Arousal. Massachusetts Institute of Technology, Term Project, mas.630 and 16.422, Spring term.

Bruseberg, A., and McDonagh-Philp, D. (2001). New products development by eliciting user experience and aspirations. International Journal of Human-Computer Studies, 55, 435-452.

Buchanan, G., Farrant, S., Jones, M., Thimbleby, H.W., Marsden, G., and Pazzani, M.J. (2001). Improving Mobile Internet Usability. WWW10, May 1-5, 2001, Hong Kong, China.

Burnett, G. and Porter, J. (2001). Ubiquitous computing within cars: designing controls for non-visual use. International Journal of Human-Computer Studies, 55, 521-531.

Bushey, R., Manuney, J.M., and Deelman T. (1999). The development of behaviour-based user models for a computer system, In: Proceedings of the Seventh International Conference on User Modeling, Banff, June 20-24, 1999; SpringerWien New York, pp.109-118.

Butts, L. and Cockburn, A. (2002). An Evaluation of Mobile Phone Text Input Methods. Proceedings of the Third Australasian User Interfaces Conference (AUIC2002), Melbourne, Australia.

Buyukkokten, O., Garcia-Molina, H., Paepcke, A. (2001). Seeing the Whole in Parts: Text Summarization for Web Browsing on Handheld Devices. 10th International WWW Conference, Hong Kong, May 1-5, 2001.

Buyukkokten, O., Garcia-Molina, H. & Paepcke, A. [2001], Accordion Summarization for End-game Browsing on PDAs and Cellular Phones, in J. A. Jacko & A. Sears (eds.), Proceedings of CHI'01 Conference on Human Factors in Computing Systems, CHI Letters 3(1), ACM Press, pp.213–20.

Buyukkokten, O., Kaljuvee, O., Garcia-Molina, H., Paepcke, A., and Winograd, T. (2002). Efficient Web Browsing on Handheld Devices Using Page and Form Summarization. ACM Transactions on Information Systems, Vol.20(1), January 2002, 82-115.

Calder, B.J., Philips, L.W., and Tybout, A.M. (1981). Designing Research for Application. Journal of Consumer Research (8:2), pp. 197-207.

Campbell, D.J. (1988). Task Complexity: A Review and Analysis. Academy of Management Review, 13(1), 40-52.

Campbell, D. & Friske, D. (1967). Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 56, 81-105.

Camponovo, G. and Cerutti, D. (2004). The spam issue in mobile business: a comparative regulatory overview. Proceedings of the International Conference on Mobile Business (m-Business 2004), 12–13 July, New York, USA, Electronic Proceedings

Camponovo, G. and Pigneur, Y. (2003). Analyzing the actor game in m-business. The First International Conference on Mobile Business. Conference Proceedings. July 8-9, 2002.

Carroll, J. (1997). Human-computer interaction: psychology as a science of design. International Journal of Human-Computer Studies. 46, 501-522.

Carroll, J.M. and Rosson, M.B. (1991). Deliberated evolution: stalking the view matcher in design space. Human-Computer Interaction, 6, 281-318.

Chau, P. and Hui, K.L. (1998). Identifying early adopters of new IT products: A case of Windows 95. Information & Management, Vol. 33, 225-230.

Chan, S., Fang, X., Brzezinski, J., Zhou, Y., Xu, S., and Lam, J. (2002). Usability for mobile commerce across multiple form factors. Journal of Electronic Commerce Research, Vol. 3(3), pp.187-199.

Chen, D. and Vertegaal, R. (2004). Using Mental Load for Managing Interruptions in Physiologically Attentive User Interfaces. Proceedings of CHI 2004, April 24-29, 2004, Vienna, Austria, pp. 1513-1516.

Chen, L., Gillenson, M.L., and Sherrell, D.L. (2002). Enticing online consumers: an extended technology acceptance perspective. Information & Management, Vol. 39, 705-719.

Chen, Y., Ma, W.-Y. and Zhang, H. (2003). Detecting Web Page Structure for Adaptive Viewing on Small Form Factor Devices, in G. Hencsey, B. White, L. Kovács & S. Lawrence (eds.), Proceedings of the Twelfth International World Wide Web Conference (WWW2003), ACM Press, pp.225–33. Also available as a Technical Report from Microsoft Research, MSR-TR-2002-120.

Cheverst, K., Davies, N., and Mitchell, K. (2002). Exploring context-aware information push. Personal and Ubiquitous Computing, 6, 276-281.

Cheverst, K., Davies, N., Mitchell, K., Friday, K., and Efstratiou, C. (2000). Developing a Context-Aware Electronic Tourist Guide: Some issues and Experiences. Proceedings of CHI'00 (The Hague, Netherlands: ACM), pp. 17-24.

Chin, W.W. (2001). PLS-Graph User's Guide, Version 3.0". Soft Modeling Inc.

Chin, W.W. (1998). The Partial Least Squares Approach for Structural Equation Modeling. Lawrence Erlbaum Associates, 1998, pg. 295-336.

Chin, W.W. (1994). PLS Graph Manual. Unpublished Manuscript.

Chin, W.W. and Gopal, A. (1995). Adoption Intention in GSS: Importance of Beliefs. Data Base Advance, Vol 26(2&3), pg. 42-64.

Chin, W. W., and Newsted, P. R. (1999). Structural Equation Modeling analysis with Small Samples Using Partial Least Squares. In Rick Hoyle (Ed.), Statistical Strategies for Small Sample Research, Sage Publications, pp. 307-341.

Chincholle, D., Chincholle, D., Goldstein, M., Nyberg, M., and Eriksson, M. (2002). Lost or Found? A Usability Evaluation of a Mobile Navigation and Location-Based Service. Proceedings from Mobile HCI 2002, LNCS 2411, 211-224.

Chittaro, L. and Dal Cin, P. (2002). Evaluating Interface Design Choices on WAP Phones: Navigation and Selection. Personal and Ubiquitous Computing, 6, 237-244.

Churchill, G.A. Jr and Suprenant, C. (1982). An investigation into the determinants of consumer satisfaction. Journal of Marketing Research, Vol. 24, August, pp. 305-14.

Cockburn, A. and Jones, S. (1996). Which way now? Analysing and easing inadequacies in WWW navigation. International Journal of Human-Computer Studies, 45, 105-129.

Cockburn, A. and McKenzie, B. (2001). What do web users do? An empirical analysis of web use. International Journal of Human-Computer Studies, 54, 903-922.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). New York: Academic Press.

Collins, C., Caputi, P., Rawstorne, P., and Jayasuriya, R. (1999). Correlates of End-User Performance and Satisfaction with the Implementation of a Statistical Software Package. Proceedings of the 10th Australasian conference on Information Systems, 1999.

Compeau, D.R., and Higgins, C.A. (1995). Application of Social Cognitive Theory to Training for Computer Skills. Information Systems Research, Vol 6(2), pg. 118-143.

Compeau, D.R., and Higgins, C.A. (1991). The Development of a Measure of Computer Self-Efficacy. Proceedings of the ASAC 1991, Niagara Falls, NY, 1991a, 34-48.

Conklin, N. (2001). Information Visualization Publication Ranking. Information Visualization Course CS5984.

Corritore, C., Kracher, B., and Wiedenbeck, S., (2003). On-line trust: concepts, evolving themes, a model. International Journal of Human-Computer Studies, 58, 737-758.

Coursaris, C., and Hassanein H. (2002). Understanding m-Commerce: A Consumer-Centric Model. Quarterly Journal of Electronic Commerce, 3(3), 247-272.

Coursaris, C., Hassanein, K., and Head, M. (2003). M-Commerce in Canada: An interaction framework for wireless privacy. Canadian Journal of Administrative Sciences, 20(1), 54-73.

Crease, M. (2005). Measuring the Effect of Distractions on Mobile User's Performance. Measuring Behavior 2005: The 5th International Conference on Methods and Techniques in Behavioural Research. Wageningen, The Netherlands, August 30-September 2, 2005. NRC 48215.

Cronin, J.J. Jr. and Taylor, S.R. (1992). Measuring service quality: a re-examination and extension. Journal of Marketing, Vol. 19, November, pp. 491-504.

CWTA (2006). Mobile Originated Text Messages in Canada. Canadian Wireless Telecommunications Association. Retrieved from http://www.txt.ca/facts.htm# on April 17, 2005.

Damper, R. I., Tranchant, M. A. and Lewis, S. M. (1996). Speech versus keying in command and control: effect of concurrent tasking. International Journal of Human-Computer Studies, 45, 337-348.

Davis, F.D., Bagozzi, R.P. and Warshaw, P.R. (1989). User acceptance of computer technology: a comparison of two theoretical models. Management Science, 35(8), 982-1003.

Davis, F.D. (1993). User acceptance of information technology: system characteristics, user perceptions and behavioural impacts. International Journal of Man-Machine Studies, 38, 475-487.

Davis, F.D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Quarterly, 13, September (1989), 319-339.

Dayal, S., Landesberg, H., and Zeisser, M. (1999). How to build trust online. Marketing Management 8(3), 64-71.

De Bruijn, O., Spence, R. and Chong, M. Y. (2002). RSVP Browser: Web Browsing on Small Screen Devices,. Personal and Ubiquitous Computing, 6(4), 245–52.

De Carolis, B., DeRosis, F., and Errore, S. (1995). A user-adapted iconic language for the medical domain. International Journal of Human-Computer Studies, 43, 561-577.

De Rosis, F., Pizzutilos, S., and DeCarolis, B. (1998). Formal description and evaluation of user-adapted interfaces. International Journal of Human-Computer Studies, 49, 95-120.

DeGroot, B. and Welie, M. (2002). Leveraging the Context of Use in Mobile Service Design. Proceedings from Mobile HCI 2002, LNCS 2411, 334-338.

Dewan, S., and Mendelson, H. (1990). User delay costs and internal pricing for a service facility. Management Science, 36(12), 1502-1517.

