

Usability Test Results

Group 6: Calendar Group:

Alexander Schlicker, Brock Dean, Hyunji Kim, Matthew Tithof, Zachary Wolbers

Usability Expert: Harriet King

April 2011, HCI Class

Table of Contents

Introduction.....	1
Test Goals.....	2
Test Cases.....	2
Adding.....	2
Editing.....	2
Deleting.....	2
Results.....	3
User’s Mood.....	3
Lostness.....	4
Number of Touches Per Task.....	5
Time Per Task.....	6
Consistency Across Tasks.....	7
User’s Rating of Application.....	8
Would User Want This Application?.....	9
User Suggestions.....	10
How Users Were Rated.....	12
What We Learned.....	12
Usability Issues.....	13
Bugs.....	13
More detail available in the User Test Plans document, including:.....	14
References.....	15

Introduction

This is the final report for usability testing of the prototype of the calendar application by Group 6. Generally the application prototype tested out well with positive ratings by the test users. Only two users out of five said they would use the application but that could be due to the abundance of existing effective calendar applications already available. Most suggestions were aimed more at default properties of Android than anything to do with the calendar application.

The calendar application is for any user wanting to track their schedule with an Android device. More information about the test plans and details is available in the User Test Plan.

The basic functionality is: [quoting Group 6] “create appointments and be able to open up these appointments to see basic information. These appointments can be edited and contain information

such as time, location, attendees, general notes, and possibly the length of the meeting. Have an available task list that can be filled with tasks that have a due date and a general note about that task.”

Test Goals

The test goals are to discover if the application functions correctly and is easy for the user.

The tests will include having the testers use the calendar application to add an appointment, change an appointment, and delete an appointment. If there is time, we will also test whether the application is easy to use by creating a stressful environment by playing annoying sounds on headphones and seeing if the user can still do the tasks easily.

Test Cases

There are three test actions distributed over three scenarios used by the tester to test the calendar application. See next section for scenario details.

Adding

Goal: user can add an appointment in less than X seconds with less than Y touches.

Quantitative Measure:

X seconds (stop watch, Tester Observer 1)

Y touches (count, Tester Observer 2)

Qualitative Measure:

Is there any “lostness”, ranked by observer from 0 to 10

What is affect of tester ? (watching and “reading”, Tester Observer 3)

What is affect of tester ? (questionnaire)

Scenario: You have been invited to a party Saturday, April 16 from 5:30 to 7:30

Scenario: Add dentist's appointment Monday, May 9 at 9:45 am, include dentist’s name and address:
Dr. Moreno at 16578 North Broadway

Scenario: Big test, add study session Monday, April 18 from 3 to 6 pm with pizza at the dog house
AND add test on Tuesday, April 19 at 8 am in Rekhi 112

Editing

Goal: user can change an appointment in less than X seconds with less than Y touches.

Quantitative Measure:

X seconds (stop watch, Tester Observer 1)

Y touches (count, Tester Observer 2)

Qualitative Measure:

Is there any “lostness”, ranked by observer from 0 to 10

What is affect of tester ? (watching and “reading”, Tester Observer 3)

What is affect of tester ? (questionnaire)

Scenario: Change the party to pickup 3 friends starting 30 minutes earlier (separate event or same event with edits)

Scenario: Change Dr. Moreno to May 12 at 3 pm

Scenario: Change study session to the ambassador

Scenario: Change test to Thursday of that week in EERC 116

Scenario: Change party to stop at store before in the afternoon (separate event or edit)

Deleting

Goal: user can delete an appointment in less than X seconds with less than Y touches.

Quantitative Measure:

X seconds (stop watch, Tester Observer 1)

Y touches (count, Tester Observer 2)

Qualitative Measure:

Is there any "lostness", ranked by observer from 0 to 10

What is affect of tester ? (watching and "reading", Tester Observer 3)

What is affect of tester ? (questionnaire)

Scenario: Broke up with your girlfriend, delete the party and the pickup of friends

