Common Industry Format for Usability Test Report v2.0
iPhone 3G ver. 4.1 for Elderly Users

Test Lead and Report Prepared by William Clasen
Test Conducted: April 4 - 6, 2011
Report Prepared: April 7 - 8, 2011

Contact:
William Clasen
(608) 921-8324
wjclasen@uwm.edu
708 Suffolk Drive
Janesville, WI 53546

Executive Summary

Many services readily utilized by seniors are migrating online. Marketing analysts have identified three common goals with respect to users of smart phones over fifty-years of age: text messaging, email, and web connectivity. A usability test was conducted targeting the iPhone 3G ver.4.1 (pictured) on the AT&T cellular network for use by elderly clientele. This document contains test methodology, efficiency / effectiveness ratings, task completion rates, time on task, errors, participant feedback, and discussion for improvements based on applicable theory.

Four participants over the age of fifty-five were selected to participate in three tasks that targeted the functions requiring evaluation. Half of the participants had previous smartphone experience with a different device. All participants were daily users of the web, text messaging, and email. Overall, the participants appeared very receptive to smartphone technology. Only a handful of modifications are required to mitigate sensory and cognitive challenges experienced by the target user group:

Change	Justification	Severity
1. Emphasize zoom and magnify in marketing, sales, and help 2. Add 'magnify' button in supporting applications to increase font size all-at-once	• Mitigation of sensory challenges that result from aging. • All four participants failed to acknowledge the zoom and magnify features to reduce eye strain. • 25% of participants failed to complete web browsing task as a result	High
3. Auto-correction default status to 'off' or provide link directly to this setting on the home screen	• Mitigation of cognitive stability challenges that result from aging. • 50% of participants indicated this feature was "irritating" or "distracting", but did not know how to turn it off	Moderate
4. Add contextual label to 'create new' icon 5. Modify contextual label for Safari to "Web Browser" 6. Change 'plus' icon to 'contact' icon for consistency across all platform applications	• Mitigation of cognitive simplicity and memorability challenges that result from aging. • 75% of users failed to recognize Safari as a web browser application • 100% of users failed to recognize the 'plus' icon to access the contact list • 75% failed to recognize the 'plus' icon in subsequent tasks after assistance	High

Introduction

Product Description

The iPhone 3G iOS ver. 4.1 (pictured below) as manufactured by Apple and operated on the AT&T cellular network is being marketed as a smartphone for mainstream users. Functions such as text messaging, email, and web connectivity offered by such devices are of current interest to smartphone marketing analysts with regard to senior citizens. (Greengard, p21)

Many commerce, health care, and government services readily utilized by seniors are migrating online. Greengard maintains that elderly clients are entirely receptive to smartphone technology so long as it useful and easy to operate (Greengard, p20). In order to achieve usefulness and ease of operation, the design must mitigate challenges presented by degradation of vision, tactile perception, and cognition as the age of clientele advances beyond fifty-five years of age. (Greengard, p20-1).

Test Objectives

This report assesses usability and discusses improvement for this device with respect to senior citizens and their common goals for functionality. In particular, findings regarding web navigation, sending text messages, and sending emails are reported. Audio, as it is unrelated to these particular functions, is not evaluated.

Methodology

Participant Profile

All four participants were chosen as a representative sample of the targeted user population. Criteria included age and previous smartphone experience. All participants are older than fifty-five years of age. Participants were segmented based on previous smartphone experience. Two participants each had two years of experience using web enabled Blackberry devices. Two participants had no previous experience with smartphones.

Prior to evaluation, each participant was asked about their average daily involvement with the following functions: hours spent accessing the web, number of text messages sent, and number of emails sent. The questions did not specify cellular devices. The goal was simply to ascertain overall awareness and experience with these functions via any available technology that they may have had. It was further determined that none of the participants had any experience with Apple devices or software.

Key	Name	Age	Gender	Smartphone Experience (Years)	Daily Web Access (Hours)	Daily Text Messages Sent	Daily Emails Sent
P1	Jerry Clasen	66	Male	2	2	12	8
P2	Carol Clasen	68	Female	2	4	10	15
P3	James Wuetrich	70	Male	None	3	0	4
P4	Dorothy Wuetrich	68	Female	None	2	3	12

Context of Product Use in Test

There is one significant difference between the evaluated context and the expected context of use. The audio capabilities of this smartphone device are not critical to evaluating web access, text messaging, and email functions of the target user group. Audio capabilities, as seen largely a component of traditional cellular phones and multimedia uses associated with the smartphone, are not evaluated in this report.

