Multi-Layered Interfaces to Improve Older Adults' Initial Learnability of Mobile Applications

ROCK LEUNG
University of British Columbia
LEAH FINDLATER
University of Washington
JOANNA MCGRENERE, PETER GRAF, and JUSTINE YANG
University of British Columbia

Mobile computing devices can offer older adults (ages 65+) support in their daily lives, but older adults often find such devices difficult to learn and use. One potential design approach to improve the learnability of mobile devices is a Multi-Layered (ML) interface, where novice users start with a reduced-functionality interface layer that only allows them to perform basic tasks, before progressing to a more complex interface layer when they are comfortable. We studied the effects of a ML interface on older adults' performance in learning tasks on a mobile device. We conducted a controlled experiment with 16 older (ages 65–81) and 16 younger participants (age 21–36), who performed tasks on either a 2-layer or a nonlayered (control) address book application, implemented on a commercial smart phone. We found that the ML interface's Reduced-Functionality layer, compared to the control's Full-Functionality layer, better helped users to master a set of basic tasks and to retain that ability 30 minutes later. When users transitioned from the Reduced-Functionality to the Full-Functionality interface layer, their performance on the previously learned tasks was negatively affected, but no negative impact was found on learning new, advanced tasks. Overall, the ML interface provided greater benefit for older participants than for younger participants in terms of task completion time during initial learning, perceived complexity, and preference. We discuss how the ML interface approach is suitable for improving the learnability of mobile applications, particularly for older adults.

Categories and Subject Descriptors: H.5.2 [**Information Interfaces and Presentation**]: User Interfaces—*Graphical User Interfaces (GUI), Screen Design*

General Terms: Design, Experimentation, Human Factors

This research was supported in part by the Canadian Institutes of Health Research (CIHR), the Natural Sciences and Engineering Research Council of Canada (NSERC), and Nokia.
Authors' addresses: R. Leung, J. McGrenere, and J. Yang, Department of Computer Science, University of British Columbia, Vancouver, Canada; L. Findlater, Information School, University of Washington, Seattle, WA; P. Graf, Department of Psychology, University of British Columbia, Vancouver, Canada; corresponding author: R. Leung: email: rockl@cs.ubc.ca.

Additional Key Words and Phrases: Age-related differences, learnability, menu design, mobile devices, multi-layered interfaces, older adults, user study

ACM Reference Format:
Leung, R., Findlater, L., McGrenere, J., Graf, P., and Yang, J. 2010. Multi-layered interfaces to improve older adults' initial learnability of mobile applications. ACM Trans. Access. Comput. 3, 1, Article 1 (September 2010), 30 pages. DOI = 10.1145/1838562.1838563.
http://doi.acm.org/10.1145/1838562.1838563.

1. INTRODUCTION

Mobile computing devices, such as smart phones, offer older adults (ages 65+) a variety of useful tools and services to age more independently, both inside and outside their homes. For example, such devices can help older adults stay more connected with loved ones and caregivers, remember important health-related information, and entertain themselves with fun and stimulating games. Despite these benefits, however, older adults have been relatively slow to adopt mobile phones and other devices [Ofcom 2006] and reportedly find them difficult to use [Kurniawan et al. 2006]. In a 2005 UK survey of 2600 mobile phone owners (ages 16+) [Ofcom 2006], 88% of all respondents could store a new contact on their phone and 81% could send a text message. However, of the older phone owners surveyed (ages 65+), only a much smaller percentage could perform these two tasks (51% and 29%, respectively). This survey also found that about a quarter of older phone owners were interested in performing the two tasks but did not know how (22% and 25%, respectively). These survey findings suggest that many older adults are not able to learn to use mobile technology in ways they desire, contributing to an existing "grey digital divide" [Millward 2003].

There are a number of reasons why older adults may have more difficulty than younger adults in learning to use existing mobile technology. Compared to desktop computers, mobile device interfaces have relatively fewer buttons that are often used to perform multiple context-dependent functions, and smaller screens that limit the amount of information shown at once. Interacting with these device interfaces thus places demands on the user's working memory (online mental capacity for storing and processing information), which declines naturally with age [Fisk et al. 2009]. Older adults also have less computer and mobile device experience in general than young adults, decreasing opportunities for positive transfer. Furthermore, the natural decline in older adults' verbal and visual-spatial working memory makes it more difficult for them to learn and remember new computer skills [Echt et al. 1998; Fisk et al. 2009].

The goal of our research is to improve the learnability of mobile device applications for older adults, in order to lower the barrier to adoption of mobile technology by this population. Although learnability is known to be an important component of usability [Nielsen 1996], there is a lack of consensus on its definition [Grossman et al. 2009]. We define a system's learnability as the degree to which it enables novices with no experience with the interface to achieve mastery in performing basic and advanced tasks on the system. There are many ways to improve the learnability of mobile device applications (e.g., provide training from an experienced user) but we focused on independent learning and how it could be facilitated by improvements in the design of the user interface.

One approach to improving initial learning of software applications is the Multi-Layered (ML) interface approach [Shneiderman 2003]. ML interfaces can be designed to support learning such that novices first learn to perform basic tasks by working in a reduced-functionality, simplified layer (version) of the interface. Once users have mastered this layer or require more advanced functionality, they can transition to increasingly complex layers and learn to perform more advanced tasks. Thus ML interfaces can provide a form of scaffolding, a technique arising from Learner-Centered Design [Quintana et al. 2002] that temporarily supports novices to engage in activities normally outside their reach. Specifically, ML interfaces reduce an application's complexity (e.g., functions, content) during the learning process, thereby helping learners focus on key elements to begin performing tasks. Because the initial layers place fewer demands on the user's working memory, ML interfaces appear to be particularly applicable for older novice users with reduced working memory capacity. ML interfaces also support self-paced learning, which is preferred by many older adults [Fisk et al. 2009]. Although ML interfaces have the potential to help older adults to learn to use mobile technology, little research has studied the learning benefits of ML interfaces for older adults and no research has evaluated ML interfaces on mobile devices to improve learnability.

In this article, we present a controlled lab study with 16 older (ages 65+) and 16 younger (ages 29–39) adults, comparing ML and full-functionality mobile address book applications across four stages of learning. The main contribution is to identify age-related differences with ML interfaces: results show the ML interface provides greater benefit overall for older participants than for younger participants in terms of task completion time during initial learning, perceived complexity, and preference. We also study ML interfaces in a number of new ways, by focusing on a mobile device application instead of traditional desktop computer programs, by using a commercial mobile device to increase ecological validity in comparison to the simulations that are commonly used in mobile HCI research, and by comparing performance across four stages of learning that are directly relevant to ML interfaces.

2. RELATED WORK

Much research has looked at how to help older people learn to use new computer technology [Rogers et al. 1996; Echt et al. 1998; Morrell et al. 2000; Freudenthal 2001; Fisk et al. 2009; Hickman et al. 2007], and a number of recommendations have been made for improving a system's interface design and learning support for older users. Recommendations include minimizing working memory demands, providing cues and aids, not overloading older learners with too much information, and not requiring learners to make complex inferences or fill in gaps of missing information [Fisk et al. 2009]. ML interfaces used for learning inherently follow these recommendations. Older adults have also been found to experience greater frustration and anxiety than young adults in learning complex tasks, so it is recommended to include both immediate feedback and support to help build the user's confidence, and motivating exercises that lead to an attainment of mastery within a reasonable period of time [Fisk et al. 2009]. Further, when the training material itself is presented on the computer, it is important to focus not only on the pedagogical aspects of the training but also on ensuring that the computer interface is designed for older adults [Hawthorn 2005]. We followed these recommendations in the design of the tasks and mobile address book application used in our study.

To our knowledge no research has evaluated ML interfaces on mobile devices to improve learnability. However, a number of studies have found that ML interfaces can improve the learnability of desktop software [Carroll and Carrithers 1984; Catrambone and Carroll 1986; Findlater and McGrenere 2007], and can improve experienced users' satisfaction in using feature-rich software [McGrenere et al. 2002]. The Training Wheels (TW) interface [Carroll and Carrithers 1984], for example blocked advanced functions of a desktop word processor for novice users. The study's evaluation showed that novices were faster and less error-prone when learning to perform a basic task (creating and printing a letter) on the reduced-functionality interface compared to the full-functionality version of the application. This result was replicated in a follow-up study [Catrambone and Carroll 1986], which also evaluated the transition from the TW interface to the full-functionality interface: participants who learned the basic task in the TW interface performed a new advanced task more quickly on the full-functionality interface than participants who had only used the full-functionality interface. This performance benefit was attributed to the TW participants spending less time making mistakes.

Little research has looked at incorporating ML interfaces into applications to help older adults learn to use computer technology. Two exceptions are a desktop software tutorial called FileTutor [Hawthorn 2005] and a ML Web portal [Dickinson et al. 2007]. FileTutor was designed to help older adults learn to use a file management application [Hawthorn 2005]. FileTutor offers learners a number of exercises that use an embedded reduced-functionality version of Windows Explorer in which advanced functions are hidden. Although this reduced-functionality interface has not been formally evaluated (e.g., no pretesting, no comparison with a control interface), the study reported that older adults (ages 60–88), who had previously failed a file management learning module, used the tutorial to successfully perform many file management tasks on Windows Explorer with few help requests.