Dillon, A. (2003). Beyond Usability: process, outcome and affect in human computer interactions. Canadian Journal of Information Science 26 (4), pp. 57-69.

Dillon, T. and Norcio, A. (1997). User performance and acceptance of a speech-input interface in a health assessment task. International Journal of Human-Computer Studies, 47, 591-602.

Dishaw, M.T. and Strong, D.M. (1999). Extending the technology acceptance model with task-technology fit constructs. Information and Management, 36(1), 9-21.

Dray, S., Siegel, D., Feldman, E., and Potenza, M. (2002). Why Do Version 1.0 and Not Release It? Conducting Field Trials of the Tablet PC. Interactions, March-April 2002, 11-16.

Duda, S., Schießl, M., Hess, J.M. (2000). Mobile usability. Mobile Economy GmbH.

Dunlop, M. D. and Crossan, A. (2000). Predictive text entry methods for mobile phones. Personal Technologies, 4(2):134-143.

Dyson, M. and Haselgrove, M. (2001). The influence of reading speed and line length on the effectiveness of reading from screen. International Journal of Human-Computer Studies. 54, 585-612.

Ebling, M., John, B.E., and Satyanarayanan, M. (2002). The Importance of Translucence in Mobile Computing Systems. ACM Transactions on Computer-Human Interaction, Vol.9(1), March 2002, 42-67.

Elliott, A. and O'Donohue, W.T. (1997). The effects of Anxiety and Distraction on Sexual Arousal in a Nonclinical Sample of Heterosexual Women. Archives of Sexual Behaviour, 26 (6), 607-624.

Engel, F. and Haakma, R. (1993). Expectations and feedback in user-system communication. International Journal of Man-Machine Studies, 39, 427-452.

Eysenck, M.W. (1984). A Handbook of Cognition Psychology. London: Lawerence Erlbaum Associates, Ltd. Publishers, 1984.

Fang, X., Chan, S., Brzezinski, J., and Xu, S. (2003). A Study of Task Characteristics and User Intention to Use Handheld Devices for Mobile Commerce. Proceedings of the Second Annual Workshop on HCI Research in MIS, Seattle, WA, December 12-13, 2003.

Fang, X. and Salvendy, G. (2000). Keyword comparison: a user-centered feature for improving web search tools. International Journal of Human-Computer Studies, 52, 915-931.

Fazio, R.H. (1995). Attitude accessibility and motivation as determinants of biased processing: A test of the MODE model. Personality and Social Psychology Bulletin, 21(7), 704-410.

Fazio, R.H. and Zanna, M.P. (1978). Attitudinal qualities relating to the strength of the attitude-behaviour relationship. Journal of Experimental Social Psychology, 14, 398-408.

Finkelstein, L., Gabrilovich, E., Matias, Y., Rivlin, E., Solan, Z., Wolfman, G., and Ruppin, E. (2002). Placing Search in Context: The Concept Revisited. ACM Transactions on Computer-Human Interaction, Vol.20(1), January 2002, 116-131.

Fishbein, M. and Ajzen, I. (1975). Belief, Attitude, Intentions, and Behaviour: An Introduction to Theory and Research, Addison-Wesley, Boston.

Fockert, J.W., Rees, G., Frith, D.C., and Lavie, N. (2001). The Role of Working Memory in Visual Selective Attention. Science 2 March 2001: 1803-1806.

Folinas, D., Vlachopoulou, M., Manthou, V. (2003). The value system of m-business. The First International Conference on Mobile Business. Conference Proceedings. July 8-9, 2002.

Fornell, C., and Larcker, D.F. (1981). Evaluating structural equation models with unobserved variables and measurement error. Journal of Marketing Research, Vol. 18, 1981, pg. 39-50.

Franken, R.E., *Human Motivation* 5ed, Wadsworth, United States, 2002, pp84-104.

Freudenthal, A. and Mook, H. (2003). The evaluation of an innovative intelligent thermostat interface: universal usability and age differences. Cognition, Technology & Work, Special Issue on Household Technologies, 5, 55-66.

Frokjaer, E., Hertzum, M., and Hornbaek, K. (2000). Measuring Usability: Are Effectiveness, Efficiency, and Satisfaction Really Correlated? Proceedings of the ACM CHI 2000 Conference on Human Factors in Computing Systems (The Hague, The Netherlands, April 1-6, 2000), pp. 345-352. ACM Press, New York.

Furnham, A. and Strbac, L. (2002). Music is as distracting as noise: the differential distraction of background music and noise on the cognitive test performance of introverts and extraverts. Ergonomics, February 2002, Vol. 45, No. 3, pp. 203-17.

Gallaugher, J., Auger, P., and BarNir, A. (2001). Revenue Streams and digital content providers: an empirical investigation. Information & Management, Vol. 38, 473-485.

Galliers, R.D. and Land, F.E. Choosing appropriate information system research methodologies, Communications of the ACM, Nov. 1987, pp. 900-902.

Gardner, C. and Amoroso, D.L. (2004). Development of an instrument to measure the acceptance of Internet Technology by Consumers. Proceedings of the 37th Hawaii International Conference on System Sciences, 2004.

Gebauer, J., and Shaw, M.J. (2004). Success factors and impacts of mobile business applications: results from a mobile e-procurement study. International Journal of Electronic Commerce, Vol. 8 No.3, pp.19-41.

Gebauer, J., Shaw, M. and Gribbins, M. (2004). Usage and Impact of Mobile Business Applications - An Assessment Based on the Concepts of Task/Technology Fit. The 10th Americas Conference on Information Systems (AMCIS), New York, New York, August 5-8, 2004.

Gebhart, J. (1994). Who Owns Information? From Privacy to Public Access; Managing Privacy: Information Technology and Corporate America. MIT Sloan Management Review, 36(1), 12-15.

Gefen, D. (2003). Tam or Just Plain Habit: A Look at Experienced Online Shoppers. International Journal of End User Computing, 15(3).

Gelderman, M. (1998). The relation between user satisfaction, usage of information systems and performance. Information & Management, 34, 11-18.

Gentner, D.R., Grudin, J.T., Larochelle, S., Norman, D.A., and Rumelhart, D.E. (1983). A glossary of terms including a classification of typing errors. Cognitive aspects of skilled typing. Cooper, W.E., Ed. New York: Springer-Verlag. 39-44.

Good, M., Spine, T.M., Whiteside, J. and George, P. (1986). User-derived impact analysis as a tool for usability engineering. Proceedings for CHI'86 Conference on Human Factors in Computing Systems, Boston, pp. 241-246. New York: Association for Computing Machinery.

Gosbee, J. and Ritchie, E. (1997). Human-Computer Interaction and Medical Software Development. Interactions, July-August, 13-18.

Gotlieb, J.B., Grewal, D. and Brown, S.W. (1994). Consumer satisfaction and perceived quality: complementary or divergent constructs? Journal of Applied Psychology, Vol. 79, No. 6, pp. 875-85.

Green, H., Yang, C., and Judge, P.C. (1998). A little privacy, please. Business Week 3569, 98-99.

Gugerty, L. (1993). The use of analytical models in human-computer-interface design. International Journal of Man-Machine Studies, 38, 625, 660.

Hair, J.F., Anderson, R.E., Tatham, R.L., and Black, W.C. (1995). Multivariate Data Analysis with Readings, 4th ed., Prentice-Hall, Englewook Cliffs, NJ, 1995.

Hall, R. (2001). Prototyping for usability of new technology. International Journal of Human-Computer Studies, 55, 485-501.

Harbluk, J.L. and Noy, Y.I. (2002). The impact of cognitive distraction on driver visual behaviour and vehicle control. TP# 13889 E, Transport Canada. Retrievd on February 24, 2006, from http://www.tc.gc.ca/roadsafety/tp/tp13889/pdf/tp13889es.pdf

Harris Interactive (2001), Consumer privacy attitudes and behaviours survey wave II, The Privacy Leadership Initiative, July 2001, http://www.understandingprivacy.org/content/library/harris2-execsum.pdf

Hart, S.G., & Staveland, L.E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P.A. Hancock and N. Meshkati (Eds.) Human mental workload (pp.139-183). Amsterdam: North-Holland.

Hassan, S. and Li, F. (2005). Evaluating the Usability and Content Usefulness of Web Sites: A Benchmarking Approach. Journal of Electronic Commerce in Organizations. Hershey: Apr-Jun 2005, Vol.3, Iss. 2, pg. 46-68.

Hassanein, K., and Head, M. (2003). Ubiquitous Usability: Exploring Mobile Interfaces within the Context of a Theoretical Model. Proceedings of the Ubiquitous Mobile Information and Collaboration Systems Workshop (UMICS 2003), The 15th Conference On Advanced Information Systems Engineering (CAiSE), Velden, Austria, June.

Hawk, S. (1993). The effects of user involvement: some personality determinants. International Journal of Man-Machine Studies, 38, 839-855.

Henderson, R., Mahar, D., Saliba, A., Deane, F., and Napier, R. (1998). Electronic monitoring systems: an examination of physiological activity and task performance within a simulated keystroke security and electronic performance monitoring system. International Journal of Human-Computer Studies, 48, 143-157.

Henderson, S. and Snyder, C. (1999). Personal information privacy: implications for MIS managers. Information & Management, Vol. 36, 213-220.

Henry, J.W. and Stone, R.W. (1995). Boosting empowerment by developing self-efficacy. Asia Pacific Journal of Human Resources, 37, 1, 52-63.

Hermann, F. and Heidmann, F. (2002). User Requirement Analysis and Interface Conception for a Mobile, Location-Based Fair Guide. Proceedings from Mobile HCI 2002, LNCS 2411, 388-392.

Hodkinson, C., Kiel, G., and Mccoll-Kennedy, J.R. (2000). Consumer web search behaviour: diagrammatic illustration of wayfinding on the web. International Journal of Human-Computer Studies, 52, 805-830.

