MIS Quarterly

RESEARCH ARTICLE

MOBILE APPLICATION USABILITY: CONCEPTUALIZATION AND INSTRUMENT DEVELOPMENT[1]

Hartmut Hoehle and Viswanath Venkatesh

Department of Information Systems, Sam M. Walton College of Business, University of Arkansas,
Fayetteville, AR 72701 U.S.A. {hartmut@hartmuthoehle.com} {vvenkatesh@vvenkatesh.us}

This paper presents a mobile application usability conceptualization and survey instrument following the 10-step procedure recommended by MacKenzie et al. (2011). Specifically, we adapted Apple's user experience guidelines to develop our conceptualization of mobile application usability that we then developed into 19 first-order constructs that formed 6 second-order constructs. To achieve our objective, we collected four datasets: content validity (n = 318), pretest (n = 440), validation (n = 408), and cross-validation (n = 412). The nomological validity of this instrument was established by examining its impact on two outcomes: continued intention to use and mobile application loyalty. We found that the constructs that represented our mobile application usability conceptualization were good predictors of both outcomes and compared favorably to an existing instrument based on Microsoft's usability guidelines. In addition to being an exemplar of the recent procedure of MacKenzie et al. to validate an instrument, this work provides a rich conceptualization of an instrument for mobile application usability that can serve as a springboard for future work to understand the impacts of mobile application usability and can be used as a guide to design effective mobile applications.

Keywords: Usability, mobile applications, survey instrument development, continued use, mobile application loyalty, mobility

Introduction

In most developed countries, mobile phone penetration rates have reached over 100 percent per capita, with individuals often owning more than one mobile phone (BBC News 2010). Google announced that 500,000 new Android-based devices are registered every day (Reisinger 2011). In conjunction with this trend, over the last 5 years or so, mobile phone technology has changed significantly, with devices and operating systems becoming more sophisticated. These developments have led to a large variety of mobile applications designed for smartphone operating systems provided by mobile operating system vendors, such as Apple, Google, and Microsoft. For instance, Apple's iStore includes over 1 million mobile applications providing consumers with access to a variety of services via iPhones. As of May 2013, the available mobile applications had been downloaded more than 50 billion times from Apple's iTunes store (iTunes 2013). These statistics accentuate that it is vital for organizations to integrate mobile applications into their channel strategies.

Despite these trends and a recognition that they miss out on business opportunities, two-thirds of all firms selling consumer goods do not have established mobile strategies (Forrester Research 2011). Integrating mobile channels into existing strategies is a major challenge for firms. First, the development of mobile applications is expensive and integrating mobile channels into existing processes requires considerable investments by firms (Deloitte 2012; Forrester Research 2011). Recent studies show that mobile applications can cost millions of dollars to develop (Deloitte 2012). Second, despite the success of some, the majority of mobile

[1]Rajiv Kohli was the accepting senior editor for this paper. John Wells served as the associate editor.

The appendices for this paper are located in the "Online Supplements" section of the *MIS Quarterly*'s website (http://www.misq.org).

applications fail outright or are not as successful as expected (Deloitte 2012). For instance, only 1 percent of all mobile applications have been downloaded more than one million times and, once downloaded, one in four mobile applications are never used again. Likewise, 80 percent of all branded mobile applications are downloaded less than 1,000 times and the most popular 10 percent of mobile applications generate 80 percent of all downloads in Apple's iTunes store (Deloitte 2012; Dredge 2011). The latest market research suggests that the lack of usability has been identified as the most important factor influencing consumers' decisions to reject mobile applications (Deloitte 2012; Forrester Research 2011; Youens 2011).

The lack of mobile application usability can be due to various reasons. For example, many mobile applications do not prioritize the most essential aspects of the application and content is ineffectively presented which in turn negatively influences user interactions (Forrester Research 2011; Youens 2011). Further, instead of developing applications that account for the unique characteristics of mobile devices (e.g., small screens, tiny input mechanisms, various screen resolutions), firms often identically replicate the content of their Internet-based applications on mobile channels (Forrester Research 2011). Presenting a large body of content on mobile devices is problematic because the application interface is overloaded with information, site links, and text (Adipat et al. 2011). As a result, users feel overwhelmed and become frustrated because the application does not emphasize one of the key usability principles for mobile applications, such as not using large buttons to facilitate the data input because interface space is sacrificed for detailed information (Deloitte 2012). To guide practitioners in developing mobile applications, leading operating system vendors (e.g., Apple, Google, Microsoft) offer guidelines that aim to help mobile application developers build better applications. However, neither do these guidelines provide points of emphasis nor do they help in the evaluation of mobile applications (Nielsen 2012a; Nielsen Norman Group 2012). For example, Apple's *mobile user experience guidelines* recommend: "Beautiful artwork also helps to build your app's brand in people's eyes" (Apple 2012). Although this suggests artwork is important, it does not establish how important it is, when it is important (i.e., types of applications), and whether a particular application has done this effectively (i.e., evaluation). Rigorously developed and comprehensive survey instruments can help practitioners in achieving better mobile application design and evaluation. This is due to the fact that such instruments support application designers in quantifying the impact of a given usability principle on the overall mobile application usability.

A review of the literature on IS and human–computer interaction (HCI) revealed no context-specific instruments for holistically measuring mobile application usability. First, we found that much research conceptualized and measured mobile application usability without proactively taking the mobile context into consideration. Instead of integrating critical contextual factors for mobile applications, such as tiny input mechanisms (Kurniawan 2008), we found that field studies typically used instruments that evolved from website usability, much like website usability before it evolved from software usability (see Venkatesh and Ramesh 2006). Second, related to the void of context-specific conceptualizations of mobile application usability, our literature review suggested that prior research treated mobile application usability as performance using metrics like response time and error rates related to tasks users performed (e.g., Burigat et al. 2008; Hummel et al. 2008; Li, Van Heck, and Vervest 2009). Although these measurement techniques account for the mobile application usability context and are useful for evaluating the interplay among individuals (e.g., users), technology (e.g., mobile application), and tasks (e.g., typing text), these features are typically studied in a piecemeal fashion and are therefore less suited for providing a comprehensive understanding of the antecedents of mobile application usability. Therefore, while a useful starting point, these studies do not provide a comprehensive view of mobile application usability, thereby potentially neglecting critical factors pertaining to various aspects of the mobile application usability context. Third, we found that prior research has used a variety of conceptually dissimilar constructs to assess mobile application usability including satisfaction with system use, ease of learning, and effectiveness of mobile applications (e.g., Gebauer et al. 2007; Lee et al. 2007; Min et al. 2009; Urbaczewski and Koivisto 2007). Associating these concepts with mobile application usability seems problematic as this practice could result in interpretational confounding (Burt 1976) that occurs when the empirical meaning of a latent variable differs from the meaning assigned by a researcher (Bollen 2007). We concur that mobile application usability leads to higher consumer satisfaction but both concepts are conceptually different and should be treated as such. The same principle applies to the concept of ease of learning and mobile application effectiveness.

Against the backdrop of these issues, we believe it is important to think from the ground up about mobile application usability as it will "contribute to our understanding of current technological and organizational problems or challenges faced by IS or other practitioners" (Straub and Ang 2011, p. 3). A context-specific and comprehensive survey instrument for measuring mobile application usability would be appropriate to start addressing the above-mentioned practical issues. Such an instrument would also advance our theoretical understanding of mobile application usability by being more precise and limiting interpretational confounding. We argue it is critical that IS and HCI research can draw on a comprehensive mobile application usability conceptualization and associated instrument that integrates context-specific factors for two major reasons. First, recent advances on theory development and knowledge creation suggest that theories and models grounded in specific contexts reveal rich insights (Alvesson and Kärreman 2007; Bamberger 2008; Brown et al. 2010; Hong et al. 2014; Johns 2006; Van der Heijden 2004; Venkatesh et al. 2012). Second, a holistic instrument integrating IT-specific antecedents addresses recent calls for more specific theoretical models that step "backward toward IT, implementation, and design factors, leading to research that is able to provide actionable advice" (Benbasat and Barki 2007, p. 213). To address these gaps, this paper develops a context-specific and fine-grained conceptualization and measurement of mobile application usability. In order to achieve this goal, we adapt Apple's user experience guidelines for mobile applications (Apple 2012), relate the derived usability constructs with the extant literature on mobile application usability, and develop and validate a survey instrument following the 10-step procedure of MacKenzie et al. (2011).

This work is expected to make key scientific and practical contributions. Having a context-specific conceptualization and measurement of mobile application usability will aid theory development in IS and HCI research alike. Specifically, our comprehensive view on mobile application usability centers around the IT artifact, and we provide conceptual clarity and develop and validate an associated survey instrument. Our mobile application usability instrument can be used as a springboard for theoretically motivated studies that allow researchers to develop a cumulative research tradition in this emerging research area. Practitioners will benefit from a comprehensive measurement instrument as our IT-centric view of mobile application usability will help them to determine the most desirable features of mobile applications by surveying current or potential customers during the design, planning, and development phases of mobile applications. During the maintenance and review phases of mobile applications, our instrument can be used to identify desirable mobile application functions and design features. This should help practitioners to design more successful mobile applications and integrate mobile applications into their channel strategies more effectively. This is particularly important as the latest market research suggests that practitioners are in need of theoretical frameworks dedicated to analyzing and overcoming usability problems of mobile applications (Nielsen 2012a). Forrester Research (2011) found that 70 percent of firms are currently reconsidering their mobile strategies and they are planning to integrate mobile channels into their distribution channels. A comprehensive framework for assessing the usability of mobile applications would help practitioners in designing successful mobile applications (Nielsen 2012a; Nielsen Norman Group 2012).

Mobile Application Usability

Mobile application usability is defined, drawing from the International Standards Organization's (ISO) definition of usability, as the extent to which a mobile application can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use (see Venkatesh and Ramesh 2006). A mobile application is an IT software artifact that is specifically developed for mobile operating systems installed on handheld devices, such as smartphones or tablet computers. Mobile applications are either preinstalled on mobile devices or can be downloaded from various mobile application stores (e.g., Apple's iTunes store). It is important to note that mobile application usability differs from mobile device usability in that the latter term commonly refers to the extent to which the operating system (e.g., iOS) is user friendly (Adipat et al. 2011). Mobile application usability has been studied in a range of contexts and researchers have used a variety of conceptualizations and measurement approaches to study the topic. Although there is a considerable amount of literature available on mobile application usability (Hess et al. 2009; Hong et al. 2004b; Palmer 2003; Thong et al. 2002; Wells, Parboteeah, and Valacich 2011; Wells, Valacich, and Hess 2011), our literature review suggests that there is a lack of theoretically sound instruments for measuring mobile application usability holistically. Specifically, based on a review of literature on mobile application usability, we identify three major issues that we discuss next.

First, we found that much research has conceptualized and measured mobile application usability without integrating critical context-specific factors that are relevant to individuals using mobile applications (e.g., the mobile application's ability to display content in horizontal and vertical mode driven by the way the user holds the mobile device). Instead, we found that most field studies used conceptualizations and instruments, typically designed for traditional personal computers and websites (see Table 1). For instance, Venkatesh and Ramesh (2006) applied a conceptualization and scales based on Microsoft's usability guidelines (MUG) to a multinational research context and tested the generalizability of the MUG instrument to mobile sites. The MUG conceptualization and scales were originally developed by Agarwal and Venkatesh (2002) based on Microsoft's white paper on usability for assessing the usability of websites (see Keeker 1997). Although Venkatesh and Ramesh's study confirmed that specific aspects of the MUG conceptualization and scales were found to be more important for mobile contexts (e.g., content, ease of use), a major limitation of their study was that the conceptualization and instrument was not customized for the context of mobile application usability. By simply examining the generalizability and differences between website and

Table 1. Prior Conceptualizations and Measurement Approaches for Mobile Application Usability

Usability Attributes	Conceptualization	Usability Evaluation Technique	Study
Context, content, community, customization, communication, connection and commerce	General, deriving mobile commerce framework from Internet-based e-commerce literature	Conceptual (non-empirical)	Lee and Benbasat (2003)
Content, ease of use, made-for-the-medium, promotion and emotion	General, understanding differences between e-commerce and m-commerce sites	Field study (empirical)	Venkatesh, Ramesh, and Massey (2003)
Effectiveness, efficiency and satisfaction	General, two alternative mobile applications were benchmarked and subsequently tested using general questionnaire representing mobile application usability	Experiment (empirical)	Hyvarinen et al. (2005)
Users were exposed to predefined tasks: usability problems were recorded related to the task performance	General, users were asked to perform tasks in two alternative user settings	Think aloud feedback (empirical)	Kaikkonen et al. (2005)
Goals (hedonic/utilitarian), emotion, hands, leg, visual, auditory, colocation interaction	General, users were asked to use mobile Internet and complete a usability diary based on the experiences	Usability diary (empirical)	Kim et al. (2005)
Satisfaction, effectiveness and efficiency	General, managers were interviewed regarding their views on mobile applications in their organizations	Qualitative interviews	Nah et al. (2005)
Content, ease of use, made-for-the-medium, promotion and emotion	General/holistic, understanding m-commerce sites usability	Experiment (empirical)	Massey et al. (2005)
Ease of use, usefulness and compatibility	General, m-commerce users were surveyed regarding m-commerce applications	Field study (empirical)	Wu and Wang (2005)
Design aesthetics, ease of use and usefulness	General, users were asked to evaluate mobile applications	Field study (empirical)	Cyr at al. (2006)
Effectiveness, contextual awareness, task hierarchy, visual attention, hand manipulation and mobility	General, users were asked to complete predefined tasks using a mobile application and were subsequently surveyed regarding their subjective views on mobile application usability	Field study (empirical)	Duh et al. (2006)
Effectiveness, efficiency and satisfaction	Specific, two alternative mobile applications were benchmarked using time, number of attempts, success rate and number of errors as success measures	Experiment (empirical)	Huang et al. (2006)
Effectiveness, efficiency and satisfaction	Specific, mobile application was tested in lab and field setting using completion time, task completeness and a single item satisfaction measure	Experiment/field study (empirical)	Nielsen et al. (2006)
Content, ease of use, made-for-the-medium, promotion and emotion	General/holistic, understanding differences between e-commerce and m-commerce sites	Field study (empirical)	Venkatesh and Ramesh (2006)
Task completion time, task duration and accuracy	Specific, three alternative mobile applications were benchmarked via two tasks using tasks completion time, task duration and task accuracy as success measures	Experiment (empirical)	Burigat et al. (2008)
Satisfaction, ease of use, ease of learn	General, three alternative applications were evaluated and subsequently evaluated using Likert type scales representing the identified usability attributes	Experiment (empirical)	Urbaczewski and Koivisto (2007)
Display, keyboard, design, customer service, ease of use, external and internal sound	General, user reviews were analyzed, coded and used in quantitative analysis to determine the influence of each usability attribute on the overall usability of mobile applications	Conceptual (non-empirical)	Gebauer et al. (2008)
Delay and error rate	Specific, mobile application was tested in lab using delay and error rates as success measures for mobile application usability	Experiment (empirical)	Hummel et al. (2008)
Effectiveness, efficiency, satisfaction, learnability, security	General, mobile application usability attributes were evaluated using Likert scale questionnaire	Field study (empirical)	Min et al. (2009)

Table 1. Prior Conceptualizations and Measurement Approaches for Mobile Application Usability (Continued)

Usability Attributes	Conceptualization	Usability Evaluation Technique	Study
Usefulness, enjoyment, ease of use	Specific, a mobile application was tested in lab session focusing on multidisplay buttons. All usability attributes were evaluated using 3-item Likert scales	Experiment (empirical)	Kim et al. (2010)
Convenience, speed and personalization	General, mobile application usability was evaluated using high-low rankings of each usability attribute	Field study (empirical)	Li, Hess et al. (2009)
Ease of use	General, mobile application usability was associated with ease of use	Field study (empirical)	Hong and Tam (2006)
User distraction, network connection quality and user mobility	General, deriving mobile application usability from prior work	Conceptual (non-empirical)	Gebauer et al. 2010
Download problems, navigation problems and comprehension problems	General, mobile application usability was determined through qualitative think aloud protocol analysis	Field study (empirical)	Benbunan-Fich and Benbunan (2007)
Complexity	General, the system complexity of mobile applications was associated with usability	Field study (empirical)	Mallat (2007)
Search time and accuracy	Specific, several mobile applications were developed and tested in laboratory research setting. Usability attributes were evaluated based on testers' performance	Experiment (empirical)	Adipat et al. (2011)
Speed and accuracy	Specific, focus on text input mechanism as part of mobile application usability	Conceptual (non-empirical)	MacKenzie and Soukoreff (2002)
Text entry rates and error rates	Specific, focus on text input mechanisms as part of mobile application usability	Experiment (empirical)	Lyons et al. (2006)
Task efficiency, task effectiveness and ease of use	Specific, alternative mobile applications were benchmarked measuring the usability attributes as success measures	Experiment (empirical)	Ziefle and Bay (2006)
Color, text and menu icons	Specific, alternative mobile applications were benchmarked measuring menu color, text and menu icons	Experiment (empirical)	Sonderegger and Sauer (2010)
Design, customer needs, satisfaction, innovativeness, feedback and efficiency	General, mobile application usability was evaluated using Likert scale questionnaire for the identified usability attributes	Field study (empirical)	Kim et al. (2012)
Icon characteristics (semantically close meaning, familiar, labeled and concrete)	Specific, alternative mobile applications were benchmarked manipulating icon characteristics	Experiment (empirical)	Leung et al. (2011)
Learnability, efficiency, memorability, error, satisfaction, effectiveness, simplicity, comprehensibility and learning performance	General, deriving mobile application usability attributes from prior studies	Conceptual (non-empirical)	Zhang and Adipat (2005)
Predictability, learnability, structure principle, consistency, memorability, familiarity	General, mobile application usability attributes were measured through single item questionnaire	Field study (empirical)	Ji et al. (2006)
Efficiency and direct usability measures	General, mobile application usability attributes were measured through single item questionnaire	Experiment (empirical)	Jokela et al. (2006)

mobile site usability, it is likely that important requirements of the mobile context were omitted (e.g., fingertip-sized buttons that are used to select features on small multi-touch screens).

Second, we found much of the research that evaluated mobile application usability was conducted in laboratory settings. These studies typically used performance measures (e.g., speed, error rates) to evaluate mobile applications usability. For example, in order to evaluate mobile application usability, Jokela et al. (2006) asked usability raters to benchmark two application interfaces using several predefined tasks (e.g., inserting a new entry in a phone book application). Next, the participants were asked to rate each application interface in terms of its efficiency and overall usability. Adipat et al. (2011) studied the hierarchy of mobile applications and conducted an experiment in which they exposed participants to multiple mobile applications. By manipulating the presentation of mobile sites and the task level, they confirmed that hierarchical text and colorful presentation techniques influence the perceived usability of mobile applications. However, we found that many of these studies evaluated mobile application usability in a piecemeal fashion, rather than attempting to holistically study the concept of mobile application usability and associated context. Consequently, although able to identify error rates and the time users take to perform a given task using a mobile application, these studies suffer from their inability to predict and explain why users took more or less time to perform a given task. A holistic instrument providing an IT-centric view of mobile applications would help in understanding the triggers that influence such performance outcomes.

Third, our literature review indicated that research on mobile application usability does not sufficiently "build upon each others work" (Keen 1980, p. 9) but instead use various conceptualizations and scales for assessing mobile application usability. For instance, Lee and Benbasat (2003) used seven design elements for e-commerce sites, namely context, content, community, customization, communication, connection, and commerce, to assess mobile application usability. Other studies combined concepts commonly seen in the technology acceptance literature (e.g., ease of use; Venkatesh 2000; Venkatesh, Morris et al. 2003) with HCI principles (e.g., design aesthetics), as well as with concepts from the marketing research discipline (e.g., satisfaction) (see Cyr et al. 2006; Cyr et al. 2009; Huang et al. 2006; Urbaczewski and Koivisto 2007). These studies often argued that satisfaction, efficiency, and effectiveness would represent a user's perception of mobile application usability. Associating such concepts with mobile application usability seems problematic as it could lead to interpretational confounding. Table 1 summarizes the literature and shows the various conceptualizations and measurement approaches used in prior studies to examine mobile application usability.

Taken together, this discussion confirms the several issues that we raised earlier. There is a lack of theoretical and methodological clarity surrounding the overall usability of mobile applications. Next, we provide a background on Apple's user experience guidelines, which is a white paper by the leading vendor of mobile applications (Apple 2012).

Apple's User Experience Guidelines

Apple's iTunes store is the most accepted mobile application store and consumers have downloaded more than 50 billion apps as of the middle of 2013. Gartner (2011) predicts that Apple will remain the single best-selling store until 2015 and the company reports that Apple holds a market share of more than 80 percent in terms of revenues made through mobile applications. Apple's first iOS was released in 2007 and, in four years, the platform has become one of the most widely accepted operating systems for smartphones (Apple 2012). Much of the success of the iPhone has been attributed to the user friendliness of the interface of iPhone applications. Apple's user experience guidelines aim to support developers by designing user friendly and successful applications for the iPhone and iPad. The guidelines particularly focus on the user interface design and place less emphasis on technical assistance during the programming and system development process (Apple 2012). In total, the guidelines include 28 sections related to the user experience of mobile applications and most sections include recommendations for developers. For instance, one section recommends using succinct and short text that users can absorb quickly and easily (Apple 2012). We found Apple's guidelines particularly suited to inform the development of a mobile application usability conceptualization. We also considered alternative user experience guidelines published by independent usability consultants and researchers (e.g., Masse 1998; Nielsen 2012a; Nielsen Norman Group 2012) but decided to focus on Apple's user experience guidelines due to the company's prevailing position in the mobile application market. We also felt that Apple's guidelines will help us to provide a relevant contribution to academics and practitioners alike. Rosemann and Vessey (2008) suggest that one systematic approach for developing relevant research is that it is "not necessarily based in theory, [but] involves examining a practical intervention using a well-established, rigorous research approach" (p. 7). We embrace this notion by employing and evaluating user experience guidelines developed by practitioners, and rigorously developing the constructs and the associated survey instrument to represent mobile application usability.

Concept and Instrument Development

MacKenzie et al. (2011) present a comprehensive construct conceptualization, measure development, and validation procedure. This procedure synthesized prior scale development literature (DeVellis 2011; Straub 1989; Straub et al. 2004), and integrated several methodological strategies for construct and scale development and validation. In total, the procedure consists of 10 steps that are shown in Figure 1.