Researchers have also designed a ML desktop computer Web portal to help older adults navigate and search the Internet [Dickinson et al. 2007]. The Web portal layered both functionality and content into two layers (the first layer contained only basic functions and the second layer contained basic plus advanced functions). This ML Web portal was compared with a commercial Web portal of a major internet service provider. Participants, 11 older adults (ages 63–87), performed basic tasks in the first visit and a combination of basic and advanced tasks in the second visit. Participants made fewer errors over the two visits using the ML Web portal and commented positively on the simplicity of its interface. Participants in the ML interface condition also completed more given tasks in the first visit than those using the commercial Web portal.

The foregoing studies provide evidence that ML interfaces can improve the learnability of new technology in terms of errors, task completion time, and number of tasks completed successfully. There is also some preliminary evidence that ML interfaces help older adults to learn new technology. However, it is not clear how well results of past studies involving desktop applications generalize to mobile applications; performing mobile application tasks through the device's small screen and limited set of buttons is very different than performing desktop applications tasks. Our study extends past work by exploring the learning effects of ML interfaces on mobile applications and involving multiple age groups to identify age-related differences in these effects.

3. EXPERIMENTAL METHODOLOGY

The goal of this experiment was to assess older and younger adults' performance using a ML mobile application across four learning phases: *Basic Task Acquisition* on an initial reduced-functionality layer, *Retention* (performing basic tasks after a break), *Transition* to full interface (performing basic tasks, but on a second more complex full-functionality layer), and *Advanced Task Acquisition* on this second layer. As stated earlier, previous research has already shown that ML interfaces can be beneficial for younger adults, and we were particularly interested in studying whether those benefits would be the same or different for older adults. Assessing performance across four phases also allowed us to develop a more complete understanding of how ML interfaces impact learning, from initial usage to more experienced usage, and in the acquisition of tasks on different layers.

In this section, we start by describing our participants. We then present the ML mobile device application used in the study, beginning with the choice of mobile application (an address book) and its functions, followed by the design of the two interface layers. Afterward, we describe the tasks that participants were asked to perform and the overall experimental procedure.

3.1 Participants

We recruited 16 older adults (ages 65–81) and 16 younger adults (ages 21–36). Participants were recruited through posters at a local university, at libraries and senior centers, as well as via postings on an online classifieds site. Participants were prescreened over the phone in order to identify individuals with limited computer experience, no or very little experience with handheld computers and no experience with smart phone capabilities beyond making phone calls. Participants were also prescreened to be free of visual and physical impairments that would prevent them from operating a mobile device. Visual acuity was tested during the study using the Snellen pocket eye chart and all participants were found to have normal or corrected-to-normal eyesight. Four older participants, two assigned to the ML interface condition and the other two assigned to the control group, could not finish the study within the allotted time and had to be replaced (we discuss the details and implications in Section 5.5). Participants received a financial honorarium for participating.

Participants were randomly assigned to one of two interface groups: ML and control interface (described in Section 3.2). A Mann-Whitney U test revealed no significant differences with respect to age and number of years education between the two older adult interface groups (age: U=30.5, p=.88; education: U=21, p=.24) nor between the younger adult interface groups (age: U=29, p=.75; education: U=28.5, p=.70). However, between the younger and older groups, Mann-Whitney U tests revealed significant differences with respect to years of education (U=50.5, p=.003; the younger participants had more years of education) and years of mobile experience (U=39.5, $p<0.001$; the younger participants had more years of mobile experience), but no difference with respect to years of computer experience (U=99.5, p=.22).

As we did not prescreen participants on cognitive abilities, all participants completed a test battery so that we could characterize the groups and check for interface-group differences. We assessed participants' verbal working memory (Reverse Digit Span Test (RDST), [Wechsler 1981]), visual-spatial working memory (Corsi Block test, [Milner 1971]), and perceptual and motor speed (Digit Symbol test, WAIS-R). We also assessed participants' attitudes towards computers and the Internet (Technology Profile Inventory (TPI), [DeYoung and Spence 2004]). A MANOVA revealed no significant difference between the interface groups regarding these measures for older participants ($F_{4,11}$=.93, p=.48) as well as for younger participants ($F_{4,11}$=.51, p=.73). However, a MANOVA did reveal a significant difference between the two age groups in regards to measured cognitive abilities ($F_{4,25}$=25, $p<.001$). Follow-up ANOVAs revealed that younger participants outperformed older participants in tests that measured visual-spatial working memory ($F_{1,28}$=36.1, $p<.001$) and perceptual and motor speed ($F_{1,28}$=79.3, $p<.001$), as expected. Table I gives an overview to the descriptive data for age, gender, education, and cognitive abilities for all four experimental groups.

3.2 Interface Conditions

To explore the effects of learning on a ML interface, we compared it to a traditional, nonlayered interface, which we used as our experimental control. The ML interface had two layers: an initial reduced-functionality layer that presented only a small subset of functions and a full-functionality layer that included all functions available in the application.

We had the following two interface conditions.

(1) ML interface: participants learned the basic task set (described in Section 3.3) on a reduced-functionality layer and learned the advanced task set on a full-functionality layer.

(2) Control (nonlayered) interface: participants learned both basic and advanced task sets on a full-functionality layer.

3.2.1 *Mobile Application.* We chose a mobile phone address book (also known as phone book, contact list, contacts application) as the experimental

Table I. Means and Standard Deviations for Participant Characteristics (*N*=32)

	ML Interface M (SD)	**Control Interface M (SD)**
Younger Participants, *N* =16		
Age	24.4 (5.1)	22.4 (1.1)
Gender	3 male, 5 female	4 male, 4 female
# Years Education	15.9 (1.5)	16.3 (1.2)
Verbal Working Memory	6.3 (1.9)	7.6 (2.3)
Visuo-Spatial Working Memory	7.3 (1.8)	7.6 (1.3)
Perceptual and Motor Speed	74.1 (8.2)	77.5 (11.2)
Attitudes towards Computer Technology	112 (13.0)	116 (13.5)
Older Participants, *N*=16		
Age	71.6 (5.5)	70.8 (4.2)
Gender	8 female	4 male, 4 female
# Years Education	11.8 (3.7)	14.8 (3.4)
Verbal Working Memory	7.0 (2.6)	5.8 (1.2)
Visuo-Spatial Working Memory	5.0 (1.2)	5.6 (1.7)
Perceptual and Motor Speed	50.0 (6.4)	47.6 (7.8)
Attitudes towards Computer Technology	110 (15.5)	110 (17.2)

Notes:
M = mean, SD = standard deviation
RDST, verbal working memory: higher score = better memory, max. score: 12
Corsi Block Test, visuo-spatial working memory: higher score = better memory, max score: 12
Digit Symbol test, perceptual and motor speed: higher score = faster
TPI, attitudes towards computer technology: higher score = more positive attitudes, max score: 150

application for this study because both the older and younger participants were likely to be familiar with this application domain, having used a paper or electronic address book. Further, older adult mobile phone users have previously expressed interest in learning to use the address book on their phones but have had difficulty with functions such as adding a new contact [Ofcom 2006].

The experimental application was based on existing commercial mobile address books. The application consisted of two main screens (shown in Figure 1): a contact list screen and a contact details screen. The contact list screen allowed users to page through their contacts, with 6 names presented at one time. Selecting a name from the list led to a screen of contact details for that individual; users returned to the contact list by pressing the "Back to list" right soft key on the contact details screen. The address book was prepopulated with 24 contacts, all actors and actresses whose names were expected to be relatively familiar to our participants.

The two application screens each provided an "Options menu" (referred to hereafter simply as *menu*) to execute a variety of task functions. In our control interface, these two menus offered a total of 24 functions, which were chosen based on a survey of common features in 9 existing mobile address book applications (from Nokia, Motorola, Sony Ericsson, Sanyo, Blackberry, and Apple phones). Functions for *editing* an individual contact's information (e.g., edit, add custom ringtone) were placed in the Contact Details menu, while functions

Fig. 1. Contact List screen (left) and Contact Details screen (right) of mobile address book application used in the experimental study.

Table II. Options menu functions for the Reduced- and Full-Functionality Layers

	List of Functions in the Two Menus		
	Reduced-Functionality Layer	Full-Functionality Layer	
Contact List Screen	Call View Contact **New Contact**[1]	*(page 1)* Call Use Number **Send Message**[2] Send Contact Mark/Unmark View Contact	*(page 2)* **New Contact**[1] SIM Phone Book Synchronization Help Settings General Info
Contact Details Screen	**Edit Contact**[1] **Delete Contact**[1]	*(page 1)* Duplicate Contact **Edit Contact**[1] **Set Voice Dial**[2] Add Picture Set Speed Dial Categories	*(page 2)* **Add Custom Ringtone**[2] Set as Default Copy to SIM Copy from SIM PTT Options **Delete Contact**[1]

[1] Used in basic task set.
[2] Used in advanced task set.

for *using* the contact information to perform a task using a contact's information (e.g., call, send text message) were placed in the Contact List menu. Other miscellaneous functions (e.g., help, settings) were also placed in the Contact List menu. The menus, which were located in the bottom left corner of the screen, were accessed using the device's left soft key. All functions related to the experimental tasks (see Section 3.3) were implemented. The other functions were either implemented or displayed a "function not simulated" message when the user tried to execute them. The contents of the two menus are shown in Table II.