Hoffman, D.L., Novak, T.P., and Peralta, M. (1999). Building consumer trust online. Communications of the ACM 42(4), 80-85.

Hong, C.C. and Hiew, P. L. (2005). Factors Influencing the Adoption of Mobile Entertainment: Empirical Evidence from a Malaysian Survey. IEEE 4th International Conference on Mobile Business, Sydney.

Hornbaek, K., Bederson, B.B., and Plaisant, C. (2003). Navigation Patterns & Usability of Zoomable User Interfaces With and Without an Overview. ACM Transactions of Computer Human Interaction (ToCHI), January-February 2003, 11-12.

Hughes, R.W. and Jones, D.M. (2001). The intrusiveness of sound: Laboratory findings and their implications for noise abatement. Noise Health, Vol. 4, No. 13, pp. 51-70.

Hughes, R.W. and Jones, D.M. (2003). Indispensable benefits and unavoidable costs of unattended sound for cognitive functioning. Noise Health, Oct-Dec, Vol. 6, No. 21, pp. 63-76.

Irani, T. (2000). Prior Experience, Perceived Usefulness, and the Web: Factors Influencing Adoption of Internet Communication Tools. Proceedings of the Southern Association of Agricultural Scientists, Lexington, KY, January.

Islam, N. and Fayad, M. (2003). Toward Ubiquitous Acceptance of Ubiquitous Computing. Communications of the ACM, 46(2), 89-92.

ISO-9241 (1998). Ergonomic requirements for office work with visual display terminals (VDTs) -- Part 11: Guidance on usability.

Kahneman, D. (1973). Attention and Effort. Englewood Cliffs, Prentice Hall.

Kaplan, D. (2000). Structural equation modeling: Foundations and extensions. Thousand Oaks: Sage Publications.

Jaaksi, A. (2002). Developing Mobile Browsers in a Product Line. IEEE Software, July-August, 73-80.

Jain, S. and Kannan, P. (2002). Pricing of Information Products on Online Servers: Issues, Models, and Analysis. Management Science, 48(9), 1123-1142.

James, C. L. & Reischel, K. M. (2001). Text input for mobile devices: comparing model prediction to actual performance. Proceedings of CHI2001, ACM, New York, pg. 365-371.

Jarvenpaa, S., Lang, K., and Tuunainen, V.K. (2004), Manifestations of Technology Paradoxes and Implications on the Experience of Mobile Technology Users. Mobile Roundtable, Austin, TX, USA March 11-12, 2004.

Jarvenpaa, S. L. and Todd, P. A. (1997). Is there a future for retailing on the Internet?. In Peterson, R. A. (ed.) Electronic Marketing and the Consumer. Thousand Oaks.

John, B.E. (1990). Extensions of GOMS analyses to expert performance requiring perception of dynamic visual and auditory information. Proceedings of CHI'90 Human Factors in Computing Systems, Seattle, pp. 107-115. New York: Association for Computing Machinery.

Jones, D.M., Alford, D., Macken, W.J., Banbury, S.P., and Tremblay, S. (2000). Interference from degraded auditory stimuli: linear effects of changing-state in the irrelevant sequence. Journal of Acoustics Society of America, September 2000, Vol. 108(3 Pt. 1), pp. 1082-8.

Jones, S., Wilikens, M., Morris, P. and Masera, M. (2000). Trust requirements in e-Business. Communications of the ACM, Vol. 43(12), pg. 80-87.

Jones, M., Marsden, G., Mohd-Nasir, N., and Boone, K. (1999). Improving Web Interaction on Small Displays. Computer Networks 31(11–6), 1129–37. Proceedings of the Eighth International World Wide Web Conference.

Juola, J. a. V., D. (2004). First time usability testing for bluetooth-enabled devices (Technical Report No. ITTC-FY2005-TR-35580-02): The University of Kansas.

Jue, H.Y. (2004). Analysis of SMS Efficiency. Department of Computer Science, School of Computing, National University of Singapore, 2004/2004.

Kaasinen, E. (2003). User needs for location-aware mobile services. Personal and Ubiquitous Computing. 7, 70-79.

Kahneman, D. (1973). Attention and Effort. Englewood Cliffs: Prentice Hall.

Kai, A., Nakano, T., and Nakagawa, S. (2000). Usability of Browser-Based Pen-touch / Speech User Interfaces for Form-Based Applications in Mobile Environment. Proceedings from ICMI 2000, LNCS 1948, 549-556.

Kaikkonen, A., Kallio, T., Kekäläinen, A., Kankainen, A., Cankar, M. (2005). Usability testing of mobile applications: A comparison between laboratory and field testing. Journal of Usability studies, 1(1), 4-16.

Kaikkonen, A. and Roto, V. (2002). XHTML in Mobile Application Development. Proceedings from Mobile HCI 2002, LNCS 2411, 344-348.

Kalakota, R. and Whinston, A.B., (1996). Frontiers of Electronic Commerce, Addison-Wesley, Reading, MA.

Kao, D. and Archer, N. (1997). Abstraction in conceptual model design. International Journal of Human-Computer Studies, 46, 125-150.

Kallinen, K. (2004). The effects of Background Music on Using a Pocket Computer in a Cafeteria: Immersion, Empotional Responses, and Social Richness of Medium. CHI 2004, April 24-29, 2004, Vienna, Austria.

Karnouskos, S. and Vilmos, A. (2004). The European Perspective on Mobile Payments. In: IEEE SympoTIC'04 Proceedings, IEEE Symposium on Trends in Communications (SympoTIC '04), Bratislava, SLO, 24-26 October 2004.

Karnouskos, S., Hondroudaki, A., Vilmos, A., and Csik, B. (2004). Security, Trust and Privacy in the Secure Mobile Payment Service. Third International Conference on Mobile Business 2004 (m-Business), New York City, USA, 12-13 July 2004.

Karahanna, E. and Straub, D. (1999). The psychological origins of perceived usefulness and ease of use. Information & Management, 35, 237-250.

Karampelas, P., Akoumianakis, D., and Stephanidis, C. (2003). User Interface Design for PDAs: Lessons and Experience with the WARD-IN-HAND Prototype. Proceedings from User Interfaces for All, LNCS 2615, 474-485.

Kaynama, S. A (2000). A proposal to assess the service quality of online travel agencies: An exploratory study. Journal of Professional Services Marketing. 21(1), 63-88.

Kerlinger, F.N. Foundations in Behavioural Research. Holt, Rinehart, and Winston. New York, NY, 1964.

Khalifa, M. and Cheng, S.K.N. (2002). Adoption of Mobile Commerce: Role of Exposure. Proceedings of the 35th Hawaii International Conference on System Sciences.

Khalifa, M. and Liu, V. (2002). Satisfaction with Internet-Based Services. Proceedings of the 35th Hawaii International Conference on System Sciences.

Kieras, D., and Poison, P. (1985) An approach to the formal analysis of user complexity. International Journal of Man-Machine Studies, 22, 365-394.

Kieras, D.E. (1988). Towards a practical GOMS model methodology for user interface design. In M. Helander, Ed. Handbook of Human-Computer Interaction, pp. 135-157. Amsterdam: North-Holland.

Kim, K., Kim, J., Lee, Y., Chae, M., and Choi, Y. (2002). An Empirical Study of the Use Contexts and Usability Problems in Mobile Internet". Proceedings of the 35th Annual Hawaii International Conference on System Sciences, Los Alamitos, CA: IEEE Computer Society Press.

Kjeldskov, J. and Graham, C. (2003). A Review of MobileHCI Research Methods. Proceedings of the 5th International Mobile HCI 2003 conference. Udine, Italy, 2003.

Kjeldskov, J., Graham, C., Pedell, S., Vetere, F., Howard, S., Balbo, S., and Davies, J. (2005). Evaluating the Usability of a Mobile Guide: The Influence of Location, Participants and Resources. Behaviour and Information Technology 24(1), 51-65.

Kjeldskov, J. and Stage, J. (2004). New Techniques for Usability Evaluation of Mobile Systems. International Journal of Human-Computer Studies, 60, 599-620.

Kohno, M. and Rekimoto, J. (2002). New Generation of IP-Phone Enabled Mobile Devices. Proceedings from Mobile HCI 2002, LNCS 2411, 319-323.

Krishnan, T.V., Bass, F.M., Jain, D.C. (1999). Optimal Pricing Strategy for New Products. Management Science, 45(12), 1650-1663.

Kwahk, J. and Han, S. (2002). A methodology for evaluating the usability of audiovisual consumer electronic products. Applied Ergonomics, 33, pp. 419-431.

Lacohee, H. and Anderson, B. (2001). Interacting with the telephone. International Journal of Human-Computer Studies, 54, 665-699.

Lamberts, H. (2002). Case Study: A PDA Example of User Centered Design. Proceedings from Mobile HCI 2002, LNCS 2411, 329-333.

Lau, T., Etzioni, O, and Weld, D. (1999). Privacy Interfaces For Information Management. Communications of the ACM, 42(10), 89-94.

Lee, Y.E., and Benbasat, I. (2003). A framework for the study of customer interface design for mobile commerce. International Journal of Electronic Commerce, 46(12), 48-52.

Lee, A. and Girgensohn, A. (2002). Design, experiences and user preferences for a web-based awareness tool. International Journal of Human-Computer Studies, 56, 75-107.

Lee, E. and Jang, S. (2002). A User Adaptive Mobile Commerce System with a Middlet Application. Proceedings from IDEAL 2002, LNCS 2412, 193-199.

Lee, M.K.O. and Turban E. (2001). A trust model for internet shopping. International Journal of Electronic Commerce 6(1), 75-91.

