The impact of distractions on the usability and intention to use mobile devices for wireless data services

Constantinos K. Coursaris a,⁎, Khaled Hassanein b, Milena M. Head b, Nick Bontis c

aDepartment of Telecommunication, Information Studies, and Media, Michigan State University, East Lansing, MI 48823, USA
bDepartment of Information Systems, McMaster University, Hamilton, ON L8S4M4, Canada
cDepartment of Strategic Market Leadership and Health Services Management, McMaster University, Hamilton, ON L8S4M4, Canada

Article history:
Available online 4 April 2012

Keywords:
Mobile
Usability
Distraction
Expectancy-confirmation theory
Satisfaction
Adoption

1. Introduction

Mobile devices are becoming increasingly popular, having reached over 5.3 billion mobile subscribers worldwide by the end of 2010 (ITU, 2010; PBC, 2008). Corresponding wireless data services that have emerged present an important evolution in information and communication technologies (ICTs) (Byoungsoo, Minseok, & Ingoo, 2009). Coupled with continuous reduction in consumers' technology fears and lower adoption costs, mobile devices have become “mainstream” around the developed world. Such devices propose increased value to consumers due to “anytime/anywhere” connectivity, communication, and data services.

Although progress has been made in terms of technological innovations, many mobile subscribers are still concerned with the usability, reliability, and security of mobile applications and services (Coursaris & Hassanein, 2002). Key usability challenges include technology issues with respect to interface attributes, such as limited screen size, limited input methods, and navigation difficulties (Persson Waye, Bengtsson, Kjellberg, & Benton, 2001). Additionally, the mobile user has to share his or her attention with the task (application) and the surrounding environment. Furthermore, individual characteristics (e.g. age, culture) may be key factors in their ability and preferences to use a mobile device.

The concept of context of use as it relates to usability emerged out of the work of several researchers (e.g., Baker & Holding, 1993; Bevan & Macleod, 1994; Coursaris & Kripintiris, 2012; Coursaris, Swierenga, & Watrall, 2008; Lee & Benbasat, 2003; Tarasewich, 2003), who suggested that many variables beyond the immediate interface might impact usability. Although the definition of context may be slightly varied, the takeaway is that usability experiments need to consider various contextual factors (Liu & Li, 2011). In particular when assessing the usability of mobile devices and services, the following factors should be considered (adapted from Hassanein & Head, 2003):

• User (e.g. prior relevant/computing experience, age, education, culture, motion).
• Environment (e.g. lighting, noise, visual distractions of other objects or people).
• Task (e.g. complexity, interactivity).
• Technology (e.g. interface design, input/output modes, device size, weight).

The results of such contextual usability studies should guide the design of mobile devices and services resulting in better user
satisfaction and consequently higher rates of adoption for such devices and services. Adherence to a rigorous research design would constrain most studies to the investigation of only one (or slightly more) of the aforementioned factors, e.g., user motion. Even though user motion is only one attribute of a potential human–computer interaction (HCI), its inherently dynamic effects on other HCI attributes (e.g., input mode) augment a user's experience significantly from that of a desktop-based system use. With user motion comes the exposure to constantly changing visual and auditory stimuli. These stimuli effectively become distractions as users interact with mobile devices and their performance and experience with the services may be impacted significantly. Designing mobile interfaces and services should be informed so that they afford users greater capacity in this respect. Thus, this burgeoning research domain is guided by the need to explore the characteristics of mobility, develop new design principles for mobile systems, propose novel mobile usability evaluation techniques, and consequently to obtain a better understanding of, hence, improve mobile use (Biela, Grillb, & Gruhna, 2010; Heea, Hamb, Parkc, Song, & Youond, 2009).

This paper explores the impact of context on the usability of mobile devices. Specifically, the paper empirically investigates the impact of distractions as well as the post-trial confirmation of users' initial expectations of performance on the usability and the subsequent effect on consumers' behavioral intention to use a mobile device for wireless data services. Distractions are ever present during everyday use of mobile devices, yet the nature and extent to which user perceptions and performance are affected by their presence remains unknown. Similarly, the relationship between pre-trial expectations from and post-trial perceptions of usability has received limited attention in the context of wireless data services. This study will contribute to theory through an extension of usability theory, by considering cognition and by additionally testing the applicability of the Expectancy-Confirmation Theory (EDT) in explaining a mobile user's evaluative process of usability. Furthermore, this study contributes to practice by providing a better understanding of contextual usability factors that influence consumer adoption of mobile devices and wireless data services and hence can inform improved design of these devices and services.

2. Theoretical development and research model

Distractions are stimuli that are irrelevant to a subject's primary task, and can come in different forms (Sanders, Baron, & Moore, 1978), for instance, external stimuli—such as changing light (visual) conditions, the sudden introduction of music (auditory) sounds—or an internal thought process. Consequently, a distraction may affect an individual's performance with a primary task due to attentional conflict, i.e., the individual’s tendency, desire or obligation to allocate attention to various competing inputs (Baron, 1986; Nicholson, Parboteeah, Nicholson, & Valacich, 2005; Sanders et al., 1978).