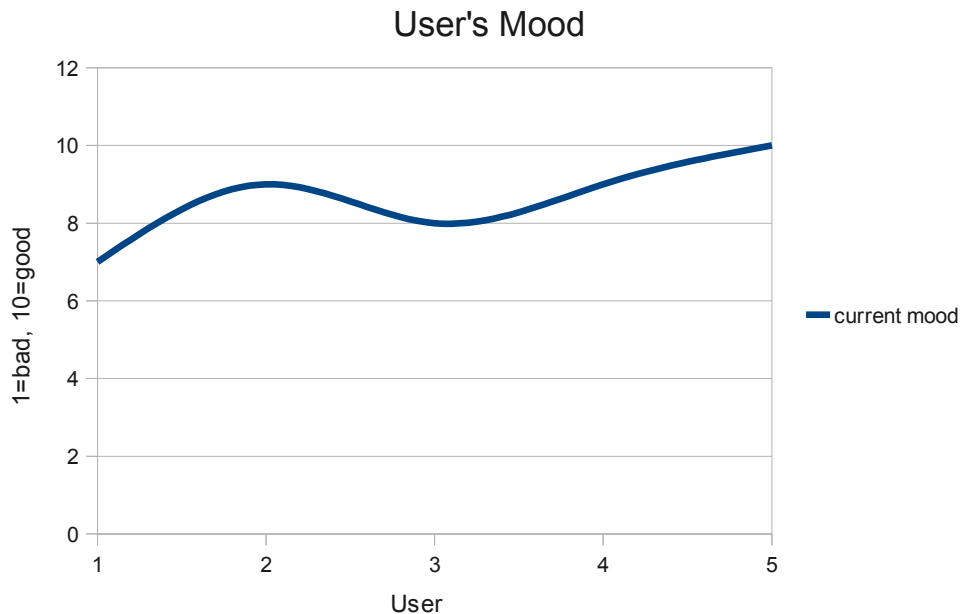
Scenario: Quit school, delete test and change study session to party! At Uphill 41 from 8 to midnight, note: bring clown costume and wig.

Results

Both pre and post questionnaires were done with results ranging from user's mood to rating the application. Many useful comments and suggestions were reported on the post questionnaire and are reported below. Quantitative measures were taken by observers that counted the number of touches and the time it took to do the tasks. Lostness was also measured by observation.

User's Mood

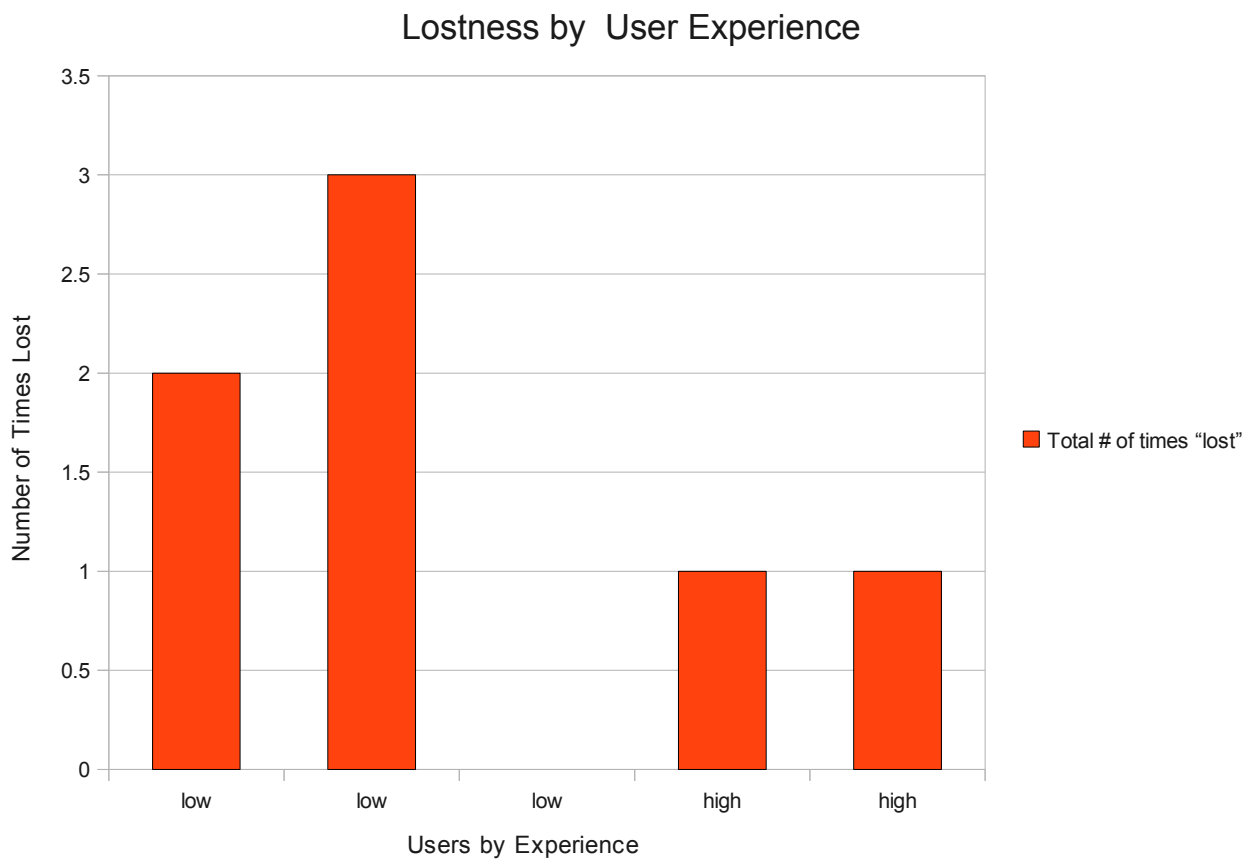
Users were asked to rate their mood in the pre questionnaire from 1, bad, to 10, good. After the testing they were asked if they had any mood change. Before the test, mood was self reported as shown in the chart below and after the test all reported no mood change, except one subject who said he was in a better mood because he wasn't bored anymore.