Lee, W.J., Kim, T.U., Chung, J.Y. (2002). User acceptance of the mobile Internet. The First International Conference on Mobile Business. Conference Proceedings. July 8-9, 2002

Legris, P., Ingham, I., and Collerette, P. (2003). Why do people use information technology? A critical review of the technology acceptance model. Information & Management, 40, 3, 191-204.

Lehrer, M. (2004). National lead markets and the design competition for 3G network applications. Journal of Business Research, Volume 57, Issue 12, December, pp. 1397-1401.

Leporini, B. and Paterno, F. (2003). Criteria for Usability of Accessible Web Sites. Proceedings from User interfaces for all, LNCS 2615, 43-55.

Liang, T.-P., and Wei, C.-P. (2004) "Introduction to the Special Issue: Mobile Commerce Applications. International Journal of Electronic Commerce 8 (3), 7-17.

Light, D. (2001). Sure, you can trust us. MIT Sloan Management Review, 43(1), 17-20.

Lim, K.H., and Benbasat, I. (2000). The Effect of Multimedia on Perceived Equivocality and Perceived usefulness of Information Systems. MIS Quarterly 24 (3), 449-471.

Lindroth, T., Nilsson, S., and Rasmussen, P. (2001). Rigour meets relevance when usability goes mobile. Proceedings of IRIS24, Ulvik, Norway, August 2001.

Liu, B., Francis, B., Salvendy, G. (2002). Applying models of visual search to menu design. International Journal of Human-Computer Studies, 56, 307-330.

Ljungberg, J., Neely, G., and Lundstrom, R. (2004). Cognitive performance and subjective experience during combined exposures to whole-body vibration and noise.

International Archives of Occupational Environmental Health, April 2004, Vol. 77, No. 3, pp. 217-21.

Loeb, M. (1981). The present state of research on the effects of noise: are we asking the right questions? Journal of Auditory Response, April, Vol. 21, No. 2, pp. 93-104.

Lowry, R. (1996). Concepts and Applications of Inferential Statistics. Retrieved on February 9, 2006from http://faculty.vassar.edu/lowry/webtext.html.

Lu, M. & Yeong, W. (1998). A framework for effective commercial Web application development. Internet Research Journal, 8(2), 166-173.

Lyytinen, K. and Yoo, Y. (2002a). Issues and Challenges in Ubiquitous Computing. Communications of the ACM, 45(12), 63-65.

Lyytinen, K. and Yoo, Y. (2002b). Research Commentary: The Next Wave of Nomadic Computing. Information Systems Research, 13(4), 377-388.

MacKenzie, I. S., Kober, H., Smith, D., Jones, T., Skepner, E. (2001). LetterWise: Prefix-based disambiguation for mobile text input. Proceedings of the ACM Symposium on User Interface Software and Technology - UIST 2001, pp. 111-120. New York: ACM.

MacKenzie, I.S. and Zhang, S.X. (1999). The Design and Evaluation of a High Performance Soft Keyboard. Proceedings of CHI '99, Pittsburgh, PA, USa, 15-20 May, 1999.

MacKenzie, S., Nonnecke, B., Riddersma, S., McQueen, C., and Meltz, M. (1994). Alphanumeric entry on pen-based computers. International Journal of Human-Computer Studies. 41, 775-792.

Maguire, M. (2001). Context of use within usability activities. International Journal of Human-Computer Studies, 55, 453-483.

Mahmood, M.A., Burn, J.M., Gemoets, L.A., and Jacquez, C. (2000). Variables affecting information technology end-user satisfaction: a meta-analysis of the empirical literature. International Journal of Human-Computer Studies, 52, 751-771.

Maitland, C.F., Van de Kar, E.A., Wehn de Montalvo, U., and Bouwman, H. (2005). Mobile information and entertainment services: Business models and service networks. International Journal of Management and Decision Making, 6(1), 47-64.

Malladi, R. and Agrawal, D. (2002). Current and Future Applications of Mobile and Wireless Networks. Communications of the ACM, 45(10), 144-146.

Malloy, A., Varshney, U., and Snow, A. (2002). Supporting Mobile Commerce Applications Using Dependable Wireless Networks. Mobile Networks and Applications, Vol. 7, 225-234.

Marcus, A. and Chen, E. (2002). Designing the PDA of the Future. Interactions. January-February, 2002, 34-44.

Mathieson, K. (1991). Predicting User Intentions: Comparing the Technology Acceptance Model with the Theory of Planned Behaviour. Information Systems Research, 2, 3 (September 1991), 173-191.

Mathieson, K. and Keil, M. (1998). Beyond the interface: Ease of use and task/technology fit. Information & Management, 34, 221-230.

McGinity, M. (1999). Flying Wireless, With a Net. Communications of the ACM. 42(12), 19-21.

McKinney, V., Yoon, K., and Zahedi, F.M. (2002). The measurement of Web-customer satisfaction: An expectation and disconfirmation approach. Information Systems Research, 13(3), 296-315.

McKnight, D., Choudhury, H., and Kacmar, C. (2002). The impact of initial consumer trust on intentions to transact with a web site: a trust building model. Journal of Strategic Systems, Vol. 11, 297-323.

McNemar,Q. (1969). Psychological statistics. Wiley, Fourth edition, New York.

McRoy, S. (1998). Achieving robust human-computer communication. International Journal of Human-Computer Studies, 48, 681-704.

MEGATAQ (1999). Methods and Guidelines for the Assessment of Telematics Applications Quality - sets out to help you answer your questions of evaluation. MEGATAQ is a project of the Telematics Engineering sector of the Telematics Application Programme of the European Commission, DGXIII.

Mehrtens, J., Cragg, P.B., Mills, A.M. (2001). A model of Internet adoption by SMEs. Information & Management, Vol. 39, 165-176.

Mendelson, H. (1985). Pricing computer services: Queueing effects. Communications of the ACM, 28(3), 312-321.

Mendelson, H. and Whang, S. (1990). Priority pricing for the M/M/1 queue. Operations Research, 38(5), 870-883.

Miles, G.E. and Howes, A. (2000). A framework for understanding human factors in web-based electronic commerce. International Journal of Human-Computer Studies, 52, 131-163.

Misra, S. K. and Wickamasinghe, N. (2004). Security of a Mobile Transaction: A Trust Model. *Springer, 4* (4).

Mittal, V., Kumar, P. and Tsiros, M. (1999). Attribute-level performance, satisfaction, and behavioural intentions over time: a consumption-system approach. Journal of Marketing, Vol. 64, April, pp. 88-101.

Mohadjer, L., Bell, B., and Waksberg, J. (1994). Accounting for item non-response bias. National Center for HealthStatistics. Hyattsville, Maryland. Retrieved on February 10, 2006, from http://www.cdc.gov/nchs/data/nhanes/nhanes3/cdrom/NCHS/MANUALS/NR_BIAS.PDF

Moon, J. and Kim, Y. (2001). Extending the TAM for a World-Wide-Web context. Information & Management, Vol. 38, 217-230.

Moreno, K.N. and Bodenhauser, G.V. (1999). Resisting Stereotype Change: The Role of Motivation and Attentional Capacity in Defending Social Beliefs. Group processes & Intergroup Relations, 2(1), 1999, 5-16.

Nagata, S. and TNO Human Factors (2003). Multitasking and Interruptions During Mobile Web Tasks. Proceedings of the Human Factors and Ergonomics Society 47th Annual Meeting, 2003.

Newell, A. and Gregor, P. (2002). Design for older and disabled people – where do we go from here? Universal Access Information Society, 2, 3-7.

Nicholson, D.B., Nicholson, J.A., Parboteeah, D.V., and Valacich, J.S. (2005). Using Distraction-Conflict Theory to Measure the Effects of Distractions on Individual Task Performance in a Wireless Mobile Environment. Proceedings of the 38th Hawaii International Conference on System Sciences – 2005.

Nielsen, J. (1993). Usability Engineering. Academic Press/AP Professional, Cambridge, MA ISBN 0-12-518406-9.

Nielsen, J. (1994). Estimating the number of subjects needed for a thinking aloud test. International Journal of Human-Computer Studies, 41, 385-397.

Nielsen, J. (1996). Designing to seduce the user. IEEE Software, September 1996, 18-20.

Nielsen, J. (1997). Let's Ask the Users: Tools, Techniques, and Concepts to Optimize User Interfaces. IEEE Software, May-June 1997, 110-111.

Nielsen, J. and Mack, R.L. Eds. (1994). Usability Inspection Methods. New York: Wiley.
Noble, R., Ruiz, K., Destefano, M., and Mintz, J. (2003).Conditions of Engagement in Game Simulation: Contexts of Gender, Culture and Age. Digital Games Research Conference 2003, 4-6 November 2003 University of Utrecht, The Netherlands.

Noirhomme-Fraiture, M. and Vanderdoncki, J. (2003). Screen Usability Guidelines for Persons with Disabilities. ACM Transactions of Computer-Human Interaction (ToCHI), January-February 2003, 25-26.

Noll, A. (1997). Internet Pricing Vs. Reality. Communications of the ACM, 40(8), 118-121.
novice performance with Jot and Graffiti. Interacting with Computers. 14, 413-433.

Nunnaly, J.C. (1978). Psychometric Theory 2nd ed., McGraw Hill, New York, 1978.

Oliver, R.L. (1997). Satisfaction: A Behavioural Perspective on the Consumer. McGraw-Hill, New York, NY.

Oliver, R.L. (1980). A cognitive model for the antecedents and consequence of satisfaction. Journal of Marketing Research, 17, 460-469.

Olshavsky, R.W. and Miller, J.A. (1972). Consumer expectations, product performance, and perceived product quality. Journal of Marketing Research. Vol.9, February, pp. 19-21.

OneUpWeb (2005). Mobile Search and Its Implications for Search Engine Marketing. Retrieved on February 10, 2006, from www.oneupweb.com/ search-marketing-results/white-papers.htm.