We closely followed these guidelines to develop the constructs and scales to assess the usability of mobile applications. Below, we discuss how we applied the 10-step proce-

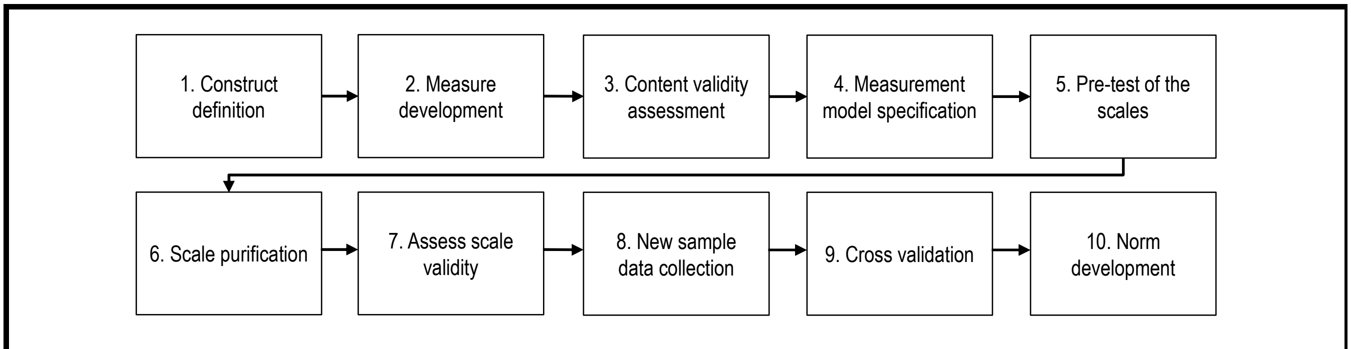


Figure 1. 10-Step Validation Procedure (Adapted from MacKenzie et al. 2011)

dure outlined by MacKenzie et al. in order to develop and validate the mobile application usability conceptualization and measures. In the discussion of each step, we describe our actions as they relate to the development of the mobile application usability conceptualization and instrument.

Step 1. Construct Definition

The first step of the 10-step procedure is to develop a conceptual definition of the constructs because the lack of a precise and detailed conceptualization of the focal constructs can cause significant measurement errors during the testing phase (DeVellis 2011; MacKenzie et al. 2011).

To inform the construct conceptualization, the lead author systematically reviewed and analyzed Apple's user experience guidelines (Apple 2012). In order to guide this process, the following three questions were posed:

(1) What are the main usability criteria suggested by Apple's guidelines?

(2) Which are the keywords associated with each usability criteria?

(3) What are the descriptions provided for each keyword?

These questions were used to systematically review and code Apple's comprehensive user experience guidelines. Strauss and Corbin's (1990) open and axial coding procedures were followed to identify conceptually similar themes discussed in Apple's guidelines. Open coding is the "analytical process through which concepts are identified and their properties and dimensions are discovered in the data" (Strauss and Corbin 1990, p. 101). Axial coding is the process of "relating categories to their subcategories, termed 'axial' because coding occurs around the axis of a category, linking categories at the level of properties and dimensions" (Strauss and Corbin 1990, p. 123). Initially, Apple's guidelines were carefully read and the content was coded using Strauss and Corbin's (1990, p. 119) line-by-line analysis. The most essential open codes identified are shown in Table 2. Next, the open codes were clustered and subcategories were formed by summarizing conceptually similar codes. The subcategories are also listed in Table 2. Then, using axial coding, the open codes were inspected for similarities and/or differences and then related into conceptual units.

For example, two open codes were identified that discussed the concept of branding in mobile applications: (1) in mobile applications, incorporate a brand's colors or images in a refined, unobtrusive way; and (2) the exception to these guidelines is your mobile application icon, which should be completely focused on your brand. Next, both open codes were conceptualized as one subcategory that was labeled as brand appropriately. Then, using axial coding, the major category was labeled as branding. Following this approach, Apple's guidelines were initially coded by the lead author and organized in a matrix, as outlined by Miles and Huberman (1999). Organizing codes in a data matrix is useful to compress coded information and it supports drawing conclusions (Miles and Huberman 1999). Next, the second author reviewed the user experience guidelines and associated coding outcomes. In a few cases where there was a disagreement between the authors, two independent judges who were unfamiliar with the study were asked to facilitate a discussion in order to reach a coding consensus. Both judges were IS researchers who held Ph.D. degrees from U.S. universities. Although both judges were familiar with HCI principles in general, neither of them considered himself/herself as a mobile application usability expert. They also indicated that they were unfamiliar with the literature on mobile application usability. We purposefully selected both judges who were unfamiliar with the literature on mobile application usability because we aimed to obtain unbiased

Table 2. Coding Matrix Adapted from Miles and Huberman (1999)

Axial Codes	Subcategory	Open Codes Derived from Apple's User Experience Guidelines
Branding	Brand appropriately	• In mobile applications, incorporate a brand's colors or images in a refined, unobtrusive way. • The exception to these guidelines is your mobile application icon, which should be completely focused on your brand.
Data preservation	Always be prepared to stop	• Mobile applications stop when people press the home button to open a different mobile application or use a device feature, such as the phone. • Mobile applications should save user data as soon as possible and as often as reasonable because an exit or terminate notification can arrive at any time. • Save the current state when stopping a mobile application.
Instant start	Start Instantly	• Mobile applications should start as quickly as possible so that people can begin using them without delay. • Display a launch image when starting mobile applications. • Avoid displaying a window or a splash screen when starting mobile applications. • When starting mobile applications, specify an appropriate status bar style.
Orientation	Handle orientation changes	• People often expect to use their mobile application in any orientation. • It is important to determine how to respond to this expectation, within the context of your mobile application and the task it enables. • Think twice before preventing your mobile application from running in all orientations. • Launch your mobile application in your supported orientation, regardless of the current device orientation. • Avoid displaying a user-interface element in your mobile application that tells people to rotate the device.
Collaboration	Enable collaboration and con-nectedness	• When appropriate your mobile application should make it easy for people to interact with others and share their location, opinions, and high scores. • People generally expect that mobile applications can share information that is important to them.
Content relevance	Focus on the primary task	• The primary focus of the mobile application is on the primary task. • To maintain the focus of the mobile application, it is important to determine what is most important in each context or screen.
	Elevate the content people care about	• People care about the experience; they do not expect to manage, consume, or create content in mobile applications. • Minimize the number and prominence of controls to decrease their weight in the user interface of your mobile application.
Search	Make search quick and rewarding	• Build indexes of your data in your mobile application so that you are always prepared for search. • In your mobile application that handles or displays a lot of data, search can be a primary function. • If you need to provide search in your mobile application, follow these guidelines to ensure that it performs well. • Live-filter local data so that you can display results more quickly in your mobile application. • It is best when your mobile application can begin filtering as soon as users begin typing, and narrow the results as they continue typing. • Your mobile applications should display a search bar above a list or the index in a list. • Mobile application users expect to find a search bar in this position, because they are accustomed to the search bar in contacts and other applications. • Your mobile application should use a tab for search only in special circumstances. • If search is a primary function in your mobile application you might want to feature it as a distinct mode.
Aesthetic graphics	Delight people with stunning graphics	• Rich, beautiful, engaging graphics draw people into a mobile application and make the simplest task rewarding. • In your mobile application, beautiful artwork also helps to build your application's brand in people's eyes. • Your mobile application should consider replicating the look of high-quality or precious materials. • When appropriate, your mobile application should create high-resolution artwork. • Your mobile application should ensure that your launch images and application icons are high quality.
Realism	Consider adding physicality and realism	• When appropriate, add a realistic, physical dimension to your mobile application. • Often, the more true to life your mobile application looks, the easier it is for people to understand how it works and the more they enjoy using it. • People instantly know what the voice memos app does, and how to use it, because it presents a beautifully rendered focal image (the microphone) and realistic controls in your mobile application.

Table 2. Coding Matrix Adapted from Miles and Huberman (1999) (Continued)

Axial Codes	Subcategory	Open Codes Derived from Apple's User Experience Guidelines
Subtle animation	Use subtle animation to communicate	• In your mobile application, animation is a great way to communicate effectively, as long as it does not get in the way of users' tasks or slow them down. • Your mobile application should communicate status and enhance the sense of direct manipulation. • Your mobile application should use subtle animations to help people visualize the results of their actions.
Control obviousness	Make usage easy and obvious	• Your mobile application should make the main function of your application immediately apparent. • Your mobile application should minimize the number of controls from which people have to choose. • Labeling controls clearly in your mobile application so that people understand exactly what they do. • Be consistent with the usage paradigms of the built-in mobile applications. • Users understand how to navigate a hierarchy of screens, edit list contents, and switch among modes using the tab bar in mobile applications. • Make it easy for people to use your mobile application by reinforcing their experience.
De-emphasis of user settings	De-emphasize settings	• Avoid including settings in your mobile application if you can. • Settings include preferred mobile application behaviors and information that people rarely want to change. • Users cannot open the settings application without first quitting your mobile application, and you do not want to encourage this action. • When you design your mobile application to function the way most of your users expect, you decrease the need for settings. • In the main user-interface of the mobile application, put options that provide primary functionality or that people want to change frequently. • Your mobile application should put options that people are unlikely to change frequently on the back of a view.
Effort minimization	Minimize the effort required for user input	• Inputting information takes time and attention, whether people tap controls or use the keyboard in your mobile application. • Balance any request for input by users with what you offer users in return in your mobile application. • Your mobile application should make it easy for users to input their choices. • In your mobile application, you can use a table view or a picker instead of a text field, because it is usually easier for people to select an item from a list than to type words. • Your mobile application should get information from the operating system, when appropriate. • Your mobile application should not force people to give you information the application can easily find for itself, such as their contacts or calendar information.
Fingertip-size controls	Make targets fingertip-size	• Your mobile application should consider that the screen size of mobile devices might vary, but the average size of a fingertip does not. • Your mobile application should give tappable elements in your application a target area of about 44 × 44 points.
Concise language	Be succinct	• Think like a newspaper editor, and strive to convey information in a condensed, headline style in your mobile application. • In your mobile application, when your user-interface text is short and direct, users can absorb it quickly and easily. • Be sure to correct all spelling, grammatical, and punctuation errors in your mobile application. • Keep all-capital words to a minimum in your mobile application.
Standardized user-interface element	Use user-interface elements consistently	• People expect standard views and controls to look and behave consistently across mobile applications. • Your mobile application should follow the recommended usages for standard user interface elements. • For a mobile application that enables an immersive task, such as a game, it is reasonable to create completely custom controls. • Avoid radically changing the appearance of a control that performs a standard action in your mobile application. • Your mobile application should use standard controls and gestures appropriately and consistently so that they behave the way people expect.

Table 2. Coding Matrix Adapted from Miles and Huberman (1999) (Continued)

Axial Codes	Subcategory	Open Codes Derived from Apple's User Experience Guidelines
User-centric terminology	Use user-centric terminology	• In all your text-based communication with users, use terminology you are sure that your users understand in your mobile application. • In particular, avoid technical jargon in the user interface in your mobile application.
Logical path	Give people a logical path to follow	• Your mobile application should make the path through the information you present logical and easy for users to predict. • Your mobile application should make sure to provide markers, such as back buttons, that users can use to find out where they are. • In most cases, give users only one path to a screen in your mobile application. • If your mobile application needs to be accessible in different circumstances, consider using a modal view that can appear in different contexts.
Top-to-bottom structure	Think top down	• Your mobile application should put the most frequently used (usually higher level) information near the top, where it is most visible and easy to reach. • As the user scans the screen from top to bottom, the information displayed should progress from general to specific and from high level to low level in your mobile application.
Short icon-labeling	Give controls short labels	• Your mobile application should give controls short labels, or use well-understood symbols, so that people can tell what they do at a glance. • When appropriate, use the built-in buttons and icons in your mobile application.

feedback on the codes derived from Apple's user experience guidelines. Table 2 shows the final matrix derived from Apple's guidelines. All generated axial codes, shown in the left column of Table 2, were then used as a base for the construct development. Specifically, we used these axial codes and related them to the extant usability literature. In some cases, the user experience guidelines matched existing constructs. For instance, Apple's guidelines suggest that the *content* of mobile applications needs to be relevant to the users and developers should consider the target audience when designing mobile applications. Prior research studying the usability of web and mobile sites has also established that the content of web and mobile applications drives users' intentions to use the applications (Venkatesh and Ramesh 2006). Thus, the concept of content was informed by practitioner guidelines and prior research. In some cases, the user experience guidelines suggested construct domains for which we could not identify existing theoretical conceptualizations or concept definitions in the literature. For instance, realism was suggested by Apple's guidelines to be an essential part of mobile applications due to the fact that users would instantly associate with the application because it would look familiar to them. Examples for realism might be address books or time and date applications in which traditional designs are used to better communicate the purpose of the application (e.g., by using a traditional address book design). Although we found some literature supporting realism in mobile applications (e.g., Kang 2007), our literature review did not reveal any theoretically motivated constructs with overlapping themes to the concept of realism.

Table 3 illustrates the process of comparing the initial construct conceptualization, codes derived from Apple's guidelines, and existing literature. The leftmost column lists the construct domains we identified based on the coding procedure explained earlier. The middle columns show the open codes we derived from Apple's guidelines and prior work that has studied the identified construct domains. The rightmost column provides prior research that has studied the construct domains that we identified. In total, our analysis of Apple's user experience guidelines led to 20 initial constructs that represented the most essential concepts outlined in the user experience guidelines. Following the guidelines of MacKenzie et al., we conceptualized each construct by identifying the construct entities as well as the entities to which each construct applies. Table 4 lists all constructs and definitions.

The next phase of the construct conceptualization step included the identification of higher-order constructs. This is an important issue for the scale development process, and this needs to be done once all constructs have been conceptualized and defined (MacKenzie et al. 2011). Constructs sharing a common theme and similar characteristics should be theoretically abstracted to a higher level and identified as such (Edwards 2001; MacKenzie et al. 2011).

To identify potential higher-order constructs, we carefully examined the constructs shown in Table 4 for conceptual similarities and reviewed the existing literature on each identified construct. During this phase, we discussed how distinctive the constructs were from each other and if eliminating any of them would restrict the domain of the construct in a significant way (MacKenzie et al. 2011). We also asked two judges to help us identifying conceptual similarities among the constructs. Both judges were IS researchers and

Table 3. Interplay between Apple User Experience Guidelines and Literature

Construct	Examples of Open Codes Derived from Apple User Experience Guidelines	Literature Examples	Prior Literature
Branding	• Incorporate a brand's colors or images in a refined, unobtrusive way. • The exception to these guidelines is your application icon, which should be completely focused on your brand.	Lowry et al. (2008) studied the effect of branding in websites on individuals' trust formation. The results showed that individuals appreciate branding efforts if incorporated effectively.	Devaraj et al. 2002; Dou et al. 2010; Lowry et al. 2008
Data preservation	• Save user data as soon as possible and as often as needed. • Save the current state when stopping.	Sarker and Wells (2003) studied the handheld device use and adoption. The findings suggested that data should be preserved on mobile devices at any given time.	Sarker and Wells 2003
Instant start	• Mobile applications should start as quickly as possible so that people can begin using them without delay. • Display a launch image.	Pousttchi and Schurig (2004) defined a set of user requirement for mobile banking applications. One recommendation was that the application should be instantly available due to the fact that users would reject long waiting periods.	Galletta et al. 2006; Pousttchi and Schurig 2004; Thong et al. 2002
Orientation	• People often expect to use their mobile application in any orientation. • It is important to determine how to respond to this expectation, within the context of your application and the task it enables.	Wobbrock et al. (2008) investigated the influence of hand postures regarding the interaction with mobile devices. The findings proposed that the orientation of a handheld device is important for the human–mobile interaction.	Wobbrock et al. 2008
Collaboration	• When appropriate, make it easy for people to interact with others and share things like their location, opinions, and high scores. • People generally expect to be able to share information that is important to them.	Oulasvirta et al. (2007) recommended that mobile applications should allow users to interact with each other. This would be particularly useful for applications aiming to provide group content.	Hess et al. 2009; Oulasvirta et al. 2005; Oulasvirta et al. 2007
Content relevance	• The primary focus of the application is on the primary task. • People care about the experience; they do not expect to manage, consume, or create content.	Agarwal and Venkatesh (2002) developed a survey instrument to test usability of websites. Content was an important factor influencing perceived website usability.	Agarwal and Venkatesh 2002; Cyr et al. 2009; Kim and Stoel 2004; Mithas et al. 2007; Sørensen and Altaitoon 2008; Tan et al. 2009; Venkatesh and Agarwal 2006; Venkatesh and Ramesh 2006; Wells et al. 2005; Wells, Valacich, and Hess et al. 2011; Xu et al. 2009
Search	• Build indexes of your data so that you are always prepared for search. • In applications that handle or display a lot of data, search can be a primary function.	Dou et al. (2010) studied the effect of brand positioning strategy using search engine marketing. The findings suggested that search features are vital to website utility.	Dou et al. 2010; Gebauer et al. 2007; Hess et al. 2009; Kurniawan 2008; Nah et al. 2010; Nah et al. 2011; Tan et al. 2009; Xu et al. 2009
Aesthetic graphics	• Rich, beautiful, engaging graphics draw people into an application and make the simplest task rewarding. • Beautiful artwork also helps to build your application's brand in people's eyes.	Wells, Parboteeah, and Valacich (2011) conducted an experiment in which color and visual graphics were manipulated. Visual graphics was found to be important for the overall website quality.	Aladwani and Palvia 2002; Cyr et al. 2009; De Wulf et al. 2006; Huizingh 2000; Kim and Stoel 2004; Parboteeah et al. 2009; Wells, Parboteeah, and Valacich 2011; Wells, Vala-cich, and Hess 2011; Zhang and von Dran 2000, 2001

Table 3. Interplay between Apple User Experience Guidelines and Literature (Continued)

Construct	Examples of Open Codes Derived from Apple User Experience Guidelines	Literature Examples	Prior Literature
Realism	• When appropriate, add a realistic, physical dimension to your application. • Often, the more true to life your application looks and behaves, the easier it is for people to understand how it works and the more they enjoy using it.	Kang (2007) studied the impact of realistic avatars on the perceived interaction with mobile phones. They recommend that mobile phone interface use is facilitated by including realistic elements.	Kang 2007
Subtle animation	• Animation is a great way to communicate effectively, as long as it does not get in the way of users' tasks or slow them down. • Communicate status and enhance the sense of direct manipulation.	Hong et al. (2004a) studied the effect of flash animation on information search performance and individuals' perceptions.	Hess et al. 2005; Hong et al. 2004a, 2007; Lim et al. 2000
Control obviousness	• Make the main function of your application immediately apparent. • Minimizing the number of controls from which people have to choose.	Seffah et al. (2006) consolidated extant literature on mobile application usability and proposed that user controls should be instantly obvious for users. This way, individuals could immediately start using a mobile application.	Jokela et al. 2006; Seffah et al. 2006; Sørensen and Altaitoon 2008; Wells, Parboteeah, and Valacich 2011; Wells, Valacich, and Hess 2011
De-emphasis of user settings	• Avoid including settings in your application if you can. • Settings include preferred application behaviors and information that people rarely want to change.	Jokela et al. (2006) developed a quantitative measure for evaluating the user interface of mobile phones. One recommendation of this study was to avoid prompting users to change settings in the user menu.	Jokela et al. 2006; Sørensen and Altaitoon 2008; Tan et al. 2009; Xu et al. 2009
Effort minimization	• Inputting information takes time and attention, whether people tap controls or use the keyboard. • Balance any request for input by users with what you offer users in return.	Tan et al. (2009) proposed several meta-categories for web-site design using repertory grid technique. The research participants suggested that surfing sites must be effortlessly in order to maintain a high usability of applications.	Gebauer et al. 2007; Jokela et al. 2006; Kurniawan 2008; Parboteeah et al. 2009; Seffah et al. 2006; Tan et al. 2009; Valacich et al. 2007; Xu et al. 2009
Fingertip-size controls	• The screen size of mobile devices might vary, but the average size of a fingertip does not. • Give tappable elements in your application a target area of about 44 × 44 points.	Kurniawan (2008) researched the effect of age on mobile application usability. The results suggested that particularly elder people appreciate larger buttons as part of the mobile interface.	Kurniawan 2008
Concise language	• Think like a newspaper editor, and strive to convey information in a condensed, headline style. • When your user-interface text is short and direct, users can absorb it quickly and easily.	De Wulf et al. (2006) argued that individuals like using websites that use concise language to communicate content.	Aladwani and Palvia 2002; De Wulf et al. 2006; Hess et al. 2009; Robbins and Stylianou 2003; Sørensen and Altaitoon 2008; Wells, Valacich, and Hess 2011
Standardized user-interface element	• People expect standard views and controls to look and behave consistently across applications. • Follow the recommended usages for standard user interface elements.	Tan et al. (2009) proposed that effective website design relies on consistency. Users should recognize the user interface structure based on other media.	Adipat et al. 2011; Gebauer et al. 2007; Tan et al. 2009; Xu et al. 2009
User-centric terminology	• In all your text-based communication with users, use terminology you are sure that your users understand. • In particular, avoid technical jargon in the user interface.	Kurniawan (2008) recommended that particularly older users would reject technical jargon used in mobile applications.	Aladwani and Palvia 2002; De Wulf et al. 2006; Huizingh 2000; Kurniawan 2008; Robbins and Stylianou 2003; Tan et al. 2009; Xu et al. 2009; Zhang and von Dran 2000; Zhang and von Dran 2001

Table 3. Interplay between Apple User Experience Guidelines and Literature (Continued)

Construct	Examples of Open Codes Derived from Apple User Experience Guidelines	Literature Examples	Prior Literature
Logical path	• Make the path through the information you present logical and easy for users to predict. • In addition, be sure to provide markers, such as back buttons, that users can use to find out where they are and how to retrace their steps.	Treiblmaier (2007) reviewed the extant literature on website usability. User-friendly websites should follow a clear and logical path making it easier to move from one site to another.	Adipat et al. 2011; Hess et al. 2005; Hong et al. 2004a, 2004c; Jokela et al. 2006; Mithas et al. 2007; Palmer 2003; Treiblmaier 2007; Wells, Valacich, and Hess 2011
Top-to-bottom structure	• Put the most frequently used (usually higher level) information near the top, where it is most visible and easy to reach. • As the user scans the screen from top-to-bottom, the information displayed should progress from general to specific and from high level to low level.	Adipat et al. (2011) studied the effect of tree-view based presentation on mobile sites. The findings suggested that individuals can follow mobile sites easier if the most essential information is presented near the top of the site.	Adipat et al. 2011; Geissler 2001; Hong et al. 2004a, 2004c; Kim et al. 2005; Mithas et al. 2007; Palmer 2003; Wells, Parboteeah, and Valacich 2011; Wells, Valacich, and Hess 2011
Short icon-labeling	• Give controls short labels, or use well-understood symbols, so that people can tell what they do at a glance. • When appropriate, use the built-in buttons and icons.	Kurniawan (2008) suggested that individuals reject lengthy labels for icons and short labels would positively influence the usability of mobile applications.	Gebauer et al. 2007; Huizingh 2000; Jokela et al. 2006; Kurniawan 2008; Robbins and Stylianou 2003

*Only two open code examples are listed above. The complete final coding matrix is shown in Table 2.

unfamiliar with the study's content. To facilitate this process, both judges were provided with 20 cards. Each card was labeled with a construct name and associated construct definition (as shown in Table 4) that was identified in the previous phase of the construct development procedure. The card sorting results were then compared with the higher-order constructs identified by both authors. After a thorough discussion between the authors, 6 second-order constructs were identified to represent aggregations of the 20 identified first-order constructs. We also considered if second-order constructs could be represented by third-order constructs, that is, higher-level abstractions (Rindskopf and Rose 1988; Wetzels et al. 2009). However, we concluded that each second-order construct represented a unique part of mobile application usability and it was thus not necessary or useful to form third-order constructs. Table 5 shows all conceptualized second-order constructs and their definitions which we discuss next.