3.2.2 *Multi-Layer Design.* We created two distinct interface layers that primarily differed in the number of available functions. Following existing ML design conventions, the Reduced-Functionality layer, which a user would learn on first, only contained relatively *basic* functions while the Full-Functionality

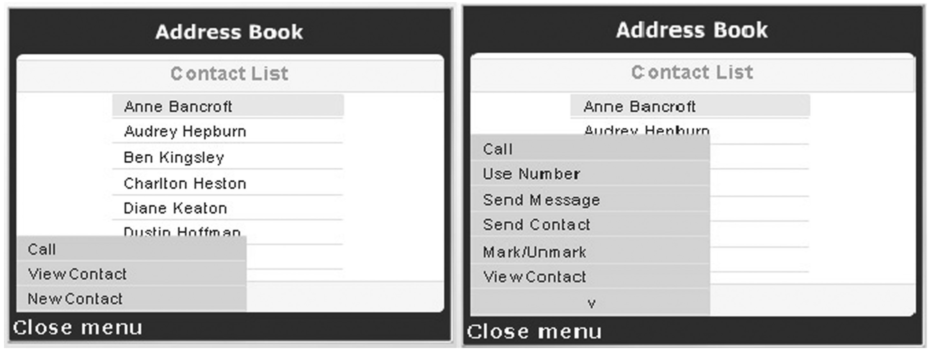

Fig. 2. Contact List screen, with its menu open: Reduced-Functionality layer (left) and Full-Functionality layer (right).

layer contained more advanced functions, in addition to the basic functions. Functions were classified as being basic or advanced based on an informal survey of mobile phone users (4 users ages 20–39 and 4 users ages 50+). Basic functions were ones that most of our surveyed mobile phone users reported learning first and thought were necessary for using the application. In contrast, advanced functions were seen by those surveyed as secondary functions that a user would learn to use after the basic functions. The Reduced-Functionality layer contained 5 functions (including all basic ones), split across the Contact List and the Contact Details menus, while the Full-Functionality layer contained 24 functions, split evenly across the two menus. Since a maximum of 6 functions could be shown at once, the menus in the Full-Functionality layer each contained two pages of functions. Table II shows the lists of functions for each layer and Figure 2 shows the Contact List screen menu for both layers.

Related functions were grouped together (e.g., *Duplicate Contact* positioned close to *Edit Contact*), which intermixed basic and advanced functions in the Full-Functionality interface and is representative of many mobile and desktop computer interface menus. As a result of this design decision, using the Full-Functionality layer sometimes required the user to page to the second page of a menu to find a basic task function. The main drawback of this approach is that the menu positions of the basic functions were not consistent across layers. An alternative approach is grouping functions by layer, with basic functions at the top of the menu followed by advanced function below. By grouping functions by layers, basic functions would remain in the same location regardless of the interface layer, which would likely help users transition between layers. However, grouping functions by layer often separates related functions, removing contextual information that helps users to interpret function names (e.g., the function *Use Number* shown in Table II is closely related to *Call*, but may be harder to correctly interpret being between the functions *New Contact* and *Send Message*). We chose the grouping-by-related-functions approach over the grouping-by-layer approach as it is representative of many user interface menus and has been used in a number of other research studies on

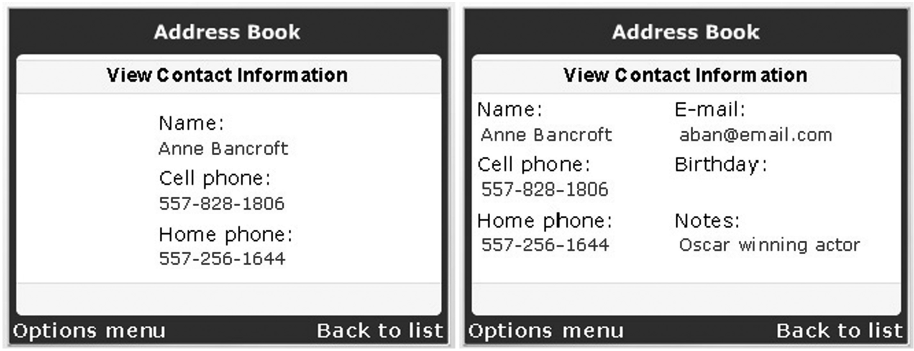

Fig. 3. Contact Details screen: Reduced-Functionality layer (left) and Full-Functionality layer (right).

multi-layered interfaces [McGrenere et al. 2002; Shneiderman 2003; Findlater and McGrenere 2007].

The only operational difference in performing the tasks on either layer was the number of steps required to navigate through the options menu to find a particular function. We define a *step* to be one device button press performed by the user. Other aspects of the tasks, such as scrolling through the contact list and entering text, required the same number of steps for both interface layers. In addition, the Full-Functionality layer's contact detail screen also had more visual complexity in the form of additional contact details (see Figure 3). Thus, any differences due to the interface conditions should be related to (i) the number of menu items, (ii) paging, and (iii) visual complexity. The application used in the study did not allow us to separately analyze the effect of each of these three ML design characteristics, but rather allowed us to evaluate the learnability of a multi-layered interface for a mobile application as a whole.

3.3 Tasks

Participants were asked to perform sets of basic and advanced tasks throughout the experiment. The basic set consisted of three tasks: adding a new contact into the address book, editing the information of a previously entered contact, and deleting a contact from the address book. The advanced set also consisted of three tasks: adding voice dialing to a contact (to phone the contact by speaking their name), sending a text message to a contact, and adding a custom ringtone to a contact (so the phone would produce a distinctive ringtone whenever that contact called). Participants repeatedly performed the same set of three tasks for a particular phase, but task-related information (e.g., contact name, phone number, text message) varied from attempt to attempt. Example wordings of the tasks given in the experiment session were "Add Meryl Streep's name and her home phone number 267-946-9907," and "Change Kevin Spacey's cell phone number to 468-243-2301."

The average number of required steps for performing basic task sets on the Reduced-Functionality and Full-Functionality layers was 75.2 steps and 90.2 steps, respectively. The average number of required steps for performing

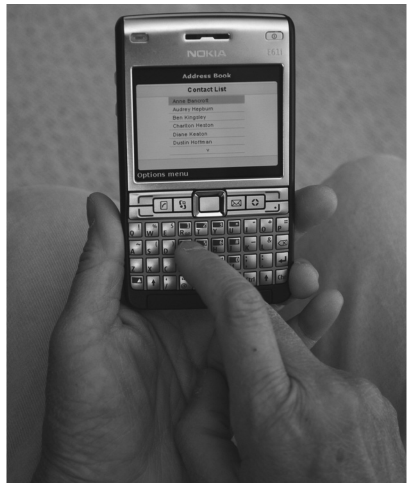

Fig. 4. An older adult holding the mobile device (Nokia E61i) used in this study.

advanced task sets on the Full-Functionality layer was 87.0 steps (see Table III for the required number of steps broken down by step type).

Depending on the experimental condition and progress through the four phases of the study, participants performed the tasks in either the Reduced- or Full-Functionality layer. Participants were not able to switch between layers on their own.

3.4 Apparatus

The experiment ran on a Nokia E61i device, shown in Figure 4. The E61i was chosen for its relatively large screen size (320x240 pixels, 5.7cm x 4.3cm) and QWERTY keyboard (button size: 0.9cm x 1.1cm). The E61i can also run Flash Lite applications (unlike many other devices with a similar form factor), which allowed us to quickly develop the address book prototype in Flash Lite 2.1. In our pilot studies, no participants reported difficulties reading the mobile application text off the device. A number of older participants reported some difficulty pressing the buttons but were comfortable using the eraser end of a pencil, which we provided in our study, to press each individual button. Four of the 16 older participants chose to use the pencil during the study, which appeared to help them to avoid pressing the buttons surrounding the target button and to increase typing accuracy; no differences in input speed were observed.

3.5 Procedure

Device Tutorial. All participants started the study session by completing an interactive tutorial, given on the device, to learn how to enter text with the keyboard and to use the soft keys and direction pad for navigation. Participants were given as much time as desired to familiarize themselves with the device buttons and could repeat the tutorial exercises. Completing the tutorial required around 5 minutes for younger participants and around 10 minutes for older participants; no participants chose to repeat the tutorial. Participants were asked to hold the device in any position they found comfortable (e.g., in hands, on table).

Basic Task Acquisition Phase. The set of basic tasks (i.e., add, edit, and delete a contact) was first described orally to participants. Participants were then asked to "learn to perform these tasks with as few extra steps as possible" (speed was not mentioned). Participants carried out a series of attempts (3 minimum) until they had mastered the tasks (see Section 3.6 for definition of mastery; there was no cutoff on the number of attempts). Each attempt consisted of performing the set of three tasks, with new contact information provided for each attempt. Participants in the ML condition completed this phase in the Reduced-Functionality layer, while those in the control condition used the Full-Functionality layer.

All participants performed the tasks in the same order. Tasks were given in written form on paper in large font (36-point Arial). No further instructions on how to use the address book were given, and help was only offered when participants were stuck for more than two minutes.

30-Minute Break from Interface. Participants completed the cognitive assessments listed in Section 3.1 and a distractor task (assembling a jigsaw puzzle). These tasks were intended to prevent rehearsal of what participants had just learned in order to permit an assessment of short-term retention.

Retention Phase. Participants were then asked to perform the basic tasks twice more (i.e., two attempts) on their assigned interface layer (Reduced-Functionality or Full-Functionality), this time "as accurately and quickly" as they could.

Advanced Task Acquisition and Transition Phases (Attempts Interleaved). All participants used the Full-Functionality layer in these two phases. The advanced set of tasks (voice dialing, text message, custom ringtone tasks) was described to participants. Participants were then asked to "learn to perform these tasks with as few extra steps as possible." No further instructions on how to perform the advanced tasks were given, and help was offered only when participants were stuck for more than two minutes. Participants carried out a series of attempts until they achieved mastery or until they had performed 10 total attempts (unlike in the Basic Task Acquisition phase, which had no cutoff).