Lostness

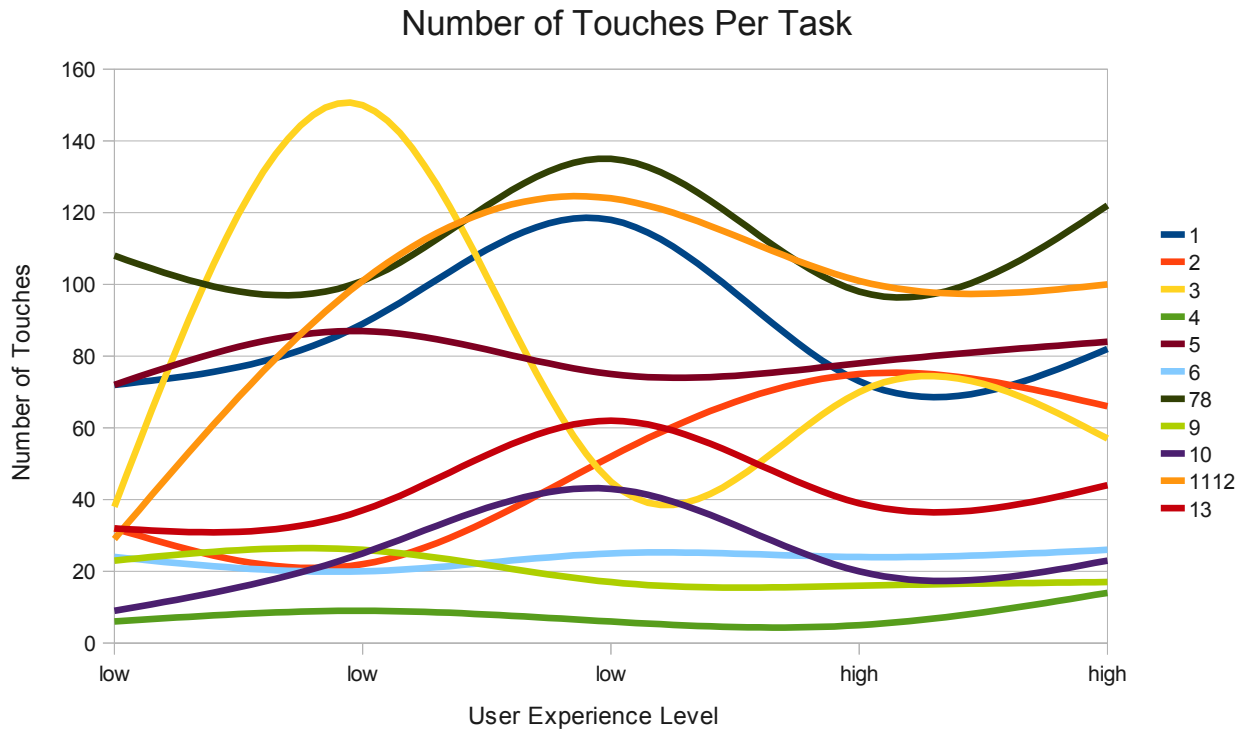
There was no direct correlation between the number of times a user was “lost” and their experience with android handheld devices. One of the low experience users was the only user with no observed lostness. Please see table below for the details showing user experience rank compared to number of times lost.

