Available online at www.sciencedirect.com
ScienceDirect
Journal of Strategic Information Systems 16 (2007) 393–412
Strategic Information Systems
www.elsevier.com/locate/jsis

Understanding user behavior with new mobile applications

Raquel Benbunan-Fich [a,*], Alberto Benbunan [b]

[a] SCIS Department, Baruch College, CUNY, Box B11-220, New York, NY 10010, United States
[b] Mobile Dreams Factory, Madrid, Spain

Received 30 November 2006; accepted 15 August 2007
Available online 5 November 2007

Abstract

This paper presents the results of an exploratory study designed to evaluate Microediciones®, a new mobile content-based application that delivers daily summaries of newspapers into cell phones. The research question underlying this study sought to understand how users, accustomed to using their handsets as talking devices, download applications and navigate through data. This question is explored from the theoretical perspective of automaticity with data collected via protocol analysis of a sample of subjects. An analysis of the problems experienced by the users provided the basis to suggest specific operational improvements to the application, bearing in mind the strategic objectives of the mobile company to obtain revenue from user subscriptions and advertising. Beyond the strategic implications for the company launching this particular service, the results of this study have significant theoretical and practical implications for researchers and developers of new mobile applications.

Keywords: Mobile applications; User behavior; Human error; Think-aloud; Protocol analysis

1. Introduction

Due to the development of wireless applications, cell phones are used for much more than just voice communication. According to a report by LogicaCMG, about one-fifth

* Corresponding author. Tel.: +1 646 312 3375.
E-mail addresses: rbfich@baruch.cuny.edu (R. Benbunan-Fich), Alberto.benbunan@mdfactory.com (A. Benbunan).

0963-8687/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.jsis.2007.08.002

of mobile phone users worldwide have downloaded into their handsets content such as directions, weather reports, stock prices and other types of information; and this percentage is expected to rise in the next few years (LogicaCMG, 2005).

In this category of content-based wireless applications, a mobile services company based in Spain – Mobile Dreams Factory – recently developed the concept and the technical platform for Microediciones®. These are daily mobile Java [J2ME] editions of newspapers that individual subscribers download into their cellular phones. These microeditions contain an abbreviated version of the most important news stories covered by newspapers. The content is presented on consecutive screens as if they were pages in the newspaper. However, unlike printed papers, the headlines function as hyperlinks to article summaries, some featuring photos as well. Ads are presented in full-page format, and inserted between content screens (Mobile Dreams Factory, 2004).

To ensure the success of a content-based application like Microediciones®, it is essential to obtain a clear understanding of user behavior in a mobile environment (Anckar and D'Incau, 2002; Lu et al., 2005). In such an environment, two main aspects are deemed critical. The first one is the technical environment of cell phones: the small screen and numeric keypad layout may impose physical constraints for content delivery and interactivity, when compared to traditional computer environments. The second critical issue is to understand how people use a mobile phone to interact with content, as the delivery of data into cell phones is likely to change how companies serve customers with new products and services (Lu et al., 2005; Siau et al., 2001).

Our research question was to understand how users, who had experience using a mobile device for *voice* services, utilized the device for *data* services. This question is explored from the theoretical perspective of automaticity (Logan, 1988) and data are collected via protocol analysis of a sample of subjects using the new mobile application for the first time.

The potential of new mobile applications, along with the technical environment of cell phones and the behavioral challenges posed by a new user base, offer a fertile area for Research and Development. Unlike other R&D endeavors with a similar purpose, this project benefited from a collaborative effort between industry and academia. At the time of this study, Mobile Dreams Factory was evaluating the adoption of a mixed business model combining advertising revenue and user subscriptions for Microediciones®. Thus, the evaluation took place in a broader strategic context where user preferences were analyzed in terms of their implications for the potential sources of profitability of the mobile application.

To present the results of this study, this paper opens with a review of literature to establish the theoretical background for the study. Then, it describes the research methodology. It follows with the data analysis and results, and then with the discussion, implications and conclusions.

2. Theoretical background

User interaction with a technological artifact is the result of the interplay among the user's cognition, the possibilities and constraints afforded by the artifact and the nature of the task. These three elements – cognition, artifact and task – define the interaction environment (Gray and Altman, 2001). As users carry out their tasks with a technological artifact, the interaction includes physical activities (or motor actions) as well as mental

activities (or cognitive processes). Accordingly, the steps necessary to complete the task can be separated into physical and cognitive components (Zhang and Norman, 1994).

Cognitive processes associated with using a technological artifact include formulating an intention and developing an action plan. The user first establishes what needs to be done, and develops a plan with the specific steps needed to reach the goal. Then, the physical steps are carried out (Norman, 1988). When using technological artifacts for the first time, the steps are performed gradually as users formulate goals and develop action plans. With practice from repeated exposures or from similar environments, individuals are able to automatically retrieve action plans from memory instead of formulating new procedures to carry out tasks. This ability to perform activities quickly, effortlessly and without conscious awareness is called *automaticity* (Kai et al., 1996; Logan, 1988).

A key factor in achieving automaticity is the user's familiarity with the technological environment. Automaticity relies on the retrieval of stored information accumulated after practice in a consistent environment (i.e. repeated exposures to the task environment). Practice is important because it increases the amount of information previously stored and available for retrieval. Consistency is also important because it ensures that the instances retrieved will be useful (Logan, 1988). Therefore, first-time users exposed to new technology applications will achieve automaticity if they can transfer their familiarity with other domains to the current environment.

Individuals think about specific activities at different levels of abstraction. Higher levels refer to *what* needs to be done, while lower levels refer to *how* it is done (Kai et al., 1996). Familiar activities are identified by users at a high (goal-oriented) level, which facilitates automaticity, while unfamiliar activities are identified at a lower (action-oriented) level, which inhibits automaticity. When both levels are available, the higher level tends to become dominant. However, the lower level becomes central when errors are made, or when the task environment is unknown (Kai et al., 1996). For instance, among novice users of a technological artifact the lower level of abstraction will prevail, as they have to carefully think about how to perform the steps to carry out the task. In contrast, higher levels of identification will be dominant for experienced users, who can draw from their familiarity with the domain and perform the lower levels automatically.