Use by the target group will be personal rather than professional and evaluated as such.

Task List

Tasks were selected to evaluate the most common goals associated with the device relative

to the target audience as represented by the participants.

Task 1	
Task	View apple.com website and navigate to [support]
State	Home screen

SSC	Participant
	   1. opens safari application (33%)
	   2. enters apple.com in address bar (33%)
	   3. activates link to support tab located on apple.com home page (33%)

Benchmark	0.5 minute. Task Failure at 2.5 minutes.
Script	Please open the internet browser and go to apple.com. Once the apple.com web page has loaded, find the support link on that page and click it.

Task 2	
Task	Send text message to contact John Doe
State	Home screen

SSC	Participant
	   1. opens message application (25%)
	   2. selects contact (25%)
	   3. types message (25%)
	   4. sends message (25%)

Benchmark	0.5 minute. Task Failure at 2.5 minutes.
Script	John Doe's contact information has been stored in the phone. Please send the following text message to John Doe: *Hello*

Task 3	
Task	Send 30 word email to contact John Doe
State	Home screen. Participant is given hard copy of 50 word email to duplicate and send.

SSC	Participant
	   1. opens email application (25%)
	   2. selects contact (25%)
	   3. types email content (25%)
	   4. sends (25%)

Benchmark	1.5 minutes. Task Failure at 7.5 minutes.
Script	John Doe's contact information has been stored in the phone. Please send the following email to John Doe: *Dear John,* *You are receiving the first email that I am sending from an IPhone. I am helping to conduct a usability study for AT&T. I hope all is well.*

Test Facility

Participants were observed performing the tasks set forth by this evaluation at their respective homes in order to replicate typical user environment.

Device Environment

All participants used the same device (Serial # 87926MD8Y7H) in order to negate hardware specific variables.

All participants were evaluated in their homes in the same general location within the AT&T cellular network in order to negate cellular signal strength variables.

The following technical specifications represent aspects of the device environment that are subject to this report:

1. Size and weight
   a. Height: 4.5 inches (115.5 mm)
   b. Width: 2.4 inches (62.1 mm)
   c. Depth: 0.48 inch (12.3 mm)
   d. Weight: 4.7 ounces (133 grams)

iPhone 3G ver. 4.1 on AT&T cellular network

2. Capacity
   a. 8GB flash drive

3. Cellular and wireless
   a. AT&T 3G cellular network
   b. GSM/EDGE (850, 900, 1800, 1900 MHz)

4. Display
   a. 3.5-inch (diagonal) widescreen Multi-Touch display
   b. 480-by-320-pixel resolution at 163 ppi
   c. 16M (full) color
   d. Support for display of multiple languages and characters simultaneously

Test Monitor Role

The test administrator greeted participants, explained each task and provide supporting material as each task is introduced, and responded to requests for assistance only if asked directly by the participant. Assistance is to be limited to the current SSC segment for the task on hand.

Design, Procedure, and Instructions

The task evaluation will consist of a performance test made of four sections:

1. *Participant greeting and background* Each participant will be personally greeted by the test monitor and made to feel comfortable and relaxed. The test monitor will record participant responses to the follow questions:

    a. What is your name?

    b. What is your age?

    c. Have you ever used a smartphone? If so, for how long?

    d. How much time per day would you estimate that you average browsing the web on cell phones, computers, TV browsers, etc.?

    e. How many instant or text messages per day would you estimate that you send on average via cell phones, computers, TV browsers, etc.?

    f. How many emails per day would you estimate that you send on average via cell phones, computers, TV browsers, etc.?

Four Section Task Evaluation Model

2. *Orientation*

   Participants will each receive a scripted, verbal introduction and orientation which will explain the purpose and goals of the test. They will also be assured that the device is what is being tested, not themselves.