LOSTNESS	experience rank				
	low	low	low	high	high
Total # of times “lost”	2	3	0	1	1



Number of Touches Per Task

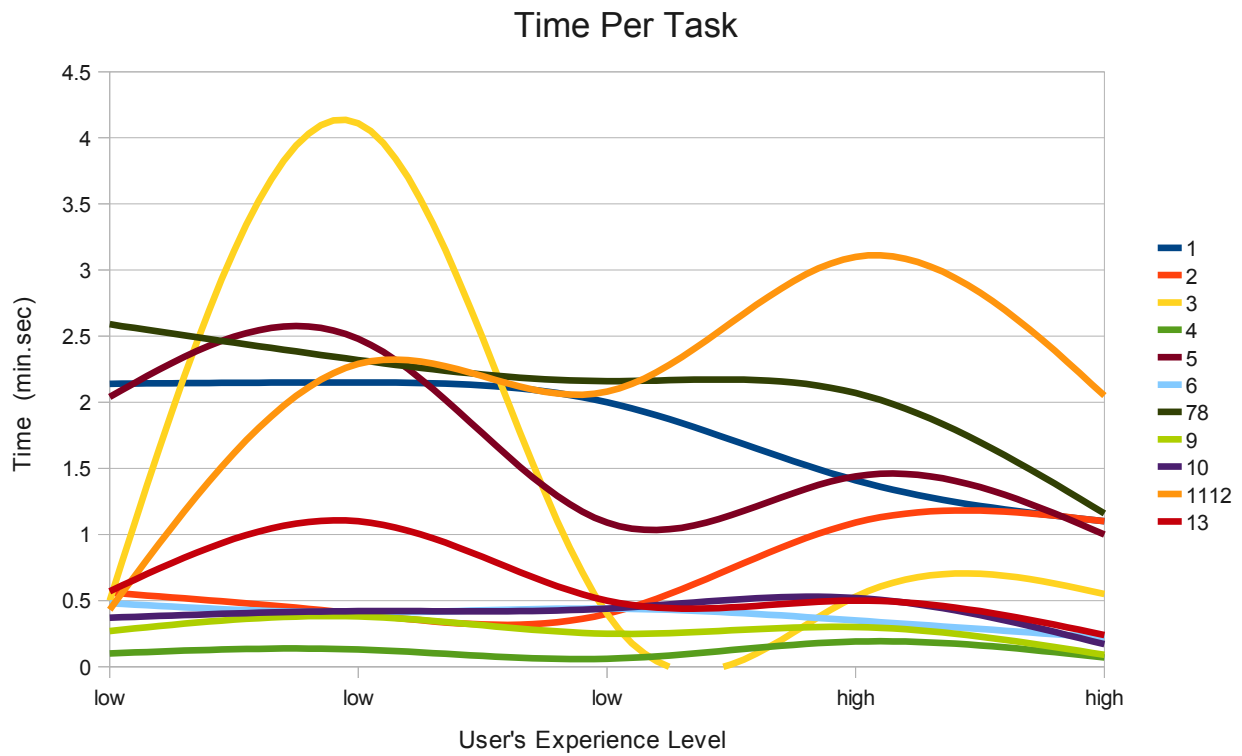
Observers counted the number of touches for each task using a handheld counter. They sat beside and behind the users and were quietly watching and not intruding on the user. The Users are here separated into low and high experience levels with android to see if there is any correlation between experience and efficiency. The complete results for touches are included in the table below with each line being one of the tasks. For detail on what the user chose to do for that task, see “Consistency” section later.



TOUCHES	experience rank				
	low	low	low	high	high
Task 1	72	89	118	73	82
Task 2	32	22	52	75	66
Task 3	38	150	45	70	57
Task 4	6	9	6	5	14
Task 5	72	87	75	78	84
Task 6	24	20	25	24	26
Task 7&8	108	101	135	98	122
Task 9	23	26	17	16	17
Task 10	9	25	43	20	23
Task 11&12	29	101	124	101	100
Task 13	32	37	62	39	44