In addition to the degree of familiarity with the domain, the nature of the instructions (provided through the interface, in the case of technological artifacts) affects successful interaction. Instructions provided at the level in which the user thinks (goal-oriented for high level of identification, and action-oriented for low levels) promote efficient processing, while mismatches between instruction types and abstraction levels are likely to result in errors (Kai et al., 1996). Hence, when someone uses new technological artifacts to perform specific tasks, the information embedded in the design of the artifact (and conveyed through the interface) is as important as the prior experience of that individual (Norman, 1988). An easy-to-use interface minimizes the need for users to plan what to do and how to do it; and it allows users to skip the formulation phase and move directly from goal to action, thus promoting automaticity.

Well designed artifacts reduce the need for users to remember large amounts of information, while poorly designed artifacts increase the demands on the users and therefore increase the likelihood of errors (Norman, 1988). In this context, two types of errors occur: Mistakes due to a mismatch between the user's understanding of the artifact and the actual design of such artifact; and slips due to a deviation between the user's intentions and the execution of the correct action sequence. Consequently, for well designed applications,

most of the errors committed by expert users (those with comparatively more knowledge) tend to be action slips, while mistakes are predominantly made by novices (Norman, 1981; Reason, 1990).

A more detailed classification of errors by Gray (2000) addresses the underlying reasons behind user errors. This taxonomy distinguishes among knowledge-based, rule-based, and skill-based errors. *Knowledge-based errors* happen because the user lacks the required knowledge. *Rule-based errors* occur when users develop (or apply) a rule that is inappropriate in the context of the application. *Skill-based* errors are those that occur when knowledge is available, and the appropriate rule is selected, but there is a slip in the execution. Slips are indicative of a mismatch between the mental instructions for carrying out an action and the physical execution of such action. This three-way classification (knowledge-based, rule-based and skill-based) is more useful from an analytical viewpoint than the traditional mistake vs. slip dichotomy, because it provides two different reasons for making mistakes (lack of knowledge and formulation of wrong rules). Thus this model offers more insight into the cognitive aspects of the user activity, and improves our understanding of user behavior with mobile applications.

2.1. Prior research on mobile applications

Mobile technology is more complex than other types of Information Technology environments, as it includes not only the hardware (handheld devices) and software (interface and applications that run in these handsets), but also the communications networks that make possible most of the mobile services (Jarvenpaa et al., 2003). In addition, mobile devices are considered *hybrid* because they usually combine voice and data features (Sarker and Wells, 2003). Some devices originally designed for voice features are increasingly used for data-related applications. Unlike information systems operating in traditional computer environments, mobile applications present visual and motor challenges to users because they are delivered through handheld devices. In the case of mobile phones, the small screen and numeric keypad layout of the handset may present physical constraints such as difficulties in reading/comprehension due to small screen displays, as well as interactivity restrictions related to the use of fine motor skills when operating the keypad (Buchanan et al., 2001; Palen and Salzman, 2002).

Despite these limitations, prior research suggests that users are able to deal with the physical requirements of new mobile environments. For example, Rahman and Muter (1999) examined reading efficiency in different types of small-display presentations. Their findings suggest that reading comprehension is not affected by the type of display because users are capable of physically adapting to smaller screen sizes by reducing their eye movements. In the area of visual and motor skills, Buchanan et al. (2001) studied three types of navigation in an application that delivered news headlines to mobile users. Each navigation scheme (horizontal scrolling, vertical scrolling and paging) required a particular keystroke and resulted in a different visual effect. Their results indicate that while vertical scrolling is more natural than horizontal scrolling, any scrolling is potentially tedious when used extensively. In contrast, paging (moving through pages) represents a trade-off, as it is less monotonous than scrolling but adds a layer of complexity to the interaction. However, their study found that user errors with each type of navigation scheme were very low and not significantly different from each other.

In addition to featuring differing keypad layout and screen sizes, mobile phones operate differently from one another, depending on the type of phone. Even if users are familiar with their handsets, they may not be well versed in using these devices for receiving data or downloading content-based applications. Through experience, users develop a set of skills and relevant knowledge that leads to effective performance with a particular technological environment (Schenk et al., 1998). However, lack of familiarity is likely to result in a wider range of challenges for new users of mobile data services, when compared to users who have experience with other mobile data services. To distinguish between these two types of users, we use the terms novice and advanced, respectively.

Overall, the results of prior empirical studies with mobile technology suggest that user performance with new applications largely results from the user's understanding and subsequent ability to adapt to the physical environment provided by the mobile IT artifact (device and application), given the requirements of the task at hand. According to the theory of automaticity, new users are likely to experience more adaptation challenges than those who, by virtue of being more advanced, can apply their prior experience with mobile applications to a relatively new but similar context. If the application is well designed, advanced users will exhibit a higher degree of automaticity than novel users, despite the fact that both types of users will be interacting with the mobile application for the first time.

An analysis of systematic errors by both types of users offers the opportunity to improve the mobile application. For example, knowledge-based errors, though more frequent for novice users, are reduced with better instructions. In contrast, rule-based errors tend to be the result of general misconceptions and occur among all users, regardless of their level of expertise. When these rule-based errors are made by a portion of the user base, more accurate descriptions of the application are warranted. Similarly, systematic skill-based errors or slips indicate the need for alternative actions, to minimize the likelihood of these slips. The occurrence of errors in systematic patterns may shed some light on the features of an IT artifact that are not well understood by users or not properly designed.

3. Research methods

In order to evaluate Microediciones®, we conducted a comparative study among new and advanced users, and examined the influence of two different types of phones. Our research study was structured in two phases. The first phase was used to test the concept and gather initial user reactions, via focus groups with individual users. Their insights led to initial improvements to the application. The second phase consisted of an in-depth investigation of individual users' interaction with the mobile application via think-aloud trials.

3.1. Microediciones®

Unlike existing wireless content-based applications which typically deliver news headlines, stock quotes, weather reports and other tidbits of information into mobile devices, Microediciones® was conceived as a self-contained information delivery service based on the content of a particular newspaper. Instead of sending headlines in piecemeal fashion, Microediciones® presents a summary of the newspaper's printed edition in an abbreviated and easy-to-read format. There are several distinctive features of this application when compared to other news headlines delivery services.

On the technical front, Microediciones® uses JAVA technology for content delivery of this commercial application. Until now this technology has mostly been used for games, rather than mobile commercial applications. In addition to utilizing this technological platform, the mobile company secured access to proprietary content by partnering with major newspapers in Spain. Another important difference between this application and its competitors is that the content is downloaded on the phone and then viewed offline. This makes it possible to access the mobile edition in places without cellular signals, such as subways or underground locations. The application also works in most multimedia-enabled phones, and thus does not require a state-of-the-art mobile phone. The third difference is the use of full-page ads inserted in the content of the microedition. Upon launching the service, revenues will be generated by user subscriptions (pay-per-download or monthly payment plans), though eventually the company would like to shift to a mixed business model based on user subscriptions and advertising revenue.