Oppermann, R. (1994). Adaptively supported adaptability. International Journal of Human-Computer Studies, 40, 455-472.

Oquist, G. and Goldstein, M. (2002). Towards an Improved Readability on Mobile Devices: Evaluating Adaptive Rapid Serial Visual Representation. Proceedings from Mobile HCI 2002, LNCS 2411, 225-240.

Ovans, A. (1999). Is your website socially savvy? Harvard Business Review, 77(3), 20-21.

O'Keefe, R., Cole, M., Chau, P., Massey, A., Montoya-Weiss, M., and Perry, M. (2000). From the user interface to the consumer interface: results from a global experiment. International Journal of Human-Computer Studies, 53, 611-628.

O'Neill, M., Palmer, A.. and Wright, C. (2003). Disconfirming user expectations of the online service experience. Internet Research: Electronic Networking Applications and Policy, Vol. 13, No. 4, pp. 281-296.

Padovani, S. and Lansdale, M. (2003). Balancing search and retrieval in hypertext: context-specific trade-offs in navigational tool use. International Journal of Human-Computer Studies, 58, 125-149.

Paelke, V., Reimann, C., and Rosenbach, W. (2003). A visualization design repository for mobile devices. ACM, 57-61.

Page, J., Zaslavsky, A., and Indrawan, M. (2004). Securing m-Business Scenarios Using the Buddy Model of Security for Agent Communities. M-Business conference, New York, July.

Palen, L. and Salzman, M. (2002). Beyond the Handet: Designing for Wireless Communications Usability. ACM Transactions on Computer-Human Interaction, Vol.9(2), June 2002, 125-151.

Palen, L., Salzman, M.C., and Youngs, E. (2001). Discovery and Integration of Mobile Communications in Everyday Life. Personal and Ubiquitous Computing, 5, 109-122.

Papadopoulou, P., Andreas, A., Kenellis, P., Martakos, D. (2001). Trust and relationship building in electronic commerce. Internet Research: Electronic Networking Applications and Policy 11(4), 322-332.

Paradis, F., Crimmins, F., Ozkan, N. (2003). A Task Oriented Approach to Delivery in Mobile Environments. Proceedings from MDM 2003, LNCS 2574, 386-390.

Paul, P. (2001). Mixed signals. American Demographics 23, 44-49.

Pedersen. P.E. (2002). The adoption of text messaging services among Norwegian teens: development and test of an extended adoption model. SNF-report no. 23/02. Foundation for Research in Economics and Business Administration, Bergen Bergen, Norway.

Pedhazur, E. J., Liora Pedhazur Schmelkin (1991). Measurement, Design, and Analysis: An Integrated Approach. Lawrence Erlbaum Associates, Hillsdale, NJ.

Pennington, R., Wilcox, H.D., and Grover, V. (2004). The role of system trust in business-to-consumer transactions. Journal of Management Information System, Winter 2003-4, Vol. 20, No. 3, pg. 197-226.

Perdue, B.C., and Summers, J.O. (1986). Checking the success of manipulations in marketing experiments. Journal of Marketing Research, Vol. 23, 1986, pg. 317-326.

Persson Waye, K. Bengtsson, J., Kjellberg, A., and Benton, S. (2001). Low frequency noise "pollution" interferes with performance. Noise Health, Vol. 4, No. 13, pp. 33-49.

Pinckard, M.J. (2002). Genderplay: Successes and Failures in Character Designs for Videogames CS 377A: Introduction to Cybernetics and the Design of Systems, Fall 2002, Stanford University. Retrieved on February 10, 2006, from http://www.gamegirladvance.com/archives/2003/04/16/genderplay_successes_and_failures_in_ character_designs_for_videogames.html.

Piolat, A., Roussey, J.Y., and Thunin, O. (1997). Effects of screen presentation on text reading and revising. International Journal of Human-Computer Studies, 47, 565-589.

Podsakoff, N. P., & Organ, D. W. (1986). Self-reports in organizational research: Problems and prospects. Journal of Management, 12(4), 531-544.

Podsakoff, N. P., Todor, W. D., Grover, R. A., & Huber, V. L. (1984). Situational moderators of leader reward and punishment behaviours: Fact or fiction. Organizational Behaviour and Human Performance, 34, 21-63.

Podsakoff, P. M., Mackenzie, S. B., & Podsakoff, N. P. (2003). Common method biases in behavioural research: A critical review of the literature and recommended remedies. Journal of Applied Psychology, 88(5), 879-903.

Polson, P. (1988). The consequences of consistent and inconsistent user interfaces. In R. Guindon (ed.), Cognitive Science and Its Applications for Human-Computer Interaction. Hillsdale, NJ: Lawrence Erlbaum, , 59-108.

Polson, P.G., Lewis, C., Rieman, J., and Wharton, C. (1992). Cognitive walkthroughs: A method for theory-based evaluation of user interfaces. International Journal of Man-Machine Studies, 36, 741-773.

Pontari, B.A. and Quirk, C. (2000). The influence of cognitive load on self-presentation: can cognitive busyness help as well as harm social performance? Journal of Personality and Social Psychology, 78(6), 1092-1108.

Porta, M. (2002). Vision-based user interfaces: methods and applications. International Journal of Human-Computer Studies, 57, 27-73.

Potosnak, K. (1989). Human Factors: When a Usability Test is not the answer. IEEE Software, July 1989, 105-106.

Pousttchi, K. (2004). An Analysis of the Mobile Payment Problem in Europe. In: Branki, C.; Unland, R.; Wanner, G. (Hrsg.): Multikonferenz Wirtschaftsinformatik (MKWI) 2004, Universität Duisburg-Essen, 9.-11. März 2004, Band 3: Mobile Business Systems, Mobile and Collaborative Business, Techniques and Applications for Mobile Commerce (TAMoCO). Essen 2004, S.260-268.

Powlik, J. and Karshmer, A. (2002). When accessibility meets usability. National Science Foundation.

Pride, W. and Ferrell, O. (2003). Marketing, 12th ed., Texas A&M University.

Quesenbery, W. (2003). Dimensions of usability. In Albers, M., & Mazur, B. (Eds.), Content and complexity: Information design in technical communication. Mahwah, NJ: Lawrence Erbaum Associates.

Rader, M. (1981). Dealing with Information Overload. Personnel Journal (pre-1986); May 1981, 60, 000005, pp. 373-375.

Rao, B., and Minakakis, L. (2003). Evolution of mobile location-based services. Communications of the ACM 46 (12), 61-65.

Reid, H. (2002). Maps on your PDA - Simple Stuff #8 – Mapping the Trade Show Floor. *Directions Magazine 2002,* http://www.directionsmag.com/article.php?article_id=199 (accessed 2 June 2004).

Rice, M. (1997). What makes users revisit a Web site. Marketing News, 31(6), 12-13.

Risden, K., Czerwinski, M.P., Munzner, T., and Cook, D.B. (2000). An initial examination of ease of use for 2D and 3D information visualizations of web content. International Journal of Human-Computer Studies. 53, 695-714.

Rivard, S., and Huff, S. (1988). Facotrs of Success for End-User Computing. Communications of the ACM, Vol. 31(5), May 1988, pp 552-561.

Rodden, K., Milic-Frayling, N., Sommerer, R. and Blackwell, A. (2003). Effective Web Searching on Mobile Devices. Proceedings of the 2003 HCI, England, 8-12 September 2003.

Rogers, E.M. (1983). Diffusion of Innovations, 3rd ed., New York: Free Press.

Ross, T. and Burnett, G. (2001). Evaluating the human-machine interface to vehcle navigation as an example of ubiquitous computing. International Journal of Human-Computer Studies, 55, 661-674.

Royce, W.W. (1970). Managing the development of large software systems: concepts and techniques. Proceedings of Western Electric Show and Convention, Westcon, Los Angeles, pp. (A/1)1-(a/1)9 (Reprinted in: Proceedings of the 11th International Conference on Software Engineering, Pittsburgh, May, pp. 328-338).

Rubin, J. (1994). Handbook of usability testing. John Wiley and Sons, New York, NY, ISBN 0-471-59403-2.

Rumelhart, D. (1980). Schemata: the building blocks of cognition, In R. Spiro, B. Bruce, and W. Brewer (eds.), Theoretical Issues in Reading Comprehension. Hillsdale, NJ: Lawrence Erlbaum, pp. 33-58.

Sarker, S. and Wells, J. (2003). Understanding mobile handheld device use and adoption. Communications of the ACM, Vol. 46, No. 12.

Savage, P. (1995). Business Telephone. Interactions. January, 1995, 32-41.

Schmidt, A., Beigl, M., and Gellersen, H. (1999). There is More to Context than Location. Computer & Graphics Journal, Vol. 23, No. 6, pp. 893-901.

Schmidt-Belz, B., Laamanen, H., Poslad, S., and Zipf, A. (2003). Location- based mobile tourist services - first user
Experiences. ENTER 2003 Conference, January 2003 (Helsinki).

Schneiderman, B. (1980). Software Psychology: Human Factors in Computer and Information Systems. Cambridge, MA: Winthrop.

Schneiderman, B. (1997). Designing information-abundant web sites: issues and recommendations. International Journal of Human-Computer Studies, 47, 5-29.

Schommer, J. (2000). An Experimental Approach for Investigating Consumers' Evaluation of Pharmacist Consultation Services. AAPS PharmSci. 2000; 2 (2):article 15. DOI: 10.1208/ps020215.

Sears, A. and Arora, R. (2002). Data entry for mobile devices: an empirical comparison of

Sears, A. and Lund, A. (1997). Creating Effective User Interfaces. IEEE Software. July-August, 21-24.

Seffah, A. and Forbrig, P. (2002). Multiple User Interfaces: Towards a Task-Driven and Patterns-Oriented Design Model. Proceedings from DSV-IS 2002, LNCS 2545, pp. 118-132.