Attending to more than one stimulus at a time requires greater mental activity on the part of an individual's working memory (Sweller, 1988, 1994), which is commonly referred to as cognitive load. Increased cognitive load may affect performance negatively by reducing the individual's attentional control, accuracy, working memory, and retrieval efficiency (Eysenk, 1984; Lavie, 2010). As a result, individuals exposed to distractions may omit procedural steps, forget to complete tasks, and take unbeneficial shortcuts (Latino, 2008, p. 10). Yet, such effects have rarely been examined in the human–computer interaction literature, and even less in the context of mobile computing and usability. Nonetheless, the effect of distractions in the use of mobile devices may have substantial consequences, ranging from short-term inconveniences (e.g., annoyance) to life-threatening situations (e.g., driving accidents).

In order to analyse the role of distractions, researchers need to pay close attention at the dyadic inverse relationship that exists between methodological rigour and relevance of findings (Lindroth, Nilsson, & Rasmussen, 2001). It can be argued that the more natural the experimental setting, the more relevant and applicable the study’s results will be. However, typically, usability studies are performed in controlled laboratory settings where external variables (e.g., distractions), are absent (Kallinen, 2004) in an attempt to uphold a rigorous methodology. By omitting distractions, however, such studies exclude factors that would typically be present in a real-world setting and therefore the external validity of these findings is limited.

2.1. Distractions and performance

The aforementioned limitation in excluding distractions from contextual usability studies arises mainly from the observation that distractions negatively affect information processing and performance (Baker & Holding, 1993; Sörqvist, 2010). 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” (note: for a comprehensive review of cognitive load, refer to (Hollender, Hofmann, Deneke, & Schmirtz, 2010). Any single distraction adds to the total cognitive stimuli (i.e., load) thereby reducing one’s capacity to process information efficiently (Miller, 1956) and effectively and, hence, potentially one’s overall performance.

Extensive literature focuses on auditory and visual distractions and their impact on performance. In this respect, it has been shown that a quiet environment results in higher efficiency, while the presence of irrelevant sound lowers mental efficiency and performance due to the obligatory cognitive process of organizing unattended information (Hughes & Jones, 2003). It is interesting to note that both noise and music hinder performance (Persson Wayne et al., 2001; Stansfeld, Haines, & Brown, 2000), but music has been shown to have a more substantial negative impact on performance compared to noise (Unemura, Honda, & Kikuchi, 1992). Additionally, increased variability of background noise results in lower performance (Hughes & Jones, 2003).

Research has also shown that visual distractions may elicit different responses from the brain than auditory distractions, yet, the negative impact on performance remains. In fact, studies have shown that it may be more difficult to return to one’s thoughts and task after certain visual rather than auditory distractions (Berti & Schroger, 2001).

Another previously explored source of distraction is motion. Ljungberg, Neely, and Lundstrom (2004) study supports the argument that the combination of a subject’s motion (e.g., walking) with the presence of any other auditory or visual distraction would impact the subject's performance negatively, as they would have an additive effect on cognitive load. Although there is a substantial body of literature on the negative effect of various forms of distractions on performance (Baker & Holding, 1993), no studies have yet explored the role of distractions in the context of wireless data services.

Yet, based on the discussion of previous findings regarding the negative effect of auditory, visual and motion-related distractions, it can be inferred that the greater the level of each of these types of distraction, the more adverse its impact on performance. 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.

2.2. Performance and usability

Watters, Duffy, and Duffy (2003), defined performance as the efficiency and effectiveness associated with a task's completion. These two utilitarian measures are included in the definition set forth for usability by the International Organization for Standardization (ISO, 1998), along with an affective measure for satisfaction described as follows:

- **Efficiency**: the level of resources 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 with using the technology.

This ISO definition of usability was chosen for this study in part because it is the international standard of measuring usability, but it should be noted that several approaches to measuring usability have previously been put forth by scholars. One of the earlier approaches was proposed by Nielsen (Nielsen, 1993), where usability was measured as the learnability, efficiency, memorability, less errors, and satisfaction involved in a user's interaction with a technology. Rubin (1994) proposed similar usability dimensions, including learnability, effectiveness, usefulness, and attitude. Quesenbery (2003) defined usability in terms of five dimensions: efficiency, effectiveness, engagement, error tolerance, and ease of learning.

All of these measures and more were identified in two qualitative reviews of empirical mobile usability studies conducted by Coursaris and Kim (2006, 2011) observing that the three constructs of efficiency, effectiveness, and satisfaction represent the core dimensions of usability. Further, the use of the ISO standard allows for consistency with other studies in the measurement of these three constructs (Brereton, 2005). Additionally, Frokjaer, Hertzum, and Hornbaek (2000) tested these three constructs of efficiency, effectiveness and satisfaction for correlation in reference to usability and found adequate discriminant validity, hence, concluded that all three constructs should be included in usability testing unless domain-specific studies suggest otherwise.