3. *Performance Test*

   The performance test consists of a series of tasks that the participants will carry out while being observed. These tasks will be conducted in the following manner:

   a. Participant will be handed a script, which the test monitor will read out loud, detailing the task to be accomplished. They will then attempt that task.

   b. During the performance test the monitor will make notes on elapsed time and participant errors. The test monitor will respond only to direct requests for assistance by verbally offering procedural help. The test monitor will document number of requests and at what stage during the task they occurred. The monitor will also make note of any unusual circumstances, such as incorrect key strokes, that may not be presented as direct measures of the SSC.

   c. After concluding the tasks, the participant will be debriefed by the test monitor.

4. *Participant Debriefing*

   Each participant will be debriefed by the test monitor. Each session will be documented by the test monitor. The question is as follows:

   a. Do you have any particular feedback or recommendations for the iPhone based on the task you just performed?

Usability Metrics

1. Measures of effectiveness include percent of task completion, quantity of errors, and frequency of assists.

2. Measures of efficiency includes time taken to achieve task.
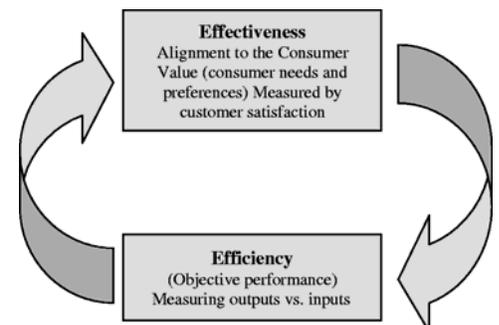
Usability metrics to determine satisfaction

Results

Data Analysis

Task completion is segmented as a percentage in the task list for the purpose of quantifying results. Assists represent each direct request from the user to the test monitor for assistance.

Task time benchmarks, as presented in the task list, are measures of time representing the duration of the task, rounded up to the nearest 0.5 minute interval, as completed by the test monitor, William Clasen. William Clasen has been a daily mainstream user of the device for two years.

Task time failure are benchmarks times multiplied by five. Task failure times represent time-out due to frustration.

Keystroke errors represent every keystroke on the keypad that did not lead directly toward task completion.

Task 1 Access Web Site

Key	Unassisted Task Effectiveness (% Task Complete)	Assisted Task Effectiveness (% Task Complete)	Assists	Task Time (min)	Keystroke Errors
P1	0	100	2	1.25	6
P2	66	100	1	1.00	0
P3	0	66	0	2.50	8
P4	0	100	2	1.75	2

P1, P3, and P4 failed to recognize the default web browser application, Safari. As a result, two out of four participants asked, "How do I get to the web?" One participant asked, "Where is Explorer?"

All participants failed to implement pinch-zoom or magnify as tools to expand text on the apple.com home page. These features were necessary to render the support link legible. Activating the support link represented the final 33% of the SSC. All participants that asked for
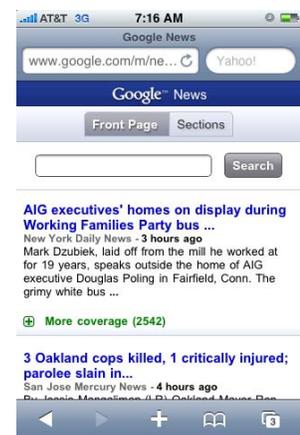
Task 1 - Web Access

assistance regarding this feature asked, "How do I zoom?" P3 did not seek assistance with the zoom feature and timed-out while randomly tapping the screen in hopes of finding the correct link. All other keystroke errors resulted from single character typographical errors and the subsequent deletion of the errant character.
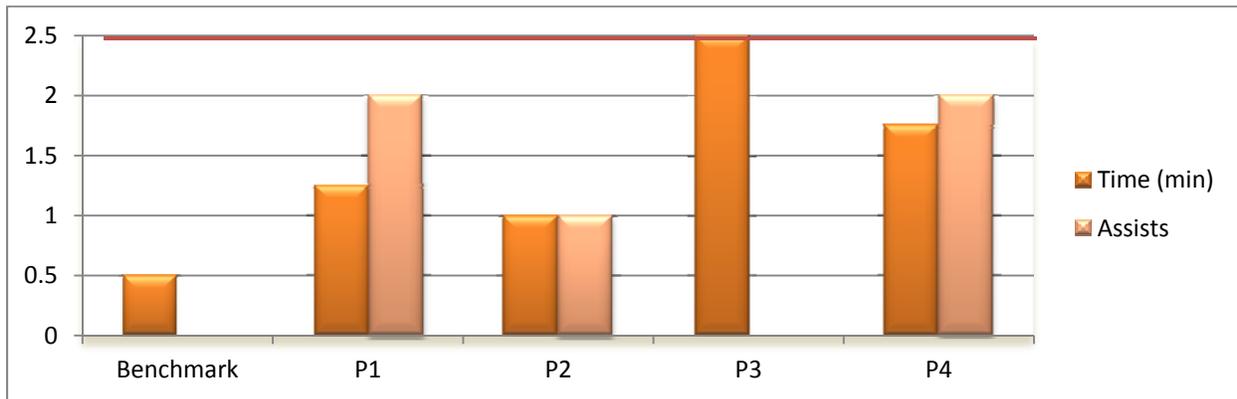