Sein, M., and Bostrom, R. (1989). The influence of individual differences in determining the effectiveness of conceptual models in training novice users. Human-Computer Interaction, 4, 197-229.

Sellen, A. J., Murphy, R., and Shaw, K. L. (2002). How Knowledge Workers use the Web, in D. Wixon (ed.), Proceedings of CHI'02 Conference on Human Factors in Computing Systems: Changing our World, Changing Ourselves, ACM Press, pp.227–34.

Shackel, B. (1991). Usability: context, framework, design, and evaluation. In B. Shackel & S. Richardson (Eds.), Human factors for informatics usability (pp. 21-38). Cambridge: Cambridge University Press.

Shankar, V., Urban, G.L., and Sultan, F. (2002). Online trust: a stakeholder perspective, concepts, implications, and future directions. Journal of Strategic Systems, Vol. 11, 325-344.

Shami, N.S., Leshed, G., and Klein, D. (2005). Context of use evaluation of peripheral displays (CUEPD). INTERACT 2005, LNCS 3585, 579-587.

Smith, D.L., Chang, J., Cohen, D., Foley, J., and Glassco, R. (2005). A simulation approach for evaluating the relative safety impact of driver satisfaction during secondary tasks. Proceedings of the 12th World Congress on ITS, 6-10 November, 2005, San Francisco.

Smith, P.A., Newman, I.A., and Parks, L.M. (1997). Virtual hierarchies and virtual networks: some lessons from hypermedia usability research applied to the World Wide Web. International Journal of Human-Computer Studies, 47, 67-95.

Sohn, C. S. (2000). Customer evaluation of Internet-based service quality and intention to re-use internet-based services. Unpublished dissertation, Southern Illinois University, IL: Carbondale.

Soukoreff, R.W. (2002). Text Entry for Mobile Systems: Models, Measures, and Analyses for Text Entry Research. M.Sc. Thesis, York Universiry, June 2002.

Soukoreff, R.W. and MacKenzie, I.S. (2004). Recent Developments in Text-Entry Error Rate Measurement. Proceedings of CHI 2004, April 24-29, 2004, Vienna, Austria.

Spreng, R.A., MacKenzie, S.B., and Olshavsky, R.W. (1996). A re-examination of the determinants of consumer satisfaction. Journal of Marketing, Vol.60, July, pp. 15-32.

Staggers, N. (1993). Impact of screen density on clinical nurses' computer task performance and subjective screen satisfaction. International Journal of Man-Machine Studies. 39, 775-792.

Staggers, N. and Norcio, A. (1993). Mental models: concepts for human-computer interaction research. International Journal of Man-Machine Studies, 38, 587-605.

Stansfeld, S., Haines, M., and Brown, B. (2000). Noise and health in the urban environment. Review of Environmental Health, Jan-Jun, Vol. 15, No. 1-2, pp. 43-82.

Stidham, S. (1992). Pricing and capacity decisions for a service facility: Stability and multiple local optima. Management Science 38(8), 1121-1139.

Storm, G. (2002). Mobile Devices as Props in Daily Role Playing. Personal and Ubiquitous Computing, 6, 307-310.

Strader, T.J. and Hendrickson, A.R. (1998). A framework for the analysis of electronic market success. Proceedings of the AIS American Conference, 360-362.

Straub, D.W. (1989). Validating Instrument in MIS Research. MIS Quearterly, Vol. 12 (2), June 1989, pg. 147-170.

Suhm, B., Myers, B., Waibel, A. (2001). Multimodal Error Correction for Speech User Interfaces. ACM Transactions on Computer-Human Interaction, Vol.8(1), March 2001, 60-98.

Sutcliffe, A.G., Ennis, M., Hu, J. (2000). Evaluating the effectiveness of visual user interfaces for information retrieval. International Journal of Human-Computer Studies. 53, 741-763.

Sweeney, M., Maguire, M., Shackel, B. (1993). Evaluating user-computer interaction: a framework. International Journal of Man-Machine Studies, 38, 689-711.

Sweller, J., Van Merrienboer, J.J.G., and Paas, F.G.W.C. (1998). Cognitive Architecture and Instructional Design. Educational Psychology Review, 10, 251-296.

Tarasewich, P. (2003). Designing mobile commerce applications. Communications of the ACM, Vol. 46, No. 12, pp. 57-60.

Tauscher, L. and Greenberg S. (1997). How people revisit web pages: empirical findings and implications for the design of history systems. International Journal of Human-Computer Studies, 47, 97-137.

Taylor, C. (2004). The text files: is it time for advertisers to jump into cell-phone messaging? Mobile Marketing Association. Posted on July 7, 2004, retrieved on February 10, 2006, from http://mmaglobal.com/modules/news/ article.php?storyid=27.

Taylor, S.A. and Baker, T.L. (1994). An assessment of the relationship between service quality and customer satisfaction in the formation of consumers' purchase intentions. Journal of Retailing, Vol. 70, No. 2, pp. 163-78.

Taylor, S., & Todd, P. (1995). Assessing IT usage: the role of prior experience. MIS Quarterly, 561-570.

Teo, H., Oh, L., Liu, C., and Wei, K. (2003). An empirical study of the effects of interactivity on web user attitude. International Journal of Human-Computer Studies, 58, 281-305.

Thimbleby, H. (2001). Permissive user interfaces. International Journal of Human-Computer Studies, 54, 333-350.

Thimbleby, H., Cairns, P, and Jones, M. (2001). Usability Analysis with Markov Models. ACM Transactions on Computer-Human Interaction, Vol.8(2), June 2001, 99-132.

Thomas, P. and Macredie, R. (2002). Introduction to The New Usability. ACM Transactions on Computer-Human Interaction, Vol.9(2), June 2002, 69-73.

Tornatzky, L.G. and Klein, K.J. (1982). Innovation characteristics and innovation adoption-implementation: a meta-analysis of findings. IEEE Transactions on Engineering Management, 29(1), 28-45.

Toyone, T., Tanaka, T., Kato, D., Kaneyama, R., and Otsuka, M. (2005). Patients' Expectations and Satisfaction in Lumbar Spine Surgery. MD Spine. 2005;30(23):2689-2694. Lippincott Williams & Wilkins.

Trumbly, J., Arnett, K., Martin, M. (1993). Performance effect of matching computer interface characteristics and user skill level. International Journal of Man-Machine Studies. 38, 713-724.

Tse, D.K. and Wilton, P.C. (1988). Models of consumer satisfaction formation: an extension. Journal of Marketing Research, Vol.25, May, pp. 204-12.

Turel, O. and Yuan, Y. (2005). Investigating the Dynamics of the M-Commerce Value System: A Comparative Viewpoint. International Journal of Mobile Communications (accepted for publication in July 2005).

Umemura, M., Honda, K., and Kikuchi, Y. (1992). Influence of noise on heart rate and quantity of work in mental work. Annuals of Physiology Anthropology, Vol. 11, No. 5, pp. 523-32.

Umlauft, M., Pospischil, G., Niklfeld, G., and Milchmayr, E. (2003). LoL@, a Mobile Tourist Guide for UMTS, Information Technology & Tourism, **5(3)**, 151-164.

Unknown (2005). The usability of PDAs for Assessment of practical performance.

Van Elteren, J.J. (2004). The relation between learning and cognitive load on a modern multimedia device. 1st Twente Student Conference on IT, Enschede 14 June 2004.

Van Ryzin, G. (2004). Expectations, Performance, and Citizen Satisfaction with Urban Services. Journal of Policy Analysis and Management, Vol. 23, No. 3, 433-448.

Van Ryzin, G. (2005). Testing the Expectancy Disconfirmation Model of Citizen Satisfaction with Local Government. Journal of Public Administration Research and Theory Advance, forthcoming.

Varshney, U. (2002). Multicast Over Wireless Networks. Communications of the ACM, 45(12), 31-37.

Varshney, U. and Vetter, R. (2000a). Emerging Mobile And Wireless Networks. Communications of the ACM, 43(6), 73-81.

Varshney, U. and Vetter, R. (2002b). Mobile Commerce: Framework, Applications, and Networking Support. Mobile Networks and Applications, Vol. 7, 185-198.

Venkatesh, V. and Davis, F.D. (2000). A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. Management Science, February 2000, Vol 46(2), pg. 186-204.

W3C (WorldWideWebConsortium) (2003). About the World Wide Web Consortium (W3C). Retrieved on April 15, 2003, http://www.w3.org/Consortium/#background

Walter, T., Bussard, L., Robinson, P., and Roudier, Y. (2004). Security and Trust Issues in Ubiquitous Environments - The Business-to-Employee Dimension. Symposium on Applications and the Internet.

Warren, J., Frankel, H., and Noone, J. (2002). Supporting special-purpose health care models via adaptive interfaces to the web. Interacting with Computers, 14, 251-267.

Waterson, S., Landay, J.A., and Matthews, T. (2002). In the Lab andOut in the Wild: Remote Web Usability Testing for Mobile Devices. Proceedings of CHI 2002, April 20-25, 2002, Minneapolis, Minnesota, USA

Watters, C., Duffy, J., and Duffy, K. (2003). Using large tables on small display devices. International Journal of Human-Computer Studies. 58, 21-37.

Wellbery, C. (2005). Patient expectations were vague and heterogeneous. Is patient satisfaction related to patient expectations? American Family Physician, October 1, 2005. Retrieved on February 10, 2006, from http://newssearch.looksmart.com/p/articles/mi_m3225/is_7_72/ai_n15798296

Westland, J.C. (1992). Congestion and network externalities in the short run pricing of information system services. Management Science, 38(7), 992-1009.

Wigdor, D., and Balakrishnan, R. (2003). TiltText: Using Tilt for Text Input to Mobile Phones. Proceedings of UIST 2003, Vancouver, BC, Canada.