Time Per Task

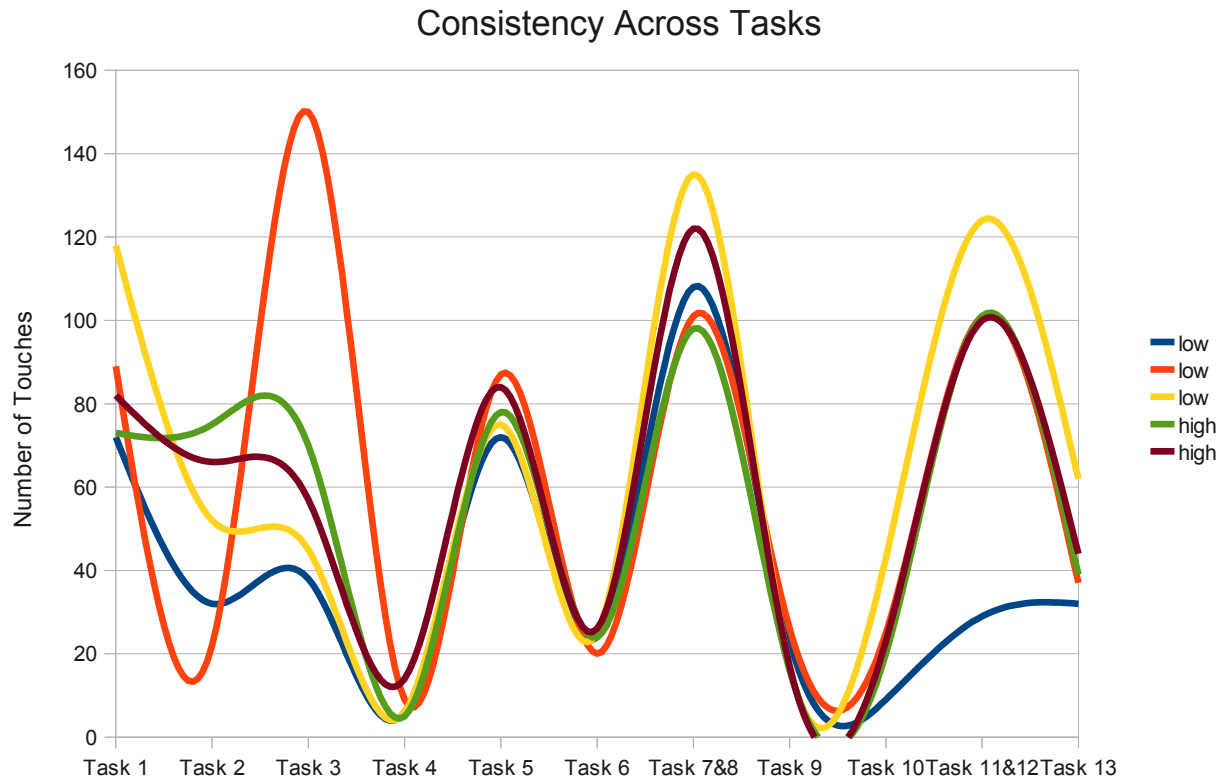
Observers timed each task using a handheld stop watch. They sat across from the users and were quietly watching and not intruding on the user. The Users are here separated into low and high experience levels with android to see if there is any correlation between experience and efficiency. The complete results for time are included in the table below with each line being one of the tasks. For detail on what the user chose to do for that task, see “Consistency” section later..



TIME	experience rank				
	low	low	low	high	high
Task 1	2.14	2.15	2.00	1.41	1.10
Task 2	0.56	0.40	0.40	1.09	1.10
Task 3	0.50	4.11	0.40	0.53	0.55
Task 4	0.10	0.13	0.06	0.19	0.07
Task 5	2.04	2.48	1.09	1.44	1.00
Task 6	0.48	0.41	0.44	0.35	0.22
Task 7&8	2.59	2.32	2.16	2.07	1.16
Task 9	0.27	0.38	0.25	0.30	0.09
Task 10	0.37	0.42	0.44	0.52	0.17
Task 11&12	0.43	2.29	2.08	3.10	2.05
Task 13	0.57	1.10	0.50	0.50	0.24

Consistency Across Tasks

Some of the tasks could be done as changes, adds, or deletes, depending on the user's choice. No instructions were given to specify. Clearly tasks 2 and 3 and tasks 11 and 12 were done in the most variety of ways. Each line is one of the users marked by their experience rank from low to high. See table below for the exact results of which action to accomplish the task the user selected.