Second-Order Construct: Application Design

Application design is the degree to which a user perceives that a mobile application is generally designed well. Prior literature on application design suggests that users are influenced by several factors when evaluating the overall design of a mobile application. For example, it is important that the application preserves data that is input by the user (Adipat et al. 2011). If users are required to input the same data twice (e.g., when switching from one screen to another), they will become frustrated and dissatisfied with the mobile application (Adipat et al. 2011; Kurniawan 2008; Tan et al. 2009). It is also essential that a mobile application is instantly ready to be used after being switched on as long wait times normally lead to user frustration (see Devaraj et al. 2002, 2006; Galletta et al. 2006). Moreover, due to the fact that mobile phones are handheld devices, systems designers should consider that the mobile application displays the information well, independent of whether the mobile device is held horizontally or vertically (Wobbrock et al. 2008). Finally, Apple's user experience guidelines suggest that well-designed applications employ subtle branding efforts. This is consistent with marketing research on website design that suggests that subtle branding efforts can increase consumer satisfaction with overall website design (Dou et al. 2010). Hence, data preservation, instant start, orientation, and branding are proposed as first-order constructs forming the design of mobile applications.

Second-Order Construct: Application Utility

Application utility is defined as the degree to which a user perceives a mobile application generally serves its purpose well. Several factors were proposed in Apple's guidelines related to this concept. Mobile applications should focus on the content that is most relevant to its users (Venkatesh and Agarwal 2006; Venkatesh and Ramesh 2006). It is essential that the user gets what he/she expects and the main purpose of the application should be emphasized (see Hess et al. 2009;

Table 4. Constructs, Construct Entities, and Construct Definitions

Construct Name	Entity (E) to which the Construct Applies and General Property (GP)	Construct Definition
Branding	E = Person; GP = perception about the branding embedded within the mobile application	The degree to which a user perceives that the mobile application integrates branding appropriately.
Data preservation	E = Person; GP = perception about the data preservation mechanisms during the use of the mobile application	The degree to which a user perceives that the mobile application preserves data automatically.
Instant start	E = Person; GP = perception about the ability to start using the mobile application instantly after launching it	The degree to which a user perceives that the mobile application starts instantly after switching it on.
Orientation	E = Person; GP = perception about the ability to run the mobile application independent of its orientation (vertical/horizontal)	The degree to which a user perceives that the mobile application displays information well independent of whether the device is held horizontally or vertically.
Collaboration	E = Person; GP = perception about the ability to collaborate with others through the mobile application	The degree to which a user perceives that the mobile application enables users to connect with other individuals.
Content relevance	E = Person; GP = perception about the content relevance of the mobile application	The degree to which a user perceives that the mobile application focuses on the most relevant content.
Search	E = Person; GP = perception about the search options of the mobile application	The degree to which a user perceives that the mobile application helps users to search for information.
Aesthetic graphics	E = Person; GP = perception about the use of aesthetic graphics as part of the mobile application	The degree to which a user perceives that the mobile application makes use of aesthetic graphics.
Realism	E = Person; GP = perception about the use of realistic icons or pictures in the mobile application	The degree to which a user perceives that the mobile application incorporates realistic icons or pictures.
Subtle animation	E = Person; GP = perception about the integration of appropriate animations as part of the mobile application	The degree to which a user perceives that the mobile application uses subtle animations effectively.
Control obviousness	E = Person; GP = perception about the controls of the mobile application	The degree to which a user perceives that the mobile application deploys controls that are immediately obvious.
De-emphasis of user settings	E = Person; GP = perception about the user settings of the mobile application	The degree to which a user perceives that the mobile application de-emphasizes user settings.
Effort minimization	E = Person; GP = perception about the effort minimization for user input in the mobile application	The degree to which a user perceives that the mobile application minimizes the effort to input data.
Fingertip-size controls	E = Person; GP = perception about the use of large control elements as part of the mobile application	The degree to which a user perceives that the mobile application deploys fingertip-size controls.
Concise language	E = Person; GP = perception about the concise language used within the mobile application	The degree to which a user perceives that the mobile application makes use of concise language.
Short icon-labeling	E = Person; GP = perception about the labeling of the icons used within the mobile application	The degree to which a user perceives that the mobile application makes use of short icon labels.
Standardized user-interface element	E = Person; GP = perception about the user-interface of the mobile application	The degree to which a user perceives that the mobile application deploys standardized user-interfaces that are commonly used by other mobile applications.
User-centric terminology	E = Person; GP = perception about the use of user-centric language in the mobile application	The degree to which a user perceives that the mobile application deploys user-centric terminology.
Logical path	E = Person; GP = perception about the logical path of the mobile application	The degree to which a user perceives that the mobile application presents information logically and predictably.
Top-to-bottom structure	E = Person; GP = perception about the top-to-bottom structure of the mobile application	The degree to which a user perceives that the mobile application displays frequently used information on the top of the application.

Table 5. Second-Order Constructs, Construct Entities, and Construct Definitions

Construct Name	Entity (E) to which the Construct Applies and General Property (GP)	Construct Definition
Application design	E = Person; GP = overall perception about the mobile application design	The degree to which a user perceives that the mobile application is generally designed well.
Application utility	E = Person; GP = overall perception about the utility of the mobile application	The degree to which a user perceives that the mobile application generally serves its purpose well.
User interface graphics	E = Person; GP = overall perception about the interface graphics of the mobile application	The degree to which a user perceives that the mobile application's user interface graphics are effectively designed.
User interface input	E = Person; GP = overall perception about the user interface input mechanisms of the mobile application	The degree to which a user perceives that the mobile application allows users to input data easily.
User interface output	E = Person; GP = overall perception about the user interface output of the mobile application	The degree to which a user perceives that the mobile application presents content effectively.
User interface structure	E = Person; GP = overall perception about the user interface structure of the mobile application	The degree to which a user perceives that the mobile application is structured effectively.

Hong et al. 2004b, 2004c; Li, Hess et al. 2009; Thong et al. 2002; Venkatesh and Ramesh 2006; Wells, Parboteeah, and Valacich 2011; Wells, Valacich, and Hess 2011). For example, prior research on mobile applications has emphasized the concept of context-awareness, which can be defined as "the ability of an application to detect and understand its situational context and to adapt its behavior in a user preferred manner accordingly" (Zhang et al. 2009, p. 29). An example of a context-aware application is a GPS-based mobile application that provides an interface for viewing map products and managing geographic data in real time. Because timely and accurate geographic information is most relevant to users of mobile GPS-based map applications (Kaikkonen et al. 2005; Kakihara and Sørensen 2002), it is important that these context-aware applications center on location-based information and do not focus on less relevant features (e.g., detailed environmental information) that could distract the user from the most relevant content of the application. Likewise, a mobile application should enable users to search for information. If bounded by limited screen size, a search function can lead to a better user experience because it helps users to better navigate an application (Campbell et al. 2009; Hess et al. 2009; Valacich et al. 2007; Wells et al. 2005). Finally, an application should also help users to share information with others and collaborate with friends and colleagues (Hess et al. 2009; Oulasvirta et al. 2007). For example, many mobile gaming applications emphasize this feature and users view this as a useful feature because it leads to a more social user experience. Thus, content relevance, search, and collaboration are proposed as first-order constructs forming application utility of mobile applications.

Second-Order Construct: User Interface Graphics

User interface graphics is defined as the degree to which a user perceives a mobile application's user interface graphics to be effectively designed. For example, it is important that animations are designed subtly and are not used extensively in mobile applications (see Hess et al. 2005; Hong et al. 2004a, 2007; Lim et al. 2000). Likewise, it would lead to a better user experience if the mobile application incorporates realistic icons or pictures (Kang 2007). For instance, many pre-installed applications, including Apple's calendar and address book functions, are designed realistically to represent traditional calendars or address books. This would allow users to instantly recognize the function of an application (Flavian et al. 2006). Finally, the graphics used in a mobile application should be designed so as to be aesthetically appealing (Aladwani and Palvia 2002; Cyr et al. 2009; Hess et al. 2005; Hong et al. 2007; Kim and Stoel 2004; Parboteeah et al. 2009; Wells, Parboteeah, and Valacich 2011; Wells, Valacich, and Hess 2011). Therefore, subtle animation, realism, and aesthetic graphics are proposed as first-order constructs forming user interface graphics.

Second-Order Construct: User Interface Input

User interface input is the degree to which a user perceives that a mobile application allows easy input of data. For example, fingertip-size controls would help users to select functions and menus in mobile applications (Kurniawan 2008). Relatively large controls would be more usable because it will be easier for users to pick desired functions. Further, mobile application controls should be immediately obvious and intuitive to use (Jokela et al. 2006; Seffah et al. 2006; Sørensen and Altaitoon 2008). Apple's guidelines suggest that users are not willing to spend much time learning mobile applications and it is essential that the user interface is clear when using an application for the first time. The effort it takes users to input data should be minimized and it is important that a mobile application offers mechanisms for easy data input. This could include drop-down menus or automated data entry support (see Hong et al. 2004a; Valacich et al. 2007; Wells et al. 2005). Finally, although user settings should be available in mobile applications, they should be designed to remain in the background of the application and users should not be frequently prompted to adjust the settings when using a mobile application (Jokela et al. 2006; Sørensen and Altaitoon 2008; Tan et al. 2009; Xu et al. 2009). Therefore, fingertip-size controls, control obviousness, effort minimization, and de-emphasis of user settings are proposed as first-order constructs forming user interface input.

Second-Order Construct: User Interface Output

User interface output is the degree to which a user perceives that a mobile application presents content effectively. Apple's guidelines suggest that mobile applications should apply user-centric terminology and avoid technical jargon. Technical terms and difficult to read text passages would lead to frustrated users and decreased user friendliness. Similarly, the displayed text should be written in a concise style and lengthy descriptions should be avoided (Hess et al. 2009). The same principle would apply to the labels of icons and thus lengthy icon descriptions should be avoided (Gebauer et al. 2007; Huizingh 2000; Jokela et al. 2006; Kurniawan 2008; Robbins and Stylianou 2003). Finally, to increase the recognizability of functions used in a mobile application, standardized user interface elements should be employed. Users will appreciate standard elements because it would make them believe that they are already familiar with the user interface (Hess et al. 2009; Thong et al. 2002; Hong et al. 2007; Vila and Kuster 2011; Wells, Parboteeah, and Valacich 2011; Wells, Valacich, and Hess 2011). Hence, user-centric terminology, concise language, short icon labeling, and standardized user interface elements are proposed as first-order constructs forming user interface output.

Second-Order Construct: User Interface Structure

User interface structure is the degree to which a user perceives that a mobile application is structured effectively. Apple's guidelines suggest that the user interface of mobile applications should be structured from top-to-bottom. Users would intuitively start searching for the most important information on the top of the screen (Li, Hess et al. 2009; Valacich et al. 2007; Wells, Valacich, and Hess 2011). Thus, it would be best to place key information at the top of the screen. Also, mobile applications require a logical and predictable path (see Adipat et al. 2011; Devaraj et al. 2002, 2006; Hess et al. 2009; Hong et al. 2004a, 2004c; Venkatesh, Ramesh, and Massey 2003; Thong et al. 2002; Wells, Parboteeah, and Valacich 2011; Wells, Valacich, and Hess 2011). Therefore, top-to-bottom structure and logical path are proposed as first-order constructs forming user interface structure.

Following MacKenzie et al., we next discuss if the identified second-order constructs should be viewed as defining characteristics of mobile application usability. This would indicate that the identified mobile application usability dimensions should be modeled as formative indicators of the mobile application usability construct (Diamantopoulos et al. 2008; MacKenzie et al. 2011; Petter et al. 2007). In order to decide a construct's directionality, MacKenzie et al. proposed considering if changes in one of the subdimensions would be associated with a change in the focal construct. Thus, in thinking through the relationship between the second-order constructs and mobile application usability, we felt that the subdimensions are defining characteristics of mobile application usability—hence, mobile application usability should be modeled formatively. For example, it is reasonable to say that an increase in the level of application design would be associated with an increase in the overall usability of a given mobile application, without necessarily being associated with any changes in the application utility or user interface graphics. Similar arguments could be made for all second-order constructs. Thus, we model mobile application usability as a function of the second-order constructs shown in Table 5.

Step 2. Measure Development

Once the constructs of interest are well defined, the next step is the creation of items (MacKenzie et al. 2011). During the item development, the codes derived from the mobile user experience guidelines were leveraged. Many sections of Apple's guidelines included information that was helpful in designing the initial items. The open codes derived from the guidelines were also helpful during the item creation process because they included keywords describing the construct domains identified in step 1. We also searched the literature for prior research that may have examined similar constructs to identify items that may be relevant to us. While doing this, we examined the identified literature for additional keywords that could be used to develop items.

Altogether, 120 items were initially developed to capture the most essential aspects of the constructs outlined in Table 4. Next, the simplicity and wording of the items were examined using a face validity check. Face validity checks are useful in situations when items are developed from scratch and have not yet been tested with individuals (DeVellis 2011; MacKenzie et al. 2011; Straub et al. 2004). The face validity check focused on the items themselves and did not ask participants to rank or respond to the items. Six individuals performed the face validity check. Three administrative staff members, two Ph.D. students, and one Master's student volunteered to participate in the face validity check. A prerequisite for participation in the face validity check was that the participant be a mobile smartphone user, which would ensure that they understood the context of the items. The participants were provided with a paper-based survey that included all the 120 items. The participants were asked to examine all items and to comment on the clarity of the questions. To identify weak items, we asked the participants to flag items whose wording was confusing or vague. In total, 41 items were identified as too vague or worded unclearly, with specific changes suggested to certain items. The authors discussed the 41 flagged items. Three of these 41 items remained in the item pool after the wording was modified. This led to a pool of 82 items that were used in the next step, the content validity check.

Step 3. Assessment of the Content Validity of the Items

Content validity is defined as the extent to which a scale represents all facets of a given construct (Lewis et al. 2005; MacKenzie et al. 2011; Straub et al. 2004). According to MacKenzie et al., researchers should consider two major components when assessing the content validity of a survey instrument: (1) Is the individual item representative of an aspect of the content domain of the construct? (2) Are the items as a set collectively representative of the entire content domain of the construct? In the IS field, content validity has been assessed infrequently by researchers developing survey instruments (MacKenzie et al. 2011; Straub et al. 2004).

MacKenzie et al. proposed using a variance analysis approach to assess content validity of the items. Although this technique is rarely used in IS research, organizational studies have used it frequently (see Hinkin and Tracey 1999; Yao et al. 2007). This procedure includes the use of a matrix in which items are placed in the rows and construct definitions are listed at the top of the columns. Then, raters would indicate the extent to which items capture the construct domain using a Likert-type scale (e.g., 1 = not at all; 5 = completely). A one-way ANOVA can be used to assess if an item's mean rating on one aspect of the construct's domain differs significantly from the item's other construct domain. One disadvantage of this approach is that raters would need "sufficient intellectual ability to rate the correspondence between items and the theoretical definitions" (MacKenzie et al. 2011, p. 306).

A slightly less comprehensive technique for assessing the content validity of new scales was proposed by Anderson and Gerbing (1991). Their approach assumes that each item represents a single construct. Thus, rather than rating each item-construct combination, respondents are only asked to assign each item to a single corresponding construct definition. Although the variance analysis approach to content validity (Hinkin and Tracey 1999) is appealing due to its preciseness (Yao et al. 2007), Anderson and Gerbing's approach seems to be particularly suited for large survey instruments (Yao et al. 2007). This is due to the fact that raters would be required to rate each item-to-construct definition in the approach suggested by Hinkin and Tracey (1999), whereas Anderson and Gerbing's procedure would only require selecting one construct per item.

As recommended by MacKenzie et al., we initially attempted to use Hinkin and Tracey's variance analysis approach to assess the content validity of our newly developed scales. Thus, we developed a matrix and organized construct definitions in the columns and placed the items in the rows (Yao et al. 2007). It is important to note that Hinkin and Tracey used this approach with four constructs and 39 items. MacKenzie et al. stressed that it is important to avoid overburdening the raters by exposing them to too many content domains at the same time. They proposed limiting the content domains to a maximum of 8 to 10 aspects at the same time. We, therefore, split the pool of items and developed several matrixes, with a view toward reducing the complexity of the rating exercise. They also suggested that researchers who construct new scales would be well-advised to employ construct definitions from related constructs in order to ensure that the new items have content that is free from external content domains. Therefore, we aimed to cluster similar construct definitions together in order to identify potentially overlapping across constructs. In particular, we combined the first-order constructs that belonged to a similar multidimensional construct in each matrix. For instance, as identified in step 1 of the survey development process, the higher-order construct application design is formed by four first-order constructs: branding, data preservation, instant start, and orientation. Therefore, these first-order constructs were combined in one matrix in order to identify potential overlap across these constructs.

Next, we asked four independent raters to evaluate and examine the matrixes. Two raters were IS researchers and two were administrative staff at an Australian university. Each rater was provided with instructions and paper-based matrixes (Table 6 illustrates a matrix used for the content validity check).

Each respondent was asked to rate how well each item (row) corresponds to each construct definition (column) on a scale from 1 to 7 (1 = strongly disagree; 7 = strongly agree). Once respondents completed the survey, we interviewed them to ask them about their experiences with the rating procedure. All respondents reported that the instructions were clear but

Table 6. Matrix Example Used for Content Validity Check

Items	Construct Definitions			
The mobile application:	The degree to which a user perceives that the mobile application deploys controls that are immediately obvious.	The degree to which a user perceives that the mobile application minimizes the effort to input data.	The degree to which a user perceives that the mobile application de-emphasizes user settings.	The degree to which a user perceives that the mobile application integrates branding appropriately.
doesn't request you to modify the user setting within the application.				
uses intuitive commands.				
uses brand colors or images in a refined and unobtrusive way.				
offers you fields to choose from so that you don't have to type in text.				
employs controls that are intuitive.				
uses controls that are immediately obvious.				
doesn't prompt you to change user settings within the application.				
quietly reminds you of the brand that runs the application.				
integrates branding effectively.				
minimizes effort for you to type in information.				
makes the main function of the application immediately apparent.				
makes it easy for you to input your choice.				
doesn't force me to watch an advertisement.				
de-emphasizes user settings.				
avoids setting up user preferences within the application.				
allows me to perform tasks without having to input data.				

they found it challenging to complete the matrixes and warned that individuals may have difficulties completing the task. For example, one administrative staff provided feedback via e-mail and stated,

> I found it difficult to link the concepts to the questions. Some of them did not correspond to each other at all and I was wondering why I should rate them. I think it would be better to simplify the questions.

Further, three out of four raters suggested that it would be more logical to pick-and-choose the most appropriate construct definition for each item (rather than rating how well each construct definition corresponds to each item). Interestingly, the approach suggested by the respondents corresponded to the content validity assessment approach described by Anderson and Gerbing (1991).[2]

[2]When queried about whether they were familiar with the Anderson and Gerbing technique or if they had used such a technique before, none of the respondents indicated familiarity.

Given the overall length of the survey and the feedback we received, we decided to use the modified matrixes, as suggested by Anderson and Gerbing. We anticipated that ranking all the initial item-to-construct definition combinations would overburden raters. The content validity survey was reorganized and the instructions were rewritten by asking raters to select only one corresponding construct definition for each item. It should be noted that the items were worded generically for the task but when they were actually administered, the application could be replaced with the name of the actual application. This approach is consistent with prior instrument development research in IS (e.g., Compeau et al. 2012). We discussed this new survey with the raters who had previously assessed the matrixes. Each rater confirmed that the task would be more meaningful to them and they indicated it would be easier to complete the survey.

Content Validity Check

MacKenzie et al. and Anderson and Gerbing suggested that content validity raters should come from the main population of interest. Therefore, we executed this step by collecting data from actual consumers recruited by a market research firm. We obtained 350 responses. Out of 350 respondents, 318 participants took a reasonable length of time to complete the survey. The excluded respondents took less than 5 minutes to complete the content validity check, possibly because they did not pay sufficient attention to the questions asked.

Appendix A summarizes the respondent demographics. As evident from Appendix A, the sample had a large proportion of students, which matched the profile of the sampling frame provided by the market research firm. Thus, it was felt that nonresponse bias was not a concern. Further, comparing early versus late responses was not felt to be useful because all responses were collected during a single weekend and no reminders were employed (Churchill 1979; Hair et al. 1998). At first glance, the high proportion of young students may seem problematic because newly developed items should be understood by the main population of interest (Compeau et al. 2012; MacKenzie et al. 2011). However, such a sample was appropriate to check the content validity of the scales for two major reasons: (1) mobile applications are primarily used by young individuals and these applications are particularly popular among students (Hampton et al. 2011), and (2) Hinkin and Tracey suggested that even samples containing exclusively students are appropriate for content validity checks because the sorting procedure requires analytical thinking skills (see Compeau et al. 2012). For these reasons, the respondent demographics were seen to be acceptable for the content validity check.

We computed two indexes from the survey data following Anderson and Gerbing. First, the proportion of substantive agreement (P_{SA}) was computed. P_{SA} indicates the proportion of respondents who assign items to their intended constructs by using the following formula:

$$P_{sa} = \frac{nc}{N}$$

Where nc is the number of respondents who assigned an item to its intended construct and N is the total number of respondents. P_{SA} values can range between 0 and 1, with a high value indicating higher agreement that the construct definition represents the items judged (Anderson and Gerbing 1991).

Second, the substantive validity coefficient (C_{SV}) was computed. C_{SV} is the extent to which respondents assign items to the posited construct rather than to any other construct. The ratio is computed using the following formula:

$$C_{sv} = \frac{nc - n0}{N}$$

Where nc is the number of respondents assigning an item to the intended construct, $n0$ is the highest number of assignment of the measure to any other construct, and N is the total number of respondents. C_{SV} values can range from -1 and 1. Positive values suggest that an item was assigned to its intended construct more than assignment to any other construct. Negative values suggest the opposite.

The results of these analyses are shown in Appendix B. We applied a threshold of .60 as a cut-off point for P_{SA} and C_{SV} values for our content analysis. When using a .60 cut-off value, the results suggest that more than 60 percent of all raters associated the items with the intended construct definitions.

Overall, the content validity ratios were high, indicating that most respondents sorted the majority of items into their posited construct domains. Out of 82 items, 11 items did not meet the .60 cut-off value. Appendix B shows that the C_{SV} values obtained for DAPR4, CRLV2, CRLV4, COOB2, ICOL1, ICOL2, ICOL3, ICOL4, TTPS2, TTPS3, and TTPS5 were lower than the threshold of .60. However, all P_{SA} values obtained for these items were above the .60 cut-off value. Thus, we carefully inspected the item wordings and compared them with the construct definitions. In some cases (e.g., CRLV2), we reworded the items slightly in order to better align the items with the construct definitions. In contrast, all short-icon labeling (ICOL 1-4) items were closely associated with concise language items (CLAN1-4) and the resulting C_{SV} values were lower than .05. These findings could be interpreted as suggesting that individuals assume that if a mobile application makes use of concise language, it will also have concise labels for icons. After careful consideration, we decided to exclude the ICOL construct. Appendix C shows the item pool base following the content validity check including 78 items for measuring the constructs derived from Apple's user experience guidelines.

Step 4. Formally Specify the Measurement Model

The next step in the survey development process is the specification of the measurement model (MacKenzie et al. 2011). This process focuses on specifying how the indicators relate to the constructs and the relationships between the first-order and second-order constructs.