3.2. Focus group

For the focus group, seven subjects (6 males and 1 female, aged between 19 and 30 years) were recruited. All of them were cell phone users and active users of text messaging. The group was introduced to the concept of Microediciones® through a PowerPoint presentation which featured screen shots of the application delivered in a typical phone. After this brief presentation, participants were given the opportunity to express their opinions and make suggestions, and were also questioned about the likelihood of becoming subscribers of this service.

Important suggestions emerged from these initial reactions. Focus group participants did not understand the process that delivers the application, or the basic navigation actions through the microedition. This confusion resulted from the fact that the PowerPoint screen shots started with the application already downloaded in the phone and did not feature any navigation instructions on the initial screen. Also, at this point in the design of the application, the screen shots of the headline pages supported only linear navigation. The sole available buttons were "next" and "back", and only the last page of the microedition allowed users to exit the application. As a result of the comments gathered from focus group participants, downloading instructions were improved and the "index" and "exit" buttons were added to all the screens.

Another issue that generated some debate among focus group participants was the proportion of ads vis-à-vis the content of the microedition. The presentation featured about 10 pages of content, four of which were full-page ads. Users complained about the high number of ads and suggested a tolerance level of about 20–25%. Finally, the majority of the participants indicated that they would subscribe to this type of service, as long as the content was consistent with the printed edition (at least in terms of headlines and editorial page). Some of them also suggested including an abbreviated version of the front page of the newspaper's printed edition as an initial screen, to provide a greater sense of connection and authenticity to the source.

3.3. Application refinement

With the information gathered from the focus group, the application was improved for individual experimentation with prospective users. The downloading process was stream-

lined and reduced to three steps (Request – Download – Launch). First, to *request* the application, the user sends a text message to a specific 4-digit number. Then, the service provider sends in return a "WAP-Push" message with a link, where the user can *download* the microedition. Once it is downloaded into the phone, the user can *launch* the application and view or read the content offline.

Content from the newspaper "El País," a national newspaper with wide circulation, was used for the trials. The news stories were summarized and arranged in consecutive screens/pages. Based on the feedback from the focus group, the first screen presented a shot of the front page of the newspaper on which the micro-edition is based. The next screen presented the newspaper's editorial page content, while subsequent pages featured headlines. To read a summary of any article, users had only to click on the appropriate headline. The sequence of the first set of screens is presented in Fig. 1.

General navigation instructions were presented on the third screen (after the front page and editorial). In addition to the back ("atras") and next ("siguiente") options, each screen reminded the user of the main navigation buttons (0 for index and 9 for exit), and provided a locator bar highlighting the current page of the microedition (see the description of a typical screen/page in Fig. 2).

3.4. Think-aloud protocol analysis: Individual trials

This phase of the study consisted of asking a sample of subjects to individually use the application and "think-aloud" while performing a task with the "El País" microedition. To facilitate a comparative examination, novice and advanced users of mobile applications from two phone brands were recruited. The outcomes of interest were found in performance data (time to complete the task and accuracy) and in the verbal protocols gathered from the users' interaction with the application.

The "think-aloud" or verbal protocol approach is grounded in the theory of humans as information processors. The basic assumption is that concurrent verbalizations produce an accurate and valid account of the subjects' cognitive processes while performing a task (Ericsson and Simon, 1996). This methodology can be used during a problem-solving episode, and/or during an instance of human computer interaction, to observe cognitive processes as they are taking place. In the Information systems literature, protocol analysis has been extensively used to understand systems development tasks (Glass et al., 1992; Vessey

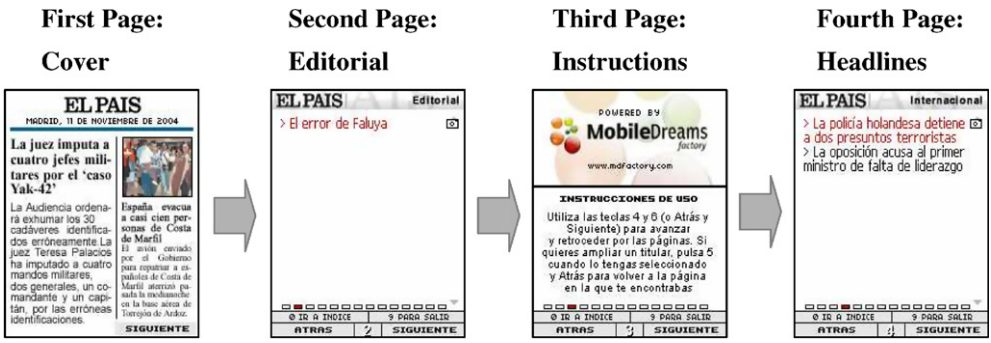

Fig. 1. Sequence of first four screens.

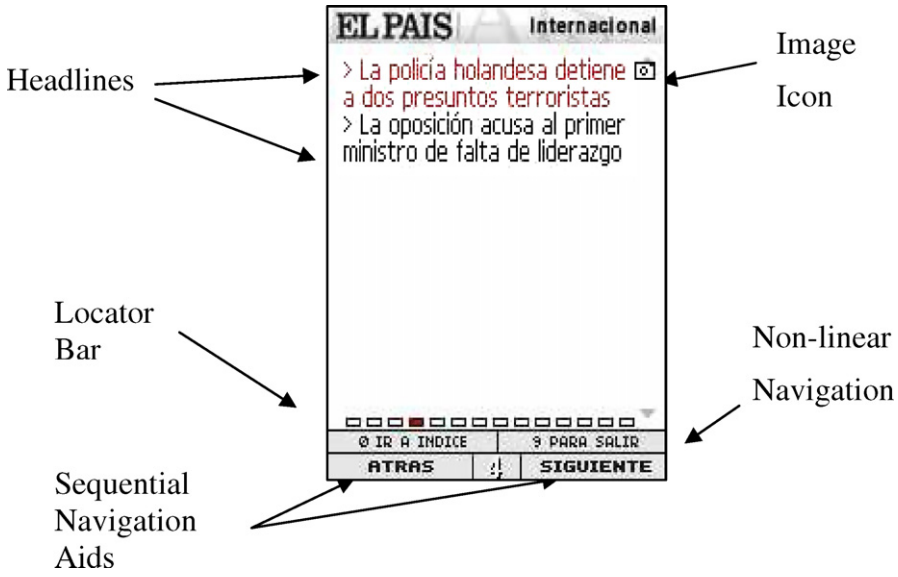

Fig. 2. Description of a screen.

and Conger, 1993) and user behavior regarding knowledge-based systems (Mao and Benbasat, 2000).