The relationships between these three constructs have previously been analyzed and demonstrated the direct impact of the two performance measures—efficiency and effectiveness—on satisfaction (for example, Churchill & Suprenant, 1982; Tse & Wilton, 1988). Applying the work by Frokjaer et al. (2000) in this context, it is argued that each discriminant performance dimension, i.e. efficiency and effectiveness, carries a respective satisfaction measure, i.e. Satisfaction with Efficiency and satisfaction with Effectiveness. As the impact of performance on satisfaction will be tested in this study by focusing on mobile devices for wireless data services, the following hypotheses are proposed:

H2a. 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.

H2b. 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.

2.3. Usability and confirmation of pre-trial expectations

As discussed previously, distractions may have effects on the usability of a mobile device. Additionally, pre-trial user perceptions, such as expectations of performance with a particular mobile application, may in fact have an impact on the post-trial perceptions of its usability and on the user's intentions of its future adoption. This argument is in line with the Expectancy-Confirmation Theory (ECT), a prominent theory from the field of consumer behavior, which was first proposed by Oliver (1980) (for a review see Oliver, 1997). As shown in Fig. 1, ECT describes the process by which a consumer assesses his/her satisfaction with a product or service. The process begins with a consumer forming expectations of performance 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 (Olshavsky & Miller, 1972). Use of the product will result in the consumer forming perceptions of the product's actual performance. Initial expectations and perceived performance are combined in a qualitative assessment. Consequently, negative confirmation (or dissatisfaction) is experienced as a result of unmet expectations, whereas positive confirmation (or satisfaction) occurs when pre-trial expectations are surpassed by the consumer's trial experience.

ECT has previously been used in various disciplines, including Information Systems (ISs) Bhattacherjee & Premkumar, 2004; McKinney, Yoon, & Zahedi, 2002. The evaluative process of satisfaction outlined by ECT provides a model for understanding and measuring end-user satisfaction with IS performance. Several studies have shown performance to be such a strong predictor of satisfaction or perceived service quality (Cronin & Taylor, 1992; Erevelles & Leavitt, 1992), with expectations being only a relatively weak predictor (Bolton & Drew, 1991a, 1991b; Boulding, Kalra, Staelin, & Zeithaml, 1993). This finding has led some scholars to argue for the exclusion of pre-trial expectations from the structural and measurement models (Hammer, 2006). Consistent with these recommendations as well as additional support by Babakus and Boller (1992) and Khalifa and Liu (2002), who found that expectations did not contribute to confirmation scores significantly, this study will omit the measurement of expectations.

Furthermore, in the domain of usability, ECT 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 ECT essentially creates a dyadic satisfaction formation process for the consumer, one for efficiency and one for effectiveness. With this dual path in mind, and given
that performance has been shown to have a positive effect on the confirmation of pre-trial expectations (Oliver, 1980, 1997; Van Ryzin, 2004), the following hypotheses are proposed:

H3a. Higher levels of perceived efficiency of a mobile device will positively influence the confirmation of user expectations of efficiency in using the mobile device for wireless data services.

H3b. Higher levels of perceived effectiveness of a mobile device will positively influence the confirmation of user expectations of effectiveness in using the mobile device for wireless data services.

If pre-trial expectations are met, confirmation occurs, and finally if expectations are surpassed, a consumer experiences positive confirmation, and finally if expectations are met, satisfaction occurs, and consequently satisfaction (Oliver, 1980; Van Ryzin, 2004). Hence, the following hypotheses are proposed:

H4a. Higher levels of confirmation 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.

H4b. Higher levels of confirmation 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.

Lastly, there is a strong effect between satisfaction and a consumer's behavioral intention to use a product (Cronin & Taylor, 1992; Gotlieb, Grewal, & Brown, 1994; Mittal, Kumar, & Tsiros, 1999; Taylor & Baker, 1994). This relationship will be explored further in the next section.

2.4. Usability and technology adoption

As outlined above, usability can impact the growth of m-Business, as poor usability may hinder adoption. Here we examine the impact of usability on the adoption of mobile devices, PDAs in the context of this study, for wireless data services. According to the Theory of Reasoned Action (TRA) Fishbein & Ajzen, 1975, a consumer’s behavioral intention is determined by his/her behavioral intention, which in turn is determined by the person’s attitude and subjective norm concerning the specific behavior, in this case the use of PDAs for wireless data services. Furthermore, in TRA, a consumer’s attitude is determined by his or her beliefs about the potential consequences of performing the behavior as well as the evaluation of those consequences.

Usability is one such belief that may directly or indirectly impact a user’s attitude towards using mobile devices (Hsu & Chiu, 2004). Therefore it can be argued that usability impacts attitude, which in turn determines behavioral intention. Similar to TRA, the Technology Acceptance Model (TAM) Davis, Bagozzi, & Warshaw, 1989 also argues that actual use of an information system (e.g. a mobile device) is impacted by the user's behavioral intention to use the technology. As subsequent studies have shown there is a strong positive relationship between attitude towards use, behavioral intentions towards use, and actual use of a technology (Venkatesh & Davis, 2000). Therefore, measuring consumer’s behavioral intention towards using a mobile device for wireless data services may suffice in predicting actual usage of this technology.