Following the preceding steps in the survey development, we developed two alternative measurement models. Although we developed the rationale for the second-order constructs in step 1 earlier, we wanted to ensure that our model was a good specification. In the first measurement model, all constructs

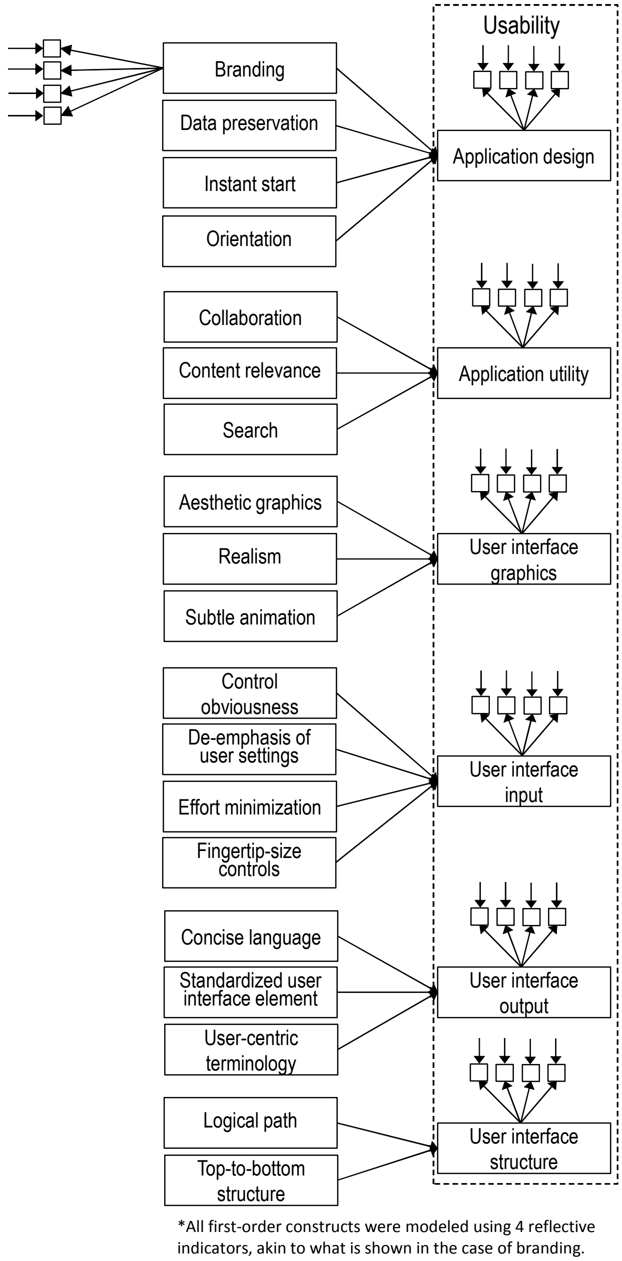


Figure 2. Measurement Model

were modeled as first-order constructs. In this case, one path between the reflective indicators and each first-order construct was set as 1 when setting up the measurement model (Hair et al. 1998; MacCallum and Browne 1993; MacKenzie et al. 2011). Our proposed measurement model (i.e., second-order model) is shown in Figure 2. Here, too, one path between the indicators and each first-order construct was set as 1. We repeated this procedure for the relationships between the first-order constructs and the second-order constructs. We also used four global reflective items to measure each second-order construct. The global items were developed based on the second-order construct definitions that were discussed earlier in step 1. In order to avoid overburdening raters, the global items were not included in the content validity assessment outlined in the previous section. All remaining steps of the instrument development procedure were followed to

Table 7. Global Items Used to Directly Measure Each Second-Order Construct

Second-Order Construct	Item Used
Application design (DES1-4)	Overall, I think *the mobile application* is designed well.
	In general, I believe that *the mobile application* has a great design.
	Generally speaking, *the mobile application* is well designed.
	I am very satisfied with the overall design of *the mobile application*.
Application utility (PURP1-4)	To me, *the mobile application* is very functional.
	Overall, I think that *the mobile application* is useful.
	Generally speaking, *the mobile application* serves its purpose well.
	In general, *the mobile application* is of value to me.
User interface graphics (INTG1-4)	Overall, I think that the graphics displayed on *the mobile application* are designed effectively.
	In general, the interface graphics of *the mobile application* are designed well.
	Generally speaking, I like the graphics displayed on the interface of *the mobile application*.
	Overall, *the mobile application* has very good user interface graphics.
User interface input (INP1-4)	In general, *the mobile application* allows me to input data easily.
	Overall, the user input mechanisms are designed effectively on *the mobile application*.
	I am very satisfied with the input mechanisms of *the mobile application*.
	Generally speaking, it is easy to type in data into *the mobile application*.
User interface output (CONT1-4)	In general, the content of *the mobile application* is presented effectively.
	Overall, I believe that *the mobile application* presents contents very well.
	Overall, I think that *the mobile application* presents content effectively.
	I am very satisfied with the way that *the mobile application* presents content.
User interface structure (STRU1-4)	Overall, I think *the mobile application* structures information effectively.
	In general, *the mobile application* is structured very well.
	I am very satisfied with the way *the mobile application* is structured.
	Generally speaking, *the mobile application* is structured nicely.

develop the global items as outlined above. Specifically, we developed items that represented the second-order construct definitions discussed earlier. We also searched for relevant literature that might have examined similar constructs in order to inform our item development process. In order to evaluate the wording of the global items, we conducted a face validity check involving the same six participants that supported us in evaluating the wording of the first-order constructs. The face validity check was conducted using the same procedure as for the first-order constructs explained earlier. Flagged items that were identified as too vague were deleted or their wording was modified. All global items used to measure the higher-order constructs are shown in Table 7.

Step 5. Conduct Pretest Data Collection

Once the measurement model is formally specified, the next step of the scale development is the pretest of the survey instrument. The psychometric properties, including the convergent, discriminant, and nomological validity, of the scales should be investigated (MacKenzie et al. 2011; Straub 1989; Straub et al. 2004). To evaluate nomological validity of the focal construct, it is important to collect data for constructs that are theoretically related (MacKenzie et al. 2011).

For the pretest, we created a survey including instructions for the participants and the items developed as discussed in step 2 earlier. All items were measured using a seven-point Likert-agreement scale (1 = strongly disagree; 7 = strongly agree). Before collecting data from a large sample, we asked two individuals to read the instructions and provide feedback on the items and survey structure. Both individuals were administrative staff at an Australian university. They read the instructions that accompanied the survey and also completed the survey. Only very minor changes (e.g., pagination) were proposed by these two individuals, and they confirmed that the instructions were clear and easy to follow.

In order to collect data from a wider audience, we surveyed a different set of U.S. consumers who were also recruited by the same market research firm that we had used earlier. Due to the exploratory nature of this phase, we aimed to collect 500 responses that were necessary to investigate the psychometric properties of the scales (Hair et al. 1998; MacCallum et al. 1999). Given that our survey instrument included 102 newly developed items, our ratio of items to responses was in range specified by MacKenzie et al. Due to the popularity of social media sites among mobile smartphone users (Hampton et al. 2011), we tailored all questions to a mobile social media application, such as Facebook. This approach of tailoring the questions to a specific application, as noted earlier, is consistent with extant research (e.g., Adipat et al. 2011; Devaraj et al. 2002, 2006; Hess et al. 2009; Hong et al. 2004a, 2004c; Thong et al. 2002; Venkatesh and Ramesh 2006; Venkatesh, Ramesh, and Massey 2003; Wells, Parboteeah, and Valacich 2011; Wells, Valacich, and Hess 2011). At the beginning of the survey, we asked two qualifying questions: (1) What is the frequency with which you access mobile social media applications via your smartphone? (2) Which mobile social media application do you use most often via your smartphone? Respondents who did not use mobile social media applications at all were disqualified from participating in the survey and we did not collect data from these individuals. Similarly, we excluded those potential respondents who did not use one of the following mobile social media applications: Facebook, Google+, LinkedIn, MySpace, or Twitter. We pursued this sampling strategy for three reasons. First, these mobile social media applications are very popular and successful, and we felt that this would allow us to understand the usability factors that can truly contribute to the outcome variables. We felt that the findings of our work should be particularly informative for application designer with less extensive funding opportunities. Second, we felt asking questions regarding a specific mobile social media application was necessary to ensure that all respondents had well-established thoughts about a specific social media application so that they could meaningfully respond to items related to the mobile application usability of a specific social media application. Third, these mobile applications pursue similar business strategies and they provide comparable value propositions to users (e.g., allowing users to create and share user-generated content). This was particularly useful because it allowed us to group the responses in the data analyses discussed below. We felt this was reasonable because the usability principles should not differ at a theoretical level. During the data collection, the online survey was programmed to carry forward the mobile social media application response that each individual provided. For instance, rather than the item being stated as "the mobile social media application allows you to connect with other people," the item was displayed as "Facebook (mobile) allows you to connect with other people."

In total, 500 responses were collected for the pretest. As with the content validity check, all responses were scrutinized for the duration that the respondents took to complete the survey. Those respondents who took too little time and/or did not correctly answer reverse-coded filler items were excluded from the sample. We applied a cut-off threshold of 8 minutes to complete the survey including all items of our instrument (113 items as shown in Appendix C, Table 7, and Table 8) and excluded those responses that were completed in less than 8 minutes. We felt that a response ratio of more than 15 questions per minute would indicate that the survey taker did not pay adequate attention to the questions. This led to 440 usable responses. Appendix A provides information on the respondent demographics. As with the content validity check, we felt that nonresponse bias was not a concern because the sample matched the profile of the sampling frame provided by the market research firm. We also felt that comparing early versus late responses was not useful because all responses were collected during a single weekend and no reminders were employed (Churchill 1979; Hair et al. 1998).

As indicated above, assessing the nomological network of focal constructs is an important part of the scale development procedure. Bagozzi (1980, p. 129) argues that we

> must also consider the relationship of the concept under investigation to other concepts in an overall context of a theoretical structure. This will… involve the use of syntactic criteria in combination with the modeling of theoretical and empirical relationships.

We followed Bagozzi's recommendation and explored the nomological network of mobile application usability. In order to assess the nomological validity of our proposed mobile application usability conceptualization, we collected data on two dependent variables, namely continued intention to use and mobile application loyalty.

Intention is a commonly studied dependent variable in IS research (Venkatesh, Morris et al. 2003). Among those with experience with a technology, continued intention to use is an appropriate dependent variable (Bhattacherjee 2001; Brown et al. 2012, 2014). We define continued intention as the degree to which a user feels he or she will keep using a mobile application (adapted from Bhattacherjee 2001). There are theoretical reasons and empirical evidence that usability is a key predictor for individuals' ongoing use of technological applications (Agarwal and Venkatesh 2002; Devaraj et al. 2002, 2006; Hong et al. 2004a; Venkatesh and Agarwal 2006; Venkatesh and Ramesh 2006; Wells, Valacich, and Hess 2011). Brand loyalty is widely used in marketing research and we felt that examining the impact of usability on an outcome of such organizational significance would help us in validating our conceptualization and scales and furthering criterion validity. We adapted the concept of brand loyalty to

Table 8. Scales Used to Measure the Outcome Variables		
Outcome Variable	**Items Used**	**Scales Adapted From**
Continued intention to use	I intend to continue using *the mobile application*.	Bhattacherjee 2001; Venkatesh and Goyal 2010
	I want to continue using *the mobile application* rather than discontinue.	
	I predict I will continue using *the mobile application*.	
	I plan to continue using *the mobile application*.	
	I don't intend to continue using *the mobile application* in future.	
	Chances are high that I will continue using *the mobile application* in future.	
Mobile application loyalty	I encourage friends and relatives to be the customers of *the mobile application*.	Johnson et al. 2006
	I say positive things about *the mobile application* to other people.	
	I will use more services offered by *the mobile application* in the next few years.	
	I would recommend *the mobile application* to someone who seeks my advice.	
	I consider *the mobile application* to be my first choice.	

Note: Mobile application was replaced by the specific brand of the social medial application (e.g., Facebook).

the context of mobile applications. We define mobile application loyalty as the degree to which a user has a deeply held commitment to rebuy or repatronize a mobile application (adapted from Johnson et al. 2006). Much practitioner-based research suggests that mobile application loyalty is a consequence of the overall usability of a given application (Deloitte 2012; Forrester Research 2011; Youens 2011). For instance, Gartner (2012) argues that mobile application loyalty is ultimately driven by user experience and the usability of the mobile application. Further, we found a considerable amount of research on website usability that established a theoretical link between web application usability and web application loyalty (Casalo et al. 2008; Cyr 2008; Cyr et al. 2006; Flavian et al. 2006). Therefore, we felt it is important to examine a broader outcome in conjunction with mobile application usability. Further, MacKenzie et al. argue

> Additional research on the construct often results in an expansion of the nomological network to include…consequences of the focal construct…learning more about a theoretical construct is a matter of elaborating the nomological network in which it occur, or of increasing the definiteness of the components (pp. 320-321).

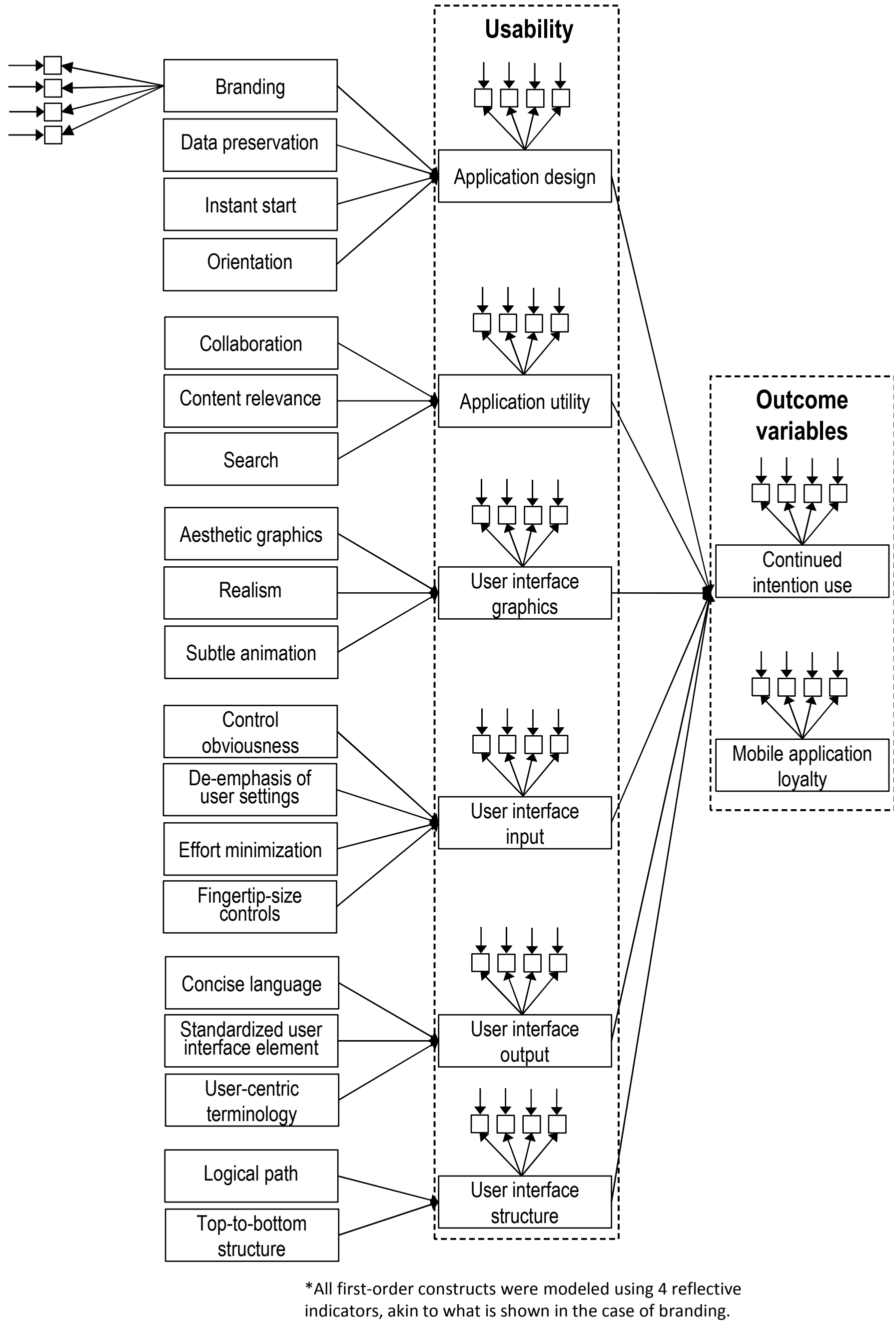
Thus, we decided to expand the nomological network of mobile application usability and examine whether mobile application usability is a predictor of consumers' mobile application loyalty. We adapted a previously validated scale to measure continued intention to use (Bhattacherjee 2001; Venkatesh and Goyal 2010). We measured mobile application loyalty by adapting a previously validated scale of brand loyalty (Johnson et al. 2006). Table 8 lists the items. Figure 3 displays the structural model.

Step 6. Scale Purification and Refinement

With the pretest data, MacKenzie et al. suggest purification and refinement of the survey instrument. This course of action involves a set of statistical tests to evaluate the measurement properties of the scales using the pretest data. The course of action that should be performed for evaluating the measurement model requires: (1) assessing goodness of fit; (2) assessing the validity of the set of indicators at the construct level; (3) assessing the reliability of the set of indicators at the construct level; (4) assessing individual indicator validity and reliability; and (5) eliminating weak indicators.

We used SPSS and AMOS to perform the statistical tests reported in this and subsequent sections. We compared the two alternative measurement models discussed in step 4: first-order versus second-order models. Consistent with prior research, we used the χ^2-difference test (see Tanriverdi 2005; Wallace et al. 2004) and comparative model fit. The results suggested that the model including the second-order constructs had a slightly lower χ^2 but was not statistically significantly different from the first-order model. However, the parsimony in predicting variables that comes with the second-order model caused us to favor this model. Further, all of the other fit statistics were better for the second-order model and in the acceptable range (see Hair et al. 1998; Straub et al. 2004). In fact, three of the six fit statistics were below the acceptable levels for the first-order measurement model. Next, the procedures outlined by MacKenzie et al. were followed to purify and refine the scales.

Appendix D includes the goodness of fit statistics of the measurement model. Overall, the goodness of fit statistics were well in line with the recommended cutoff values (Hair



Figue 3. Structural Model

et al. 1998; MacKenzie et al. 2011; Straub et al. 2004), thus supporting the validity of the measurement model. We then assessed the validity of the indicators at the construct level by examining the average variance extracted (AVE) of all first-order constructs. The results confirmed that all AVEs were in excess of .70, which is higher than the recommended threshold of .50 (MacKenzie et al. 2011). We also examined the reliability of the scales using Cronbach's alpha. All Cronbach's alpha values ranged from .75 to .85 and were above the desired threshold of .70. These results are shown in Appendix E. Following MacKenzie et al., we assessed the unique proportion of variance in the second-order construct accounted for by each first-order construct and Fornell and Larcker's (1981) construct reliability index. Appendix F shows these results. Most first-order constructs explained a considerable amount of variance in the respective second-order constructs. Notably, content relevance explained 35 percent of unique variance in application utility. All construct reliability index scores were above .77 and higher than the recommended threshold of .70 (Fornell and Larcker 1981).

Next, we inspected the item loadings for the reflectively measured first- and second-order constructs, and assessed the weights of each first-order construct on the respective second-order construct. Appendix G shows these results. All items loaded highly on the intended constructs, with item-to-construct loadings between .65 and .93 and only four loadings lower than .70 (UCT1, TTPS3, DUS4, SANM1), thus supporting convergent validity. The weights of the first-order formative constructs on the respective second-order constructs were significant ($p < .001$ in all cases), thus indicating that each first-order construct contributed substantially to the corresponding second-order construct. These findings were positive because a well-distributed weight structure is desired when using formatively measured constructs (Cenfetelli and Bassellier 2009; Diamantopoulos et al. 2008; Petter et al. 2007). We also examined the variance inflation factors (VIFs) for all first-order constructs forming the second-order constructs. All VIF values were less than 4, indicating that multicollinearity was not a major issue (see Petter et al. 2007). The bottom part of Appendix G shows the results of the item loadings obtained for the reflectively measured second-order constructs. As discussed earlier in step 4, we used four global items to identify each second-order construct. The findings confirmed that the items loaded highly on the intended second-order construct, with loadings ranging between .70 and .93.

The structural model results are shown in Table 9. Usability explained 42 percent of variance in continued intention to use. The R^2 was lower for mobile application loyalty (17%). Although encouraging that usability predicted mobile application loyalty, the lower variance explained is to be expected given that the determinants of loyalty are manifold (e.g., affective commitment) and go beyond the usability of mobile applications (Johnson et al. 2006). All paths between the second-order constructs and continued intention to use were significant. Most influential was application design (.36), followed by application utility (.28) as well as user interface graphics (.28). Three of the six second-order constructs were significant predictors of mobile application loyalty, with application design and application utility being the strongest, and user interface graphics having a weak effect. Due to these promising findings, we did not feel that any items needed to be removed.

Step 7. Conduct New Sample Data Collection (Study 1)

Once the scales are pretested, refined, and problematic indicators are eliminated, data should be collected from a new sample in order to reexamine the purified scales. This is an essential step that is frequently ignored in the IS field (MacKenzie et al. 2011).

Similar to the pretest, we collected data from a new sample consisting of U.S. consumers using mobile social media applications. To ensure that we collected data from a new sample, we hired a different market research firm from the one used for the content validity check and pretest. The new firm conducted the data collection using the same survey instrument that we used for the pretest. Following the procedure outlined in step 5, we screened respondents who did not qualify based on their usage patterns, time spent on the survey, and response to reverse-coded items. In total, 408 usable responses were collected. Appendix H provides the information on the respondent demographics. As in the case of the pretest, we did not compare early versus late responses because all responses were collected during a single weekend and no reminders were employed (Churchill 1979; Hair et al. 1998). The sample was well distributed across income, jobs, and age groups. Men were over-represented but the sample corresponded to the sampling frame provided by the market research firm.

Step 8. Assess Scale Validity

The next step is to assess scales validity (MacKenzie et al. 2011). A major goal during this step is to evaluate if the items used to assess the focal construct are

> (1) accurate representations of the underlying construct (through experimental manipulation or comparing groups known to differ on the construct), (2) adequately capture the multidimensional nature of the construct, (3) are distinguishable from the indicators of other constructs (discriminant validity), and (4) are related to the measures of other constructs specified in the construct's theoretical network (nomological validity)" (MacKenzie et al. 2011, p. 317).

MacKenzie et al. suggested that experimental manipulation checks and video tape observations work well if benchmarking different outcome scenarios (e.g., performance measures). In contrast, they noted that individuals' internal states (e.g., attitudes, anger, frustration, values) would be less suited to experimental manipulation. Given that prior research has

Table 9. Pretest: Structural Model Results

	Continued Intention to Use	Mobile Application Loyalty
R^2	.42	.17
Application design	.36***	.28***
Application utility	.28***	.25*
User interface graphics	.28***	.13*
User interface input	.15*	.05
User interface output	.22***	.04
User interface structure	.24***	.07

*p < 0.05, **p < 0.01, and ***p < 0.001

pointed out that usability is a subjective judgment made by users (e.g., Devaraj et al. 2002; Devaraj et al. 2006; Hess et al. 2009; Hong et al. 2004a, 2004c; Thong et al. 2002; Wells Parboteeah, and Valacich 2011; Wells, Valacich, and Hess 2011), we concluded that an experimental manipulation was impractical to validate our newly developed instrument. Also, a literature review did not reveal any meaningful theoretical patterns across groups that could be used to validate the scales. Moreover, given the number of constructs being examined here and the length of the paper even as it stands now, we felt that we could not develop hypotheses comparing different groups in this paper.