Figure 1 - Task 1 Time and Assists

All participants exceeded the benchmark time of 0.5 minutes for experienced mainstream users. P1 and P2, both experienced smartphone users, asked more immediate questions while requesting assistance and completed the task faster than P3 and P4, non-experienced smartphone users. During the task debriefing, P2 and P3 expressed frustration with the duration of the task. P1 and P2 suggested that the zoom feature was not very intuitive.

Task 2 Send Text Message

Key	Unassisted Task Effectiveness (% Task Complete)	Assisted Task Effectiveness (% Task Complete)	Assists	Task Time (min)	Keystroke Errors
P1	25	100	1	1.00	4
P2	25	100	1	0.75	2
P3	0	100	2	2.00	10
P4	0	100	2	1.50	8

P3 and P4 did not know how to access the text messaging application on the device. The test monitor noticed that these two participants held the phone close as if strained to read the text labels that accompanied icons. P3, after 25 seconds and 4 random keystrokes, and P4, after 15 seconds and 2 random keystrokes, asked the test monitor how to access the application. P1 and P2 accessed the application directly from the home page rather than browsing through the contacts.

Task 2 - Text Message

None of the participants recognized the 'plus' icon as the way to access the contact list. Each participant averaged 15 seconds searching the screen before asking the test monitor how to access the contact list. Similar observations were made regarding the 'create new icon'. All four participants expressed frustration regarding the icon as unintuitive. P2 suggested that the contacts icon on the home page was more recognizable.

Keystroke errors, other than those reported above, were due to typographical errors and subsequent correction. P2 stated the keypad responsiveness, "seemed slow". P1 and P3 stated that the autocorrect feature, "was annoying". P4 offered an unsolicited statement that the autocorrect feature, "was handy", though it appeared to take P4 more time to evaluate the output.
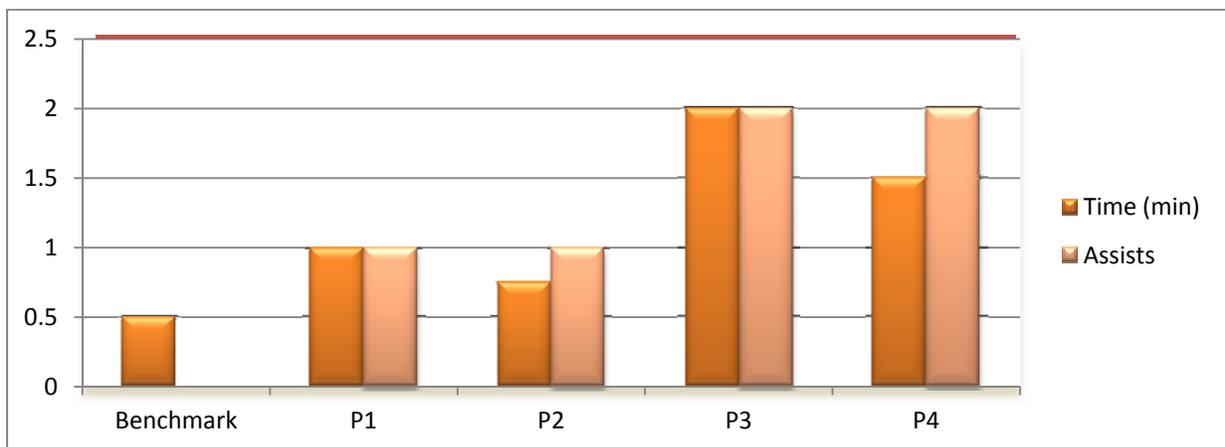

Figure 2 - Task 2 Time and Assists

All participants exceeded the benchmark time of 0.5 minutes for experienced mainstream users. P1 and P2, both experienced smartphone users, completed the task faster than P3 and P4,

non-experienced smartphone users. All participants completed the task well before the task failure time of 2.5 minutes.