A plethora of studies suggest a strong effect between satisfaction and a consumer's behavioral intention to use a product (e.g. Mittal et al., 1999; Taylor & Baker, 1994). Several studies have validated the positive influence of higher levels of user satisfaction on intention to use and actual use of information systems (DeLone & McLean, 1992; Rai, Lang, & Welker, 2002; Seddon, Staples, Patnayakuni, & Bowtell, 1999). Hence, the following hypotheses are proposed linking the two dimensions of user satisfaction identified in subsection 2.3 above (i.e. Satisfaction with Efficiency and satisfaction with effectiveness) with a user’s intention to use a mobile device for wireless data service:

H5a. 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.

H5b. 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.

Our proposed research model and hypotheses are outlined in Fig. 2.

3. Methodology

3.1. Experiment 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: Environmental distractions in the form of auditory and visual stimuli). This design allowed for any resulting differences 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 stimuli.
auditory cues in the environment. This approach was used in previous usability studies by Umemura et al. (1992) and Byoungsoo et al. (2009), while a variant (2 × 3 factorial design) was employed by Kjeldskov and Stage (2004).

Experiment tasks involved the use of four PDA applications adapted from often-cited earlier studies: sending text messages (James & Reischel, 2001), scheduling an appointment in the calendar, updating the address book (Lindroth et al., 2001), and searching the Web on a PDA (Rodden, Milic-Frayling, Sommerer, & Blackwell, 2003). The PDA selected, RIM's Blackberry with a QWERTY keyboard input, is a fair representative of typical PDAs currently available in the market in terms of the supported functionality and general form factor. The four tasks selected create the most value for consumers, second only behind voice communication (Jarvenpaa, Lang, & Tuunanen, 2004), and represent a mix of accessing and authoring activities. All tasks were randomized within- and between-applications.

A total of 93 participants were recruited for the study, with a minimum of 20 subjects in each of the four treatments. 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 that lasted approximately 45 min. Participants progressed through the following experiment procedure: pre-test survey, instructions, training, controlled lab experiment, and post-test survey. Participants were not allowed to interact with others during the experiment in an attempt to isolate the conditions that were being tested and to increase the realism of the task.

Similar to Byoungsoo et al. (2009) and Nicholson et al. (2005), the four groups of subjects 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 (environment distractions) as the changing parameters. To study the effect of user motion on performance, subjects were asked to complete the tasks either being seated or while walking in a controlled environment (i.e. large room in a building), staying within the boundaries of a complex path outlined on the ground (i.e. non-linear), and walking at a steady pace.

To further study the effect of auditory and visual distractions on performance, subjects were asked to undertake the tasks either in the absence or presence of background auditory stimuli (in this study music and speech) and visual stimuli (in this study the presence of and motion by five actors hired for this study). Again, distractions in this study were either the isolated or combined effects of user motion and/or visual and auditory stimuli in the environment, which served as the manipulation of the exogenous construct (distraction) of the research model. The factors and their levels that defined the treatment conditions in this study are shown in Table 1.

### 3.2. Subjects

The 93 participants recruited for this study were native English speakers, approximately of equal gender distribution, and covering a broad range for age and education. Subjects were recruited from a major Canadian university and included students, staff, and faculty. This strategy was aimed at soliciting a convenience sample that in fact displayed representative mobile user characteristics in terms of the control variables (i.e. age, gender, education). Of the 93 participants recruited, 87 usable questionnaires were collected. This group exhibited an average age of 28, 97% were at least college-educated, and the women to men ratio was 60/40 – none of the participants had any prior experience with wireless data services. ANOVA tests found no significant differences for subjects in the various treatment groups in terms of subject characteristics (i.e. age, gender, education). Therefore, randomization of assignment across groups was successful in terms of the control variables. Details regarding the participants' demographics are included in Table 2.

### 3.3. Instrument scales and validity

The questionnaire used for data collection contains scales that measure the various constructs shown in the research model and are provided in the Appendix. All scales were adapted from prior studies (i.e. distraction measured by the TLX scale from Hart and Staveland (1988)), (Efficiency by Butts and Cockburn (2002), MacKenzie and Zhang (1999), McHaney, Hightower, and White (1999)), (Effectiveness by Chittaro and Dal Cin (2002)), (ECT constructs by Oliver (1980), Spreng, MacKenzie, and Olshavsky (1996)), (BI by Venkatesh and Davis (2000)), which had established their reliability and validity, thereby satisfying content validity. In accordance with the advice of Fishbein and Ajzen (1975) and Davis (1989) 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.

While no follow up invitation was extended to solicit participants, and the entire study period lasted 7 days, we tested for the existence of temporal bias. 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 confirmations), usability (i.e. efficiency, effectiveness, and satisfaction), 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). Thus, no temporal bias exists in the data collection. 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).