Instead, we felt that benchmarking our new conceptualization and instrument against the usability conceptualization and instrument based on MUG (Agarwal and Venkatesh 2002; Venkatesh and Agarwal 2006; Venkatesh and Ramesh 2006) would further our case in support of nomological validity. We use the MUG scales for our work as an illustration only and we do not wish to critique the instrument but use it for the following reasons. First, the MUG scales were drawn from usability guidelines developed for websites (see Agarwal and Venkatesh 2002) and it is likely that critical contextual factors pertaining to mobile application usability were not included as a result of this. Second, the MUG scales were the most comprehensive conceptualization of application usability that was previously validated in a field study focusing on mobile application usability. Third, we used Google Scholar and found that the MUG studies (i.e., Agarwal and Venkatesh 2002; Venkatesh and Agarwal 2006; Venkatesh and Ramesh 2006; Venkatesh, Ramesh, and Massey 2003), were cited in excess of 1,000 times. This provided us with confidence that the instruments were well accepted. Given that we could not compare the model across groups, we felt that such a benchmarking was a good alternative. For the competing model, we adapted the scales employed by Venkatesh and Ramesh (2006), shown in Appendix I. We initially used the new sample to assess the fit of the measurement model. As listed in Appendix D, all fit indexes met the recommended thresholds, thus indicating a good model fit. Appendix J shows that the AVEs were all above .70, which is above the recommended cutoff value. The results confirmed that the AVEs for each construct exceeded the squared correlation of the construct with other constructs (Fornell and Larcker 1981). Appendix J also shows that the reliabilities, assessed using Cronbach's alphas, ranged from .70 to .83 for all scales and were above the threshold of .70. Following MacKenzie et al., we also computed the unique proportion of variance that each first-order construct explained in the associated second-order construct. Appendix K shows that most first-order constructs explained a significant proportion of variance in the second-order constructs. Similar to the pretest, content relevance explained more than 30 percent of unique variance in application utility. The reliability scores for all scales exceeded .77 and were higher than the recommended threshold of .70 (Fornell and Larcker 1981).

Next, we inspected the item loadings for the reflectively measured first-order constructs. We also assessed the weights of each first-order construct on the respective second-order construct. Appendix L shows these results. All item loadings ranged between .62 and .91. Only four items were lower than .70 (SANM1, DUS4, UCT1, and TTPS3). These findings indicated a high level of convergent validity. The weights for the formative constructs were all significant and ranged between .24 and .58, thus indicating that each first-order construct contributed substantially to the corresponding second-order construct. Further, all VIFs were less than 4, indicating that multicollinearity was not an issue (Petter et al. 2007). Also, the item loadings for the reflective indicators of the second-order constructs exceeded .70.

Next, we tested the structural model using the two outcome variables: continued intention to use and mobile application loyalty. Table 10 shows these results. Usability explained 41 percent of variance in the continued intention to use construct. Consistent with the pretest results (Table 9), the R^2 was lower in explaining mobile application loyalty (16%). All structural paths from the second-order constructs to continued intention to use were significant and the path coefficients ranged be-

Table 10. Study 1: Structural Model Results

	Continued Intention to Use	Mobile Application Loyalty
R^2	.41	.16
Application design	.35***	.30***
Application utility	.33***	.26***
User interface graphics	.30***	.14*
User interface input	.19**	.08
User interface output	.24***	.02
User interface structure	.23***	.05
MUG		
R^2	.28	.13
Content	.30***	.25***
Ease of use	.20**	.05
Promotion	.10	.23***
Made-for-the-medium	.20**	.10
Emotion	.15*	.23***

*p < 0.05, **p < 0.01, and ***p < 0.001

tween .19 and .35. Comparing these results with the pretest, we concluded that the findings for the measurement and structural models were stable and consistent across the two samples. The MUG results are also shown in Table 10. In comparison to our model, MUG is far less predictive, with the variance in intention and loyalty being only 28 percent and 13 percent respectively.

Step 9. Cross-validation (Study 2)

The next step in the scale development was to cross-validate our results. This step is best performed using a new sample in order to assess the stability of the scales (MacKenzie et al. 2011).

We collected data from a new sample. The second sample comprised 412 U.S. consumers using mobile social media applications. We employed the same market research firm that we used for study 1. Care was taken not to invite any study 1 participants. Following steps 5 and 7, we excluded individuals who were not qualified to participate in the survey (e.g., being unfamiliar with mobile social media applications), those who spent too little time on the survey, and those who responded incorrectly to the reverse-coded items. We collected 412 usable responses. Appendix H shows the respondent demographics. As in the pretest and study 1, we felt comparing early versus late responses was not useful because all responses were collected during a single weekend and we did not send out reminders to invited participants (Churchill 1979; Hair et al. 1998). Also similar to the pretest and study 1, men were over-represented in this sample but this was in line with the sampling frame that we were provided. We followed the same approach as discussed for the pretest and study 1 to assess the psychometric properties of the scales.

Appendix D shows that all fit indexes were in line with recommended cutoff values, indicating that the measurement model fit was good. All AVEs ranged from .70 and .74, and exceeded the .50 threshold. The results confirmed that the AVEs for each construct exceeded the squared correlation of the construct with other constructs (Fornell and Larcker 1981). Further, all Cronbach's alphas ranged from .74 to .82 and were above the recommended .70 threshold. Appendix M shows these results.

Next, we examined the unique proportion of variance that each first-order construct explained in the associated second-order construct. The first-order constructs explained a significant proportion of variance in the second-order constructs. Appendix N shows these results. Fornell and Larcker's (1981) construct reliability scores ranged between .78 and .89, thus exceeding the recommended threshold of .70.

Appendix O shows that the item-to-construct loadings for the first-order constructs ranged between 69. and .91. Only the loading of SANM1 was lower than .70 and the majority of item loadings exceeded .80. All first-order construct weights were significant, and ranged between .24 and .60. This suggested that all first-order constructs contributed substantially to the intended second-order constructs. The VIFs of the formative first-order constructs were all lower than 4, thus suggesting multicollinearity was not a problem. Appendix O

Table 11. Study 2: Structural Model Results

	Continued Intention to Use	Mobile Application Loyalty
R^2	.47	.19
Application design	.39***	.34***
Application utility	.35***	.28***
User interface graphics	.31***	.13*
User interface input	.17**	.07
User interface output	.25***	.03
User interface structure	.25***	.06
MUG		
R^2	.26	.15
Content	.26***	.28***
Ease of use	.19**	.04
Promotion	.08	.21***
Made-for-the-medium	.21**	.07
Emotion	.13*	.25***

*p < 0.05, **p < 0.01, and ***p < 0.001

also shows that the global items used to measure the second-order constructs loaded cleanly onto the intended constructs and the loadings exceeded .70 in all cases.

The structural model results are shown in Table 11. Usability explained 47 percent of the variance in continued intention to use. All six second-order constructs were significant, with application design, application utility, and user interface graphics being the strongest predictors. Interestingly, these three strongest predictors of continued intention to use were the only significant predictors of mobile application loyalty. The pattern of results was thus highly consistent with the pretest and study 1. Table 11 also reports the results of the prediction using the MUG constructs. As in study 1, the variance explained by the MUG constructs was much lower.

Step 10. Norm Development

The final step includes developing norms for the new scales. This is an important step because it helps in interpreting the findings and guides future research (MacKenzie et al. 2011).

Generally, it is useful to estimate the distribution of the population of interest before developing sampling strategies (DeVellis 2011; MacKenzie et al. 2011). Hence, we studied the U.S. user population of social media sides published by each provider (e.g., Facebook) and found that younger individuals use social media applications. For instance, 56 percent of all U.S. Facebook users are between 18 and 35 years old (Insidefacebook 2011). Therefore, our survey strategy matched the population of social media applications and we felt it was less meaningful to aim for a representative sample of the entire U.S. population including a large proportion of older adults. Another consideration for choosing the most suitable sample was our resource constraints. Representative samples are difficult to obtain and we instructed the market research firms to provide a sampling frame that matched the user statistics on social media users in the United States. As recommended by MacKenzie et al., we also carefully inspected all four samples for the shape of the distribution and examined the skewness and kurtosis; we found no significant issues.

Another important aspect when developing norms for scales is the sample size. The sample should be large enough to conclude that the scales are truly stable (Lee and Baskerville 2003; MacKenzie et al. 2011). Beyond the content validity check, we validated our scales using three independent samples, each comprising more than 400 individuals. We hired two different market research firms to conduct the data collection. In total, we surveyed more than 1,200 individuals who use mobile social media applications and felt this was a reasonably large sample to draw meaningful conclusions.

When developing norms for new scales, it is important to consider that the scales could vary across research contexts and time (Ancona et al. 2001; Johns 2006; MacKenzie et al. 2011). Johns (2006) specifically noted that a particular context of study can lead to different results, such as relationships going from significant to nonsignificant. Others have argued that the concept of time is important, and could impact the stability of theories and associated scales (Ancona et al. 2001; Harrison et al. 1998; MacKenzie et al. 2011). Due to practical constraints, we only tested our conceptualization and survey instrument in the context of mobile social media applications in the U.S. using cross-sectional data. Future studies could extend the scope of this work and develop further norms for our usability instrument using samples in other countries to understand cross-cultural differences and also by collecting longitudinal data to understand differences in the importance of different usability elements over time. Still, given our results, we believe it is reasonable to say that the scales are stable in the context of our work: current users of mobile social media applications in the United States.

Discussion

The current study developed and validated a mobile application usability conceptualization and survey instrument following the 10-step procedure proposed by MacKenzie et al. (2011). We adapted Apple's user experience guidelines for mobile applications to develop our conceptualization and instrument. We conceptualized and used 19 first-order constructs to form 6 second-order constructs. The scale development process included item generation and a content validity assessment using 318 U.S. consumers. Once we specified the measurement model, we collected three waves of data from over 1,200 U.S. consumers of social media applications. The scales were tested and found to be reliable and valid. The fit indexes of the proposed measurement model were good in all three samples. We also found that our conceptualization of mobile application usability was a good predictor of continued intention to use and mobile application loyalty, thus supporting the nomological and predictive validity of our scales. Finally, we benchmarked our model against a model based on Microsoft's usability guidelines (Venkatesh and Ramesh 2006), and found that our context-specific model explained more variance in continued intention to use and mobile application loyalty (see Hong et al. 2014). Based on our findings, we proposed norms to help future studies in applying our scales to different research environments.

Theoretical Contributions and Implications

Despite mobile application usability being an important emerging concept in IS, our literature review revealed that there is a lack of theoretical and methodological clarity on holistically evaluating mobile application usability. Our work addresses this issue by providing a context-specific and fine-grained conceptualization of mobile application usability and offers a reliable and valid instrument. This advances the existing body of knowledge in several ways.

First, our work highlights the importance of the user context in IS and HCI studies and provides an example for context-specific theorizing related to an IT artifact. Our literature review suggested that prior research primarily drew on scales developed for website usability for measuring mobile application usability. Although it is certainly reasonable to use conceptualizations and instruments for website usability as a starting point (e.g., Venkatesh and Ramesh 2006), such instruments may not capture important factors relevant to the mobile application usability context. To address this issue, we developed and validated several constructs that were unique to the mobile application context. For instance, the fingertip-size control construct comprised several items that focused on the button size of mobile applications. Apple's guidelines suggest that fingertip-size controls should be considered for the usability of mobile applications because users would input data via touch screen interfaces. Our results confirmed that the fingertip-size control construct substantially contributed to the user interface input construct. Likewise, web-based applications are commonly displayed on fixed computer screens. Therefore, it seems to be unnecessary for web-based applications to be displayed in horizontal and vertical modes. In contrast, mobile applications are displayed on smartphones and users often change the way they hold the smartphone when using a mobile application. Due to this, application orientation becomes critical in this context and our studies confirmed that the application orientation construct substantially contributed to the application design construct. By applying website usability instruments, these factors might remain undetected because the user interface principles of web-based applications fundamentally differ from their mobile counterparts (e.g., web-based applications are normally controlled using a mouse cursor rather than using one's fingertips in combination with touch screens). When benchmarking the MUG instrument against our instrument developed specifically for the mobile application usability context, we found that our instrument explained more variance in the outcome variables, thus highlighting the importance of context in theory development in the IS and HCI research. As such, our work follows several calls for developing more precise and context-specific theories (Alvesson and Kärreman 2007; Bamberger 2008; Brown et al. 2010; Hong et al. 2014; Johns 2006; Van der Heijden 2004; Venkatesh et al. 2012) because there is "a general tendency to seek causal explanations at lower rather than higher levels of analysis, a tactic referred to unflatteringly as explanatory reductionism" (Johns 2006, p. 403).

Second, given that our conceptualization and scales are a more comprehensive and accurate representation of mobile application usability compared to prior research, future research can use it. Our literature review indicated that researchers have often pursued a pick-and-choose strategy and combined various theoretical constructs to measure mobile application usability. We believe that such a strategy is no longer necessary because our instrument provides clarity in terms of the underlying constructs of the overall usability of mobile applications. Instead of combining theoretically unrelated constructs to measure mobile application usability, researchers could use our entire instrument or relevant parts (especially all or some of the second-order constructs) to investigate all specific aspects of mobile application usability. For example, HCI research has investigated the concept of user interfaces to a great extent (Hess et al. 2009; Hong et al. 2004b; Palmer 2003; Thong et al. 2002; Wells, Parboteeah, and Valacich 2011; Wells, Valacich, and Hess 2011). Our conceptualization includes constructs that could be used to study user interaction and interface design of mobile applications, namely interface input and user interface output. User interface input specifies how well a given mobile application allows users to input data easily. Four first-order constructs form user interface input (i.e., control obviousness, de-emphasis of user settings, effort minimization, and fingertip-size controls). User interface output specifies how well a given mobile application presents content effectively. Three first-order constructs were identified to assess user interface output (i.e., concise language, standardized user interface element, and user-centric terminology). Our results confirmed that each first-order construct significantly contributed to the associated second-order constructs. Therefore, HCI research could use our conceptualization and instrument to investigate how user interface design (including user interface input and output mechanisms) performs in terms of relating to dependent variables typically studied in HCI, such as response time and error rates. Our work also has implications that go beyond HCI research and we believe a large audience will be interested in leveraging our instrument. For instance, there is considerable interest in understanding the interplay between personality traits and HCI (Devaraj et al. 2008; Junglas et al. 2008; McElroy et al. 2007). We think that our instrument is an ideal candidate for exploring this relationship in more depth. Specifically, future research could leverage the second-order constructs and investigate the interaction effects between mobile application usability, personality traits (e.g., agreeableness, conscientiousness, extraversion, neuroticism, openness), and use of mobile applications. Such a study could reveal findings that contribute to a variety of streams in IS (e.g., HCI, IS adoption and use), psychology, and marketing.

Third, brand loyalty is frequently used in the marketing literature to explain why consumers are committed to a brand or a company (Brakus et al. 2009; Johnson et al. 2006; Melnyk et al. 2009; Wagner et al. 2009). Traditionally, marketing research has paid less attention to technocentric topics and studies have explained brand loyalty as a function of affective commitment, perceived value, and brand equity (e.g., Johnson et al. 2006). We adapted the concept of brand loyalty to mobile applications and explored the nomological network of mobile application usability. We found empirical evidence that mobile application loyalty was significantly influenced by usability. Our mobile application usability conceptualization explained about 15 percent of the variance in mobile application loyalty in each of the three samples. Due to practical constraints in our work and our focus on mobile application usability, we did not include other possible predictors of mobile application loyalty. Future research should certainly conduct such an investigation with the goal of comparing the previously known predictors to a model based on usability, followed by a study with an eye toward integrating this varied set of predictors into a cohesive whole.

Fourth, we believe that our survey development exemplar contributes to measurement theory. MacKenzie et al. (2011) provided comprehensive instrument validation guidelines for IS in particular and behavioral and social sciences research in general. We applied their recommendations and inform future research aiming to develop instruments. Overall, we did not encounter major issues by following MacKenzie et al. In some steps, we had to deviate from their recommendations due to practical considerations. For example, during the content validity check, we initially followed their preferred Hinkin and Tracey's (1999) content validity assessment approach. Given the feedback we obtained during the face validity check, we decided to employ Anderson and Gerbing's (1991) approach because we felt that ranking all item-construct pairs would overburden raters. In examining studies that applied Hinkin and Tracey's approach, we found it to be used almost exclusively when there were fewer than 50 items and 4 to 6 construct definitions (e.g., Yao et al. 2007). Although we did not conduct a group comparison (e.g., mobile social media application user versus non-user) because developing the theoretical bases for expected differences was beyond the scope of this paper, we believe that other researchers will encounter such practical constraints as well. Instead, we benchmarked our new constructs and scales against a conceptualization and instrument based on Microsoft's usability guidelines. Future work could also consider this option. It is important to note that during all stages of the instrument development process, we asked participants to provide us feedback and this was generally positive, thus adding to the credibility of our scales, the validity of the overall procedure proposed by MacKenzie et al., and the changes we made to their procedure. We also believe that, although MacKenzie et al. provide an excellent step-by-step guide for validating instruments, the item generation process is not discussed in depth due to the focus of their paper. Therefore, we believe that we complement the work of MacKenzie et al. in that we exemplify how to leverage coding procedures during the item generation step. Future research aiming to develop instruments could learn from our work and codify alternative information sources, such as qualitative interview transcripts for capturing individuals' views on technological artifacts.

Finally, our mobile application usability conceptualization and instrument can be used as a springboard for future research. As noted earlier related to step 10, researchers could replicate our work in new contexts or test the stability of the scales over time (Ancona et al. 2001; Johns 2006). For example, studies could investigate existing mobile applications besides social media applications, such as mobile news, mobile marketing, and mobile entertainment. Testing our conceptualization of mobile application usability and associated instrument in a new context is critical in order to see if the conceptualization is comparable in terms of predicting the outcomes in the context of interest. If our conceptualization predicts the outcomes fairly well in alternative contexts, this would suggest that the context would not include attributes that affect our conceptualization of mobile application usability (see Alvesson and Kärreman 2007). Future studies could also use our conceptualization and instrument in combination with other theories, such as IS continuance theory (Bhattacherjee 2001), IS success model (DeLone and McLean 1992), unified theory of acceptance and use of technology (Venkatesh, Morris et al. 2003), and task-technology fit (Goodhue and Thompson 1995), to study why individuals use mobile applications. Such theoretically motivated studies could also explain mobile application use in mobile commerce environments and inter- and intra-organizational information sharing facilitated by mobile information applications (Kohli 2007). Our work can also serve as a starting point for design science studies in the context of consumer adoption and use of mobile applications. For example, it would be interesting to investigate how users differentiate usability of mobile applications on different mobile device formats including iPhones and iPads (or alternative brands, such as Samsung's Galaxy phone versus its tablet counterpart). Such studies could leverage our mobile application usability conceptualization and explore if mobile device usability conceptually differs from mobile application usability. The findings would be particularly useful for organizations offering mobile applications on various mobile devices. Further, much practitioner-based literature has recognized that there is a push toward touch-based applications for new laptop generations running Windows 8 (Ovide 2013). For example, *The Wall Street Journal* recently reported that over 100,000 applications soon would be available for Microsoft's new Windows 8 operating system that users can operate via touch-based interfaces (Ovide 2013). Recent reports suggested that users are disappointed with the usability of the operating system. For instance, the Nielsen Norman Group (Nielsen 2012b) evaluated the usability of Windows 8 and reports several usability problems from a user perspective including cognitive overload triggered by double desktops and error prone gestures. Given this, it would be interesting to measure how individuals rate such applications in terms of the usability criteria we identified (e.g., fingertip-sized controls). Finally, it is worthwhile to note that our sampling strategy focused exclusively on mobile application users who had prior experience with mobile social media applications. Future studies should examine the impact of user experience on mobile application usability at the individual level.

Limitations and Future Research

One limitation concerns the self-reported measure of continued intention to use. As noted by Burton-Jones and Straub (2006), technology use has been conceptualized and measured variously, including as breath of use (e.g., Saga and Zmud 1994), variety of use (e.g., Igbaria et al. 1997), and extent of use (e.g., Venkatesh and Davis 2000). The same principles apply to studies using continued intention to use as a proxy for individuals' ongoing technology use behavior. By adapting previously validated scales from Bhattacherjee (2001) and Venkatesh and Goyal (2010), we purposefully focused on the extent of their use. Both papers have been cited in excess of 1,500 times, thereby underscoring their impact. By aligning our conceptualization of the continued intention to use construct with these papers, we expect that our findings speak to a broad audience in the IS community. However, being aware of the variety of conceptualizations of technology use in IS research (see Burton-Jones and Straub 2006; Sykes and Venkatesh 2015), we note that the generalizability of our findings is constrained by the limitations of the continued intention to use construct.

All items used in our work did not focus on explaining structural elements of use (see Burton-Jones and Straub 2006) as we did not ask users about particular tasks they perform on mobile social media applications. One of our major goals was to explain why individuals continue to use mobile social media applications in general. We suggest that future studies adapt our scales to specific tasks that users perform on mobile applications, such as chatting with friends or posting news on mobile social media applications, chatting with friends on mobile instant messaging applications, or transferring money on mobile banking applications. A potential starting point for such studies could be an augmented task–technology fit model (Goodhue and Thompson 1995) that evaluates the fit between tasks (e.g., chatting with friends) that users' perform on mobile applications and technological artifacts (e.g., mobile social media applications) used to perform the task. It is also important to acknowledge that smartphones and associated applications continue to develop and user interface design will adapt to the technological progress. As a consequence of these developments, mobile application vendors, including Apple, will likely release updated versions of their user interface guidelines. Future research on mobile application usability should monitor these developments and, if necessary, develop instruments based on updated interface guidelines.

Practical Implications

Our findings have important implications for practitioners because we identified critical usability elements in the context of mobile applications. Our instrument could be used by practitioners to study to-be-developed, to-be-implemented, and existing mobile applications. For example, companies could use our mobile application usability conceptualization and associated instrument in all phases of the system development life cycle including the investigation phase, analysis and design phase, application implementation phase and application maintenance stage.

During the investigation phase, when practitioners determine the scope of the application and conduct a feasibility analysis, our coding matrix could be used as a checklist for the required usability elements with which a given mobile application needs to conform. Our codes are a comprehensive reflection of Apple's guidelines and the coding matrix could be particularly useful in situations where multiple team members discuss the mobile application scope and conduct a feasibility analysis. For example, the codes derived for the concise language first-order construct emphasize precise application feature descriptions. The mobile application development team might decide to consult editors in order to ensure that the text written for the application is precise and adequate. Other codes, including subtle animation or aesthetic graphics, indicate that it might be useful to consult with graphic designers regarding the integration of these elements. Many other usability elements could be discussed with system developers (e.g., interface structure, control obviousness, fingertip-size controls) as they will be able to estimate the time and costs associated with developing the application.

During the analysis and design phase, our instrument could be used to acquire information about the most desirable features of a to-be-developed mobile application. If designing mobile applications from scratch, it might be useful to adapt our scales and survey potential users regarding their expectations related to the most important usability elements of the mobile application. Another way to use our instrument will be to design a prototype and invite potential users to test it. Once potential users have tested the prototype, our instrument could be used to obtain feedback from them. Given the iterative nature of the analysis and design phase, the feedback obtained could be incorporated into the next application prototype and the survey administration could be repeated.