3.4.1. Selection procedures

Participants for the protocol analysis phase were recruited from the target user population. Before their recruitment, subjects filled out a questionnaire requesting demographic and general information. Answers to these questions were used to classify subjects into novice or advanced users of mobile content-based applications. Subjects were classified as novices if they had never used their phones for data-related applications. Alternatively, subjects were placed in the advanced category if they had previously used their phones for wireless applications such as information retrieval, web-browsing or any other data services.

Users were also classified according to the type of cellular phone they used. Only subjects with Nokia and Siemens phones were selected to participate. Within each phone group, participants used the same phone model (Nokia 3200 and Siemens C65). We decided to focus this study on these two phones because according to confidential reports from local mobile carriers and public sources, these were the two most popular brands used in Spain at the time of the study (Noticiasdot.com, 2004).

3.4.2. Subjects

A total of 23 subjects participated in individual protocol analysis sessions: 12 in the novice category (six using each phone type), and 11 in the advanced category (6 using Nokia and 5 using Siemens). Their ages ranged from 20 to 34; 61% were males. All were native Spanish speakers, regular cell phone users and full-time employees. These characteristics match the demographic profile of the target population for Microediciones® and are also consistent with the younger readership segment of the newspaper El País.

None of these subjects had participated in the focus group conducted in the first stage, nor had they used Microediciones® before this trial.

3.4.3. Task

Participants were asked to download Microediciones® in their mobile phones and answer the following five "scavenger-hunt" (information search) questions. The questions were presented in a separate worksheet that offered space for writing down the answers.

1. Give the date and name of the newspaper featured in the microedition.
2. What is the title of the editorial? What is it about?
3. What is shown in the picture of the international section?
4. How many ads did you see? Name the companies and where did you see the ads.
5. How did Lupe Valdez become sheriff of Dallas, Texas? Indicate the page where you found this information.

3.4.4. Experimental procedures

The subjects were instructed to think out loud while performing the task. Since one of the objectives of the experiment was to investigate whether users could act without assistance, no specific instructions were given. The subjects had to figure out the process, having been told only how to send the initial request message.

Each session was videotaped with two cameras: one pointed at the phone and the other pointed at the subject. The two video feeds were synchronized in order to match the subjects' comments with specific actions and screen shots. After completing the task, the subjects were interviewed to glean their perceptions about the mobile application. The entire experimental session, including instructions, task execution and exit interview, took approximately 30 min. Task-related activities (downloading, reading and closing the microedition) lasted on average 12 min.

4. Data analysis and results

Quantitative and qualitative data were collected from each subject. Quantitative indicators consisted of objective measures of performance, such as accuracy (number of correct answers to the "scavenger-hunt" questions), and duration (time in minutes required to complete the task). Qualitative data encompassed the participants' verbal protocols as recorded in the videotapes. These verbalizations were transcribed for analysis, along with the appropriate screen/page of the microedition and a time stamp.

To analyze the verbal protocols, we scanned the transcripts to identify instances of problem episodes or anomalies. Consistent with Benbunan-Fich (2001), we defined error instances as events where a problem, misunderstanding or difficulty occurred in the user's interaction with the mobile application. Since typical concurrent verbalizations sometimes consist of incomplete sentences, the unit of analysis is a sentence describing the difficulty as verbalized by the user. In a few instances, it was necessary to piece together a continuous sequence of sentences (or phrases) to obtain complete the description of the problem. We use the term *segment* to indicate a single sentence (or sequence of phrases) reporting a particular difficulty experienced by a user. We then counted the number of error segments from the verbalization data of each user.

4.1. Quantitative results

Performance data regarding the number of correct responses, time to complete the task and error counts were tabulated according to user type and phone type. Table 1 shows the averages of performance data for each combination of type of phone and type of user, and for the full sample.

Since subjects were either advanced or novices with respect of their use of wireless applications, and since they used one of two phones, contingency table analyses were conducted to compare these categories in each of the performance measures. However, due to the small sample size, these results were analyzed with non-parametric statistics. Table 2 shows the results of the statistical analyses for each phone type and user type.

No significant performance differences are found between the two phone groups on any of the variables (i.e. none of the Z statistics are significant). However, we encounter some operational differences between the two types of phones from the analysis of the verbal protocols. Both devices differ substantially in where the application is placed after the download, and Siemens' users took a little longer to find the application in their phones and open it. This slowed down some users and may account for the slightly higher task times in this group (12.43 in Siemens vs. 11.17 in Nokia). However, this time difference is not significant ($Z = 1.17$, n.s.).

In contrast, the results by type of user do show significant differences for all the variables. The average number of errors experienced by advanced users is significantly lower than those experienced by novice users (3.73 vs. 9.08; $p < .0001$). In addition, the average number of correct answers is significantly higher for advanced users than for novice users (4 vs. 1.83; $p < .0001$). Consequently, the average time required to complete the task is significantly lower for advanced users than for novice users (8.11 vs. 15.29; $p < .0001$). It is noteworthy that the interaction of advanced users with the application is not entirely error-free.

Table 1
Performance data by condition

Average (st. dev)	Novice/Nokia ($n = 6$)	Novice/Siemens ($n = 6$)	Advanced/Nokia ($n = 6$)	Advanced/Siemens ($n = 5$)	Full sample
Number of errors	9 (2.37)	9.16 (2.56)	3.5 (1.05)	4 (0.7)	6.5 (3.26)
Correct answers	1.83 (0.75)	2 (0.75)	4 (0.89)	4 (1.0)	2.9 (1.36)
Task time	14.29 (0.86)	15.75 (1.18)	8.04 (0.62)	8.2 (0.65)	11.77 (3.72)

Each cell shows means and standard deviations in parentheses.

Table 2
Non-parametric analysis of performance data

	Type of phone			Type of user		
	Nokia	Siemens	Z	Novice	Advanced	Z
Number of errors	6.25 (3.36)	6.82 (3.28)	0.43	9.08 (2.35)	3.73 (0.90)	4.06^{***}
Correct answers	2.92 (1.38)	2.82 (1.40)	0.16	1.83 (0.72)	4.0 (0.89)	3.87^{***}
Task time (min)	11.17 (3.34)	12.43 (4.15)	1.17	15.13 (1.31)	8.11 (0.61)	4.04^{***}

Significance level: $^{***}p < .0001$.