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Participant demographics.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measure</td>
<td>Items</td>
</tr>
<tr>
<td>Gender</td>
<td>Male</td>
</tr>
<tr>
<td></td>
<td>Female</td>
</tr>
<tr>
<td>Age</td>
<td>&lt;21</td>
</tr>
<tr>
<td></td>
<td>21–25</td>
</tr>
<tr>
<td></td>
<td>26–30</td>
</tr>
<tr>
<td></td>
<td>31–35</td>
</tr>
<tr>
<td></td>
<td>36–40</td>
</tr>
<tr>
<td></td>
<td>41–45</td>
</tr>
<tr>
<td></td>
<td>46–50</td>
</tr>
<tr>
<td></td>
<td>51–55</td>
</tr>
<tr>
<td></td>
<td>&gt;55</td>
</tr>
<tr>
<td>Education</td>
<td>Graduate studies or degree</td>
</tr>
<tr>
<td></td>
<td>Undergraduate studies</td>
</tr>
<tr>
<td></td>
<td>Some college studies</td>
</tr>
<tr>
<td></td>
<td>High school</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Treatment conditions (i.e. independent variables) for dissertation study.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experiment independent variables</td>
<td>USER</td>
</tr>
<tr>
<td></td>
<td>STATIC</td>
</tr>
<tr>
<td>Environment</td>
<td>Quiet and alone</td>
</tr>
<tr>
<td>Auditory and visual stimuli</td>
<td>Group III</td>
</tr>
</tbody>
</table>
Study data was then tested for significant differences between those collected from completed questionnaires (N = 87) and those from incomplete questionnaires (N = 6). In terms of expectations, usability (i.e. efficiency, effectiveness, and satisfaction), 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 (Wilks’ Lambda = 0.754; significance level = 0.978). Also, no statistically significant difference was found in participant-specific traits (Wilks’ Lambda = 0.946; significance level = 0.180). Thus, no non-response bias was present for the data.

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. 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 3 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.

The factor loadings for the total set of items used in this study are summarized in Table 4 (along with their cross-loadings). Hair, Anderson, Tatham, and Black (1995) suggest that an item is significant if its factor loading is greater than 0.5 to ensure construct validity. Adherence to this criterion required the modification of only one scale (Distraction, measured by the TLX scale) through the removal of two items: TLX5 and TLX6. TLX, or Task Load Index, was initially used by Hart and Staveland (Hart & Staveland, 1988) to capture study participants’ cognitive load. 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).

Upon further testing it was shown that non-response, temporal, and common method biases were not present in our data set.

Results of tests for convergent validity (Bagozzi, Yi, & Phillips, 1991), discriminant validity (Bagozzi et al., 1991; Fornell & Larcker, 1981), construct means and Cronbach’s alpha can be found in Table 5. All constructs had adequate reliability (Carmines & Zeller, 1979) and internal consistency well above the 0.7 threshold (Nunally, 1978). Cronbach z-values were satisfactory for our constructs (0.844–0.956) and constructs’ AVE exceeded the 0.5 benchmark for convergent validity (Fornell & Larcker, 1981). The square root of the variance shared between a construct and its items was greater than the correlations between the construct and any other construct in the model (see Table 6) suggesting discriminant validity (Fornell & Larcker, 1981). Discriminant validity was confirmed by verifying that all items load highly on their corresponding factors and did not cross-load on other factors (see Table 4 above).

### Table 3

<table>
<thead>
<tr>
<th>Construct</th>
<th>Correlation with BI (n = 93)</th>
<th>Correlation with BI (n = 46)</th>
<th>Absolute difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>TLX</td>
<td>−0.448</td>
<td>−0.504</td>
<td>0.056</td>
</tr>
<tr>
<td>ExpEffi</td>
<td>0.335</td>
<td>0.360</td>
<td>0.025</td>
</tr>
<tr>
<td>ExpEffE</td>
<td>0.368</td>
<td>0.425</td>
<td>0.057</td>
</tr>
<tr>
<td>PerEffi</td>
<td>0.564</td>
<td>0.590</td>
<td>0.036</td>
</tr>
<tr>
<td>DistEffi</td>
<td>0.428</td>
<td>0.430</td>
<td>0.002</td>
</tr>
<tr>
<td>SatEffi</td>
<td>0.569</td>
<td>0.578</td>
<td>0.009</td>
</tr>
<tr>
<td>PerEffE</td>
<td>0.332</td>
<td>0.239</td>
<td>0.093</td>
</tr>
<tr>
<td>DistEffE</td>
<td>0.260</td>
<td>0.286</td>
<td>0.026</td>
</tr>
<tr>
<td>SatEffE</td>
<td>0.520</td>
<td>0.560</td>
<td>0.040</td>
</tr>
</tbody>
</table>

a Correlation of factor scores between exogenous variables (independent and dependent variables: TLX, EXP EFFI, EXP EFFE, PER EFFI, DIS EFFI, SAT EFFI, PER EFFE, DIS EFFE, SAT EFFE) with the right-most endogenous variable (BI) using total data.

b Correlation of factor scores using paired data.