Given that the application implementation phase is mostly technical in nature, our instrument may be less useful there but we believe that our instrument will be particularly valuable during the application maintenance and review stages. For instance, during the maintenance phase, user requests and complaints could be used and checked against our coding matrix. If the complaints deviate from the usability requirements listed in the matrix, practitioners could use the comprehensive usability guidelines offered by vendors and develop strategies to overcome these problems. Bug reports might be an alternative source for identifying potential weaknesses and they could also be benchmarked against the codes identified in our work. During the review stage of the application development lifecycle, our comprehensive mobile application usability instrument could be leveraged for detailed usability testing in a laboratory research environment or field studies. The findings would reveal how developers can fine-tune their applications in terms of continuous improvement. The findings could then be used to evaluate the usability of the existing mobile applications and designers could take actions based on the criteria identified. Specifically, as demonstrated in this research, significant first-order constructs in combination with important second-order constructs can guide practitioners on how they can improve their current mobile applications.

Conclusions

Due to the widespread diffusion of mobile devices, consumers expect user-friendly and well-designed mobile applications from service providers in various industries. So far, little systematic help has been offered to evaluate existing mobile applications or in designing new mobile applications. Our work developed a conceptualization and associated survey instrument based on Apple's general user experience guidelines that can aid such an endeavor. Our work also serves as an exemplar that uses the procedure proposed by MacKenzie et al. (2011), albeit with some modifications that we found to be necessary. The conceptualization and instrument of mobile application usability is an important contribution for IS and HCI research because it helps theory development in various research areas, such as mobile technology adoption research, mobile user interface evaluation, and mobile application development.

Acknowledgments

We gratefully thank the senior editor, Rajiv Kohli, for his support and guidance throughout the review process. We would also like to thank the associate editor, John Wells, and the anonymous reviewers who provided valuable input that greatly helped to improve the paper. Last, but not least, we thank Jan DeGross for her efforts in typesetting this manuscript.

References

Adipat, B., Zhang, D., and Zhou, L. 2011. "The Effect of Tree-View Based Presentation Adaptation on Mobile Web Browsing" *MIS Quarterly* (35:1), pp. 99-121.

Agarwal, R., and Venkatesh, V. 2002. "Assessing a Firm's Web Presence: A Heuristic Evaluation Procedure for the Measurement of Usability," *Information Systems Research* (13:2), pp. 168-186.

Aladwani, A. M., and Palvia, P. C. 2002. "Developing and Validating an Instrument for Measuring User-Perceived Web Quality," *Information & Management* (39:6), pp. 467-476.

Alvesson, M., and Kärreman, D. 2007. "Constructing Mystery: Empirical Matters in Theory Development," *Academy of Management Review* (32:4), pp. 1265-1281.

Ancona, D. G., Okhuysen, G. A., and Perlow, L. A. 2001. "Taking Time to Integrate Temporal Research," *Academy of Management Review* (26:4), pp. 512-529.

Anderson, J. C., and Gerbing, D. W. 1991. "Predicting the Performance of Measures in a Confirmatory Factor Analysis with a Pretest Assessment of their Substantive Validities," *Journal of Applied Psychology* (76:5), pp. 732-740.

Apple. 2012. "iOS Developer Library—User Experience Guidelines" (available online at http://developer.apple.com/library/ios/documentation/userexperience/conceptual/mobilehig/MobileHIG.pdf; accessed June 14, 2012).

Bagozzi, R. P. 1980. *Causal Models in Marketing*, New York: John Wiley.

Bamberger, P. 2008. "From the Editors Beyond Contextualization: Using Context Theories to Narrow the Micro–Macro Gap in Management Research," *Academy of Management Journal* (51:5), pp. 839-846.

BBC News. 2010. "Over 5 Billion Mobile Phone Connections Worldwide," July 9 (available online at http://www.bbc.co.uk/news/10569081; accessed July 14, 2011).

Benbasat, I., and Barki, H. 2007. "Quo Vadis, TAM?," *Journal of the Association for Information Systems* (8:4), pp. 211-218.

Benbunan-Fich, R., and Benbunan, A. 2007. "Understanding User Behavior with New Mobile Applications," *The Journal of Strategic Information Systems* (16:4), pp. 393-412.

Bhattacherjee, A. 2001. "Understanding Information Systems Continuance: An Expectation-confirmation Model," *MIS Quarterly* (25:3), pp. 351-370.

Bollen, K. 2007. "Interpretational Confounding Is Due to Misspecification, Not to Type of Indicator: Comment on Howell, Breivik, and Wilcox," *Psychological Methods* (12:2), pp. 205-218.

Brakus, J. J., Schmitt, B. H., and Zarantonello, L. 2009. "Brand Experience: What Is It? How Is It Measured? Does It Affect Loyalty?," *Journal of Marketing* (73:5), pp. 52-68.

Brown, S. A., Dennis, A. R., and Venkatesh, V. 2010. "Predicting Collaboration Technology Use: Integrating Technology Adoption and Collaboration Research," *Journal of Management Information Systems* (27:2), pp. 9-54.

Brown, S. A., Venkatesh, V., and Goyal, S. 2012. "Expectation Confirmation in Technology Use," *Information Systems Research* (23:2), pp. 474-487.

Brown, S. A., Venkatesh, V., and Goyal, S. 2014. "Expectation Confirmation: A Test of Six Competing Models," *MIS Quarterly* (38:3), pp. 729-756.

Burigat, S., Chittaro, L., and Gabrielli, S. 2008. "Navigation Techniques for Small-Screen Devices: An Evaluation on Maps and Web Pages," *International Journal of Human-Computer Studies* (66:2), pp. 78-97.

Burt, R. S. 1976. "Interpretational Confounding of Unobserved Variables in Structural Equation Models," *Sociological Methods and Research* (5:1), pp. 3-52.

Burton-Jones, A., and Straub, D. W. 2006. "Reconceptualizing System Usage: An Approach and Empirical Test," *Information Systems Research* (17:3), pp. 228-246.

Campbell, D., Wells, J. D., and Valacich, J. S. 2009. "Understanding Online Customer Relationships: B2C Relationship Stage Theory," *AIS Transactions on Human–Computer Interaction* (1:4), pp. 108-132.

Casalo, L., Flavian, C., and Guinaliu, M. 2008. "The Role of Perceived Usability, Reputation, Satisfaction and Consumer Familiarity on the Website Loyalty Formation Process," *Computers in Human Behavior* (24:2), pp. 325-345.

Cenfetelli, R. T., and Bassellier, G. 2009. "Interpretation of Formative Measurement in Information Systems Research," *MIS Quarterly* (33:4), pp. 689-707.

Churchill, Jr., G. A. 1979. "A Paradigm for Developing Better Measures of Marketing Constructs," *Journal of Marketing Research* (16:1), pp. 64-73.

Compeau, D., Marcolin, B., Kelley, H., and Higgins, C. 2012. "Generalizability of Information Systems Research Using Student Subjects—A Reflection on Our Practices and Recommendations for Future Research," *Information Systems Research* (23:4), pp. 1093-1109.

Cyr, D. 2008. "Modeling Web Site Design Across Cultures: Relationships to Trust, Satisfaction, and E-loyalty," *Journal of Management Information Systems* (24:4), pp. 47-72.

Cyr, D., Head, M., and Ivanov, A. 2006. "Design Aesthetics Leading to M-Loyalty in Mobile Commerce," *Information & Management* (43:8), pp. 950-963.

Cyr, D., Head, M., Larios, H., and Pan, B. 2009. "Exploring Human Images in Website Design: A Multi-Method Approach," *MIS Quarterly* (33:3), pp. 539-566.

De Wulf, K., Schillewaert, N., Muylle, S., and Rangarajan, D. 2006. "The Role of Pleasure in Web Site Success," *Information & Management* (43:4), pp. 434-446.

Deloitte. 2012. "So Many Apps – So Little to Download" (available online at http://www.mondaq.com/x/192692/IT+internet/So+Many+Apps+So+Little+To+Download; accessed March 15, 2012).

DeLone, W. H., and McLean, E. R. 1992. "Information Systems Success: The Quest for the Dependent Variable," *Information Systems Research* (3:1), pp. 60-95.

Devaraj, S., Easley, R. F., and Crant, J. M. 2008. "How Does Personality Matter? Relating the Five-Factor Model to Technology Acceptance and Use," *Information Systems Research* (19:1), pp. 93-105.

Devaraj, S., Fan, M., and Kohli, R. 2002. "Antecedents of B2C Channel Satisfaction and Preference: Validating E-commerce Metrics," *Information Systems Research* (13:3), pp. 316-333.

Devaraj, S., Fan, M., and Kohli, R. 2006. "Examination of Online Channel Preference: Using the Structure–Conduct–Outcome Framework," *Decision Support Systems* (42:2), pp. 1089-1103.

DeVellis, R. 2011. *Scale Development: Theory and Applications* (3rd ed.), Thousand Oaks, CA: Sage Publications.

Diamantopoulos, A., Riefler, P., and Roth, K. 2008. "Advancing Formative Measurement Models," *Journal of Business Research* (61:12), pp. 1203-1218.

Dou, W., Lim, K. H., Su, C., Zhou, N., and Cui, N. 2010. "Brand Positioning Strategy Using Search Engine Marketing," *MIS Quarterly* (34:2), pp. 261-279.

Dredge, S. 2011. "Most Branded Apps Are a Flop Says Deloitte. But Why?" (available online at http://www.guardian.co.uk/technology/appsblog/2011/jul/11/branded-apps-flopping; accessed January 25, 2012).

Duh, H. B. L., Tan, G. C. B., and Chen, V. H. 2006. "Mobile Usability: Usability Evaluation for Mobile Device: A Comparison of Laboratory and Field Tests," in *Proceedings of the 8th Conference on Human–Computer Interaction with Mobile Devices and Services*, Helsinki, Finland, September 12-15, pp. 4-16.

Edwards, J. R. 2001. "Multidimensional Constructs in Organizational Behavior Research: An Integrative Analytical Framework," *Organizational Research Methods* (4:2), pp. 144-192.

Flavian, C., Guinaliu, M., and Gurrea, R. 2006. "The Role Played by Perceived Usability, Satisfaction and Consumer Trust on Website Loyalty," *Information & Management* (43:1), pp. 1-14.

Fornell, C., and Larcker, D. 1981. "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error," *Journal of Marketing Research* (18:1), pp. 39-50.

Forrester Research. 2011. "How Mature Is Your Mobile Strategy?" (available online at http://blogs.forrester.com/thomas_husson/10-10-19-how_mature_is_your_mobile_strategy; accessed March 15, 2012).

Galletta, D. F., Henry, R. M., McCoy, S., and Polak, P. 2006. "When the Wait Isn't so Bad: The Interacting Effects of Website Delay, Familiarity, and Breadth," *Information Systems Research* (17:1), pp. 20-37.

Gartner. 2011. "Gartner Says Worldwide Mobile Application Store Revenue Forecast to Surpass $15 Billion in 2011" (available online at http://www.gartner.com/it/page.jsp?id=1529214; accessed January 25, 2012).

Gartner. 2012. "Top 10 Consumer Mobile Applications for 2012" (available online at http://www.gartner.com/it/page.jsp?id=1230413; accessed March 24,2012).

Gebauer, J., Shaw, M. J., and Gribbins, M. L. 2010. "Task Technology Fit for Mobile Information Systems," *Journal of Information Technology* (25:3), pp. 259-272.

Gebauer, J., Shaw, M. J., and Subramanyam, R. 2007. "Once Built Well, They Might Come: A Study of Mobile E-Mail," College of Business Working Paper 07-0117, University of Illinois at Urbana-Champaign, last updated October 5, 2007.

Gebauer, J., Tang, Y., and Baimai, C. 2008. "User Requirements of Mobile Technology: Results from a Content Analysis of User Reviews," *Information Systems and e-Business Management* (6:4), pp. 361-384.

Geissler, G. L. 2001. "Building Customer Relationships Online: The Web Site Designers' Perspective," *Journal of Consumer Marketing* (18:6), pp. 488-502.

Goodhue, D. L., and Thompson, R. L. 1995. "Task–Technology Fit and Individual Performance," *MIS Quarterly* (19:2), pp. 213-236.

Hair, Jr., J. F., Anderson, R. E., Tatham, R. L., and Black, W. C. 1998. *Multivariate Data Analysis with Readings* (5th ed.), Englewood Cliffs, NJ: Prentice-Hall.

Hampton, K. N., Goulet, L. S., Rainie, L., and Purcell, K. 2011. "Social Networking Sites and Our Lives," PEW Research Internet Project, June 16 (available online at http://www.pewinternet.org/Reports/2011/Technology-and-social-networks.aspx; accessed August 18, 2011).

Harrison, D. A., Price, K. H., and Bell, M. P. 1998. "Beyond Relational Demography: Time and the Effects of Surface- and Deep-Level Diversity on Work Group Cohesion," *Academy of Management Journal* (41:1), pp. 96-107.

Hess, T. J., Fuller, M. A., and Campbell, D. 2009. "Designing Interfaces with Social Presence: Using Vividness and Extraversion to Create Social Recommendation Agents," *Journal of the Association for Information Systems* (10:12), pp. 889-919.

Hess, T. J., Fuller, M. A., and Mathew, J. 2005. "Involvement and Decision-Making Performance with a Decision Aid: The Influence of Social Multimedia, Gender, and Playfulness," *Journal of Management Information Systems* (22:3), pp. 15-54.

Hinkin, T. R., and Tracey, J. B. 1999. "An Analysis of Variance Approach to Content Validation," *Organizational Research Methods* (2:2), pp. 175-186.

Hong, S.-J. and Tam, K. Y. 2006. "Understanding the Adoption of Multipurpose Information Appliances: The Case of Mobile Data Services," *Information Systems Research* (17:2), pp. 162-179.

Hong, W., Chan, K. Y., Thong, J. Y. L., Chasalow, L, and Dhillon, G. 2014. "A Framework and Guidelines for Context-Specific Theorizing in Information Systems Research," *Information Systems Research*, (35:1), pp. 111-136.

Hong, W., Thong, J. Y. L., and Tam, K. Y. 2004a. "Does Animation Attract Online Users' Attention? The Effects of Flash on Information Search Performance and Perceptions," *Information Systems Research* (15:1), pp. 60-86.

Hong, W., Thong, J. Y. L., and Tam, K. Y. 2004b. "The Effects of Information Format and Shopping Task on Consumers' Online Shopping Behavior: A Cognitive Fit Perspective," *Journal of Management Information Systems* (21:3), pp. 151-188.

Hong, W., Thong, J. Y. L., and Tam, K. Y. 2004c. "Product Listing Pages: The Effects of Presentation Mode and Information Format on Consumers' Online Shopping Performance and Perceptions," *International Journal of Human–Computer Studies* (61:4), pp. 481-503.

Hong, W., Thong, J. Y. L., and Tam, K. Y. 2007. "How Do Web Users Respond to Non-Banner-Ads Animation? The Effects of

Task Type and User Experience," *Journal of the American Society for Information Science and Technology* (58:10), pp. 1467-1482.

Huang, S. C., Chou, I. F., and Bias, R. G. 2006. "Empirical Evaluation of a Popular Cellular Phone's Menu System: Theory Meets Practice," *Journal of Usability Studies* (1:2), pp. 91-108.

Huizingh, E. K. R. E. 2000. "The Content and Design of Web Sites: An Empirical Study," *Information & Management* (37:3), pp. 123-134.

Hummel, K. A., Hess, A., and Grill, T. 2008. "Environmental Context Sensing for Usability Evaluation in Mobile HCI by Means of Small Wireless Sensor Networks," in *Proceedings of the 6th International Conference on Advances in Mobile Computing and Multimedia*, Linz, Austria, November 24-26, pp. 302-306.

Hyvarinen, T., Kaikkonen, A., and Hiltunen, M. 2005. "Placing Links in Mobile Banking Applications," in *Proceedings of the 7th International Conference on Human–Computer Interaction with Mobile Devices & Services*, Salzburg, Austria, September 19-22, pp. 63-68.

Igbaria, M., Zinatelli, N., Cragg, P., and Cavaye, A. L. M. 1997. "Personal Computing Acceptance Factors in Small Firms: A Structural Equation Model," *MIS Quarterly* (21:3), pp. 279-305.

Insidefacebook. 2011. "US Facebook Users by Age and Gender" (available online at http://www.insidefacebook.com/2010/07/06/facebooks-june-2010-us-traffic-by-age-and-sex-users-aged-18-44-take-a-break-2/; accessed June 21, 2011).

iTunes. 2013. "Apple iTunes Store" (available online at http://www.apple.com/itunes/; accessed May 26, 2013).

Ji, Y. G., Park, J. H., Lee, C., and Yun, M. H. 2006. "A Usability Checklist for the Usability Evaluation of Mobile Phone User Interface," *International Journal of Human–Computer Interaction* (20:3), pp. 207-231.

Johns, G. 2006. "The Essential Impact of Context on Organizational Behavior," *Academy of Management Review* (31:2), pp. 386-408.

Johnson, M. D., Herrmann, A., and Huber, F. 2006. "The Evolution of Loyalty Intentions," *Journal of Marketing* (70:2), pp. 122-132.

Jokela, T., Koivumaa, J., Pirkola, J., Salminen, P., and Kantola, N. 2006. "Methods for Quantitative Usability Requirements: A Case Study on the Development of the User Interface of a Mobile Phone," *Personal and Ubiquitous Computing* (10:6), pp. 345-355.

Junglas, I. A., Johnson, N. A., and Spitzmüller, C. 2008. "Personality Traits and Concern for Privacy: An Empirical Study in the Context of Location-Based Services," *European Journal of Information Systems* (17:4), pp. 387-402.

Kaikkonen, A., Kallio, T., Kekäläinen, A., Kankainen, A., and Cankar, A. 2005. "Usability Testing of Mobile Applications: A Comparison Between Laboratory and Field Testing," *Journal of Usability Studies* (1:1), pp. 4-16.

Kakihara, M., and Sørensen, C. 2002. "Mobility: An Extended Perspective," in *Proceedings of the 35th Annual Hawaii International Conference on System Sciences*, January 3-6, Los Alamitos, CA: IEEE Computer Society Press, pp. 1756-1766.

Kang, S. H. 2007. "The Impact of Digital Iconic Realism on Anonymous Interactants' Mobile Phone Communication," in *Proceedings of Computer and Human Interaction*, *ACM Student Research Competition*, San Jose, CA, April 28–May 3, pp. 2207-2212.

Keeker, K. 1997. "Improving Web-Site Usability and Appeal: Guidelines Compiled by MSN Usability Research" (available online at http://msdn.microsoft.com/archive/default.asp?url=/archive/en-us/dnarplan/html/IMPROVINGSITEUSA.asp; accessed May 10, 2012).

Keen, P. 1980. "MIS Research: Reference Disciplines and a Cumulative Tradition," in *Proceedings of the First International Conference on Information Systems*, Philadelphia, December 8–10, pp. 9-18.

Kim, H., Kim, J., and Lee, Y. 2005. "An Empirical Study of Use Contexts in the Mobile Internet, Focusing on the Usability of Information Architecture," *Information Systems Frontiers* (7:2), pp. 175-186.

Kim, K., Proctor, R. W., and Salvendy, G. 2012. "The Relation Between Usability and Product Success in Cell Phones," *Behaviour & Information Technology* (31:10), pp. 969-982.

Kim, S., Lee, I., Lee, K., Jung, S., Park, J., Kim, Y. B., Kim, S. R., and Kim, J. 2010. "Mobile Web 2.0 with Multi-Display Buttons," *Communications of the ACM* (53:1), pp. 136-141.

Kim, S., and Stoel, L. 2004. "Apparel Retailers: Web Site Quality Dimensions and Satisfaction," *Journal of Retailing and Consumer Services* (11:2), pp. 109-117.

Kohli, R. 2007. "Innovating to Create IT-Based New Business Opportunities at United Parcel Service," *MIS Quarterly Executive* (6:4), pp. 199-210.

Kurniawan, S. 2008. "Older People and Mobile Phones: A Multi-Method Investigation," *International Journal of Human–Computer Studies* (66:12), pp. 889-901.

Lee, A. S., and Baskerville, R. L. 2003. "Generalizing Generalizability in Information Systems Research," *Information Systems Research* (14:3), pp. 221-243.

Lee, C. C., Cheng, H. K., and Cheng, H. H. 2007. "An Empirical Study of Mobile Commerce in Insurance Industry: Task–Technology Fit and Individual Differences," *Decision Support Systems* (43:1), pp. 95-110.

Lee, Y. E., and Benbasat, I. 2003. "A Framework for the Study of Customer Interface Design for Mobile Commerce," *International Journal of Electronic Commerce* (46:12), pp. 48-52.

Leung, R., McGrenere, J., and Graf, P. 2011. "Age-Related Differences in the Initial Usability of Mobile Device Icons," *Behaviour & Information Technology* (30:5), pp. 629-642

Lewis, B. R., Templeton, G. F., and Byrd, T. A. 2005. "A Methodology for Construct Development in MIS Research," *European Journal of Information Systems* (14:4), pp. 388-400.

Li, T., Van Heck, E., and Vervest, P. 2009. "Information Capability and Value Creation Strategy: Advancing Revenue Management through Mobile Ticketing Technologies," *European Journal of Information Systems* (18:1), pp. 38-51.

Li, X., Hess, T., McNab, A., and Yu, Y. 2009. "Culture and Acceptance of Global Web Sites: A Cross-Country Study of the Effects of National Cultural Values on Acceptance of a Personal Web Portal," *DATA BASE for Advances in Information Systems* (40:4), pp. 62-87.

Lim, K. H., Benbasat, I., and Ward, L. M. 2000. "The Role of Multimedia in Changing First Impression Bias," *Information Systems Research* (11:2), pp. 115-136.

Lowry, P. B., Vance, A., Moody, G., Beckman, B., and Read, A. 2008. "Explaining and Predicting the Impact of Branding Alliances and Web Site Quality on Initial Consumer Trust of E-Commerce Web Sites," *Journal of Management Information Systems* (24:4), pp. 199-224.

Lyons, K., Starner, T., and Gane, B. 2006. "Experimental Evaluations of the Twiddler One-Handed Chording Mobile Keyboard," *Human–Computer Interaction* (21:4), pp. 343-392.

MacCallum, R. C., and Browne, M. W. 1993. "The Use of Causal Indicators in Covariance Structure Models: Some Practical Issues," *Psychological Bulletin* (114:3), pp. 533-541.

MacCallum, R. C., Widaman, K. F., Zhang, S., and Hong, S. 1999. "Sample Size in Factor Analysis," *Psychological Methods* (4:1), pp. 84-99.

MacKenzie, I. S., and Soukoreff, R. W. 2002. "Text Entry for Mobile Computing: Models and Methods, Theory and Practice," *Human–Computer Interaction* (17:2-3), pp. 147-198.

MacKenzie, S. B., Podsakoff, P. M., and Podsakoff, N. P. 2011. "Construct Measurement and Validation Procedures in MIS and Behavioral Research: Integrating New and Existing Techniques," *MIS Quarterly* (35:2), pp. 293-334.

Mallat, N. 2007. "Exploring Consumer Adoption of Mobile Payments–A Qualitative Study," *The Journal of Strategic Information Systems* (16:4), pp. 413-432.

Masse, R. E. 1998. "Guide to Usability for Software Engineers" (available online at http://webcache.googleusercontent.com/search?q=cache:http://lte-projects.umd.edu/guse/; accessed April 2, 2012).