Wirtz, J. and Bateson, J.E.G. (1999). Consumer satisfaction with services: Integrating the environment perspective in services marketing into the traditional disconfirmation paradigm. Journal of Business Research, 44(1), 55-66.

Woodroof, J. and Burg, W. (2003). Satisfaction/dissatisfaction: are users predisposed? Information & Management, 40, 317-324.

Wright, P. and Lickorish, A. (1994). Menus and memory load: navigation strategies in interactive search tasks. International Journal of Human-Computer Studies, 40, 965-1008.

Wu, B. and Newell, S.J. (2003). The impact of noise on recall of advertisements. Journal of Marketing Theory and Practice, Spring 2003, Vol. 11, No. 2., pp. 56-65.

Xu, J. and Croft, W. (2000). ACM Transactions on Information Systems. Vol 18(1), 79-112.

Yang, Z. and Jun, M. (2002). Consumer perception of e-service quality: From Internet purchaser and non-purchaser perspectives. Journal of Business Strategies, Huntsville, Spring 2002, Vol. 19(1), 19-42.

Yi, Y. (1990). A critical review of consumer satisfaction. In Zeithaml, V.A. (Ed.), Review of Marketing, Vol. 4, American Marketing Association, Chicago, IL, pp. 68-123.

Yuan, Y. and Zheng, W. (2005). From Stationary Work Support to Mobile Work Support: A Theoretical Framework. IEEE Proceedings of International Conference on Mobile Business (ICMB 2005) Sidney, Australia, July 11-13, 2005. pp.315-321.

Zaltman, G. and Stiff, R. (1973). Theory of diffusion, in: S. Ward, T.S.Robertson (Eds.), Consumer Behaviour: Theoretical Sources, Prentice-Hall, Englewood Cliffs, NJ, 416-468.

Zmijewska, A., Lawrence, E., and Steele, R. (2004). Classifying m-Payments - a User-centric Model. 3rd International Conference on Mobile Business 2004, New York.

APPENDIX A. Measurement Instrument

Evaluating The Usability of Mobile Devices (Pre-Test Survey)

Instructions: Please fill in the blank or circle the answer that best matches your opinion.

1. **What is your gender?** (Male) (Female)
2. **What is your age?** __________
3. **What is the highest level of education you have attained to date?**

(High school graduate or less) (Attending/attended college/university 1-3 years)

(Graduated from 4 year college/university) (Attending/attended graduate/postgraduate study or degree)

Instructions: For each statement below, please circle the number to the right that best matches your opinion on each statement.	Strongly Disagree	Moderately Disagree	Slightly Disagree	Neither Agree Nor Disagree	Slightly Agree	Moderately Agree	Strongly Agree
4. I expect that learning how to use a mobile device for wireless data services will be easy	1	2	3	4	5	6	7
5. I expect that using a mobile device for wireless data services will be easy	1	2	3	4	5	6	7
6. I expect that using a mobile device for wireless data services will be fast	1	2	3	4	5	6	7
7. I expect that a mobile device will be user friendly for wireless data services	1	2	3	4	5	6	7
8. I expect that I will be able to complete all wireless data services on a mobile device successfully	1	2	3	4	5	6	7

Evaluating The Usability of Mobile Devices (Post-Test Survey)

9. How much mental and perceptual activity was required (e.g., thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy/simple/forgiving (i.e. LOW) or demanding/complex/exacting (i.e. High)?

    LOW...HIGH
    1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

10. How much physical activity was required (e.g., pushing, pulling, turning, controlling, activating, etc.)? Was the task easy/slow/slack/restful (i.e. LOW) or demanding/brisk/strenuous/laborious (i.e. HIGH)?

    LOW...HIGH
    1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

11. How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred? Was the pace slow and leisurely (i.e. LOW) or rapid and frantic (i.e. HIGH)?

    LOW...HIGH
    1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

12. How hard did you have to work (mentally and physically) to accomplish your level of performance?

    LOW...HIGH
    1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

13. How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?

    LOW...HIGH
    1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

14. How insecure, discouraged, irritated, stressed and annoyed (i.e. LOW) versus secure, gratified, content, relaxed and complacent (i.e. HIGH) did you feel during the task?

    LOW...HIGH
    1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

Instructions: For each statement below, please circle the number to the right that best matches your opinion on each statement.	Strongly Disagree	Moderately Disagree	Slightly Disagree	Neither Agree Nor Disagree	Slightly Agree	Moderately Agree	Strongly Agree
15. Learning how to use the mobile device for wireless data services was easy	1	2	3	4	5	6	7
16. Using the mobile device for wireless data services was fast	1	2	3	4	5	6	7
17. The mobile device was user friendly for wireless data services	1	2	3	4	5	6	7
18. Using the mobile device for wireless data services was easy	1	2	3	4	5	6	7

Instructions: For each statement below, please circle the number to the right that best matches your opinion on each statement starting as: *"Compared to my expectations of efficiency, the"* (note: *efficiency* refers to the level of resources consumed in performing a task, e.g. time and effort)	MUCH WORSE THAN I EXPECTED			ABOUT WHAT I EXPECTED			MUCH BETTER THAN I EXPECTED
19. Learnability (that is, the degree to which it was easy to learn how to use) the mobile device for wireless data services was…	1	2	3	4	5	6	7
20. The time required to use the mobile device for wireless data services was…	1	2	3	4	5	6	7
21. The user friendliness of the mobile device for wireless data services was…	1	2	3	4	5	6	7
22. Ease of use of the mobile device for wireless data services was…	1	2	3	4	5	6	7

Instructions: For each statement below, please circle the number to the right that best matches your opinion on each statement starting as: *"Thinking about my experience with the efficiency of this device for wireless data services, I feel ..."* (note: *efficiency* refers to the level of resources consumed in performing a task, e.g. time and effort)							
23. Terrible (1)..Delighted (7)	1	2	3	4	5	6	7
24. Very displeased (1)..Very pleased (7)	1	2	3	4	5	6	7
25. Very dissatisfied (1)......................................Very satisfied (7)	1	2	3	4	5	6	7
26. Frustrated (1)...Contented (7)	1	2	3	4	5	6	7

Instructions: For the statement below, please circle the number to the right that best matches your opinion.	Strongly Disagree	Moderately Disagree	Slightly Disagree	Neither Agree Nor Disagree	Slightly Agree	Moderately Agree	Strongly Agree
27. I was able to complete all wireless data services on the mobile device successfully	1	2	3	4	5	6	7

Instructions: For the statement below, please circle the number to the right that best matches your opinion starting as: *"Compared to my expectations of effectiveness, the..."* (*effectiveness* refers to the ability of users to complete tasks using the system)	MUCH WORSE THAN I EXPECTED			ABOUT WHAT I EXPECTED			MUCH BETTER THAN I EXPECTED
28. Ability to complete all wireless data services on the mobile device successfully was ...	1	2	3	4	5	6	7

Instructions: Please circle the number to the right that best matches your opinion on each statement starting as: *"Thinking about my experience with the effectiveness of this device for wireless data services, I feel…"* (*effectiveness* refers to the ability of users to complete tasks using the system)							
29. Very dissatisfied (1)……………………………Very satisfied (7)	1	2	3	4	5	6	7
30. Frustrated (1)……………………………………Contented (7)	1	2	3	4	5	6	7
31. Terrible (1)…………………………………………Delighted (7)	1	2	3	4	5	6	7
32. Very displeased (1)……………………………Very pleased (7)	1	2	3	4	5	6	7

Instructions: For each statement below, please circle the number to the right that best matches your opinion on each statement.	Strongly Disagree	Moderately Disagree	Slightly Disagree	Neither Agree Nor Disagree	Slightly Agree	Moderately Agree	Strongly Agree
33. Given that I had access to the mobile device, I predict that I would use wireless data services in the near future	1	2	3	4	5	6	7
34. Assuming I had access to the mobile device, I intend to use wireless data services in the near future	1	2	3	4	5	6	7

APPENDIX B. Additional Figures and Tables

TABLE A.B1 DESCRIPTIVE STATISTICS FOR SMS TASK COMPLETION TIME SORTED BY TREATMENT

SMS	Treat.	N	Mean	Std. Dev.	Std. Error	95% Confidence Interval for Mean		Min.	Max.
						Lower Bound	Upper Bound		
1	1	17	39.80	19.74	4.79	29.65	49.95	19.24	105.58
	2	19	42.50	18.88	4.33	33.40	51.60	25.66	111.27
	3	23	41.23	17.69	3.69	33.58	48.88	19.31	83.77
	4	17	53.66	23.09	5.60	41.78	65.53	26.79	102.89
	Total	76	44.01	20.06	2.30	39.42	48.59	19.24	111.27
2	1	17	26.71	11.50	2.79	20.80	32.62	13.23	57.53
	2	19	25.37	5.87	1.35	22.54	28.20	17.28	36.32
	3	23	28.37	9.47	1.97	24.28	32.46	14.99	51.23
	4	17	32.93	12.04	2.92	26.74	39.12	13.22	61.59
	Total	76	28.27	10.06	1.15	25.97	30.57	13.22	61.59
3	1	17	34.51	10.72	2.60	28.99	40.02	19.70	63.57
	2	19	37.75	8.89	2.04	33.47	42.04	20.88	57.92
	3	23	39.58	10.91	2.27	34.86	44.30	26.51	75.88
	4	17	46.43	17.49	4.24	37.43	55.42	26.58	89.63
	Total	76	39.52	12.69	1.46	36.62	42.42	19.70	89.63
4	1	17	21.19	4.49	1.09	18.88	23.50	14.60	30.00
	2	19	26.65	7.57	1.74	23.00	30.29	14.99	42.70
	3	23	27.14	9.79	2.04	22.90	31.37	13.35	59.25
	4	17	29.17	11.00	2.67	23.51	34.82	17.46	63.32
	Total	76	26.14	8.94	1.03	24.09	28.18	13.35	63.32
Avg	1	17	30.55	9.74	2.36	25.54	35.56	16.93	57.36
	2	19	33.07	8.84	2.03	28.81	37.33	20.12	59.04
	3	23	34.08	9.38	1.96	30.02	38.14	21.46	60.16
	4	17	40.55	12.87	3.12	33.93	47.16	23.38	63.83
	Total	76	34.48	10.61	1.22	32.06	36.91	16.93	63.83