4. Results

The structural model shown in Fig. 2 was tested using the variance-based Partial Least Square (PLS) method. 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 does not have expectations of normality (Casaló, Flavián, & Guinalíu, 2008). This is important for this research due to the varying individual thresholds that participants may display with respect to distraction tolerance. Second, PLS allows for multiple measures for each construct, so 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 (Casaló et al., 2008; Khalifa & Liu, 2002). Another strength of PLS was that it requires small to medium sample sizes (Carmines & Zeller, 1979; Compeau & Higgins, 1995). The minimum sample size for a PLS analysis should be the larger of (i) 10 times the number of items for the most complex construct; or (ii) 10 times the largest number of independent variables impacting a dependent variable. In our model, the most complex construct has 4 items and the largest number of independent variables estimated for a dependent variable is only 3. Thus, our sample size of 87 is more than adequate for PLS. The PLS result model is shown in Fig. 3. Overall, the model demonstrated high explanatory power. The R-square of the Intention to Use construct was 0.39, or 39% of the variance in user intentions to adopt mobile devices for wireless data services was explained with our model. The R-square values for the rest of the endogenous variables in the core usability model (highlighted in Fig. 3) exceed the 10% benchmark recommended by Falk and Miller (Falk & Miller, 1992), with sole exception of the construct for perceived effectiveness (R-square = 8%). This value does not necessarily pose a threat to the model’s validity. Particularly in behavioral 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). Low R-square values have also been reported in many technology adoption studies (e.g. Davis et al., 1989; Moon & Kim, 2001). An additional explanation for this low R-square value may be that the Perceived Effectiveness construct was associated with only one construct (Distractions). Relative to multi-relationship models, these single or few-relationship associations often provide lower R-square values (Nunally, 1978). In the event that additional non-correlating constructs are introduced as antecedents to this construct with low R-square value (i.e. perceived effectiveness), the score would probably increase considerably.
From the original 10 hypotheses, 9 were supported and one was not supported. Table 7 presents the validation of these hypotheses in more detail.

Reviewing the above results, four sets of conclusions may be drawn. First, it was theorized that incremental cognitive load contributes to Behavioral Intention (H5a) was not supported (p-value < 0.001), and H5b (p = 0.098; p-value < 0.001).

Second, regarding usability, both performance and satisfaction were decoupled into two respective components of Efficiency and Effectiveness. This decoupling obtained statistical support: H2a received strong support for the positive effect of Perceived Efficiency on Satisfaction with Efficiency (β = 0.515; p-value < 0.001), and H2b received strong support for the positive effect of Perceived Effectiveness on Satisfaction with Effectiveness (β = 0.226; p-value < 0.001).

Third, the applicability of the ECT in mapping the user’s evaluative process of a mobile device for wireless data services was supported as both the strong direct effects of performance on satisfaction were shown (reported in previous paragraph) as was the mediating role of confirmation of pre-trial expectations between performance and satisfaction (for efficiency: β = 0.335; p-value < 0.001; effectiveness: β = 0.509; p-value < 0.001).

Lastly, adoption of mobile devices was explored by measuring the behavioral intention of users upon obtaining hands-on experience with wireless data services and their consequent level of satisfaction. From the two hypotheses proposed, only Satisfaction with Effectiveness (H5a) was shown to be statistically significant (β = 0.508; p-value < 0.001) in impacting the aforementioned behavioral intention; the path from Satisfaction with Effectiveness to Behavioral Intention (H5b) was not supported (β = 0.098; p-value = 0.501).

The control variables in this study (i.e. age, gender, education) were also 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. Any changes in R-square

---

**Table 4**

Matrix of loadings and cross-loadings.

<table>
<thead>
<tr>
<th>ITEM</th>
<th>TLX</th>
<th>PerEffi</th>
<th>ConEffi</th>
<th>SatEffi</th>
<th>PerEffi</th>
<th>ConEffi</th>
<th>SatEffi</th>
<th>BI</th>
</tr>
</thead>
<tbody>
<tr>
<td>TLX1</td>
<td>0.872</td>
<td>-0.463</td>
<td>-0.264</td>
<td>-0.411</td>
<td>-0.281</td>
<td>-0.107</td>
<td>-0.289</td>
<td>-0.146</td>
</tr>
<tr>
<td>TLX2</td>
<td>0.724</td>
<td>-0.443</td>
<td>-0.374</td>
<td>-0.396</td>
<td>-0.199</td>
<td>-0.179</td>
<td>-0.212</td>
<td>-0.105</td>
</tr>
<tr>
<td>TLX3</td>
<td>0.823</td>
<td>-0.374</td>
<td>-0.255</td>
<td>-0.394</td>
<td>-0.157</td>
<td>-0.010</td>
<td>-0.253</td>
<td>-0.069</td>
</tr>
<tr>
<td>TLX4</td>
<td>0.918</td>
<td>-0.468</td>
<td>-0.309</td>
<td>-0.471</td>
<td>-0.270</td>
<td>-0.097</td>
<td>-0.291</td>
<td>-0.173</td>
</tr>
</tbody>
</table>

**Note:** TLX – NASA Task Load Index; ConEffi – Confirmation of Efficiency Expectations; ConEffi – Confirmation of Effectiveness Expectations; PerEffi – Perceived Efficiency; SatEffi – Satisfaction with Efficiency; PerEffe – Perceived Effectiveness; SatEffe – Satisfaction with Effectiveness; BI – Behavioral Intention to Use.

**Table 6**

Correlation Matrix and Discriminant Validity Assessment.