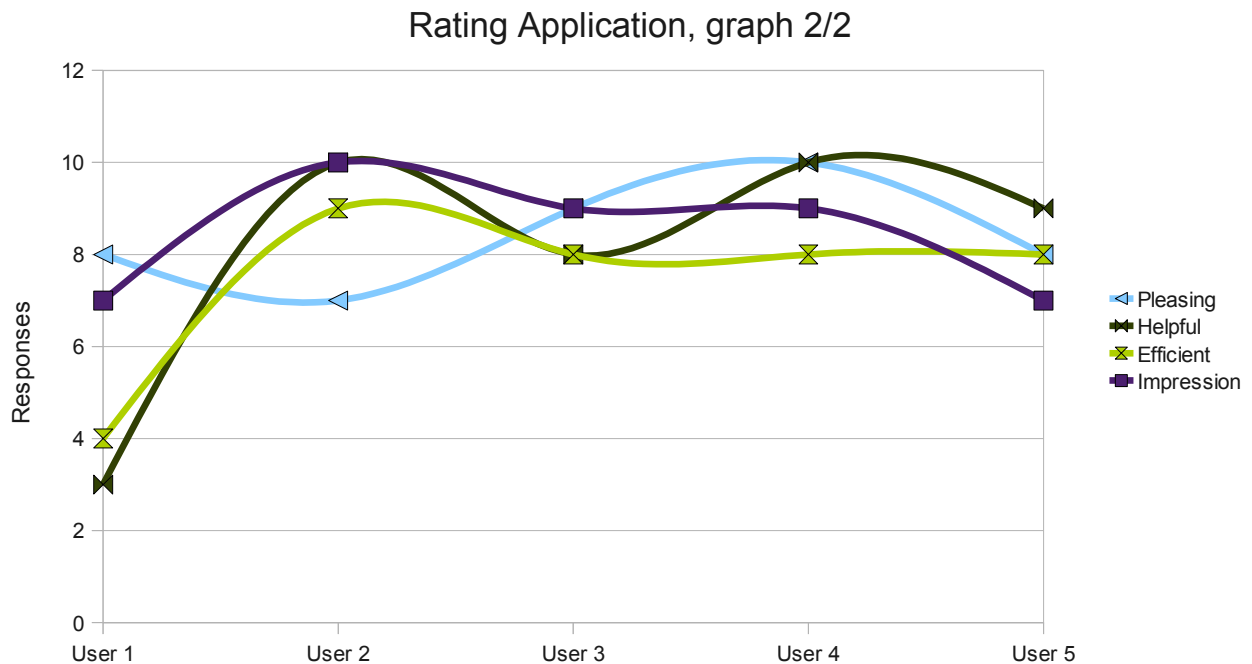
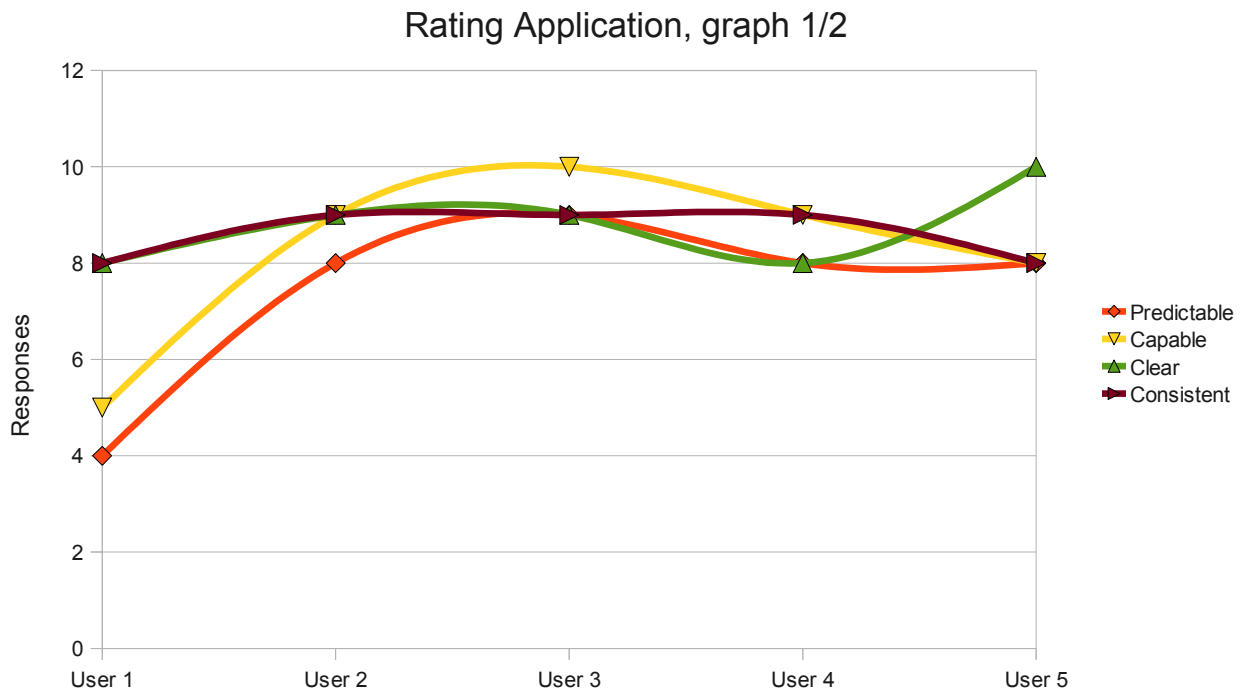