4.2. Qualitative results

To understand the nature of these quantitative results, the errors harvested from each user's verbal protocols were further analyzed. This analysis of error segments was done in iterative steps using the open-coding technique by Strauss and Corbin (1998). Instead of an a priori coding scheme for errors, we inferred the nature of the problems from the verbalization data. Through successive iterations, we classified segments into major categories. In the first pass through the transcripts, a segment was identified as an error. In the second pass, a segment was assigned a descriptive label (i.e. wrong key, failed to scroll, technical problem while downloading). And in the third pass, major error categories were identified by grouping similar labels.

Error segments were placed in three categories (Download, Navigation and Comprehension) according to the activities required by the task. Download errors occurred when users experienced difficulties in getting the application into their phones. Navigation issues occurred when user either did not move as expected within the application, or did not access available information. Comprehension errors occurred when users did not understand how the application's interface was designed, and therefore either committed specific errors or failed to answer the questions of the task correctly. Only 22 cases of download problems were found. The most frequent category of problems was navigation with 68 instances, closely followed by comprehension with 60. Table 3 shows the list of problems in each category.

Although it is the most critical error, download errors were least frequent. Problems in this area were mostly technical in nature and experienced primarily by novice users (who had not used a wireless data application before). Aside from the instances in which the handset was not configured to access the Internet, other errors in this category refer to a mishandling of the return-push message from the service provider. Some users either

Table 3
List of problems by category

Key	Problem description	No. of cases	Novices	Advanced
	Download problems	22	18	4
D1	Internet connection not configured	5	2	3
D2	Missed WAP push message with downloading instructions	6	5	1
D3	Did not download the application	4	4	0
D4	Could not find the application on the phone after downloading it	7	7	0
	Navigation problems	68	50	18
N1	Unsure how to proceed after first page	11	11	0
N2	Did not drill down	17	12	5
N3	Did not scroll down	4	4	0
N4	Returned to the index after each question	16	10	6
N5	Accidentally exited the application	7	4	3
N6	Pressed the wrong key	13	9	4
	Comprehension problems	60	46	14
C1	Did not know how to view pictures	15	12	3
C2	Can not identify sections (editorial, international, etc.)	11	11	0
C3	Can not identify pages (page locator bar not understood)	11	11	0
C4	Did not find all of the ads	23	12	11
	Total	150	114	36

ignored this message or did not know how to locate the return message (when it was not placed in the inbox). Only one of these problems was specifically related to the type of phone being used. Most novice Siemens' users had trouble locating where the application was stored after downloading it. In general, most of the difficulties in this category were attributable to the inexperience of the users in downloading applications into their mobile phones. We illustrate some of these problems with excerpts of their verbalizations translated into English, along with the key from Table 3 to identify the problem.

> *"I'm receiving an error message that my internet gateway is not configured. What does it mean?"* [D1]
> *"I don't find it... I'm looking for the link in my inbox and didn't find it."* [D2]
> *"I don't know how to handle applications in this phone... I've never used it for this before. It should work like a computer... but where should I look...for Microediciones?"* [D3]
> *"Where is the Microediciones? Uhm... I think I...I just downloaded it but I can not find it."* [D4]

Navigation issues ranged from how to move between pages of content to how to access specific information within each page. Since the initial instructions about how to navigate the microedition appeared on the third screen, some users had trouble figuring out how to proceed after the front page.

> *"OK...here is the first page...it looks like a real newspaper, but the font is kind of small. Now I want to continue... I don't know how... I see...the button [next] will take me to the next page."* [N1]
> *"I am waiting for the next page of Microediciones to appear automatically...* [pause] *Oh, I guess that I should click on something because it doesn't appear by itself."* [N1]

After this study, the navigation instructions were placed on the second screen.

In addition, some novice users did not understand that the application features two levels of content – headlines and summary – and thus sometimes failed to drill down the headlines to access the summary of the newspaper article. As a result, most novice users could not answer questions that required them to read a summary of an article.

> *"I am trying to answer the question about the editorial...* [pause] *The title is right here but I can not tell what is it about without reading it first!"* [N2]
> *"Am I supposed to guess what this editorial is about?...* [pause] *I can only see the title."* [N2]

In addition to failing to drill down, a few users did not scroll down. This is particularly evident among those who were able to access the summary of the news articles but did not provide complete or valid answers to the content-related questions (what is the editorial about? And how did Lupe Valdez become sheriff?). In order to provide a complete answer, they had to scroll down to read the entire summary of the article.

> *"I don't know... I can not answer this question [Lupe Valdez question]."* [N3]

Another important navigation finding was that both novice and advanced users moved within the microedition in a non-linear fashion by returning often to the news index, instead of using the "next" key to access the following page of the microedition.

> *"I'm going to the index to try and find the answer to the next question."* [N4]

In a few instances, users accidentally exited the application or pressed the wrong key (i.e. clicked on a keypad button that was not the intended one). For example, sometimes users tried to return to the index after viewing an article summary and found that the index button did not work. We catalogued this problem as "pressing the wrong key", because they "0" key was not available when they tried it.

> *"I am fine navigating... but suddenly I'm kicked out.. I think that I've pressed exit."* [N5]
>
> *"I'm clicking on 0 to go back to the index, but it doesn't work... I'll click again. Oh! It seems that from here I can only go back. [After going back]... Now, I can go to the index. It says 0 for index. OK it is working now."* [N6]

The last category of problems – comprehension – resulted when subjects did not understand the symbols, icons or labels they encountered in the application. Some users, for example, failed to recognize that the camera icon indicated the availability of an image to accompany the news item, and therefore could not answer the image-related question of the task.

> *"It says here [in the task question] that there is a picture somewhere... I don't see any picture."* [C1]

Other users did recognize the camera icon as a symbol for an image but did not know how to view the picture. Viewing the picture required the user to access the content of the article first. The icon was available at the level of the headline; however the picture was only visible after accessing the summary of the news item. We classified this problem as a comprehension issue because failure to access a picture indicated that the user did not understand how the picture icon was available within the content of the article.

Other comprehension problems were related to the section identifiers and to the locator bar (which indicates the relative placement of the viewed page in the microedition). Sometimes, users failed to associate the titles in the upper right corner with section identifiers.