<table>
<thead>
<tr>
<th>Items</th>
<th>TLX</th>
<th>PerEffi</th>
<th>ConEffi</th>
<th>SatEffi</th>
<th>PerEffi</th>
<th>ConEffi</th>
<th>SatEffi</th>
<th>BI</th>
</tr>
</thead>
<tbody>
<tr>
<td>TLX</td>
<td>0.839</td>
<td>0.256</td>
<td>0.820</td>
<td>0.835</td>
<td>0.503</td>
<td>0.721</td>
<td>0.655</td>
<td>0.897</td>
</tr>
<tr>
<td>PerEffi</td>
<td>-0.526</td>
<td>0.607</td>
<td>0.835</td>
<td>0.503</td>
<td>0.721</td>
<td>0.655</td>
<td>0.897</td>
<td></td>
</tr>
<tr>
<td>ConEffi</td>
<td>0.358</td>
<td>0.765</td>
<td>0.897</td>
<td>0.757</td>
<td>0.673</td>
<td>0.854</td>
<td>0.919</td>
<td></td>
</tr>
<tr>
<td>SatEffi</td>
<td>0.503</td>
<td>0.721</td>
<td>0.655</td>
<td>0.503</td>
<td>0.721</td>
<td>0.655</td>
<td>0.897</td>
<td></td>
</tr>
<tr>
<td>PerEffi</td>
<td>-0.275</td>
<td>0.414</td>
<td>0.340</td>
<td>0.373</td>
<td>0.548</td>
<td>0.569</td>
<td>0.669</td>
<td></td>
</tr>
<tr>
<td>ConEffi</td>
<td>0.122</td>
<td>0.309</td>
<td>0.269</td>
<td>0.321</td>
<td>0.410</td>
<td>0.411</td>
<td>0.540</td>
<td></td>
</tr>
<tr>
<td>SatEffi</td>
<td>-0.315</td>
<td>0.465</td>
<td>0.352</td>
<td>0.513</td>
<td>0.940</td>
<td>0.976</td>
<td>0.976</td>
<td></td>
</tr>
<tr>
<td>BI</td>
<td>-0.154</td>
<td>0.226</td>
<td>0.093</td>
<td>0.178</td>
<td>0.561</td>
<td>0.976</td>
<td>0.976</td>
<td></td>
</tr>
</tbody>
</table>

Notes: 1. The diagonal elements in bold (the square root of the average variance extracted) should exceed the inter-construct correlations below and across them for adequate discriminant validity. 2. Analysis on single-item constructs (i.e. PerEffi and ConEffi) was not performed.

---
values may indicate an impact of an independent construct on 
dependents one(s) Carmines & Zeller, 1979. Overall, the fully con-
trolled model improved the R-square values for all dependent con-
structs except for Satisfaction with Efficiency, which remained 
unchanged. Results showing moderate or considerable impacts in-
clude the following:

- Age had a considerable impact on distractions: older subjects 
  were more negatively impacted compared to younger ones.
- Gender had a considerable impact on distractions and on 
  behavioral intention: women were more impacted by distrac-
  tions and had lower intentions to use wireless data services 
  than men.

Additionally, the path coefficients and significance levels 
between the control variables and the dependent constructs were 
reviewed. A strong beta coefficient (0.237) and corresponding t-
value (2.611) supported that women may be affected more by 
distractions than men (p-value < 0.01).

5. Conclusions

This paper proposed a new model to further understanding of 
mobile usability. Specifically, distractions and user expectations 
were examined for their potential impacts on usability dimensions 
and, ultimately, on behavioral intention to use mobile services. Implications for both theory and practice are described next.

5.1. Implications for theory

From a theoretical point of view, this work contributes to usability research by providing a better understanding of the 
impacts of auditory, visual and motion distractions on the use of 
mobile devices for wireless data services. We found that such dis-
tractions do have a significant negative impact on the perceived 
efficiency and effectiveness of mobile device use, a finding that is 
consistent with research findings regarding the negative effect of 
distractions on driving performance (Ljungberg et al., 2004). While 
controlled laboratory studies help to ensure experimental rigor, 
academics must remember that usability may be greatly affected 
by context of use. This is particularly true for mobile devices, 
where distractions are more likely to occur while users are ‘on 
the move’ in the environment. Our experimental design attempted 
to approximate real-world scenarios for the tested product and 
services, while upholding a rigorous methodology, and can serve 
as an example for future empirical research on mobile usability.

With respect to ECT, and in alignment with (Babakus & Boller, 
1992) who studied satisfaction with utility services, we observed 
performance as a stronger predictor of satisfaction than the confir-
mation of pre-trial expectations was, this time in the context of a 
mobile device for wireless data services. This study also reinforces 
the positive relationship between usability and satisfaction previ-
ously validated in screen web experiences and website loyalty 
(Casaló et al., 2008) this time supported in the context of wireless 
data services.