To study the effect on performance of transitioning from a reduced-functionality to a full-functionality layer, participants were also asked to perform the basic task set four times (i.e., four attempts). We interleaved Transition phase attempts (of basic tasks) and the Advanced Task Acquisition phase attempts (of advanced tasks) (i.e., A(dvanced Task Acquisition) T(ransition) A T A T A T A A A. . .). We used this task set order instead of having all Transition attempts before Advanced Task Acquisition attempts so that participants in the ML condition would have minimal familiarity with the Full-Functionality layer for both types of attempts. This simulates the expected use of a ML interface, where users would transition to the more complex layer when they want to perform more advanced tasks, but would continue to use basic functions.

Posttask Interview. Finally, we conducted a semi-structured interview with each participant.

A number of older participants were noticeably stuck for more than two minutes while learning to perform tasks. These older participants, an equal number from both interface conditions, received help from the experimenter. Specifically five older participants from the ML interface condition required help a total of 15 times (Basic Task Acq. Phase: 2 times, Adv. Task Acq. Phase: 13 times) and five older participants from the control condition needed help a total of 16 times (Basic Task Acq. Phase: 7 times, Adv. Task Acq. Phase: 9 times). We note that participants in the control condition received more help while learning basic tasks and those in the ML condition received more help while learning advanced tasks, which was expected as those were the phases when participants first performed tasks on the more complex Full-Functionality layer.

Study sessions could run to a maximum length of 4 hours, to allow participants ample time to complete the study but also to prevent participants and the experimenter from getting overly fatigued. Younger participants generally took 2 hours to complete the study, while older participants took 3–4 hours.

3.6 Measures

Quantitative performance measures to assess the learnability of the two interfaces included the total number of attempts before mastery, the number of steps (i.e., button presses) to perform a task, and task completion times. We also calculated the number of *extra steps* (i.e., errors) for a task, which was the total number of steps taken by a participant to complete a task minus the minimum number of steps required to complete that task.

Mastery was defined as being able to perform a task set twice in a row, using no more than 20% additional steps over the minimum number of steps required. Based on our pilot study data, 20% additional steps seemed like a reasonable mastery threshold that required users to perform a task without many extra steps but would allow for some flexibility. To increase our confidence that the user had mastered the task set and was not simply lucky, we required the user to perform the task set under the mastery threshold twice in a row.

Subjective quantitative measures, gathered by a questionnaire with 6-point Likert scale items, focused on the perceived ease of learning, confidence in application use, perceived application complexity, and perceived workload.

3.7 Study Design

A 3-factor mixed design was used: 2 age groups (younger, older; between-subjects) by 2 interface conditions (ML, control; between-subjects) by required attempts (this value differed by phase as described next; within-subjects). For the Basic Task Acquisition and Advanced Task Acquisition phases, we only analyzed the initial 3 attempts. Participants were asked to perform at least 3 attempts, and although some participants mastered the given task set in 10 or more attempts, some mastered the set within the 3 attempts and thus did not have performance data for attempts 4–10 for us to analyze. We analyzed the 2 attempts performed in the Retention phase and the 4 attempts performed in the Transition phases. An equal number of participants from each age group were randomly assigned to either interface condition.

We also analyzed the subjective questionnaire data using a 2 (age group) x 2 (interface) x 2 (phase: Basic Task Acquisition vs. Advanced Task Acquisition) mixed design.

3.8 Hypotheses

We tested the following hypotheses.

H1 Basic Task Acquisition. The ML interface's Reduced-Functionality layer (compared to the control) better helps users to master the basic task set, in terms of fewer extra steps, shorter task completion times, and fewer attempts to reach mastery.

H2 Retention. The ML interface's Reduced-Functionality layer (compared to the control) helps users to better perform the basic task set mastered 30 minutes previously, in terms of fewer extra steps and shorter task completion times.

H3 Transition. Transitioning from the Reduced-Functionality layer to the Full-Functionality layer negatively affects basic task set performance (compared to no transition required in control interface) in terms of more extra steps and longer task completion times.

H4 Advanced Task Acquisition. The ML interface's Reduced-Functionality layer (compared to the control) better helps users to master the advanced task set on the Full-Functionality layer in terms of fewer extra steps, shorter task completion times, and fewer attempts to reach mastery.

H5 Greater Benefit of ML Interface for Older Adults. The performance benefits provided by the ML interface are greater for older adults than for younger ones in the Basic Task Acquisition, Retention and Advanced Task Acquisition phases.

3.9 Data Analysis and Treatment

We used ANOVAs to test our hypotheses. We report effects which were significant ($p < .05$) or which represent trends ($.05 <= p < .10$). Whenever a statistically significant interaction was found, we followed up with post-hoc pairwise comparisons, using a Bonferroni correction to protect against Type I error. In addition, Greenhouse-Geisser corrections were used when sphericity was an issue; using this correction can result in degrees of freedom that are not whole numbers. We also report the partial eta-squared (η_p^2) statistic, a measure of effect size; to interpret this statistic, 0.01, 0.06, and 0.14 are considered small, medium, and large effect sizes, respectively [Cohen 1988].

A small percentage of the participant interactions with the mobile application was not logged due to a technical issue, resulting in the loss of extra step and completion time data for 1.6% of tasks. This small percentage of missing data was spread broadly over participants in each of our experimental conditions and fully discarding incomplete participant data (i.e., listwise deletion) would have resulted in a loss of 12% of performance data. To make use of data from all participants, we imputed the missing data using scores from the attempt immediately preceding the affected attempt. Although it is known that this single imputation procedure can increase errors in significance testing, we chose this imputation procedure because it is simple to implement and more suitable in cases where only a very small portion of the data is missing [McKnight et al. 2007; Scheffer 2002]. To validate our data treatment approach, we ran ANOVAs both with missing data imputed and again with participants' data fully discarded whenever data were missing, and found that our method for dealing with missing data did not alter the pattern of effects. Thus, we report results from our ANOVAs on complete participant data sets with missing data imputed, but note with an * effects and trends that were found to be no longer statistically significant ($p > .05$ and $p > .10$, respectively) when incomplete participants' data sets were discarded.

4. RESULTS

4.1 Performance

4.1.1 *Overall Summary.* We summarize participants' performance data (i.e., extra steps, task completion times) below to visualize how each experimental group performed in the four phases. Almost all participants took fewer than 10 attempts to master each of the two task sets, so we present performance data up to the 10th attempt for the Basic and Advanced Task Acquisition phases. For those participants who achieved mastery in fewer than 10 attempts, we used the completion time and extra step values from their last attempt as a conservative approximation of what their performance would have been on further attempts (these approximations were not used in our ANOVAs).

The visual summaries of the overall data shown in Figure 5 reveal a number of interface- and age-related differences, all of which were either statistically significant or showed an informative trend in the analysis results (reported in the following sections). Specifically, younger participants using the ML interface initially took on average fewer extra steps in the Basic Task Acquisition and Retention phases than the younger participants in the control group. Younger participants' task completion times were not affected by interface condition in any of the phases. In contrast, older participants using the ML interface initially took less time, in addition to fewer extra steps, in the Basic Task Acquisition and Retention phases than the older participants in the control group. In the Advanced Task Acquisition phase, younger participants performed similarly regardless of interface condition, as did older