ACTION	experience rank				
	low	low	low	high	high
Task 1	a	a	a	a	a
Task 2	c	c	c	a	a
Task 3	d	c,a,d	c	c	a
Task 4	a	d	d	d	d
Task 5	c	a	a	a	a
Task 6	a	c	c	c	c
Task 7&8	c	a	a	a	a
Task 9	c	c	c	c	c
Task 10	d	c	c	c	c
Task 11&12	c	d	d,a	d,c,a,c	a,d
Task 13	a	a	a	a	a

a = add, d = delete, c = change

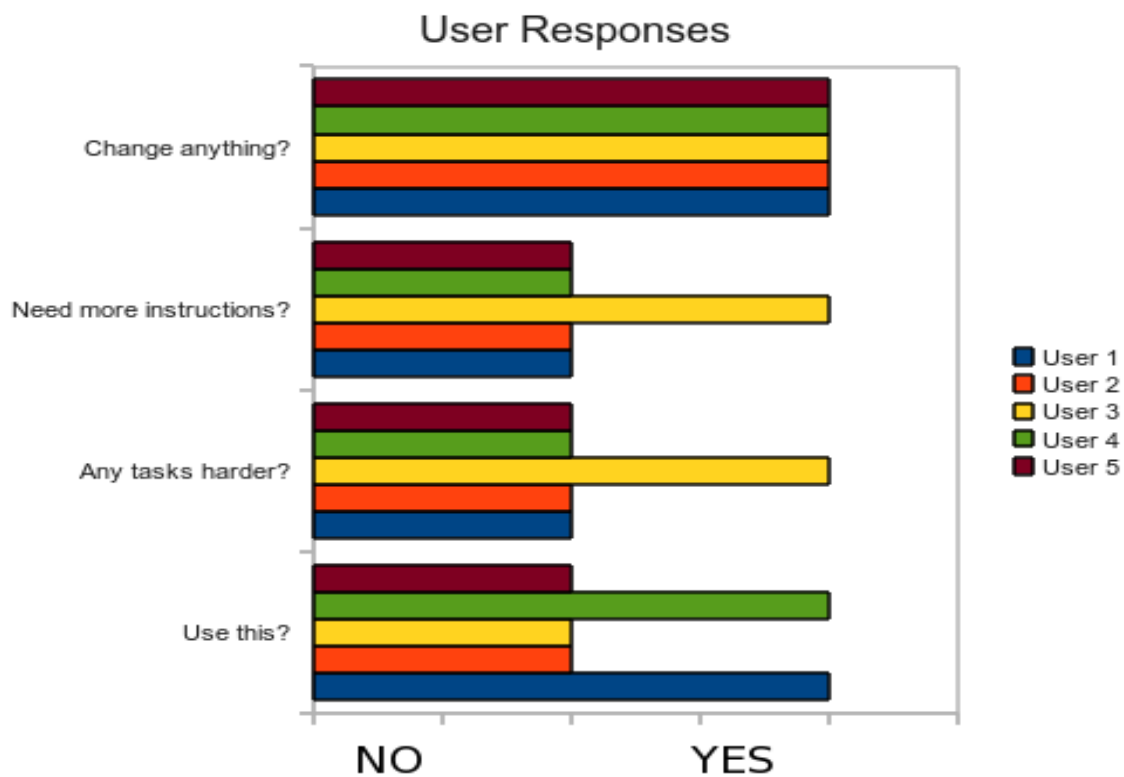
User's Rating of Application

The post questionnaire asked users to rate the application on qualitative measures on a 10 point scale. The ratings were consistently high, except User 1 who scored the application lower in general. More detail follows about user suggestions and impressions in later sections of this report.



Would User Want This Application?

Test users were asked if they would use this application as well as what changes they would like to see. Testers were also asked if they needed or wanted any more instructions. Remember that 2 minutes were devoted to a guided tour of the android and the application by a member of Group 6 to acclimate the test user to the calendar. The last question was whether any tasks were harder than others. For the details of user suggestions in response to these and other questions, please see the next section of "User Suggestions".



User Suggestions

Users gave suggestions on the post questionnaire in almost every category, quoted below:

Predictability

- the app would leave me confused sometimes
- four bugs, 24-hour inconsistency
- it was confusing to “+” sign for drawing
- for most part it worked as I expected
- a few things happened that I wasn’t suspecting

Capability

- sometimes the keyboard doesn’t push the letter I want it to
- would like g-cgl like agenda view
- functions are simple
- a few things I wanted to do I wasn’t able to do

Clarity

- easy way to get to the help
- I was unfamiliar with gestures
- the main menu was clear however the “?” Gesture was not noticeable without a tutorial
- some gestures tripped me up a bit
- I knew how to get to every thing and was easy to get to

Consistency

- you can add or delete in many different ways
- certain things did not surprise me once I got the controls
- a few times I tried to switch things around and acted differently each time

Pleasing Look

- looks very attractive
- would rate higher if fond is user-definable or app is skin-able
- the number of dates, months off by one
- looked very professional, followed android style
- everything looked good I would put location above time

Helpfulness

- when adding an event sometimes doesn't add to the right date
- useful for someone who has a busy schedule
- it would be very helpful for a very busy person

Efficiency

- kind of hard to work accurately
- would like a long-hold and drag to add event
- no access to the closest events except to know the dates
- click hold on blank areas to add, don't always default to current date
- mostly it was droid acting up taking a while to react to gestures

Overall Impression

- a fun, easy, interactive way to use a calendar
- better access to events would be nice
- it was a good product just a few bugs

Would you use this application?