During the debriefing, P1 and P2 felt that the process overall was "simple". P3 offered that the task made him feel, "empowered". P4 suggested, "I could get used to this."

Task 3 Send Email

Key	Unassisted Task Effectiveness (% Task Complete)	Assisted Task Effectiveness (% Task Complete)	Assists	Task Time (min)	Keystroke Errors
P1	100	100	0	3.00	8
P2	100	100	0	2.00	4
P3	25	100	2	5.50	20
P4	25	100	1	4.50	12

P1 and P2, presumably from having had previous experience emailing from smartphone devices, completed the task without assistance. P3 and P4 both recognized the mail application, but failed to recall that the 'plus' icon provided access to the contact list. P3, after 10 seconds and 4 random keystrokes, and P4, after 15 seconds and 6 random keystrokes, asked the test monitor how to access the contact list. The remaining keystroke errors were due to typographical errors and subsequent correction. Again, P1 and P3 expressed frustration regarding the autocorrect feature.
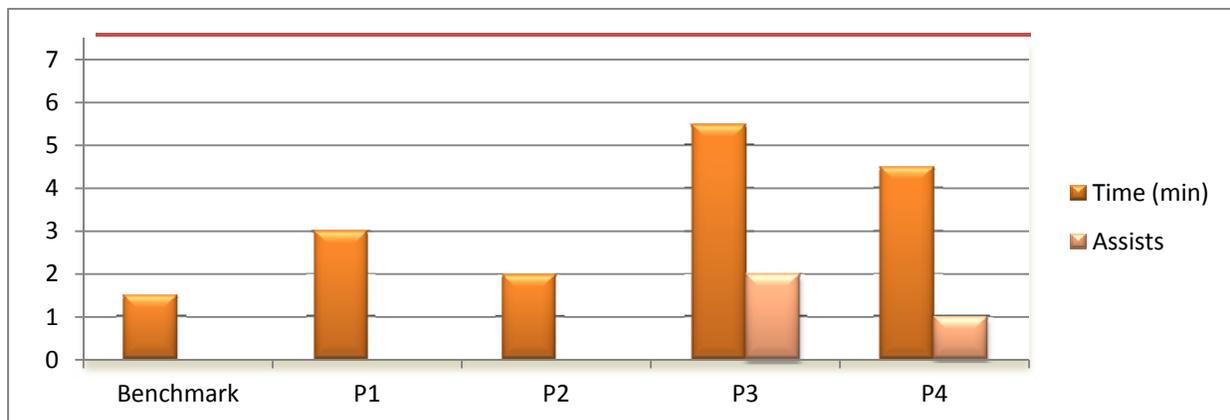

Task 3 - Email

Figure 3 - Task 3 Time and Assists

All participants exceeded the benchmark time of 2.5 minutes for experienced mainstream users. P1 and P2, both experienced smartphone users, again completed the task faster than P3 and P4, non-experienced smartphone users. All participants completed the task well before the task failure time of 7.5 minutes. During the debriefing, all participants commented that the access to the contact list in the mail program was not intuitive. P1 and P2 suggested that the contact icon on the home page be utilized instead of the 'plus' icon.

Recommendations / Discussion

Aside from some sensory and cognitive challenges, the participants expressed eagerness regarding the device. 25% of the participants felt "empowered" by the tasks. The following guidelines will offer support for assisting elderly clients with cellular web-based interaction based on the aforementioned observations.