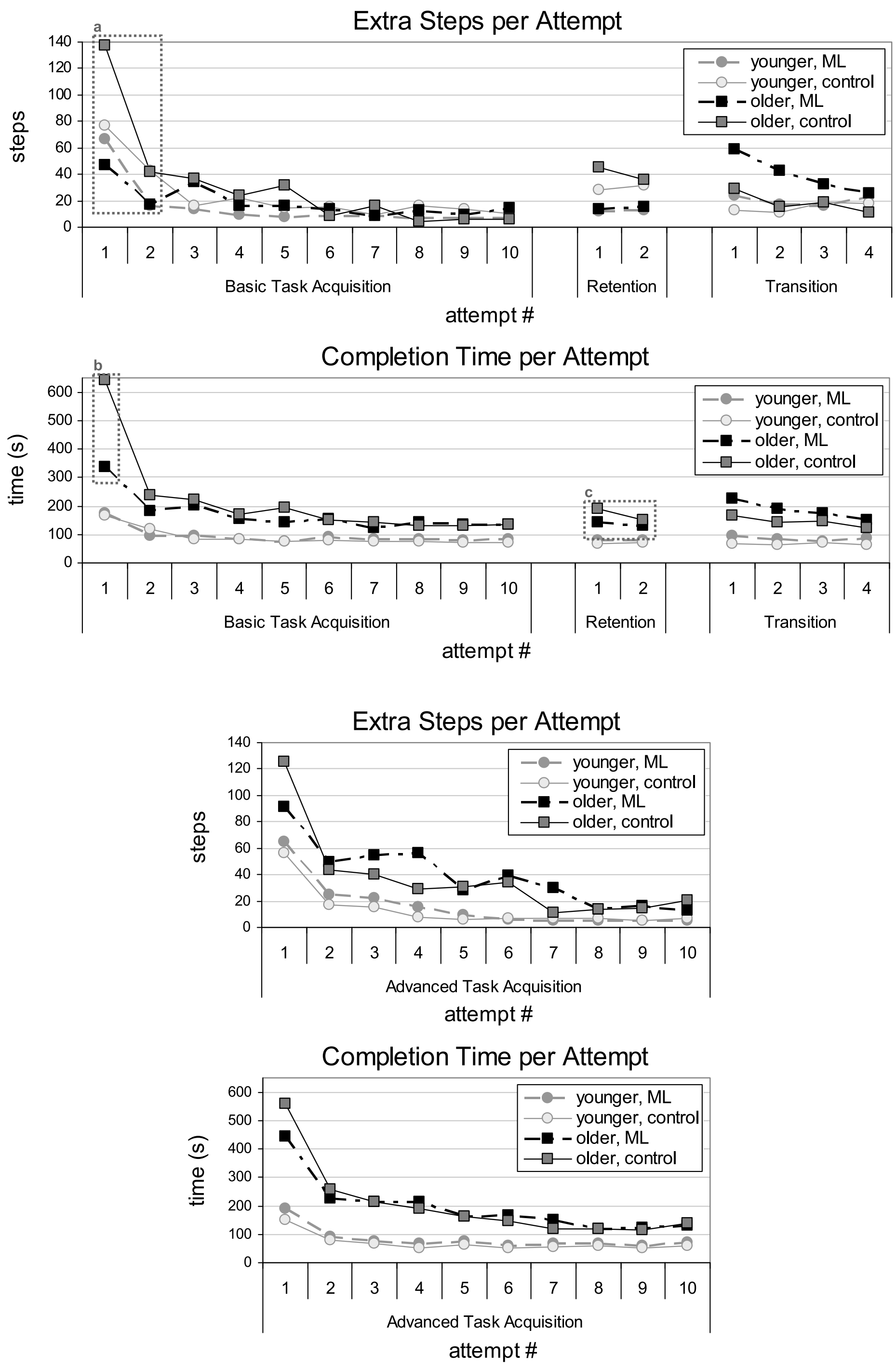

Fig. 5. Extra steps and time data for attempts in all four ML learning phases. Callout areas a, b, and c are referred to in the running text (N=32).

participants. As expected, older participants also consistently took more time over all four phases than the younger participants to complete task sets. The following sections present the statistical analysis of the performance data for each of the four phases.

4.1.2 *ML Interface Helped Initial Basic Task Acquisition, Particularly for Older Users.* Participants in the ML condition, compared to those in the control condition, performed fewer extra steps in the first three attempts of mastering the basic task set (main effect of interface*, $F_{1,28}$=5.7, p=.024, η_p^2=.17), supporting H1. As mentioned earlier, only the first three attempts were analyzed, since not all participants completed more than three attempts.

A trend in the data suggested that the ML interface may have provided a greater performance benefit for older adults in terms of fewer extra steps (3-way interaction of interface, age, and attempt*, $F_{1.2,34.2}$=2.8, p=.0996, η_p^2=.09). Inspection of the data showed that older participants in the ML condition took significantly fewer extra steps than those in the control condition for attempts 1 and 2 (attempt 1: 47 vs. 138 steps; attempt 2: 17 vs. 42 steps; shown in Figure 5(a)). Younger participants in the ML condition also took significantly fewer extra steps than those in the control condition but only in attempt 2 (16 vs. 43 steps). No significant differences were found between interface conditions for the other attempts. These findings offer support for H5, but they need to be interpreted cautiously and require more work to substantiate.

Analysis of task completion times showed that older participants benefited more from the ML interface than did younger participants. We found a 3-way interaction among interface, age and attempt on task completion time ($F_{1.3,36.1}$=6.7, p=.009, η_p^2=.19). Post-hoc pairwise comparisons revealed that in the first attempt, older participants in the ML condition took significantly less time than older participants in the control condition (338s vs. 641s, p=.002; shown in Figure 5(b)). No other significant differences were found for older participants in the second or third attempt. Younger participants in both interface conditions took similar amounts of time in their first three attempts to perform the basic task set. These findings support H5, but do not support H1. As expected older users took significantly more time overall than younger ones to complete the basic task set (main effect of age, $F_{1,28}$=34.3, p<.001, η_p^2=.55).

Although participants using the ML interface mastered basic task sets in fewer attempts on average than those using the control interface (shown in Figure 6, left), no statistically significant interface- or age-related differences were found in the number of attempts participants took to master the basic task set. This finding provides no support to H1.

4.1.3 *ML Interface Helped Participants Retain Mastery of Tasks.* The ML interface helped both age groups to retain their mastery of the task sets. Specifically, participants using the ML interface required significantly fewer extra steps, than those in the control group, to perform the basic task sets 30 minutes after mastering them (main effect of interface, $F_{1,28}$=25.0, p<.001, η_p^2=.47). This finding supports H2. No significant interaction of interface and age was found, offering no support to H5.

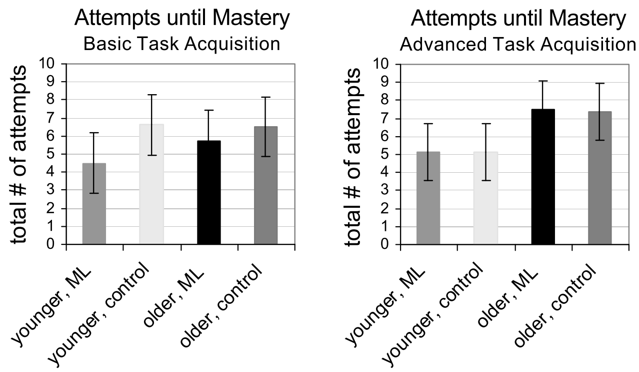

Fig. 6. Left: Number of attempts until mastery in the Basic Task Acquisition (N=32). Right: Advanced Task Acquisition phases (N=26).

A trend in the data suggested that older participants in the ML condition took less time to complete the Retention phase tasks than those in the control condition (interaction of interface and age on completion time*, $F_{1,28}$=3.8, p=.060, η_p^2=.12). Inspection of the data (shown in Figure 5(c)) revealed that older participants in the ML condition took less time than those in the control condition (137sec vs. 170sec) and there was only a minimal difference between interface conditions for younger participants. These findings offer support for H5, but need to be interpreted cautiously. As expected, older users took significantly more time than younger ones to complete the basic task set (main effect of age, $F_{1,28=}$57.4, p<.001, η_p^2=.67).

4.1.4 *Transition in Using ML Interface Led to a Negative Impact on Performance.* In the Transition phase, we looked at how participants in the ML interface condition performed on the Full-Functionality layer after learning on the Reduced-Functionality layer. We compared the performance of participants in the ML interface to those in the control who did not have to transition to another interface layer.

As hypothesized, the ML condition had a negative impact on the completion of basic tasks after participants transitioned to the Full-Functionality layer. ML participants took more extra steps and more time to complete basic tasks on the Full-Functionality layer than the control participants (main effect of interface on extra steps, $F_{1,28}$=5.8, p=.023, η_p^2=.17; large main effect of interface on completion time, $F_{1,28}$=7.3, p=.012, η_p^2=.21). These findings support H3. Although Figure 5 shows a larger impact of interface for the older participants compared to the younger participants, no significant interaction between interface and age on performance (i.e., extra steps, task completion times) was found. However, we did find that older participants took significantly more steps and were slower than younger participants as expected (main effect of age on extra steps*, $F_{1,28}$=4.5, p=.043, η_p^2=.14; main effect of age on completion time, $F_{1,28}$=69.2, p<.001, η_p^2=.71).

4.1.5 *ML Did Not Help Nor Hinder Advanced Task Acquisition.* Performance during the first three attempts for learning the advanced task sets (on the Full-Functionality layer) was similar for both interface conditions. As mentioned earlier, only the first three attempts were analyzed as participants performed a minimum of three attempts. No significant effects of interface condition were found on extra steps taken or task completion times, offering no support for H4 or H5. However, we note that older participants took more extra steps and more time to complete the advanced task sets during the first three attempts (main effect of age on extra steps, $F_{1,28}$=18.0, $p<.001$, η_p^2=.67; main effect of age on completion time, $F_{1,28}$=57.3, $p<.001$, η_p^2=.39).

Participants in the two interface conditions mastered the advanced task set in similar number of attempts. One younger participant (in the control condition) and five older participants (2 in ML condition, 3 in control condition) did not master the advanced task set within 10 attempts. Of those who did master the task set within 10 attempts, the ML participants required slightly more attempts to do so (M=5.8, SD=1.9) than the control condition (M=5, SD=1.3; see Figure 6, right); however, no significant effect of interface was found, offering no support to H4.

4.1.6 *Summary of Hypothesis Testing.*

- H1 partially supported. The ML interface's Reduced-Functionality layer (compared to the control) helped participants to master basic task sets in significantly fewer extra steps (attempts 1–3). However, our ML interface did not help participants to master basic task sets in fewer attempts nor helped younger participants to master these task sets in less time.
- H2 supported. The ML interface's Reduced-Functionality layer (compared to the control) helped participants to better retain their task mastery in terms of significantly fewer steps.
- H3 supported. Transitioning from the Reduced-Functionality to Full-Functionality layer in the ML interface (compared to no transition when using the control) negatively impacted participants' performance in terms of significantly fewer extra steps and lower task completion times.
- H4 not supported. We found no impact of interface on number of extra steps, task completion time or number of attempts in acquiring advanced tasks.
- H5 partially supported. The ML interface's Reduced-Functionality layer (compared to the control) helped older participants significantly more than younger ones to master basic task sets in less time (attempts 1–3). Trends in the data suggest that the ML interface's Reduced-Functionality layer (compared to the control) helped older participants more than younger ones to master basic task sets in fewer steps (attempts 1–2); and better retain their mastery of basic tasks in terms of task completion times.