Massey, A. P., Khatri, V., and Ramesh, V. 2005. "From the Web to the Wireless Web: Technology Readiness and Usability," in *Proceedings of the 38th Annual Hawaii International Conference on System Sciences*, January 3-6, Los Alamitos, CA: IEEE Computer Society Press, pp. 32-45.

McElroy, J. C., Hendrickson, A. R., Townsend, A. M., and DeMarie, S. M. 2007. "Dispositional Factors in Internet Use: Personality Versus Cognitive Style," *MIS Quarterly* (31:4), pp. 809-820.

Melnyk, V., van Osselaer, S. M. J., and Bijmolt, T. H. A. 2009. "Are Women More Loyal Customers than Men? Gender Differences in Loyalty to Firms and Individual Service Providers," *Journal of Marketing* (73:4), pp. 82-96.

Miles, M. B., and Huberman, A. M. 1999. *Qualitative Data Analysis: A Sourcebook of New Methods* (2nd ed.), Thousand Oaks, CA: Sage Publications.

Min, Q., Li, S., and Zhong, Q. 2009. "An Empirical Study of M-Commerce Adoption from Usability Perspective," in *Proceedings of the 8th International Conference on Mobile Business*, Dalian, China, June 27-28, pp. 215-220.

Mithas, S., Ramasubbu, N., Krishnan, M. S., and Fornell, C. 2007. "Designing Web Sites for Customer Loyalty Across Business Domains: A Multilevel Analysis," *Journal of Management Information Systems* (23:3), pp. 97-127.

Nah, F. F. H., Eschenbrenner, B., and DeWester, D. 2011. "Enhancing Brand Equity Through Flow and Telepresence: A Comparison of 2D and 3D Virtual Worlds," *MIS Quarterly* (35:3), pp. 731-747.

Nah, F. F. H., Hong, W., Chen, L. Q., and Lee, H. H. 2010. "Information Search Patterns in E-Commerce Product Comparison Services," *Journal of Database Management* (21:2), pp. 26-40.

Nah, F. F. H., Siau, K., and Sheng, H. 2005. "The Value of Mobile Applications: A Utility Company Study," *Communications of the ACM* (48:2), pp. 85-90.

Nielsen, C. M., Overgaard, M., Pedersen, M. B., Stage, J., and Stenild, S. 2006. "It's Worth the Hassle! The Added Value of Evaluating the Usability of Mobile Systems in the Field," in *Proceedings of the 4th Nordic Conference on Human–Computer Interaction,* Oslo, Norway, October 14-18, pp. 272-280.

Nielsen, J. 2012a. "Mobile Usability Update," Nielsen Norman Group, September 26 (available online at http://www.useit.com/alertbox/mobile-usability.html; accessed May 13, 2012).

Nielsen, J. 2012b. "Windows 8—Disappointing Usability for Both Novice and Power Users," Nielsen Norman Group, November 10 (available online at http://www.nngroup.com/articles/windows-8-disappointing-usability/; accessed August 18, 2014).

Nielsen Norman Group. 2012. "Usability of Mobile Websites & Applications: 237 Design Guidelines for Improving the User Experience of Mobile Sites and Apps" (available online at http://www.nngroup.com/reports/mobile/; accessed May 13, 2012).

Oulasvirta, A., Raento, M., and Tiitta, S. 2005. "Context Contacts: Re-designing SmartPhone's Contact Book to Support Mobile Awareness and Collaboration," in *Proceedings of the 7th International Conference on Human–Computer Interaction with Mobile Devices & Service*, Salzburg, Austria, September 19-22, pp. 167-174.

Oulasvirta, A., Petit, R., Raento, M., and Tiitta, S. 2007. "Interpreting and Acting on Mobile Awareness Cues," *Human–Computer Interaction* (22:1), pp. 97-135.

Ovide, S. 2013. "Microsoft to Name New Software Windows 8.1, and It's Free," *Wall Street Journal*, May 14 (available online at http://blogs.wsj.com/digits/2013/05/14/microsoft-to-name-new-software-windows-8-1-and-its-free/; accessed June 9, 2013).

Palmer, J. W. 2003. "Web Site Usability, Design, and Performance Metrics," *Information Systems Research* (13:2), pp. 151-167.

Parboteeah, D., Valacich, J. S., and Wells, J. D. 2009. "The Influence of Website Characteristics on a Consumer's Urge to Buy Impulsively," *Information Systems Research* (20:1), pp. 60-78.

Petter, S., Straub, D. W., and Rai, A. 2007. "Specifying Formative Constructs in Information Systems Research," *MIS Quarterly* (31:4), pp. 623-656.

Pousttchi, K., and Schurig, M. 2004. "Assessment of Today's Mobile Banking Applications from the View of Customer Requirements," in *Proceedings of the 37th Annual Hawaii International Conference on System Sciences,* January 5-8, Los Alamitos, CA: IEEE Computer Society Press, pp. 5-15.

Reisinger, D. 2011. "Google: 500,000 Android Devices Activated Each Day," CNET Magazine, June 28 (available online at http://news.cnet.com/8301-13506_3-20074956-17/google-500000-android-devices-activated-each-day/; accessed July 14, 2011).

Rindskopf, D., and Rose, T. 1988. "Some Theory and Applications of Confirmatory Second-Order Analysis," *Multivariate Behavioral Research* (23:1), pp. 51-67.

Robbins, S. S., and Stylianou, A. C. 2003. "Global Corporate Web Sites: An Empirical Investigation of Content and Design," *Information & Management* (40:3), pp. 205-212.

Rosemann, M., and Vessey, I. 2008. "Towards Improving the Relevance of Information Systems Research to Practitioners: The Role of Applicability Checks," *MIS Quarterly* (32:1), pp. 1-22.

Saga, V. L., and Zmud, R. W. 1994. "The Nature and Determinants of IT Acceptance, Routinization, and Infusion," in *Diffusion, Transfer and Implementation of Information Technology*, L. Levine (ed.), Amsterdam: North-Holland, pp. 67-86.

Sarker, S., and Wells, J. D. 2003. "Understanding Mobile Hand-Held Device Use and Adoption," *Communications of the ACM* (46:12), pp. 35-40.

Seffah, A., Donyaee, M., Kline, R. B., and Padda, H. K. 2006. "Usability Measurement and Metrics: A Consolidated Model," *Software Quality Journal* (14:2), pp. 159-178.

Sonderegger, A., and Sauer, J. 2010. "The Influence of Design Aesthetics in Usability Testing: Effects on User Performance and Perceived Usability," *Applied Ergonomics* (41:3), pp. 403-410.

Sørensen, C., and Altaitoon, A. 2008. "Organisational Usability of Mobile Computing-Volatility and Control in Mobile Foreign Exchange Trading," *International Journal of Human–Computer Studies* (66:12), pp. 916-929.

Straub, D. 1989. "Validating Instruments in MIS Research," *MIS Quarterly* (13:2), pp. 147-169.

Straub, D., and Ang, S. 2011. "Editor's Comments: Rigor and Relevance in IS Research: Redefining the Debate and a Call for Future Research," *MIS Quarterly* (35:1), pp. iii-xi.

Straub, D., Boudreau, M. C., and Gefen, D. 2004. "Validation Guidelines for IS Positivist Research," *Communications of the Association for Information Systems* (13:24), pp. 380-427.

Strauss, A., and Corbin, J. 1990. *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*, Thousand Oaks, CA: Sage Publications.

Sykes, T. A., and Venkatesh, V. 2015, "Explaining Post-Implementation Employee System Use and Job Performance: Impacts of the Content and Source of Social Network Ties," *MIS Quarterly*, forthcoming.

Tan, F. B., Tung, L. L., and Xu, Y. 2009. "A Study of Web-Designers' Criteria for Effective Business-to-Consumer (B2C) Website Using the Repertory Grid Technique," *Journal of Electronic Commerce Research* (10:3), pp. 155-177.

Tanriverdi, H. 2005. "Information Technology Relatedness, Knowledge Management Capability, and Performance of Multi-business Firms," *MIS Quarterly* (29:2), pp. 311-334.

Thong, J. Y. L., Hong, W., and Tam, K. Y. 2002. "Understanding User Acceptance of Digital Libraries: What Are the Roles of Interface Characteristics, Organizational Context, and Individual Differences?," *International Journal of Human–Computer Studies* (57:3), pp. 215-242.

Treiblmaier, H. 2007. "Web Site Analysis: A Review and Assessment of Previous Research," *Communications of the Association for Information Systems* (19:1), pp. 806-843.

Urbaczewski, A., and Koivisto, M. 2007. "Measuring Mobile Device Usability as a Second Order Construct in Mobile Information Systems," in *Proceedings of the 13th Americas Conference on Information Systems*, Keystone, CO, August 9-12.

Valacich, J. S., Parboteeah, D., and Wells, J. D. 2007. "Not All Interface Characteristics are Created Equal: The Online Consumer's Hierarchy of Needs," *Communications of the ACM* (50:9), pp. 84-90.

Van der Heijden, H. 2004. "User Acceptance of Hedonic Information Systems," *MIS Quarterly* (28:4), pp. 695-704.

Venkatesh, V. 2000. "Determinants of Perceived Ease of Use: Integrating Control, Intrinsic Motivation, and Emotion into the Technology Acceptance Model," *Information Systems Research* (11:4), pp. 342-365.

Venkatesh, V., and Agarwal, R. 2006. "Turning Visitors into Customers: A Usability-Centric Perspective on Purchase Behavior in Electronic Channels," *Management Science* (52:3), pp. 367-382.

Venkatesh, V., and Davis, F. D. 2000. "A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies," *Management Science* (46:2), pp. 186-204.

Venkatesh, V., and Goyal, S. 2010. "Expectation Disconfirmation and Technology Adoption: Polynomial Modeling and Response Surface Analysis," *MIS Quarterly* (34:2), pp. 281-303.

Venkatesh, V., Morris, M. G., Davis, F. D., and Davis, G. B. 2003. "User Acceptance of Information Technology: Toward a Unified View," *MIS Quarterly* (27:3), pp. 425-478.

Venkatesh, V., and Ramesh, V. 2006. "Web and Wireless Site Usability: Understanding Differences and Modeling Use," *MIS Quarterly* (30:1), pp. 181-206.

Venkatesh, V., Ramesh, V., and Massey, A. P. 2003. "Understanding Usability in Mobile Commerce," *Communications of the ACM* (46:12), pp. 53-56.

Venkatesh, V., Thong, J. Y. L., and Xu, X. 2012. "Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology," *MIS Quarterly* (36:1), pp. 157-178.

Vila, N., and Kuster, I. 2011. "Consumer Feelings and Behaviours Towards Well Designed Websites," *Information & Management* (48:4-5), pp. 166-177.

Wagner, T., Hennig-Thurau, T., and Rudolph, T. 2009. "Jeopardize Loyalty?," *Journal of Marketing* (73:5), pp. 69-85.

Wallace, L., Keil, M., and Rai, A. 2004. "How Software Project Risk Affects Project Performance: An Investigation of the Dimensions of Risk and An Exploratory Model," *Decision Sciences* (35:2), pp. 289-321.

Wells, J. D., Fuerst, W. L., and Palmer, J. W. 2005. "Designing Consumer Interfaces for Experiential Tasks: An Empirical Investigation," *European Journal of Information Systems* (14:3), pp. 273-287.

Wells, J. D., Parboteeah, V., and Valacich, J. S. 2011. "Online Impulse Buying: Understanding the Interplay between Consumer Impulsiveness and Website Quality," *Journal of the Association for Information Systems* (12:1), pp. 32-56.

Wells, J. D., Valacich, J. S., and Hess, T. J. 2011. "What Signal Are You Sending? How Website Quality Influences Perceptions of Product Quality and Purchase Intentions," *MIS Quarterly* (35:2), pp. 373-396.

Wetzels, M., Odekerken-Schroder, G., and Van Oppen, C. 2009. "Using PLS Path Modeling for Assessing Hierarchical Construct Models: Guidelines and Empirical Illustration," *MIS Quarterly* (33:1), pp.177-195.

Wobbrock, J. O., Myers, B. A., and Aung, H. H. 2008. "The Performance of Hand Postures in Front- and Back-of-Device

Interaction for Mobile Computing," *International Journal of Human–Computer Studies* (66:12), pp. 857-875.

Wu, J-H., and Wang, S. C. 2005. "What Drives Mobile Commerce? An Empirical Evaluation of the Revised Technology Acceptance Model," *Information & Management* (42:5), pp. 719-729.

Xu, Y., Tung, L. L., and Tan, F. B. 2009. "Attributes of Website Usability: A Study of Web Users with the Repertory Grid Technique," *International Journal of Electronic Commerce* (13:4), pp. 99-128.

Yao, G., Wu, C-H., and Yang, C-T. 2007. "Examining the Content Validity of the WHOQOL-BREF from Respondents' Perspective by Quantitative Methods," *Social Indicators Research* (85:3), pp. 483-498.

Youens, R. 2011. "7 Habits of Highly Effective Apps" GIGACOM, July 16 (available online at http://gigaom.com/2011/07/16/7-habits-of-highly-effective-apps/; accessed March 24, 2012).

Zhang, D., and Adipat, B. 2005. "Challenges, Methodologies, and Issues in the Usability Testing of Mobile Applications," *International Journal of Human–Computer Interaction* (18:3), pp. 293-308.

Zhang, D., Adipat, B., and Mowafi, Y. 2009. "User-Centered Context-Aware Mobile Applications—The Next Generation of Personal Mobile Computing," *Communications of the Association for Information Systems* (24:1), Article 3.

Zhang, P., and Von Dran, G. M. 2000. "Satisfiers and Dissatisfiers: A Two-Factor Model for Website Design and Evaluation," *Journal of the American Society for Information Science* (51:14), pp. 1253-1268.

Zhang, P., and Von Dran, G. M. 2001. "User Expectations and Rankings of Quality Factors in Different Web Site Domains," *International Journal of Electronic Commerce* (6:2), pp. 9-33.

Ziefle, M., and Bay, S. 2006. "How to Overcome Disorientation in Mobile Phone Menus: A Comparison of Two Different Types of Navigation Aids," *Human–Computer Interaction* (21:4), pp. 393-433.

About the Authors

Hartmut Hoehle is an assistant professor of Information Systems in the Sam M. Walton School of Business at the University of Arkansas. He received a Ph.D. in Information Systems from Victoria University of Wellington, New Zealand. He was a lecturer at the School of Accounting and Business Information Systems, Australian National University. Stemming from his professional experiences gained while working at Deutsche Bank, he is particularly interested in how services and products can be distributed through electronically mediated channels in a retail context. His work has appeared or is forthcoming in *MIS Quarterly*, *European Journal of Information Systems*, *Journal of the Association for Information Science and Technology*, *Decision Support Systems*, *Journal of Computer Information Systems,* and *International Journal of Electronic Finance*. He has presented at the International Conference on Information Systems and the European Conference on Information Systems. He has also served as a reviewer for leading journals including *MIS Quarterly*.

Viswanath Venkatesh is a Distinguished Professor and Billingsley Chair in Information Systems at the University of Arkansas, where he has been since June 2004. Prior to joining Arkansas, he was on the faculty at the University of Maryland and received his Ph.D. at the University of Minnesota. His research focuses on understanding the diffusion of technologies in organizations and society. His work has appeared or is forthcoming in leading journals in information systems, organizational behavior, psychology, marketing, and operations management. Recently, he was recognized as one of the most influential scholars in business and economics (highlycited.com). His articles have been cited about 40,000 times and about 10,000 times per Google Scholar and Web of Science, respectively. He developed and maintains a web site that tracks researcher and university research productivity (http://www.myvisionsresearch.com/ISRankings). He has published a book titled *Road to Success: A Guide for Doctoral Students and Junior Faculty Members in the Behavioral and Social Sciences* (http://vvenkatesh.com/book).

MIS Quarterly

RESEARCH ARTICLE

MOBILE APPLICATION USABILITY: CONCEPTUALIZATION AND INSTRUMENT DEVELOPMENT

Hartmut Hoehle and Viswanath Venkatesh
Department of Information Systems, Sam M. Walton College of Business, University of Arkansas,
Fayetteville, AR 72701 U.S.A. {hartmut@hartmuthoehle.com} {vvenkatesh@vvenkatesh.us}

Appendix A

Content Validity Check and Pretest: Respondent Demographics

		Content Validity Check		Pretest	
Demographic	**Category**	**n = 318**	**%**	**n = 440**	**%**
Gender	Men	179	56	236	53.6
	Women	139	44	204	46.4
Age groups	Under 20	48	15	45	10.2
	20-29	210	66	305	69.3
	30-39	54	17	70	15.9
	40-49	4	1	16	3.6
	50-59	0	0	3	.7
	60 or older	2	1	1	.2
Income (Annual, in USD)	0-10,000	38	12	37	8.4
	10,000-19,000	42	13	56	12.7
	20,000-29,000	40	13	46	10.5
	30,000-39,000	41	13	53	12
	40,000-49,000	28	9	37	8.4
	50,000-74,000	44	14	70	15.9
	75,000-99,000	43	14	56	12.7
	100,000-150,000	30	9	28	6.4
	Over 150,000	12	4	57	13
Job	ICT	37	12	69	15.7
	Banking and Finance	12	4	13	3
	Insurance, Real Estate and Legal	3	1	5	1.1
	Government and Military	6	2	11	2.5
	Medical Healthcare	0	0	3	0.7
	Construction and Engineering	10	3	11	2.5
	Retail and Wholesale	1	0	5	1.1
	Education	17	5	28	6.4
	Marketing and Advertising	18	6	31	7.0
	Student	167	53	162	36.8
	Other	47	15	102	23.2

Appendix B

Proportion of Substantive Agreement and Substantive Validity Coefficients Based on the Content Validity Survey

Construct Name	Label	PSA	CSV	Construct Name	Label	PSA	CSV	Construct Name	Label	PSA	CSV
Branding	BRAN1	.87	.82	Aesthetic graphics	AEST1	.92	.87	Concise language	CLAN1	.78	.66
	BRAN2	.87	.83		AEST2	.89	.83		CLAN2	.88	.83
	BRAN3	.89	.85		AEST3	.93	.88		CLAN3	.80	.69
	BRAN4	.89	.83		AEST4	.91	.85		CLAN4	.80	.69
Data preservation	DAPR1	.60	.87	Realism	REAL1	.87	.80	Short icon labeling	ICOL1	.43	.05
	DAPR2	.89	.86		REAL2	.89	.82		ICOL2	.41	-.01
	DAPR3	.90	.86		REAL3	.89	.82		ICOL3	.35	-.14
	DAPR4	.68	.56		REAL4	.90	.84		ICOL4	.25	-.43
Instant start	STAR1	.88	.83	Subtle animation	SANM1	.79	.75	Standardized user-interface element	SUI1	.91	.86
	STAR2	.74	.63		SANM2	.91	.88		SUI2	.90	.84
	STAR3	.88	.85		SANM3	.92	.89		SUI3	.89	.83
	STAR4	.88	.84		SANM4	.92	.89		SUI4	.90	.83
Orientation	ORIE1	.89	.85	Control obviousness	COOB1	.89	.84	User-centric terminology	UCT1	.87	.81
	ORIE2	.87	.82		COOB2	.76	.58		UCT2	.89	.84
	ORIE3	.90	.85		COOB3	.89	.85		UCT3	.89	.84
	ORIE4	.89	.82		COOB4	.81	.66		UCT4	.88	.84
Collaboration	COLL1	.85	.75	De-emphasis of user settings	DUS1	.91	.86	Logical path	LP1	.91	.85
	COLL2	.90	.84		DUS2	.90	.85		LP2	.82	.76
	COLL3	.88	.81		DUS3	.91	.86		LP3	.83	.76
	COLL4	.89	.85		DUS4	.87	.92		LP4	.82	.76
Content relevance	CRLV1	.77	.63	Effort minimization	EMM1	.86	.80	Top-to-bottom structure	TTPS1	.85	.76
	CRLV2	.61	.34		EMM2	.92	.87		TTPS2	.61	.37
	CRLV3	.82	.70		EMM3	.88	.83		TTPS3	.65	.42
	CRLV4	.49	.09		EMM4	.72	.59		TTPS4	.85	.79
Search	SEAR1	.89	.85	Fingertip-size controls	FTSC1	.89	.84		TTPS5	.68	.53
	SEAR2	.90	.86		FTSC2	.89	.84		TTPS6	.82	.73
	SEAR3	.87	.82		FTSC3	.91	.87				
	SEAR4	.90	.85		FTSC4	.64	.46				

Appendix C

Initial Item Pool Created Based on the Analysis of the Content Validity Check

AEST1	The mobile application uses beautiful artwork.
AEST2	The mobile application uses rich, beautiful, and engaging graphics that draw you into the application.
AEST3	The mobile application uses stunning graphics.
AEST4	The mobile application benefits from beautiful and engaging graphics.
BRAN1	The mobile application uses brand colors or images in a refined and unobtrusive way.
BRAN2	The mobile application doesn't force me to watch an advertisement.
BRAN3	The mobile application quietly reminds you of the brand that runs the application.
BRAN4	The mobile application integrates branding effectively.
CLAN1	The mobile application uses as few words as possible without losing the meaning.
CLAN2	The mobile application uses concise language.
CLAN3	The mobile application brings the main message across in a few words.
CLAN4	The mobile application uses precise and concise text.
COLL1	The mobile application helps you to share information with other people.
COLL2	The mobile application allows you to connect with other people.
COLL3	The mobile application supports collaboration with others.
COLL4	The mobile application helps you to interact with others.
COOB1	The mobile application makes the main function of the application immediately apparent.
COOB2	The mobile application uses intuitive commands.
COOB3	The mobile application uses controls that are immediately obvious.
COOB4	The mobile application employs controls that are intuitive.
CRLV1	The mobile application emphasizes the content you want to find.
CRLV2	The mobile application emphasizes the content that is important to you.
CRLV3	The mobile application emphasizes the content you care about.
CRLV4	The mobile application elevates the content that is relevant to you.
DAPR1	The mobile application automatically saves your data when you close the application.
DAPR2	The mobile application doesn't require you to manually save your data when you quit the application.
DAPR3	The mobile application saves the data automatically and you can re-start where you left previously.
DAPR4	The mobile application allows you to quit the application and restart at the same stage when re-entering it.
DUS1	The mobile application avoids setting up user preferences within the application.
DUS2	The mobile application de-emphasizes user settings.
DUS3	The mobile application doesn't prompt you to change user settings within the application.
DUS4	The mobile application doesn't request you to modify the user setting within the application.
EMM1	The mobile application makes it easy for you to input your choice.
EMM2	The mobile application minimizes effort for you to type in information.
EMM3	The mobile application offers you fields to choose from so that you don't have to type in text.
EMM4	The mobile application allows me to perform tasks without having to input data.
FTSC1	The mobile application uses fingertip-size controls.
FTSC2	The mobile application makes use of fingertip-size buttons.
FTSC3	The mobile application uses large-size controls.
FTSC4	The mobile application uses small controls that require you to aim carefully before you tap it.
LP1	The mobile application gives users a logical path to follow.
LP2	The mobile application follows a logical path.
LP3	The mobile application provides users a logical path to follow.
LP4	The mobile application uses a predictable path.
ORIE1	The mobile application doesn't prompt you to change the orientation of the screen (move the device).