> *"The headlines don't show any sections. I don't understand this question! [picture shown in the international section]."* [C2]

In other instances, the page indicator in the locator bar was not clear for some users.

> *"There are no page numbers here... the article is in the next to last section. So, the page is the next to last page, but I don't know the number."* [C3]

The comment above reflects how this particular user conceived the microedition as comprising sections (accessible from the index) instead of sequentially ordered pages. This notion is consistent with the non-linear navigation strategy adopted by the majority of the users.

Because they chose to utilize non-linear and index-driven navigation, most users explored each page by returning to the news index instead of hitting the "next" key. As a result, they missed the ads, which were inserted between pages and could be viewed only if the user moved sequentially from one page to the next. Although skipping the ads may not have been problematic for the users, we classified this as an error because it prevented them from answering correctly the ad-related question of the task. Also, given the mobile company's objective of generating revenues through advertising, the fact that users knowingly (or unknowingly) can opt not to view the ads is noteworthy.

> *"Which ads? I did not see any...is there a section for ads. All I could see were headlines."* [C4]

In order to further analyze the data, each of the 150 error instances was independently classified as knowledge-based, skill-based and rule-based errors. To assure the reliability of this classification, two independent coders classified these errors (Inter-coder reliability = 86%). Disagreements between the coders were reconciled by discussion. As a result of the consensus agreement, download problems were all classified as knowledge-based because they required specific understanding of how to download a data application on the phone. Most of the comprehension problems were also classified as knowledge-based, except for the one indicating that the ads were missed. Missing the ads is closely related to index-based navigation identified in the navigation category. Both of these problems (index-based navigation and missing ads as a result) were classified as rule-based errors because non-linear navigation is assumed to be the rule, rather than the linear navigation that the designers of the mobile application intended.

The remaining issues in the navigation category were either classified in the knowledge-based category (no drill down, no scroll down, not sure how to proceed after first page and pressed the wrong key) or in the skill-based category (accidentally exited the application). There was, however, one category of errors that had to be split into knowledge-based and skill-based, according to the evidence gathered from the verbal protocols. Sometimes users pressed the wrong key because they did not know which key to use (knowledge-based error) or because they inadvertently hit a key that was not consistent with their stated intention (skill-based error). Table 4 shows a breakdown of errors by type of user along with a test of differences between proportions. For this test, the total number of errors in each category was divided into two groups based on whether novices or advanced users

Table 4
Types of error by user

Type of error	Novice (%)	Advanced (%)	Binomial test of differences between proportions (Z)
Knowledge-based errors	85	15	7.06***
Skill-based errors	56	44	0.80
Rule-based error	60	40	0.63

*** Significant at $p < .0001$.

committed them. The only significant difference is found in knowledge-based errors. Novice users commit more errors of this type than advanced users.

5. Discussion of results

One of the most notable findings of this study is the discrepancy between the non-linear browsing behavior demonstrated by the users and the linear navigation expected by the designers. It may be argued that non-linear user behavior resulted from the requirements of the question-and-answer format of the task, which created a unique setting for the use of the mobile application. However, exit interviews revealed that this behavior was not necessarily a consequence of task requirements (e.g. searching for answers to specific questions). Participants indicated that when reading news online, they often search for news items of interest in a non-linear way. This finding indicates that the "underlying conceptual model" (Norman, 1999) of the application – emulating a printed sequential newspaper – was not understood by the users, as they tend to associate any form of online news delivery with web browsing rather than reading a printed newspaper.

Although user performance measures (accuracy and duration) were similar within each type of phone, there was significant variation between new and more advanced users. The former needed more time to complete the task, could not find the answers to some questions and experienced more problems. Due to the nature of the "think-aloud" method, however, task duration can not be considered representative of actual user efficiency because verbalization takes time. In this case, time measures provided a benchmark to determine the magnitude of the problems encountered by the users. In general, users who experienced more problems or were more uncertain about how to proceed took more time to complete the task than their counterparts. The loss of precision was compensated for by learning exactly what kinds of problems users faced.

When segmenting the errors by type according to Gray's (2000) classification, we found that most of the knowledge-based errors were made by novice users. Consistent with the theory of automaticity, encountering this kind of mobile application for the first time presented a completely new environment for users who had no prior experience with similar content-based applications. Although instructions were provided, in most instances these instructions were unclear, misplaced or overlooked by the users. Overall, these novice users needed more precise and better-placed instructions, as well as reminders of navigational aids, in order to compensate for their lack of knowledge about mobile data applications.

There were also cases of slips or skill-based errors for both user groups. In several instances, users accidentally exited the application or pressed a wrong/unintended key. In hindsight, it is possible that the unintended exit slips occurred as a result of repeated exposure to the corresponding navigation button (9 for exit). The presence of this reminder on each page of the microedition could have given the impression that the exit button was supposed to be a frequently used one. Advanced users also experienced some slips due to pressing wrong keys. It is possible that as they were performing the task automatically, they tried to transfer their practice with other mobile applications to the current environment – and found that the required keys are not consistent.

Rule-based errors were similar for both novice and advanced users. In general, users consistently returned to the news index and navigated the application non-sequentially. This approach indicates that non-linear navigation is assumed to be the rule, instead of

the linear navigation that designers of the mobile application intended. As a consequence, users could not answer correctly the question about the number of ads. As noted previously, developers had placed the ads in between pages of content, so they would be viewed as viewers moved sequentially from one page to the next – a strategy that reflected the practices of most newspaper readers. However, feedback gathered in the focus group indicated the need for navigational flexibility, as users cognitively relate microeditions to a web-based news delivery service, regardless of its mobile nature. Despite this early hint, developers did not abandon the printed newspaper conceptual model as the basis for incorporating ads into the mobile application.

In sum, novice users exhibit a lack of automaticity because they have no prior experience with similar mobile applications. For novice users, low level (action-oriented) instructions are crucial, while high level (goal-oriented) instructions are not as useful. In contrast, advanced users tend to exhibit automaticity as long as they can transfer their experiences with other mobile environments to the current one. Thus, for advanced users, high level (goal-oriented) instructions are recommended. The use of consistent design guidelines, and similar keystrokes for applications currently in use as well as new ones, is also advisable. Further, our results suggest that regardless of the level of familiarity with mobile applications, users are likely to experience rule-based errors when there is a conceptual mismatch with the designers' model.

A summary of findings, along with the corresponding recommendations, is presented in Table 5.