4.2 Performance—Additional Analysis

We performed an additional exploratory analysis to gain insight into which types of steps (i.e., button presses associated with a particular aspect of the

Table III. Minimum Number of Steps (Average) Required to Complete the Basic and Advanced Task Sets, Classified by Step Type

	Minimum Number of Steps		
	Basic Task Set		Advanced Task Set
Step Type	Reduced Func. Layer	Full Func. Layer	Full Func. Layer
Contact Navigation	23.8	23.8	43.6
Menu	9.0	24.0	16.0
Function	6.0	6.0	12.2
Typing	36.4	36.4	15.2
TOTAL	**75.2**	**90.2**	**87.0**

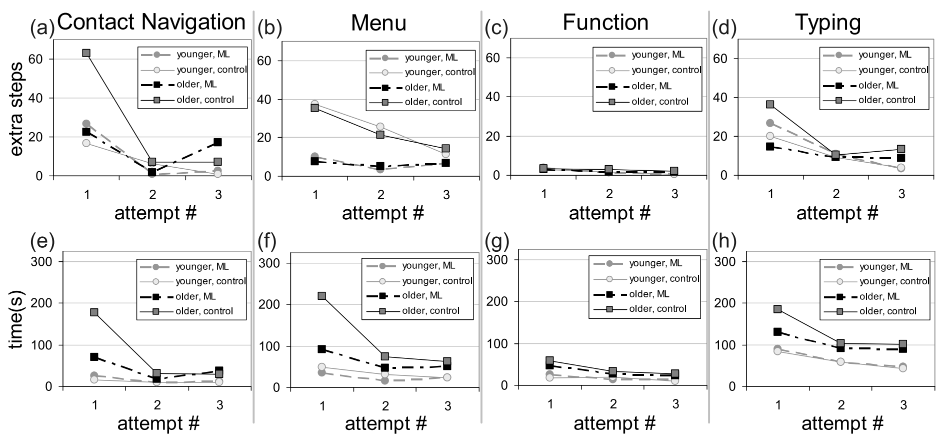

Fig. 7. Extra steps and completion times data for first three attempts in Basic Task Acquisition phase, separated by step type (N=32).

interface) contributed most to differences in performance. We classified all possible steps into one of four types: (1) *contact navigation*: navigating through the contact list and data fields of an individual contact; (2) *menu*: opening and navigating through the options menu, and executing a function in the menu; (3) *function*: function-specific steps (e.g., choosing a ringtone, saving changes); and, (4) *typing*: entering and correcting text. Table III shows extra steps and task completion times broken down by these four categories.

Separating performance data (i.e., extra step, task completion times) into the four types provides insight into which parts of the task were more difficult during the Basic Task Acquisition phase (mean data shown in Figure 7). As expected based on the differences in Table III, participants using the ML interface took fewer extra menu steps than those in the control group (Figure 7, graph b)). More interesting, however, was that completion times differed for the two age groups: younger participants spent similar lengths of time on menu steps regardless of interface condition, whereas older participants using the ML interface took less time using menus than those in the control group. Thus menu-related extra steps appear to negatively affect task completion time for older participants but not for younger participants.

The performance data separated by step type also suggest that, in comparison to the control interface, the ML interface helped older participants navigate through contact lists/details and type text. These differences were unexpected because contact navigation and typing required the exact same sequence and number of steps in both interface conditions. Older participants using the ML interface, compared to the control, took fewer extra steps and less time on both contact navigation (Figure 7, graphs (a) and (e)) and typing (Figure 7, graphs (d) and (h)).

Looking across the four phases, older and younger participants generally performed a similar number of extra function- and typing-related steps but older participants consistently took more time than younger ones. This was likely due to differences in older and younger participants' perceptual and motor speed.

4.3 Perceived Learning Experience

At the end of the session, participants used a 6-point Likert scale to rate how much they agreed with statements related to learning on their assigned interface. Interface- and age-related differences were found on perceived application complexity, ease in remembering function location, and frustration.

Although participants generally disagreed that they "felt overwhelmed by the complexity of the address book program," older participants found the ML interface to be significantly less complex than the control interface (a significant 2-way interaction of interface and age: $F_{1,28}$=4.2, p=.049, η_p^2=.13). Older participants rated the ML interface (1.4 out of 6) lower than the control interface (2.3) on complexity (significant difference found in post-hoc pair-wise comparisons, p=.011) while younger participants found the complexity of both interfaces to be similar (1.8 and 1.8, respectively).

Participants in the control group appeared to find it easier to remember function locations for both basic and advanced functions while those in the ML interface group found remembering function locations more difficult after transitioning to the Full-Functionality layer. Specifically, participants in the ML and control conditions generally agreed after mastering the basic task set that "it was easy to remember where all the [needed] functions ... were located" (4.6 and 4.3 out of 6, respectively); however, after mastering the advanced set of tasks, participants in the ML condition dropped to neutral while ratings for participants in the control condition remained positive (3.6 and 4.4, respectively; a significant 2-way interaction of interface and phase, $F_{1,28}$=8.9, p=.006, η_p^2=.24; significant difference found in post-hoc pairwise comparisons, p=.011). This difference may be due to interference between the two layers that participants in the ML condition had to learn.

Although all participants disagreed that "performing the basic task set was frustrating," a trend in the data suggested participants in the ML condition found performing the basic task set less frustrating compared to participants in

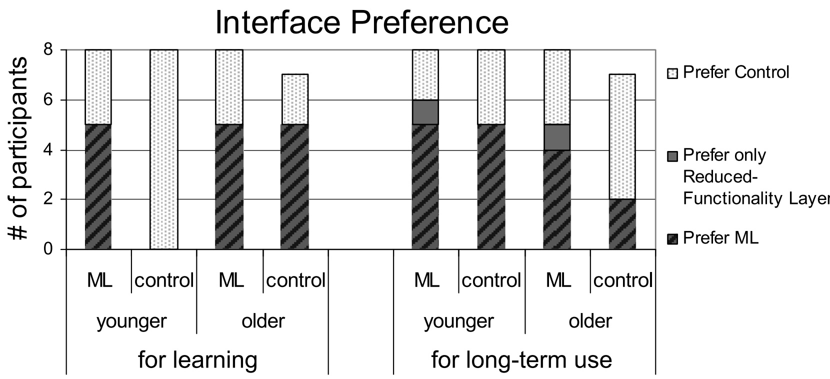

Fig. 8. Participants' preferred interface for learning and for long-term use (N=31).

the control condition (1.8 and 2.4 out of 6, respectively; main effect of interface, $F_{1,28}$=4.2, p=.0502, η_p^2=.13).

4.4 Overall Preference

In the post-task interview, we described both the ML and control address book interfaces to participants and asked participants to indicate their preference. Specifically, participants were asked to choose an address book interface (including one that consisted only of the Reduced-Functionality layer) they would most like to use if they had to: (a) learn to perform the task sets over again, and (b) use the application long-term. They were also asked to explain why they chose a particular interface. Figure 8 shows a summary of participant preferences. Note that one older participant in the control condition completed almost the entire study but did not have time to state his interface preferences in the allotted time, and thus we summarize data for only 15 older participants.

4.4.1 *Most Older Participants Prefer ML Interface for Learning, Mixed Preference for Younger Participants.* The majority of older participants (10/15) reported they would prefer to use a ML interface if they had to learn the task sets over again. The main reasons for this preference were that the Reduced-Functionality layer was perceived as being "simpler" and "easier," and that learning on a simpler version allowed them to become comfortable before moving to a more complex one. Interestingly, of those in the control condition, the majority of the older participants (5/7) indicated a preference for using the ML interface for learning, while none of the younger participants did; thus older participants perceived value in learning on ML interfaces, even without having used one, whereas younger participants did not. One older participant who preferred learning on the ML interface explained, "learning [the Full-Functionality layer] ... is more complex, so it's more difficult to get ... the hang of how the thing works ... going from simpler to complex [is] a natural progression of learning" (P19, ML).

By contrast, some older participants (5/15) reported they would prefer learning on a nonlayered interface over the ML interface. The primary reason for this preference was that the nonlayered interface did not require the user to learn how to use one layer and then later relearn another. Having different menu content for each layer also meant that menu items would likely be in different positions in the menu depending on the layer, which a number of older participants felt would be confusing. One older participant commented, "if you mastered these [functions in Reduced-Functionality layer] and then go to these [in Full-Functionality layer], then you've got two things to worry about" (P26, control condition).

The majority of younger participants in the ML interface condition (5/8) reported that they would prefer learning with the ML interface, while none of the younger participants in the control condition reported this preference. Younger participants in the ML interface condition preferred learning on the ML interface as it allowed them to quickly learn to perform basic tasks. Younger participants in the control condition reported preferring the nonlayered control interface because the address book interface was overall easy to learn to use and they preferred learning the functions all at once without having to relearn when transitioning to a new layer.

4.4.2 *Mixed Interface Preference for Long-Term Use.* Older participants reported a mix of preferred interfaces for long-term use. Over half of the older adults (8/15) would prefer using a nonlayered interface over the long run. The main reason for this preference was that having all functions in one layer made it easier to find a function. In fact, some participants commented that it would be better if all functions were in one menu even if more scrolling and menu pages were required, as opposed to one menu for the contact list screen and another for the contact details screen. Six of the older participants would prefer to use a ML interface over the long run, as long as the Reduced-Functionality layer had all the functions that they would commonly use. P27 (ML condition), an older participant, commented that she "would rather have [the ML interface, where the Reduced-Functionality layer] shows commonly used functions ... if I set [the Reduced-Functionality layer] up with my personal preferences, then that's probably what I would use most of the time." Further, we note that over half of the older participants in each of the interface conditions (ML or Reduced Functionality layer only: 5/8; control: 5/7) expressed a preference for the interface they had used, so the interface used in the study may have had some influence on their long-term preference. However, while neither interface was preferred by a large majority of the older participants for long-term use, their comments offer important design-related insights.