ORIE2	The mobile application works well independent of how you hold your mobile device.
ORIE3	The mobile application flips the content over if you change the orientation of the device (horizontal/vertical).
ORIE4	The mobile application works well independent of whether you hold your device horizontally or vertically.
REAL1	The mobile application uses realistic icons or pictures (e.g., trashcan) to help you understand the functions better.
REAL2	The mobile application helps you to understand functions by labeling them with realistic icons or pictures (e.g., trashcan).
REAL3	The mobile application uses real-life icons or pictures to illustrate the functionality (e.g., trashcan for deleting items).
REAL4	The mobile application uses realistic icons or pictures (e.g., trashcan) to get the message across.
SANM1	The mobile application uses animations effectively to communicate content.
SANM2	The mobile application uses animations appropriately.
SANM3	The mobile application doesn't overuse animations.
SANM4	The mobile application uses subtle animation to communicate content.
SEAR1	The mobile application narrows down the results as you are typing, when searching for information.
SEAR2	The mobile application helps you to search for information via a search bar.
SEAR3	The mobile application displays a search bar when you have to look for information.
SEAR4	The mobile application makes searching for information easy.
STAR1	The mobile application launches quickly and allows you to instantly start using it.
STAR2	The mobile application takes a lot of time to open.
STAR3	The mobile application doesn't require much time to open.
STAR4	The mobile application is instantly "ready to go" right after switching it on.
SUI1	The mobile application has buttons and icons that are similar to other applications.
SUI2	The mobile application has buttons and icons that I have used in other applications.
SUI3	The mobile application uses buttons and icons that you have seen in other applications.
SUI4	The mobile application uses standard icons that you already know from other applications.
TTPS1	The mobile application puts the most frequently used information near the top.
TTPS2	The mobile application displays the most important information on the top of the screen.
TTPS3	The mobile application lists the most essential information on the top of the screen.
TTPS4	The mobile application lists the most frequently used operations at the very top.
TTPS5	The mobile application arranges the least often used operation on the bottom.
TTPS6	The mobile application places the most frequently used operation at the top.
UCT1	The mobile application uses terminology that you understand.
UCT2	The mobile application avoids technical jargon.
UCT3	The mobile application doesn't use technical terms.
UCT4	The mobile application uses terminology that is comprehensible.

Appendix D

Model Fit for Pretest, Study 1 and Study 2

	Pretest	Study 1	Study 2
GFI (≥ .90)	.92	.93	.94
RMSEA (≤ .06)	.05	.04	.04
SRMR (≤ .08)	.06	.05	.05
CFI (≥ .95)	.96	.96	.97
NFI (≥ .90)	.92	.92	.93
TLI (≥ .80)	.87	.91	.88

Appendix E

Pretest: Reliabilities, AVEs, and Correlations

	Cron.α	Mean	SD	1	2	3	4	5	6	7	8	9	10	11
1. Gender	NA	NA	NA	NA										
2. Age	NA	NA	NA	-.13*	NA									
3. Income	NA	NA	NA	.17**	.15*	NA								
4. Application design	.82	4.71	1.55	-.13*	-.12*	.05	.74							
5. Application utility	.85	4.42	1.38	-.15*	-.13*	.04	.16*	.73						
6. User interface graphics	.82	4.17	1.28	-.12*	-.12*	.07	.07	.10	.75					
7. User interface input	.84	4.20	1.30	-.15*	-.16*	.02	.05	.07	.19**	.77				
8. User interface output	.75	4.34	1.32	-.14*	-.19**	.05	.04	.05	.20**	.19**	.80			
9. User interface structure	.77	4.57	1.28	-.13*	-.07	.02	.08	.08	.15*	.20**	.17**	.73		
10. Mobile application loyalty	.80	4.99	1.60	.15*	.05	.08	.40***	.35***	.20**	.13*	.12*	.10	.70	
11. Continued intention to use	.75	5.07	1.71	-.07	-.14*	.04	.44***	.37***	.38***	.20**	.30***	.35***	.38***	.71

*p < 0.05, **p < 0.01, and ***p < 0.001

Appendix F

Pretest: Unique Proportion of Variance in the Second-Order Construct Explained by Each First-Order Construct

Second-Order Construct	First-Order Construct	R^2	Fornell and Larcker's Construct Reliability Index
Application design	Branding	.06	.87
	Data preservation	.08	.86
	Instant start	.07	.82
	Orientation	.05	.82
Application utility	Collaboration	.06	.89
	Content relevance	.35	.83
	Search	.05	.81
User interface graphics	Aesthetic graphics	.14	.84
	Realism	.10	.81
	Subtle animation	.17	.81
User interface input	Control obviousness	.03	.83
	De-emphasis of user settings	.05	.79
	Effort minimization	.11	.84
	Fingertip-size control	.07	.83
User interface output	Concise Language	.03	.82
	Standardized user interface element	.05	.80
	User-centric terminology	.08	.79
User interface structure	Logical path	.08	.77
	Top-to-bottom structure	.14	.80

Appendix G

Pretest: Item Loadings and Weights

Construct Name	Error	Loadings	Weights on 2nd Order	Construct Name	Error	Loadings	Weights on 2nd Order	Construct Name	Error	Loadings	Weights on 2nd Order
First-Order Constructs											
Branding (BRAN1-4)	0.71	0.84	.35***	Aesthetic graphics (AEST1-4)	0.77	0.88	.43***	Concise language (CLAN1-4)	1.04	0.87	.25***
	0.74	0.91			0.73	0.84			1.09	0.85	
	0.78	0.82			0.85	0.75			1.20	0.79	
	0.73	0.92			1.07	0.88			0.83	0.75	
Data preservation (DAPR1-4)	1.00	0.82	.35***	Realism (REAL1-4)	0.77	0.84	.37***	Standardized user-interface element (SUI1-4)	0.91	0.75	.33***
	0.85	0.84			0.75	0.83			1.02	0.78	
	0.87	0.88			0.68	0.80			1.03	0.80	
	1.01	0.89			0.73	0.75			0.85	0.84	
Instant start (STAR1-4)	1.02	0.78	.29***	Subtle animation (SANM1-4)	1.05	0.69	.44***	User-centric terminology (UCT1-4)	1.03	0.65	.29***
	1.05	0.79			0.86	0.88			0.87	0.83	
	1.02	0.83			0.85	0.84			0.78	0.84	
	1.00	0.84			0.89	0.82			0.89	0.78	
Orientation (ORIE1-4)	1.01	0.92	.35***	Control obviousness (COOB1-4)	1.04	0.79	.22***	Logical path (LP1-4)	1.05	0.75	.38***
	1.04	0.84			1.09	0.82			1.03	0.77	
	0.86	0.77			1.18	0.84			1.04	0.78	
	0.89	0.73			1.04	0.83			0.83	0.73	
Collaboration (COLL1-4)	0.74	0.86	.34***	De-emphasis of user settings (DUS1-4)	1.05	0.80	.28***	Top-to-bottom structure (TTPS1-6)	1.03	0.84	.43***
	0.73	0.93			0.87	0.83			1.05	0.93	
	0.83	0.86			0.77	0.83			0.82	0.68	
	0.85	0.91			0.72	0.66			0.84	0.77	
Content relevance (CRLV1-4)	1.02	0.82	.67***	Effort minimization (EMM1-4)	1.02	0.83	.44***		1.04	0.73	
	0.88	0.84			0.89	0.86			1.02	0.80	
	0.83	0.83			0.89	0.82					
	1.02	0.80			1.05	0.83					
Search (SEAR1-4)	1.02	0.82	.29***	Fingertip-size controls (FTSC1-4)	0.82	0.84	.31***				
	0.89	0.87			0.89	0.87					
	0.80	0.78			0.93	0.78					
	0.74	0.73			0.94	0.79					
Second-Order Constructs											
Application design (DES1-4)	1.02	0.84	NA	User interface graphics (INTG1-4)	1.02	0.78	NA	User interface output (CONT1-4)	0.80	0.80	NA
	1.03	0.83			1.00	0.82			0.80	0.80	
	0.88	0.93			0.84	0.84			1.00	0.80	
	0.84	0.80			0.87	0.83			1.00	0.80	
Application utility (PURP1-4)	0.85	0.82	NA	User interface input (INP1-4)	1.03	0.78	NA	User interface structure (STRU1-4)	0.70	0.80	NA
	0.91	0.80			1.05	0.83			1.00	0.70	
	0.92	0.84			0.78	0.74			1.10	0.70	
	1.03	0.78			0.82	0.73			1.00	0.80	

*p < 0.05, **p < 0.01, and ***p < 0.001

Appendix H

Study 1 and Study 2: Respondent Demographics

		Study 1		Study 2	
Demographic	Category	n = 408	%	n = 412	%
Gender	Men	280	68.6	275	66.7
	Women	128	31.4	137	33.3
Age groups	Under 20	42	10.3	40	9.7
	20-29	140	34.3	144	35.0
	30-39	101	24.8	103	25.0
	40-49	74	18.1	70	17.0
	50-59	25	6.1	29	7.0
	60 or older	26	6.4	26	6.3
Income (Annual, in USD)	0-10,000	12	2.9	10	2.4
	10,000-19,000	58	14.2	60	14.6
	20,000-29,000	70	17.2	77	18.7
	30,000-39,000	30	7.4	23	5.6
	40,000-49,000	20	4.9	20	4.9
	50,000-74,000	25	6.1	26	6.3
	75,000-99,000	125	30.6	120	29.1
	100,000-150,000	40	9.8	44	10.7
	Over 150,000	28	6.9	32	7.8
Job	ICT	70	17.2	65	15.8
	Banking and Finance	22	5.4	20	4.9
	Insurance, Real Estate and Legal	28	6.9	25	6.1
	Government and Military	40	9.8	42	10.2
	Medical Healthcare	35	8.6	37	9.0
	Construction and Engineering	22	5.4	25	6.1
	Retail and Wholesale	40	9.8	41	10.0
	Education	12	2.9	15	3.6
	Marketing and Advertising	17	4.2	19	4.6
	Student	48	11.8	50	12.1
	Other	74	18.1	73	17.7
Social media preference	Facebook	210	51.5	206	50
	LinkedIn	130	31.9	128	31.1
	Twitter	30	7.4	35	8.5
	My Space	18	4.4	19	4.6
	Google+	20	4.9	24	5.8
Access to mobile sites	Application on phone	370	90.7	372	90.3
	Web browser	38	9.3	40	9.7
Primary phone use	iPhone	220	53.9	225	54.6
	BlackBerry	70	17.2	69	16.7
	Android	44	10.8	42	10.2
	Windows Mobile	28	6.9	26	6.3
	Symbian	12	2.9	11	2.7
	Other	34	8.3	39	9.5

Appendix I

Construct Definitions and MUG Scales Used by Venkatesh and Ramesh (2006)

Construct	Definition	Items Used
Content	"assesses the informational and transactional capabilities of a mobile application"	Overall, the mobile social media application offers content that is relevant to the core audience.
		The mobile social media application uses media appropriately and effectively to communicate the content.
		I think that the mobile social media application provides the appropriate breadth and depth of content.
		In general, the mobile social media application provides current and timely information.
Ease of use	"relates to the cognitive effort required in using a mobile application"	The mobile social media application offers clear and understandable goals.
		Overall, the mobile social media application is well structured and organized.
		The mobile social media application provides clear and understandable results and feedback regarding your progress.
Emotion	"taps into affective reactions invoked by a Web site"	The mobile social media application offers you an element of challenge.
		The mobile social media application provides an interesting story line.
		The mobile social media application ties to individuals, within and outside the organization, who have credibility.
		The mobile social media application allows you to control the flow of information.
Made-for-the-medium	"relates to tailoring a Web site to fit a particular user's needs"	The mobile social media application offers you the opportunity to be part of an online group or community.
		The mobile social media application treats you as a unique person and respond to your specific needs.
		The mobile social media application reflects the most current trend(s) and provides the most current information.
Promotion	"captures the advertising of a Web site on the Internet and other media"	The mobile social media application understands to incorporate advertisements.

Reference

Venkatesh, V., and Ramesh, V. 2006. "Web and Wireless Site Usability: Understanding Differences and Modeling Use," *MIS Quarterly* (30:1), pp. 181-206.

Appendix J

Study 1: Construct Reliability and Correlation Matrix

	Cron .α	Mean	SD	1	2	3	4	5	6	7	8	9	10	11
1. Gender	NA	NA	NA	NA										
2. Age	NA	NA	NA	-.14*	NA									
3. Income	NA	NA	NA	.20**	.16*	NA								
4. Application design	.80	4.74	1.50	-.13*	-.13*	.04	.73							
5. Application utility	.82	4.44	1.28	-.08	-.07	.03	.20**	.71						
6. User interface graphics	.83	4.28	1.30	-.13*	-.05	.05	.04	.07	.74					
7. User interface input	.77	4.30	1.31	-.16*	-.17**	.01	.02	.02	.20**	.75				
8. User interface output	.75	4.37	1.30	-.10	-.20**	.04	.03	.05	.22***	.21**	.77			
9. User interface structure	.73	4.61	1.30	-.07	-.08	.01	.07	.07	.14*	.22***	.19**	.72		
10. Mobile application loyalty	.75	5.04	1.51	.16*	.06	.05	.44***	.39***	.22***	.15*	.13*	.07	.71	
11. Continued intention to use	.78	4.98	1.59	-.08	-.17**	.02	.43***	.37***	.40***	.22***	.32***	.38***	.40***	.74

*p < 0.05, **p < 0.01, and ***p < 0.001

Appendix K

Study 1: Unique Proportion of Variance in the Second-Order Construct Explained by Each First-Order Construct

Second-Order Construct	First-Order Construct	R^2	Fornell and Larcker's Construct Reliability Index
Application design	Branding	.06	.87
	Data preservation	.09	.88
	Instant start	.08	.82
	Orientation	.06	.83
Application utility	Collaboration	.06	.87
	Content relevance	.34	.86
	Search	.05	.80
User interface graphics	Aesthetic graphics	.12	.86
	Realism	.10	.83
	Subtle animation	.17	.81
User interface input	Control obviousness	.02	.84
	De-emphasis of user settings	.06	.80
	Effort minimization	.12	.84
	Fingertip-size control	.07	.82
User interface output	Concise Language	.03	.81
	Standardized user interface element	.06	.83
	User-centric terminology	.07	.78
User interface structure	Logical path	.08	.78
	Top-to-bottom structure	.14	.80

Appendix L

Study 1: Item Loadings and Weights

Construct Name	Error	Loadings	Weights on 2nd Order	Construct Name	Error	Loadings	Weights on 2nd Order	Construct Name	Error	Loadings	Weights on 2nd Order
First-Order Constructs											
Branding (BRAN1-4)	0.74	.88	.33***	Aesthetic graphics (AEST1-4)	0.80	.89	.40***	Concise language (CLAN1-4)	1.05	.83	.24***
	0.78	.87			0.78	.84			1.03	.85	
	0.80	.80			0.83	.78			1.17	.78	
	0.77	.91			1.10	.90			0.85	.76	
Data preservation (DAPR1-4)	1.03	.87	.34***	Realism (REAL1-4)	0.79	.83	.34***	Standardized user-interface element (SUI1-4)	0.88	.78	.32***
	0.87	.85			0.74	.88			0.84	.79	
	0.88	.91			0.65	.82			1.05	.83	
	1.04	.89			0.78	.77			0.88	.88	
Instant start (STAR1-4)	1.03	.79	.26***	Subtle animation (SANM1-4)	1.02	.65	.40***	User-centric terminology (UCT1-4)	1.01	.64	.25***
	1.02	.80			0.83	.89			0.89	.82	
	1.03	.82			0.83	.84			0.79	.83	
	0.82	.85			0.88	.85			0.80	.79	
Orientation (ORIE1-4)	1.04	.90	.33***	Control obviousness (COOB1-4)	1.05	.82	.26***	Logical path (LP1-4)	0.93	.75	.35***
	1.00	.88			1.02	.83			1.04	.79	
	0.88	.80			1.10	.85			1.03	.80	
	0.91	.74			1.04	.83			0.89	.75	
Collaboration (COLL1-4)	0.73	.83	.33***	De-emphasis of user settings (DUS1-4)	1.09	.83	.31***	Top-to-bottom structure (TTPS1-6)	1.07	.80	.40***
	0.74	.91			0.80	.85			1.03	.91	
	0.85	.85			1.02	.85			0.80	.65	
	0.88	.87			0.75	.62			0.82	.73	
Content relevance (CRLV1-4)	1.06	.85	.58***	Effort minimization (EMM1-4)	1.05	.85	.37***		1.00	.79	
	0.89	.88			0.85	.84			1.00	.84	
	0.85	.89			0.88	.83					
	1.04	.83			1.04	.82					
Search (SEAR1-4)	1.02	.81	.31***	Fingertip-size controls (FTSC1-4)	0.84	.83	.30***				
	0.83	.83			0.91	.85					
	0.89	.77			0.95	.79					
	0.75	.78			0.97	.80					
Second-Order Constructs											
Application design (DES1-4)	1.05	.83	NA	User interface graphics (INTG1-4)	1.00	.80	NA	User interface output (CONT1-4)	0.82	.78	NA
	1.04	.82			1.04	.81			0.77	.80	
	0.85	.91			0.83	.83			1.03	.78	
	0.83	.77			0.85	.82			1.00	.74	
Application utility (PURP1-4)	0.89	.84	NA	User interface input (INP1-4)	1.02	.70	NA	User interface structure (STRU1-4)	0.77	.80	NA
	0.89	.83			1.01	.82			1.03	.78	
	0.90	.82			0.84	.80			1.03	.77	
	1.05	.77			0.84	.76			1.01	.80	

*$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$

Appendix M

Study 2: Construct Reliability and Correlation Matrix

	Cron.α	Mean	SD	1	2	3	4	5	6	7	8	9	10	11
1. Gender	NA	NA	NA	NA										
2. Age	NA	NA	NA	-.07	NA									
3. Income	NA	NA	NA	-.13*	.14*	NA								
4. Application design	.77	4.84	1.42	-.05	-.10	.10	.70							
5. Application utility	.75	4.35	1.30	-.07	-.04	.05	.22***	.74						
6. User interface graphics	.74	4.20	1.32	-.05	-.03	.07	.05	.04	.72					
7. User interface input	.78	4.41	1.28	-.07	-20**	.02	.07	.02	.24***	.71				
8. User interface output	.80	4.30	1.25	-.02	-.20**	.01	.08	.07	.24***	.22***	.70			
9. User interface structure	.82	4.48	1.20	-.03	-.10	.04	.10	.02	.15*	.25***	.22***	.72		
10. Mobile application loyalty	.80	4.92	1.47	.15*	.05	.03	.40***	.40***	.25***	. 17**	.17**	.04	.73	
11. Continued intention to use	.77	4.90	1.50	-.04	-.15*	.07	.42***	.40***	.44***	.24***	.35***	.41***	.44***	.71

*p < 0.05, **p < 0.01, and ***p < 0.001

Appendix N

Study 2: Unique Proportion of Variance in the Second-Order Construct Explained by Each First-Order Construct

Second-Order Construct	First-Order Construct	R^2	Fornell and Larcker's Construct Reliability Index
Application design	Branding	.07	.87
	Data preservation	.08	.89
	Instant start	.05	.83
	Orientation	.05	.84
Application utility	Collaboration	.05	.86
	Content relevance	.35	.86
	Search	.04	.82
User interface graphics	Aesthetic graphics	.11	.85
	Realism	.11	.84
	Subtle animation	.16	.82
User interface input	Control obviousness	.02	.81
	De-emphasis of user settings	.05	.87
	Effort minimization	.12	.85
	Fingertip-size control	.06	.84
User interface output	Concise Language	.04	.80
	Standardized user interface element	.05	.83
	User-centric terminology	.06	.79
User interface structure	Logical path	.08	.78
	Top-to-bottom structure	.15	.82

Appendix O

Study 2: Item Loadings and Weights

Construct Name	Error	Loadings	Weights on 2nd Order	Construct Name	Error	Loadings	Weights on 2nd Order	Construct Name	Error	Loadings	Weights on 2nd Order
First-Order Constructs											
Branding (BRAN1-4)	0.77	0.85	0.35***	Aesthetic graphics (AEST1-4)	0.69	0.85	0.41***	Concise language (CLAN1-4)	1.02	0.80	0.25***
	0.75	0.87			0.75	0.84			1.04	0.85	
	0.82	0.82			1.02	0.79			1.04	0.75	
	0.80	0.91			1.04	0.91			0.85	0.76	
Data preservation (DAPR1-4)	1.04	0.87	0.29***	Realism (REAL1-4)	0.74	0.83	0.33***	Standardized user-interface element (SUI1-4)	0.87	0.82	0.34***
	0.89	0.87			0.79	0.88			0.87	0.79	
	0.90	0.92			0.80	0.85			1.02	0.83	
	1.02	0.89			0.85	0.79			0.94	0.85	
Instant start (STAR1-4)	1.02	0.80	0.24***	Subtle animation (SANM1-4)	1.02	0.69	0.42***	User-centric terminology (UCT1-4)	1.04	0.71	0.26***
	1.04	0.80			0.86	0.89			0.89	0.84	
	1.02	0.83			0.80	0.83			0.80	0.80	
	0.82	0.86			0.85	0.84			0.82	0.79	
Orientation (ORIE1-4)	1.04	0.90	0.35***	Control obviousness (COOB1-4)	1.02	0.74	0.28***	Logical path (LP1-4)	0.88	0.75	0.35***
	1.02	0.89			1.00	0.83			0.89	0.75	
	0.91	0.80			1.04	0.80			1.02	0.82	
	0.93	0.76			1.03	0.84			0.90	0.75	
Collaboration (COLL1-4)	0.77	0.80	0.32***	De-emphasis of user settings (DUS1-4)	1.05	0.85	0.32***	Top-to-bottom structure (TTPS1-6)	1.04	0.82	0.37***
	0.77	0.89			0.82	0.87			1.02	0.88	
	0.85	0.88			1.02	0.85			1.01	0.70	
	0.88	0.85			0.77	0.89			0.85	0.82	
Content relevance (CRLV1-4)	1.02	0.88	0.60***	Effort minimization (EMM1-4)	1.04	0.85	0.35***		0.97	0.80	
	0.85	0.88			0.85	0.85			0.95	0.84	
	0.85	0.82			0.89	0.83					
	0.98	0.83			1.04	0.84					
Search (SEAR1-4)	1.04	0.81	0.30***	Fingertip-size controls (FTSC1-4)	0.84	0.90	0.32***				
	0.83	0.84			0.90	0.87					
	0.91	0.79			0.95	0.79					
	0.78	0.80			0.97	0.77					
Second-Order Constructs											
Application design (DES1-4)	1.04	0.83	NA	User interface graphics (INTG1-4)	1.04	0.80	NA	User interface output (CONT1-4)	0.90	0.80	NA
	1.00	0.83			1.00	0.81			0.82	0.74	
	0.84	0.90			0.84	0.82			1.00	0.77	
	0.84	0.80			0.95	0.82			0.85	0.77	
Application utility (PURP1-4)	0.88	0.80	NA	User interface input (INP1-4)	1.04	0.80	NA	User interface structure (STRU1-4)	0.78	0.78	NA
	0.84	0.75			1.04	0.75			1.00	0.80	
	0.92	0.78			0.90	0.82			1.00	0.77	
	1.04	0.80			0.86	0.76			1.04	0.82	

*p < 0.05, **p < 0.01, and ***p < 0.001