5.1. Implications for theory and research

Our findings contribute to the theory of automaticity (Logan, 1988) by showing the extent to which users can effectively transfer knowledge from other mobile applications to completely new ones. By extending the theory of automaticity with Gray's error taxonomy, this study examined not only the types of difficulties that first-time users encounter with mobile applications, but also the potential causes of such problems. When analyzed in the context of automaticity, the three-way classification of errors (knowledge-based, skill-based and rule-based) offers a more complete understanding of the interaction between first-time users and new mobile applications.

According to our findings, first-time users of Microediciones® who were familiar with other mobile applications were able to perform *as if* they were experienced with it, which validates the theory of automaticity. However, although the previous experience of these advanced users was instrumental in reducing the number of difficulties, their interactions

Table 5
Error types and recommendations

Type of error	Examples	Recommendations
Knowledge-based errors	Download problems Failed to drill down, failed to scroll down	Improve instructions Provide navigational aids
Skill-based errors	Unintended exits, pressing the wrong key	Remove exit button reminder
Rule-based error	Index-based navigation assumed to be the rule instead of sequential navigation	Rethink ad placement strategy given preference for non-linear navigation

were not entirely free from errors. These advanced users experienced two main types of errors: rule-based and skill-based. While rule-based errors were due to the mismatch between sequential and index-based navigation, skill-based errors were likely the result of inconsistencies between the application at hand and the mobile applications from which they were transferring their experiences. As these advanced users interact with Microediciones® and perform almost automatically, they may be trying to transfer the skills acquired in other applications to the current environment. In some instances, the keystrokes needed in other applications do not match the keys required in the current mobile application. As the development of content-based applications progresses, designers should develop standards of consistency to promote more efficient transferability of experiences between applications.

Another implication of the theory of automaticity is the need to develop instructions consistent with the level of abstraction at which the user operates (i.e. high level for advanced users and low level for novice users). Because novice users operate at a low level of abstraction, they need more detailed step-by-step instructions and navigational reminders – particularly the first time they use the application. In contrast, advanced users benefit from high-level instructions. In order to avoid frustrations after repeated interactions, however, users should be given the option to skip these directions altogether, or to move from low level to high level instruction after an initial period of usage.

For researchers, this project provides evidence of how industry and academia can establish a productive partnership to inform practice and conduct cutting-edge research in innovative environments. As suggested by Lyytinen and Yoo (2003), the fast pace at which new tools and solutions are developed blurs the distinction between technical (or developmental) research and behavioral research. Therefore, "new alliances need to be forged and IS researchers should be actively involved in studies where technologies are being built and tried out – not after the fact when they enter the market." (Lyytinen and Yoo, 2003, p. 387).

As IS researchers, we are uniquely positioned to undertake these studies. By combining an understanding of new technological artifacts with the right tools to gather and analyze user behavior, IS researchers are the natural liaison between developers and end users. Furthermore, researchers can also identify particular misconceptions or mismatches between designers' intentions and users' actions that contribute to the success or failure of new mobile applications.

5.2. Implications for developers and for mobile companies

The initial focus group was useful in understanding typical user behavior and refining the application accordingly. Although this forum did not include any direct interaction between the users and the application, it was informative enough to reveal the need for detailed instructions and navigational aids. One developer was surprised about the need to provide navigational aids and instructions. He did not think it was necessary because the application was "straightforward and simple".

The results of the protocol analysis trials also generated important insights for developers. While user performance measures were similar for both types of phones, the analysis of the verbal protocols revealed different kinds of problems for both groups – suggesting that the technical characteristics of mobile device cannot be completely ignored. Whenever

possible, mobile service providers need to identify the type of phone from which the user is sending the initial microedition request message, and need to provide instructions tailored to that type of device.

While developers designed the application to emulate a printed newspaper delivered electronically, user behavior revealed in the trials suggests that users cognitively associate a microedition with browsing a web page instead of reading a printed newspaper. This insight has important strategic and operational implications. At the strategic level, the current advertising strategy to insert full-page ads in between two pages of content means that these ads will be viewed only by users who navigate linearly from page to page. If users prefer to navigate in a non-sequential way through the microedition, then developers must devise a more effective ad placement strategy. Operationally, the study suggests that developers must provide efficient navigation schemes which prevent users from skipping the ads, while enabling them to feel comfortable moving within the application. As previous research into commercial mobile applications has shown, understanding the user is critical to provide a better user experience and driving revenue sources (Anckar and D'Incau, 2002; Venkatesh et al., 2003).

From a business perspective, this study highlights the importance of considering the sources of profitability as well as user preferences in order to implement successful mobile applications (Lu et al., 2005). A dual strategic and operational focus will result in the development of wireless data services that can be readily adopted by the users and bring additional sources of revenue to mobile content providers. In this particular case, the mobile company should further analyze the strategic implications of a mixed revenue model combining advertising and subscriptions, vis-à-vis a single model based on either advertising or user subscriptions. It is also advisable to study customer willingness to pay for mobile services in Spain, as it has been shown that this may differ significantly across countries (Kim et al., 2004; Rangone and Renga, 2002).

5.3. Limitations

Before drawing general conclusions, we must acknowledge the limitations of this study. First, the subjects are all drawn from the same country and the same culture. Members of other cultures may exhibit different kinds of behavior when interacting with mobile applications and technological innovations (Sarker and Wells, 2003; Venkatesh et al., 2003). In addition, due to the demographics of our sample, we were able to study only a specific age range of users (20–34 years). Younger or older users may have different experiences with mobile devices (Palen and Salzman, 2002) and by extension with wireless services (Sarker and Wells, 2003).

The second limitation is the relatively small size of the sample. Other studies have documented the trade-off between large samples and the need to conduct intensive individual trials based on qualitative analysis of verbal protocols (Vessey and Conger, 1993). Small samples, which are typical of qualitative prototype evaluation and think-aloud studies, provide researchers with rich qualitative data in which more powerful statistical inferences may not be possible (Benbunan-Fich, 2001).

Third, this project is focused on a particular task, which created a unique setting for the use of the mobile application. We assigned a scavenger-hunt task to provide comparable situation for all the subjects and to study their usage behavior with search objectives known to the researchers. Some of the problems reported here result directly from specific

questions posed by the task. Also, some of the findings are due to specific design decisions made by the developers of Microeditions®. Despite these limitations, this study contributes to advancing our understanding of mobile user behavior with new content-based wireless applications.