The majority of younger participants (10/16) indicated a preference for the ML interface for long-term use. The main reason for this preference was they could imagine the ML interface helping them be more efficient in performing common tasks. In contrast, five younger participants preferred the nonlayered control interface for long-term use because this interface allowed access to all functions without requiring any switching between layers.

5. DISCUSSION

5.1 ML Interfaces Help Users Learn Basic Mobile Application Tasks

As hypothesized, our study found that ML interfaces can help both younger and older adults to learn to perform task on a mobile application. Specifically, our ML interface's Reduced-Functionality layer (compared to the control) helped participants during initial attempts to learn to perform basic task sets in fewer extra steps. This finding is consistent with findings of past studies on desktop ML interfaces [Carroll and Carrithers 1984; Findlater and McGrenere 2007] and Dickinson et al. [2007], which focused on older adults. We also found that ML interfaces helped users to perform tasks (in terms of fewer steps) that they had mastered 30 minutes earlier.

Although we hypothesized that learning on the ML interface's Reduced-Functionality layer would improve participants' performance in learning the advanced task set on the Full-Functionality layer, our results did not support this hypothesis. Based on responses in our semi-structured interview, most of our older participants appeared to have difficulties forming a mental model with either interface of how the menus worked. Thus, they could not use such a model to help them learn the advanced task set. Although our younger participants seemed to be better able to form a mental model of the options menus, their performance on advanced tasks was similar on the two interfaces, perhaps because the tasks were too easy for them to master.

Our ML interface's lack of benefit for learning advanced tasks stands in contrast to Catrambone and Carroll's [1986] finding that their participants were able to more quickly perform an advanced task after acquiring a basic task on the reduced-functionality Training Wheels interface. This discrepancy may be due to differences in ML interface design as the Training Wheels interface blocked but still showed advanced functions, which helps users to transition from their initial interface layer to a full-functionality interface. The discrepancy in findings could also be due to differences between the studies, such as the type of application, tasks, and learning process prescribed in the two studies. For example, our study had participants repeatedly perform the same type of short mobile address book tasks until mastery, while Catrambone and Carroll had participants perform longer word processing tasks (one basic and one advanced) on a desktop computer until completion. More work is needed to better understand the types of basic and advanced tasks, application and learning process that would most benefit from ML interfaces.

5.2 ML Interfaces Offer Performance and Preference Benefits to Older Users

Prior to this study, we expected that the ML interface would benefit older adults more than younger ones. Although we did not find a benefit on all measures, we did find evidence in participants' performance data that ML interfaces help older adults more than younger ones during the initial learning process. For example, older participants using the ML interface were able to learn to perform the basic task set for the first time in 90 fewer extra steps and 5 fewer minutes, on average, than those in the control condition. Even after mastering the basic task set and then taking a break, older participants using the ML interface were able to perform the basic task set in 32 fewer extra steps and 47 fewer seconds, on average, than those in the control condition.

We predict that this initial performance benefit of ML interfaces would reduce one of the barriers older adults experience in adopting mobile technology. Making errors on a new interface often requires a high cost for recovery, which can be particularly frustrating for older adults [Birdi and Zapf 1997]. It is not clear from the literature how many attempts and how much time an older adult will spend on learning new technology before abandoning it. However, it is likely that the first few attempts are critical for novices to gain or lose confidence in their ability to learn to use the device. Minimizing the number of extra steps and time required to perform tasks on new technology will reduce frustration and increase the chance of technology adoption by older adults.

Participants' subjective data were in line with the quantitative performance results. Most older participants preferred learning basic tasks on the ML interface because it was perceived as simpler and easier to use, which is consistent with findings by Dickinson et al. [2007]. By contrast, younger participants did not prefer the ML interface for learning. Many older participants also found that learning on a simpler layer first before progressing to a more complex layer was a more natural and comfortable way to learn. Participants in the ML condition were also less frustrated during the learning process than those in the control condition.

The reduced complexity of the initial ML interface layer seems to have improved the initial learnability of our mobile address book application for older adults. Many older participants commented that when using the Full-Functionality layer, they often forgot which menu a function was located in. We frequently observed older participants closing a menu when using the Full-Functionality layer and immediately reopening the same menu to read through its items; this was not observed in the Reduced-Functionality layer where menus only had 2–3 items. Older participants, as well as some of the younger ones, commented that entering text required considerable mental effort and that it was challenging to remember which button to press to switch to the appropriate text entry mode. The reduced complexity of the initial ML layer appeared to make it substantially easier for older participants to learn to use the menus and other parts of the interface (e.g., entering text). Younger participants did not appear to experience this benefit; they performed similarly regardless of interface, perhaps because the mobile application and tasks were so simple that it did not matter which interface they used. This age-related difference is consistent with past studies on younger and older adults' performance on mobile device tasks [Ziefle and Bay 2005; 2004] and is likely due in part to differences in visuo-spatial working memory between our two age groups.

Since ML interfaces provide an initial benefit to learning basic tasks on a reduced-functionality layer and no hindrance to learning advanced tasks on a full-functionality layer, ML interfaces can be used to increase the overall learnability of the application, particularly for users who do not want or need to learn to perform advanced tasks. For example, many of our older participants commented that they normally would not need voice dialing and other advanced features in their daily lives; thus, a ML interface with basic task functions in its reduced-functionality layer would be more learnable for these participants compared to users who also need to use advanced tasks. Further, participant feedback did not suggest that providing a third layer would add any value.

5.3 Simplicity Valued

When choosing an interface to use, we found that older participants generally chose the one they perceived as being simpler. However, simplicity appeared to carry different meanings for learning a new application compared to using it long term. The ML interface was preferred for its simpler Reduced-Functionality layer. In this case, simpler referred to a reduced amount of information (e.g., available functions) in the interface to learn, while trying to perform a new task. For long-term use, older participants still chose the interface that they perceived as being simpler, but their interface preferences were mixed. Those who preferred the ML interface for long-term use wanted all commonly used functions in the Reduced-Functionality layer and felt that not having the "fancy stuff" in this layer would make the ML interface simpler. In contrast, older participants who preferred the nonlayered control interface felt that having all functions in one layer made it simpler to search for a desired function; one participant commented that he preferred scrolling through a long options menu over remembering where, amongst several menus, a function was located.

A design approach that arose from our participant interviews and may meet the needs of both groups of older adults would be a personalized ML interface that allows users to choose which functions to place into the reduced-functionality layer. Users could place commonly used functions or all functions that they would ever want to use into the reduced-functionality layer, as they desired. McGrenere et al. [2002] evaluated this approach implemented on a desktop word processing application, which was found to help participants better navigate menus and learn the application. McGrenere et al. also found that participants who favoured a simpler interface were willing to take the time to personalize it. Based on these findings, we expect that older users can benefit from a customizable ML mobile interface.

5.4 ML Mobile Application Design

We have found that designing a ML mobile application requires addressing similar challenges as designing a ML interface for the traditional desktop platform. For example, designers for both mobile and desktop platforms need to carefully determine how many layers to implement and which functions should be included in the different layers. We chose to start with a two-layered application as no previous study had formally evaluated the learning benefits of a ML mobile application. For this application, we surveyed existing users to determine which functions they thought they should learn first and placed these functions in the initial layer. Alternative criteria for selecting initial layer functions include choosing the most commonly-used functions, the simplest functions, or functions that are the easiest to learn.

While designing ML mobile and desktop applications have similar challenges, working within the constraints unique to mobile user interfaces requires special design considerations. For example, once the function sets for each layer have been determined, the mobile application designer needs to decide how the functions will be shown to the user on the mobile device's small screen. In our application, users access functions through menus and we needed to determine how to order the functions in these menus. This ordering is particularly important on mobile applications as the device's small screen size limits how many functions can be shown at once; if there is not enough space to show all functions, then functions may need to be place on different screens/pages. We chose to group related functions together, which is common in many mobile and desktop applications. As discussed in Section 3.2.2, an alternative approach would have been to group functions by layer so that basic functions are at the top of the menu, followed by advanced functions, which may better help users to transition from the first layer to the second.

There are also several options for setting the visibility of functions that are not used in a layer. For example, advanced functions in an initial layer can be hidden or marked as being disabled [Findlater and McGrenere 2007]. Hiding functions that are not used in a layer was the best approach for our application as it minimized paging and visual complexity in the initial layer's menu.

The designer also needs to implement some mechanism for the user to switch between the multiple layers. Although we did not implement this in our application, this mechanism could be implemented by adding a function at the end of each menu for switching to another layer. Another approach would be to dedicate a button on the device for cycling through the layers. The designer needs to ensure that controls for switching between layers do not significantly interfere (e.g., visually) with performing regular tasks.

5.5 Limitations and Generalizability

The research presented in this paper has a number of limitations. The study involved 16 older and 16 younger participants, which is a relatively small sample; a larger sample size is needed to see whether or not the trends that we identified are trustworthy. Our older participants were also generally well educated, had computer experience, and thus may not be representative of the older adult population. Future work is needed to see whether our findings hold for larger, more diverse samples of older adults. In addition, this study did not examine the added complexity in ML interfaces of switching between layers; our participants were not able to switch between layers on their own.