TABLE A.B2: DESCRIPTIVE STATISTICS FOR SMS WPM SORTED BY TREATMENT

SMS	Treat.	N	Mean	Std. Dev.	Std. Error	95% Confidence Interval for Mean		Min.	Max.
						Lower Bound	Upper Bound		
1	1	17	14.79	5.03	1.22	12.20	17.37	4.57	25.26
	2	19	13.18	3.32	0.76	11.58	14.78	7.03	19.39
	3	23	14.92	4.49	0.94	12.98	16.86	7.08	25.26
	4	17	10.93	4.20	1.02	8.77	13.08	4.71	18.46
	Total	76	13.56	4.50	0.52	12.53	14.59	4.57	25.26
2	1	17	16.14	4.37	1.06	13.89	18.39	8.78	25.85
	2	19	14.86	3.21	0.74	13.31	16.40	9.33	19.77
	3	23	14.51	3.93	0.82	12.81	16.21	7.47	24.00
	4	17	12.42	5.06	1.23	9.82	15.03	6.69	27.69
	Total	76	14.50	4.26	0.49	13.52	15.47	6.69	27.69
3	1	17	18.70	4.94	1.20	16.16	21.24	11.51	30.95
	2	19	17.19	4.48	1.03	15.03	19.35	10.74	28.20
	3	23	16.80	3.68	0.77	15.21	18.40	9.44	25.00
	4	17	14.78	4.36	1.06	12.53	17.02	9.57	24.00
	Total	76	16.87	4.45	0.51	15.85	17.89	9.44	30.95
4	1	17	18.65	4.34	1.05	16.42	20.88	12.43	26.40
	2	19	15.67	4.34	1.00	13.58	17.76	9.94	24.86
	3	23	15.32	4.46	0.93	13.39	17.25	5.90	24.92
	4	17	13.76	3.94	0.96	11.73	15.79	5.91	20.67
	Total	76	15.80	4.53	0.52	14.77	16.84	5.90	26.40
Avg	1	17	17.07	4.36	1.06	14.83	19.31	10.78	27.11
	2	19	15.23	3.45	0.79	13.56	16.89	10.18	22.78
	3	23	15.39	3.58	0.75	13.84	16.94	8.48	23.97
	4	17	12.97	3.94	0.96	10.95	15.00	8.03	21.99
	Total	76	15.18	3.99	0.46	14.27	16.09	8.03	27.11

TABLE A.B3: DESCRIPTIVE STATISTICS FOR THE CALENDAR TASK SORTED BY TREATMENT

	N	Mean	Std. Dev.	Std. Error	95% Confidence Interval for Mean		Min.	Max.
					Lower Bound	Upper Bound		
1	17	158.20	60.40	14.65	127.15	189.26	65.91	295.71
2	19	153.07	64.76	14.86	121.85	184.28	21.29	293.15
3	23	152.24	75.43	15.73	119.62	184.85	42.09	327.15
4	12	137.55	47.05	13.58	107.65	167.44	61.96	224.57
Total	71	151.40	63.98	7.59	136.26	166.55	21.29	327.15

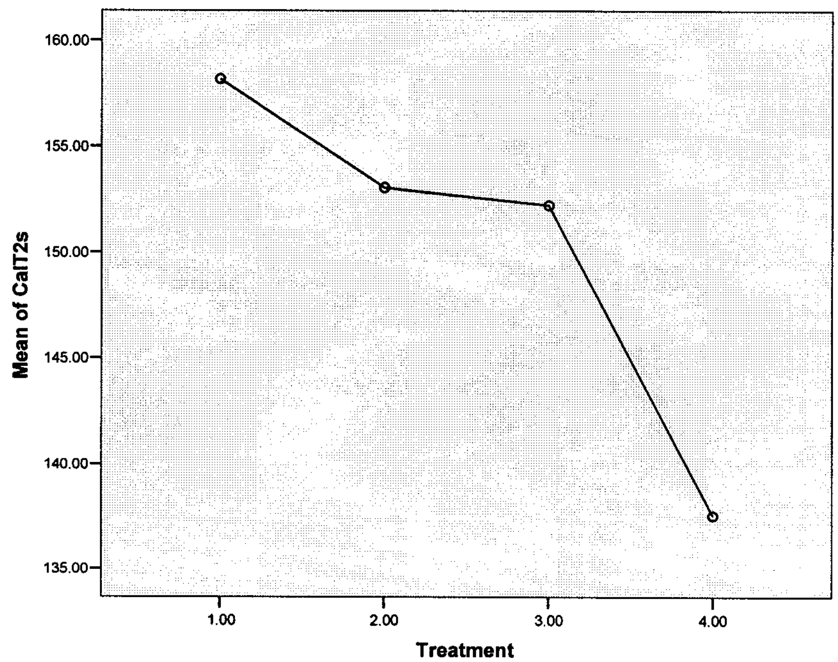


FIGURE A.B1: CALENDAR TASK COMPLETION TIME BY TREATMENT

TABLE A.B4: DESCRIPTIVE STATISTICS FOR THE ADDRESS BOOK TASK SORTED BY TREATMENT

	N	Mean	Std. Dev.	Std. Error	95% Confidence Interval for Mean		Min.	Max.
					Lower Bound	Upper Bound		
1	17	129.18	32.17	7.80	112.64	145.73	80.71	179.71
2	19	133.59	44.86	10.29	111.97	155.21	67.48	216.36
3	23	129.49	43.78	9.13	110.56	148.42	33.01	208.26
4	14	135.10	37.69	10.07	113.34	156.86	85.62	203.98
Total	73	131.56	39.77	4.65	122.28	140.84	33.01	216.36

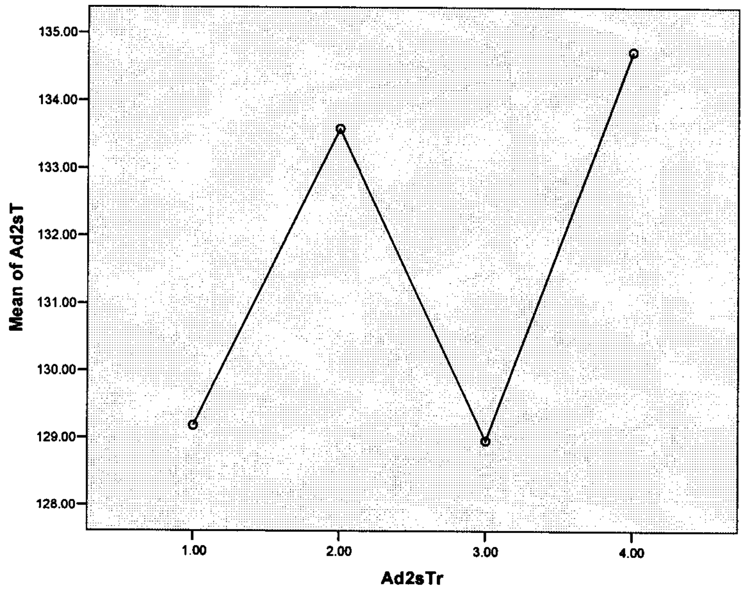


FIGURE A.B2: ADDRESS BOOK TASK COMPLETION TIME BY TREATMENT

TABLE A.B5: DESCRIPTIVE STATISTICS FOR WEB COMPLETION TIME BETWEEN TREATMENT CONDITIONS

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean Lower Bound	95% Confidence Interval for Mean Upper Bound	Minimum	Maximum
1.00	14	95.7857	43.47798	11.61998	70.6823	120.8892	38.00	184.00
2.00	17	108.1176	39.63093	9.61191	87.7413	128.4940	40.00	205.00
3.00	17	94.3529	43.56883	10.56699	71.9519	116.7540	48.00	213.00
4.00	11	117.9091	66.81535	20.14559	73.0219	162.7963	35.00	237.00
Total	59	103.0508	47.27193	6.15428	90.7317	115.3700	35.00	237.00

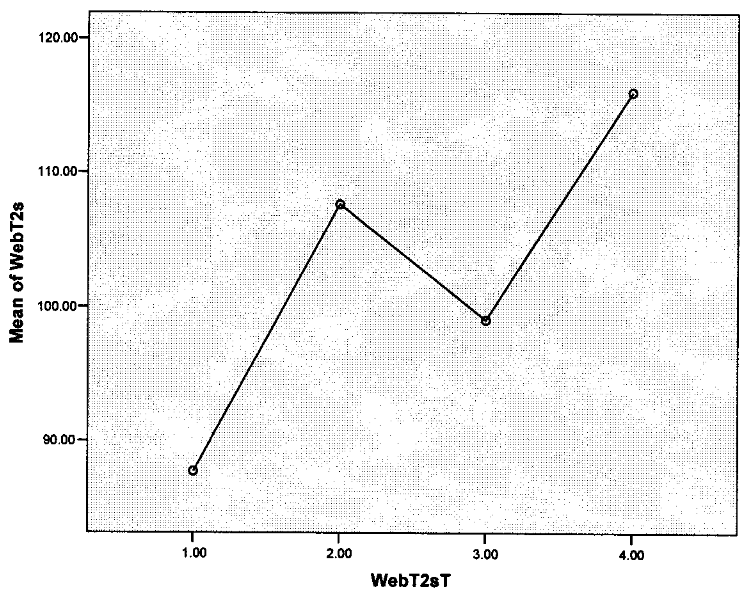


FIGURE A.B3: WEB TASK COMPLETION TIME BY TREATMENT