6. Conclusions

We conducted an exploratory study of user behavior with a new content-based mobile application. One contribution of this study to the field is in its examination of the theory of automaticity in conjunction with a comprehensive taxonomy of errors encompassing mental and physical aspects of user's actions. With this approach, researchers can obtain a more comprehensive view of how users interact with novel IT-artifacts in order to accomplish a particular task. With respect to the particular mobile application evaluated in this study – Microediciones® – perhaps the most important insight is that users cognitively associate it with web-based news delivery regardless of its mobile nature. This finding has important operational and strategic implications for the mobile company launching this application.

The emergence of new content-based mobile applications poses a number of challenges for their adoption and commercial success. Due to their technical and behavioral novelty, this new generation of applications delivered through mobile devices provides a new frontier for IS researchers. Those who wish to explore this new area will probably benefit from establishing partnerships with industry while these mobile applications are being developed. This project is the result of one of such partnerships. We encourage other researchers to look for this type of mutually beneficial alliances.

Acknowledgements

The authors are very grateful to Shoshana Altschuller, Diane Covkin, Radhika Jain, Marios Koufaris, Karl Lang, Kannan Mohan and Gregory Truman for their useful comments and suggestions to improve this paper. We also thank the reviewers of the Mobile Technology Conference and the editors and reviewers of this journal for their invaluable feedback.

References

Anckar, B., D'Incau, D., 2002. Value creation in mobile commerce: findings from a consumer survey. Journal of Information Technology Theory and Applications 4 (1), 43–65.

Benbunan-Fich, R., 2001. Using protocol analysis to evaluate the usability of a commercial web site. Information and Management 39, 151–163.

Buchanan, G., Farrant, S., Jones, M., Thimbleby, H., Marsden, G., Pazzani, M., 2001. Improving mobile Internet usability. In: Proceedings of the 10th International Conference on World Wide Web (WWW 01), May 1–5, Hong Kong, China, pp. 673–680.

Ericsson, K.A., Simon, H.A., 1996. Protocol Analysis: Using Verbal Reports as Data, second ed. Oxford University Press, Cambridge, MA.

Glass, R.L., Vessey, I., Conger, S., 1992. Software tasks: intellectual or clerical?. Information and Management 23 183–191.

Gray, W.D., 2000. The nature and processing of errors in interactive behavior. Cognitive Science 2, 205.

Gray, W.D., Altman, E.M., 2001. Cognitive modeling and human computer interaction. In: Karwoswki, W. (Ed.), International Encyclopedia of Human Factors, vol. 1. Taylor & Francis Ltd., New York, NY, pp. 387–391.

Jarvenpaa, S., Lang, K.R., Takeda, Y., Tuunainen, V.K., 2003. Mobile commerce at crossroads. Communications of the ACM 46 (12), 41–45.

Kai, L.H., Benbasat, I., Todd, P.A., 1996. An experimental investigation of the interactive effects of interface style, instructions and task familiarity on user performance. ACM Transactions on Computer–Human Interaction 3 (1), 1–37.

Kim, J., Lee, I., Lee, Y., Choi, B., Hong, S.-J., Tam, K.Y., Naruse, K., Maeda, Y., 2004. Exploring e-business implications of the mobile internet: a cross-national survey in Hong Kong, Japan and Korea. International Journal of Mobile Communications 2 (1), 1–21.

Logan, G.D., 1988. Toward an instance theory of automatization. Psychological Review 95 (4), 492–527.

LogicaCMG Press Release, 2005. Mobile content market set to triple to more than 7.6 billion euros within a year, July 6, <http://www.logicacmg.com/pressroom>.

Lu, J., Yao, J.E., Yu, C-S., 2005. Personal innovativeness, social influences and adoption of wireless Internet services via mobile technology. Journal of Strategic Information Systems 14 (3), 245–268.

Lyytinen, K., Yoo, Y., 2003. Research commentary: the next wave of nomadic computing. Information Systems Research 13 (4), 377–388.

Mao, J-Y., Benbasat, I., 2000. The use of explanations in knowledge-based systems: cognitive perspectives and process-tracing analysis. Journal of Management Information Systems 17 (2), 153–180, Fall.

Mobile Dreams Factory. 2004. A newsstand in your mobile. Press Release, November 10, 2004, <http://www3.mobilemonday.net/mm/pressrelease.php?id=221>.

Norman, D.A., 1981. Categorization of action slips. Psychological Review 88, 1–15.

Norman, D.A., 1988. The Psychology of Everyday Things. Basic Books, New York.

Norman, D.A., 1999. Affordances, conventions and design. Interactions 6 (3), 38–43.

Noticiasdot.com, 2004. Nokia, Siemens y Sony/Ericsson lideres de ventas en Espaňa, <http://www.noticiasdot.com/publicaciones/2004/1004/2910/noticias291004/noticias291004-11.htm>.

Palen, L., Salzman, M., 2002. Beyond the handset: designing for wireless communications usability. ACM Transactions on Computer–Human Interaction (TOCHI) 9 (2), 125–151.

Rahman, T., Muter, P., 1999. Designing an interface to optimize reading with small display windows. Human Factors 1 (1), 106–117.

Rangone, A., Renga, F., 2002. Mobile Internet: an empirical study of B2C WAP Applications in Italy. Electronic Markets 12 (1), 27–37.

Reason, J., 1990. Human Error. Cambridge University Press, New York, NY.

Sarker, S., Wells, J.D., 2003. Understanding mobile handheld device use and adoption. Communications of the ACM 46 (12), 35–40.

Schenk, K.D., Vitalari, N.P., Davis, S.K., 1998. Differences between novice and expert systems analysts: what do we know and what do we do? Journal of Management Information Systems 15 (1), 9–50.

Siau, K., Lim, E.P., Shen, Z., 2001. Mobile commerce: promises, challenges, and research agenda. Journal of Database Management 12 (3), 4–14.

Strauss, A.L., Corbin, J.M., 1998. Basics of Qualitative Research, second ed. Sage Publications, Thousand Oaks, CA.

Venkatesh, V., Ramesh, V., Massey, A., 2003. Understanding usability in mobile commerce. Communications of the ACM 46 (12), 53–56.

Vessey, I., Conger, S., 1993. Learning to specify information requirements: the relationship between application and methodology. Journal of Management Information Systems 10 (2), 177–201.

Zhang, J., Norman, D.A., 1994. Representations in distributed cognitive tasks. Cognitive Science 18, 87–122.