Our findings may be more applicable to seniors who are better with using such mobile devices and computer technology. As stated in Section 3.1, four older participants who could not complete the study were replaced. One of these participants (in the ML group) almost completed the Basic Task Acquisition phase but became noticeably tired and uncomfortable using the device; this discomfort may have been due to arthritis in her hands, which she reported in the study session. The other three participants (1 in the ML group, 2 in the control group) were able to complete the first two learning phases but took so long to master the basic tasks that they did not have time in the 4-hour study session to complete the last two learning phases. These participants appeared to perform similarly to other participants in the first few attempts but continuously made typing errors in later attempts, which took them much longer to reach our mastery criterion and move on to the next phase of the study. From our observations, these typing errors seemed to be caused by fatigue, as well as carelessness (e.g., saving a contact's information without verifying that no typing errors were made). We speculate that the effect of age on our results may have been stronger if these participants were able to complete the study and their data were included in the analysis.

Finally, we expect our results to generalize to many other mobile devices. Other mobile devices, even those with different form factors (e.g., smaller screen with fewer buttons than the E61i, large touch screen), use multiple function menus and have relatively similar ways of navigating from one application screen to another. We expect that one of the differences among mobile devices that may most influence learning effort for performing new tasks is the device's input method, particularly for text entry, which can vary widely (e.g., single button press vs. multi-tap, physical vs. on-screen keyboard). Our results are also expected to generalize to many other mobile applications. The address book's interface developed for this study required the use of many standard mobile interactions, such as soft keys and a direction pad for navigation, and menus (1/screen) for browsing and executing functions. The application itself is also relatively simple compared to many existing mobile applications (e.g., calendar, internet browser); we expect that ML interfaces would provide greater initial learning benefits to both older and younger novices on more complex mobile applications.

6. FUTURE WORK

One area of future work is to evaluate the ML interface approach in the context of learning to use different types of mobile applications, particularly those that are unfamiliar to older adults. As stated in previous sections, the ML mobile application evaluated in our study was relatively simple; existing mobile applications are generally more complex, have more features (e.g., home screen and application launcher), offer a wide variety of interaction methods (e.g., multitouch gestures for navigation and interacting with content), or present more detailed, composite content (e.g., GPS-enabled navigation application). Future work is needed to study how ML interfaces can be used to layer functions or content of different types of existing mobile applications and evaluate the effectiveness of ML interfaces on the application's learnability. Alternatively, one could also evaluate the effect of using ML interfaces to learn to use mobile applications that involve concepts and tasks that users are not familiar with (e.g., social networking tools). We expect that ML mobile interface with an initial reduced-functionality layer will provide more benefits to older adults for learning to use unfamiliar mobile applications.

Future work is also needed to study the use of a ML mobile application over time. For example, a longitudinal study could help us to understand why and how often users switch between layers, and measure the effect of the switching overhead on an application's learnability and long-term use. Comments made by the older participants suggest that they would primarily use the mobile address book in one layer and not switch much between layers; research is needed to confirm whether this generally holds true for this and other ML mobile applications. A longitudinal study could also be used to explore the effects of personalizing the reduced-functionality layer on longer term use.

Further, assessing the learnability of ML mobile applications using different methods may be useful. We used a specific criterion for assessing mastery (i.e., perform task set, twice in a row, using no more than 20% additional steps over the minimum number of steps required) but there are other ways of defining mastery and assessing learnability [Grossman et al. 2009] that may better help us to find differences in learning benefits.

7. CONCLUSION

This paper presents the findings of an experimental study that explored the effects of a multilayered (ML) interface for helping older adults to learn mobile applications. We found that our ML mobile address book provided participants an initial learnability benefit in terms of fewer extra steps taken. The ML application also helped participants to better retain the ability to perform the tasks they had mastered 30 minutes previously. When participants transitioned to a full functionality interface layer, they experienced a temporary decrease in basic task performance as they needed to relearn the function menus. However, no negative impact was found on learning advanced tasks in a full-functionality layer. We also found that our ML application helped our older participants more than our younger ones to perform initial basic tasks in less time. Further, the majority of older participants preferred learning on the ML interface and found it less complex than the non-layered control interface. Given the initial performance benefit, the overall preference for the ML interface for learning, and lack of major drawbacks, the ML interface appears to be a suitable design approach for improving mobile applications for older adults and lowering barriers for adoption.

ACKNOWLEDGMENTS

The authors wish to thank Karyn Moffatt, Garth Shoemaker and Tony Tang for their constructive feedback. We are also grateful for the helpful comments and suggestions from the editors and anonymous reviewers.

REFERENCES

BIRDI, K. S. AND ZAPF, D. 1997. Age differences in reactions to errors in computer-based work. *Behav. Inform. Technol. 16*, 6, 309–19.

CARROLL, J. M. AND CARRITHERS, C. 1984. Training wheels in a user interface. *Commun. ACM. 27*, 800–806.

CATRAMBONE, R. AND CARROLL, J. M. 1986. Learning a word processing system with training wheels and guided exploration. In *Proceedings of CHI/GI*. ACM, New York, NY, 169–174.

COHEN, J. 1988. *Statistical Power Analysis for the Behavioral Sciences*. 2nd Ed. Erlbaum, Hillsdale, N. J.

DEYOUNG, C. G. AND SPENCE, I. 2004. Profiling information technology users: En route to dynamic personalization. *Comput. Human Behav. 20*, 55–65.

DICKINSON, A., SMITH, M. J., ARNOTT, J. L., NEWELL, A. F., AND HILL, R. L. 2007. Approaches to Web search and navigation for older computer novices. In *Proceedings of CHI*. ACM, New York, NY, 281–290.

ECHT, K. V., MORRELL, R. W., AND PARK, D. C. 1998. Effects of age and training formats on basic computer skill acquisition in older adults. *Educ. Gerontology 24*, 1, 3–25.

FINDLATER, L. AND MCGRENERE, J. 2007. Evaluating reduced-functionality interfaces according to feature findability and awareness. In *Proceedings of INTERACT*. Springer Berlin, 592–605.

FISK, A. D., ROGERS, W. A., CHARNESS, N., CZAJA, S. J., AND SHARIT, J. 2009. *Designing for Older Adults: Principles and Creative Human Factors Approaches, 2nd Ed*. CRC Press, London.

FREUDENTHAL, D. 2001. The role of age, foreknowledge, and complexity in learning to operate a complex device. *Behav. Inform. Technol. 20*, 1, 23–35.

GROSSMAN, T., FITZMAURICE, G., AND ATTAR, R. 2009. A survey of software learnability: Metrics, methodologies, and guidelines. In *Proceedings of the 27th International Conference on Human Factors in Computing Systems (CHI'09)*. ACM, New York, NY, 649–658.

HAWTHORN, D. 2005. Training wheels for older users. In *Proceedings of the 19th Conference of the Computer-Human Interaction Special Interest Group (CHISIG) of Australia on Computer-Human Interaction*. CHISIG of Australia, Narrabundah, 1–10.

HICKMAN, J., ROGERS, W., AND FISK, A. 2007. Training older adults to use new technology. *J. Gerontology*, 62B (Special Issue I), 77–84.

KURNIAWAN, S., MAHMUD, M., AND NUGROHO, Y. 2006. A study of the use of mobile phones by older persons. In *Extended Abstracts of CHI*. ACM, New York, 989–994.

MCGRENERE, J., BAECKER, R. M., AND BOOTH, K. S. 2002. An evaluation of a multiple interface design solution for bloated software. In *Proceedings of CHI*. ACM, New York, 164–170.

MCKNIGHT, P. E., MCKNIGHT, K. M., SIDANI, S., AND FIGUEREDO, A. J. 2007. *Missing Data: A Gentle Introduction*. Guilford Press, New York.

MILLWARD, P. 2003. The 'grey digital divide': Perception, exclusion, and barriers of access to the Internet for older people. *First Monday, 8*, 7. http://www.firstmonday.org/issues/issue8_7/millward/index.html.

MILNER B. 1971. Interhemispheric differences in the localization of psychological processes in man. *Brit. Med. Bull., 27*, 272–277.

MORRELL, R., PARK, D., MAYHORN, C., AND KELLEY, C. 2000. Effects of age and instructions on teaching older adults to use ELDERCOMM, an electronic bulletin board system. *Educ. Gerontology 26*, 221–235.

NIELSEN, J. 1996. Usability metrics: Tracking interface improvements. *IEEE Softw. 13*, 6, 12–13.

OFCOM. 2006. *Media Literacy Audit: Report on Media Literacy amongst Older People*. Ofcom, London, UK.

QUINTANA, C., CARRA, A., KRAJCIK, J., AND SOLOWAY, E. 2002. Learner-Centered Design: Reflections and New Directions. In *Human-Computer Interaction in the New Millennium*, Carroll J., Ed. Addison-Wesley, Reading, MA, 605–626.

ROGERS, W., MEAD, S., WALKER, N., AND CABRERA., E. 1996. Training older adults to use automatic teller machines. *Hum. Fact. 38*, 3, 425–433.

SCHEFFER, J. 2002. Dealing with missing data. *Resear. Lett. Inform. Math. Sci. 3*, 153–160.

SHNEIDERMAN, B. 2003. Promoting universal usability with multi-layer interface design. In *Proceedings of CUU*. ACM, New York, 1–8.

WECHSLER, D. 1981. *WAIS-R Manual: Wechsler Adult Intelligence Scale-Revised*. The Psychological Corporation, New York.

ZIEFLE, M. AND BAY, S. 2004. Mental models of a cellular phone menu: Comparing older and younger novice users. In *Proceedings of the 6th International Conference on Human Computer Interaction with Mobile Devices (MobileHCI)*. Springer-Verlag Berlin, 25–37.

ZIEFLE, M. AND BAY, S. 2005. How older adults meet complexity: Aging effects on the usability of different mobile phones. *Behav. Inform. Technol. 24*, 5, 375–389.

Received November 2009; revised February 2010, May 2010; accepted May 2010