Sensory design principles revolve around visibility of fonts, pointers, and icons in consideration of the elderly market. The font size used on both the screen and keypad should support visual limitations that result from aging (Nielsen, 2002). Nielsen further stipulates that a minimum twelve point font should be utilized. The iPhone 3G provides users with ways to mitigate sensory restraints
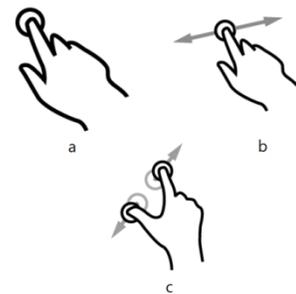
Marketing and help should emphasize zoom feature

and increase overall usability. As a result of this test, it appears that the existence of these features is not apparent. It is recommended that marketing to the target group emphasizes access to the

Recommend magnify icon

pinch-zoom and magnify features. As a further suggestion, an icon representing a magnifying glass (pictured) could be included in predominantly text driven applications as a way to call attention to the magnify feature. Alternatively, the icon could serve to increase on screen font size all at once.

In order to maximize usability, the cognitive load required to carry out tasks must be reduced to a minimum with regard to elderly clients (Greengard, p21). Cognitive load refers to
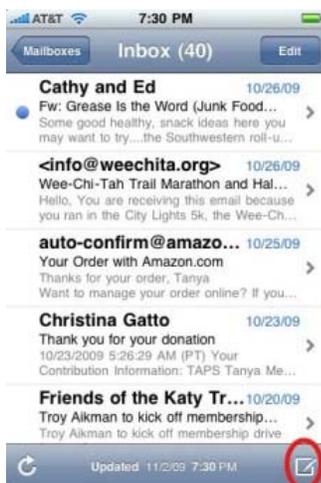
stability, simplicity, memorability, and predictability in determining the usefulness of design

principles (Heim, pp198-208).

In terms of stability, the elderly require interfaces that are forgiving to the point of allowing

multiple formats for data entry such as telephone and credit card numbers (Nielsen, 2002). During

the debriefing for tasks 3 and 4, 50% of participants indicated that the inherent auto-correct

feature that provides suggestions while typing was "irritating" or

"distracting". When marketing to elderly users, special attention

should be given toward a solution that will make the location of the setting to turn off this feature

(pictured) more readily accessible. Alternatively, the default setting could be stet to 'off'.

Mainstream users will be more adept at finding the setting to turn it back on if desired. Forgiving

interfaces require that error messages be eliminated, if possible, due to obscurity and imprecision

that may result in cognitive disconnect (Nielsen, 2002).

In terms of simplicity, windows should utilize a single document

interface (SDI) and short contextual menus if a static interface is

unattainable (Heim, pp368-69, 376-86). The iPhone 3G presents an fairly

straight forward SDI. Icons presented contextual dilemmas for all of the

elderly participants, however. In task 2, all participants failed to

immediately recognize the 'create new' icon. This function (pictured) is

represented by a relatively small square containing a hash mark. The

**'Create New' icon not
identifiable**

users expressed frustration regarding the meaning of the icon. In order

to increase usability for seniors, it is recommended that this icon be given a

contextual label.   Along similar analysis, three out of four users did not recognize

the Safari application icon (pictured) as access to an internet browser. More issues

**Browser icon not
recognizeable**

regarding icon recognition are described below.

Memorability requires events to be pronounced (Nielsen, 2002). As the test results indicate, fifty percent of users cannot be expected to recall icon identification even after the function of the icon is illustrated. 25% of participants suggested that the contacts icon visible on the home screen was more recognizable. In the same vein, predictability requires the interface to be familiar (Greengard, p20). In order to overcome the obstacles apparent to accessing the contact list, it is recommended that the 'plus' icon be replaced with the more commonly recognized contact list icon as pictured.

Contact icons

In order to maintain utility (Heim, p198) any smartphone device offered to the elderly should incorporate broadband internet access. In order to satisfy the demands for web access, text messaging, and email by users over fifty-five years of age, the interface must mitigate sensory and cognitive abilities as age progresses. (Greengard, p21) Should phone and multi-media marketing analysis be desired, audio usability testing may be required.

Bibliography

Nielsen, J. 2002. Usability for Senior Citizens [online]. Available: http://www.useit.com/alertbox/seniors.html (accessed March 12, 2011).

Greengard, Samuel. "Facing an Age-Old Problem" Communications of the ACM. Sep. 2009: 20-22 .

Heim, Steven. The Resonant Interface. New York: Pearson Education, Inc. 2008. Print.