- An efficient tool to add dates
- Google's calendar app does what I need unless there are new gestures it doesn't have
- I'm not busy enough
- with a little tweaking the gestures would speed up calendar browsing
- I usually don't use calendars/planners

Were any tasks harder than others?

- The gestures, mostly with the directions
- no, most were pretty straight forward

Is there anything you would change?

- Maybe making the keyboard a little bigger if possible
- different color schemes
- fine tune gestures (cancel/submit on finger lift)
- switch location and time's positions, if you delete numbers in time reset to zero

How Users Were Rated

Based on the pre questionnaire, users were assigned a low, medium, high experience level with android handheld devices. There were 3 rated as low experience and 2 as high experience and no users were rated medium experience.

Below is the table of answers from the user.

PRE QUESTIONNAIRE					
Has an android?	no	yes	no	yes	no
Used handheld before?	yes	yes	yes	yes	no
How much time using handheld?	hours	years	hours	years	months
Currently use a calendar?	yes	yes	no	yes	no
What kind of calendar?	handheld	Online & handheld	no	online	no
EXPERIENCE RATING	low	high	low	high	low

What We Learned

The team from Group 6 that developed the calendar application prototype were very helpful and attentive to the testing. It was a pleasure to work with them because of their professionalism. One test subject was a no-show with lame excuses but we did reschedule him and it worked out without too much aggravation.

There are a few changes that would be improvements if done again. For the test setup, it would be good to have an open ended interview of the test user at the end so they can say whatever and the usability tests capture everything the test might notice.

It might have been informative to ask test users if there were any things they would change about the test setup. The testers generally seemed very comfortable and one wanted to stay but we asked him to leave when testing was completed so the next test user had some privacy.

I had planned to play an obnoxious sound on headphones for the user to measure if their performance with the application changed under pressure, much like in the real world when stressing to enter an appointment while standing in the grocery store line answering that important call back from the nurse at the doctor's office and juggling your groceries and wallet. Something that would simulate real life factors on the usability of the calendar application. But I did not run those double tests. We just did the regular usability tests.

Users were confused by the scenario printed sheet, thinking that the task was above the scenario when it was the other way around. As the designer of the test, I would not have noticed that, but remembered to point out the correct order to users after having one user ask about it.

Another scenario ambiguity was in scenario two: the dentist. Was the user supposed to schedule the golf with the boss also? We instructed users to not include the golf outing on the calendar. That could have been clearer in the scenario and instructions.

We had to remember to clear the calendar between test users. It was very helpful to have group 6 manage the phone and details of the calendar application. The user was willing to give the word when they began and ended a task which helped the time keeper. The counter needed to be reset and turned on when beginning a task, that was missed once and the user re-started.

The test users took between 40 and 55 minutes to complete the test process. We did offer our test users candy to enjoy and take with them.

Usability Issues

- If you are ahead or behind in the calendar, it still jumps to today when do an add.
- Can't find things added to edit sometimes
- Some users deleted and re-added rather than edit
- Some users "forgot" the gestures and went out to the widget screen to add with the plus button
- One test user did gesture to add task number 2
- One test user used the menu for everything
- scrolling up and down and side to side in the agenda looked like the plus sign if done quickly, thus confusing the gesture controls

Bugs

1. Jumped to March 16 ?
2. even if in correct later date then edit it jumped date to today
3. add is still broken: just start date, after datepicker popup month is show as 1 behind when it isn't
4. month of April shows as month 3
5. check if end time is before start then does it ignore your add or edit (didn't seem to show up)
6. plus gesture has to be vertical line to start

More detail available in the User Test Plans document, including:

1. INSTRUCTIONS AND SCENARIOS for user
2. Quantitative and Qualitative Measurement Lists
3. Test Set Up, Checklist for test
4. Timetable of Test
5. Scheduled Testers
6. PRE QUESTIONNAIRE
7. POST QUESTIONNAIRE
8. Instructions to Users/Testers
9. LOG SHEET

References

Questionnaire adapted from

<http://ucs.ist.psu.edu/CaseStudy/Garden-com/Evaluation/SessionScript.pdf>

Instructions to user adapted from

<http://ucs.ist.psu.edu/CaseStudy/Garden-com/Evaluation/SessionScript.pdf>