5.2. Implications for practice

For practice, this study’s results have direct implications for 
designers and retailers of mobile devices. By decomposing satisfac-
tion we offer relevant insight as to which performance dimension 
(i.e. efficiency) becomes critical in personal use decisions for a mo-
bile device. Learnability, ease of use, and time required to complete 
a task are prevalent dimensions in the decision making process of
using a mobile device for wireless data services, while successful task completion seems to be less relevant. Thus, complex interfaces that offer enhanced capabilities while having a toll on the efficiency of a mobile device may deter a consumer from using it. This observation is in agreement with the findings of Rust, Thompson, and Hamilton (2006), who also call upon device developers to avoid “feature fatigue”, i.e. overwhelming users by adding device functionality that leads to increased use complexity and product dissatisfaction. Such dissatisfaction often leads to product returns, customer attrition, and/or negative effects on brand equity. If repeat business, i.e. increasing the life-time value of the customer, is the goal, manufacturers are better off producing a device that is simple and easy to use than a “powerful” all-in-one device (Marcus, 2003).

Additionally, strong beta coefficients and corresponding t-values indicate that women may be affected more by distractions than men, which is in agreement with the results of Bruni’s (2004) work, who examined the impact of instant messaging on task performance. If women are less robust to distractions than men in the context of using a mobile device for wireless data services, it is likely that such applications, both professional and leisurely, will not be as popular with women. Hence, manufacturers could arguably capture a market opportunity by designing different interfaces or offering special tools/accessories for mobile devices geared towards increasing usability for women. Similarly, usability and accessibility become imperative design considerations for devices aimed at older users, who were found to be more susceptible to distractions.

Furthermore, the differences found in this study among user groups in terms of performance, satisfaction, and intention to use wireless data services highlight the importance of targeted marketing communications, thereby creating realistic product expectations for each user group. In addition, businesses providing decision aids, such as recommendation agents that help identify a user’s real needs, may help increase the prominent importance of usability in the purchase decision (Chin, 1998; Chin & Gopal, 1995; Rust et al., 2006). Additionally, closing the gap between pre- and post-use consumer preferences may lead to higher product satisfaction, repeat business, and favorable effects on brand equity.

5.3. Limitations

As with all experimental studies, there are limitations for this study, which can prompt future research in this area. First, the study’s tasks were simulated in a laboratory setting. Thus, any

<table>
<thead>
<tr>
<th>Table A1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Construct items and their factor loadings.</td>
</tr>
</tbody>
</table>

TLX – NASA Task Load Index (Fishbein & Ajzen, 1975), scales used were Semantic Differential 1–20 (Low–High)

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 or leisurely (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)

TLX5a 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)

TLX6a 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?

PerEff: Perceived Effectiveness (Chen & Vertegaal, 2004)

PerEff1 I was able to complete all wireless data services on the mobile device successfully

ConEff: Confirmation of pre-trial Expectations with Effectiveness (Marcus, 2003; Oliver, 1997)

ConEff1 Learnability (that is, the degree to which it was easy to learn how to use) of the mobile device for wireless data services was...

ConEff2 The time required to use the mobile device for wireless data services was...

ConEff3 The user friendliness of the mobile device for wireless data services was...

ConEff4 Ease of use of the mobile device for wireless data services was...

SatEff: Satisfaction with Effectiveness (Oliver, 1997)

SatEff1 Thinking about my experience with the effectiveness of this device for wireless data services, I feel... Terrible (1) ... Delighted (7)

SatEff2 Thinking about my experience with the efficiency of this device for wireless data services, I feel... Very displeased (1) ... Very pleased (7)

SatEff3 Thinking about my experience with the effectiveness of this device for wireless data services, I feel... Very dissatisfied (1) ... Very satisfied (7)

SatEff4 Thinking about my experience with the efficiency of this device for wireless data services, I feel... Frustrated (1) ... Contented (7)

Bi: Behavioral Intention (Sanders et al., 1978)

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

* Denotes items removed from the subsequent analysis.
sense of urgency or other contextual responses that a user may experience in a real-setting may 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. Second, the experiment was carried out using one particular mobile device (RIM’s Blackberry) with one particular interface input mode (QWERTY keyboard). Results from this study should be validated across multiple mobile devices and interface input modes. Third, the experiment was conducted in a Canadian context and should not be generalized to other cultures before further validation.

5.4. Future research

Given the interdisciplinary nature of this study, several opportunities emerge for future research. First, a study could collect 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 (self-reported) and objective (device) data is needed. A more focused investigation for performance-related attributes (e.g., time, learnability, error rate, success rate, etc.) is also warranted. By conducting mobile usability studies that triangulate between self-reported, observed, and device data, analyses would be richer and with greater external validity (i.e. closer to the ‘real world’).

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. Beyond gender, studies could explore the effects of cultural traits on the motivations, adoption, and user experience with wireless data services, as was the recent effort by Lee, Kim, Choi, and Hong (2010) with a focus on the mobile Web. Marketing practices would be better informed through the results of such studies, guiding the gender-specific and culturally-appropriate tailoring of marketing communication messages for wireless data services.

Acknowledgments

The authors would like to thank the journal editor, Dr. Tennyson, and the reviewers for their detailed, constructive feedback that strengthened the presentation and contributions of this manuscript.

Appendix A

See Table A1.

References


