

CS 455/855

Mobile Computing

Evaluating Mobile Software

Dr. Orland Hoeber

orland.hoeber@uregina.ca

<http://www.cs.uregina.ca/~hoeber/cs455/2018F>

- R. Harison, D. Flood, and D. Duce, Usability of mobile applications: literature review and rationale for a new usability model, Journal of Interaction Science, 1:1, 2013.
- <http://journalofinteractionscience.springeropen.com/articles/10.1186/2194-0827-1-1>

Why Evaluate?

- Evaluations are an important part of the software development lifecycle that are often overlooked
 - common attitude: I built it, therefore it is great
- even the most carefully designed software must be evaluated
 - were the requirements properly captured?
 - can real users understand what is shown to them?
 - can real users know what must be done to complete their tasks?
 - is the software easy to use?
 - is the software easy to learn?
 - is the software useful?
- in many cases, there are alternative designs that we would like to compare to one another

Evaluation Challenges

- Evaluating mobile software is especially challenging
 - context of use
 - the software might be used very differently in many different mobile contexts
 - while sitting
 - while lying down
 - while walking
 - in low-light conditions
 - in short bursts of usage
 - while doing other things
 - difficulty of observing use
 - because of the small screen size and the use while in motion, it can be difficult to accurately observe the use of the software

Evaluation Methods

- There are a wide range of evaluation methods from the field of Human-Computer Interaction that can be applied to evaluating mobile software
 - ▣ expert reviews
 - ▣ usability tests
 - ▣ controlled laboratory evaluations
 - ▣ field trials
 - ▣ longitudinal studies
- Some of these are quite difficult to implement well because of the mobility of the user/participant in the study

Expert Reviews

- A small group of experts thoroughly test the applicability of the tool
 - ▣ mobile-specific reviews
 - attention to context of use of the software
 - conducted by mobile experts
 - ▣ task-specific reviews
 - ensure that the system supports the key tasks
 - conducted by a combination of mobile and domain experts
- Performed in small teams (2-3) of experts

Usability Tests

- Evaluations based on participants performing a prescribed set of tasks with the system
 - ▣ focus on “The Five E’s”: effective, efficient, engaging, error-tolerant, and easy to learn
 - ▣ measuring efficiency and effectiveness is easy
 - time to task completion
 - error rates
 - ▣ the others require observation of the users as they perform the tasks:
 - the difficulties they have
 - which features they use
 - level of comfort with the system
 - level of satisfaction in completing the task

Controlled Laboratory Tests

- Controlled experiments focus on evaluating a set of alternative interfaces or features
 - driven by the scientific method
 - goal: test hypotheses and support theories
 - use multiple treatments of independent variables (alternatives)
 - show statistically significant differences on one or more dependent variables
 - outcome: validated or rejected hypotheses

- challenging to design around the mobile context of use
 - difficult to control some of the factors
 - observations and data collection of time/accuracy may be difficult

Field Trials

- Field trials evaluate the system in the natural environment of the user, with real data
- The researcher is involved closely in the study
 - ▣ observations
 - ▣ helping the participants to overcome difficulties (mimic expert use)
- Time scales vary
 - ▣ a few hours
 - focus on the ability of the user to perform actual work tasks with actual data in a real work environment
 - ▣ a few weeks
 - focus on the ability for the user to become effective after the initial curiosity and learning curve has been overcome
- Data collected is primarily qualitative
 - ▣ interviews, focus groups, questionnaires

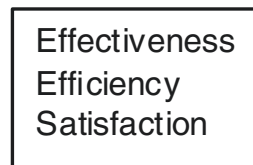
Longitudinal Studies

- Longitudinal studies evaluate the system in the natural environment of the user, with real data, over extended periods of time
- Unlike field trials, participants use the software on their own
 - ▣ can be run remotely with many participants
 - ▣ may implement remote usage logging
- Allows for the evaluation of various elements:
 - ▣ learnability
 - ▣ continued motivation to use the app
 - ▣ use within the full range of mobile contexts
- Data collected is a mix of qualitative and quantitative
 - ▣ questionnaires
 - ▣ usage log analysis
 - ▣ interviews or focus groups with select participants

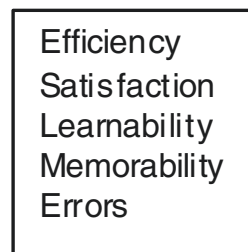
PACMAD Usability Model

(Harrison, Flood, & Duce, 2013)

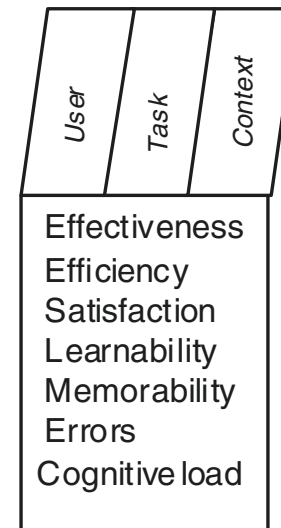
- PACMAD: “People At the Centre of Mobile Application Development”
 - considers the mobile context of use and the user’s *cognitive load* as an important considerations for the usability of mobile apps



ISO



Nielsen



PACMAD

Factors of Usability

- User
 - ▣ ability to see and interact with the app
 - ▣ previous experience
- Task
 - ▣ focus on supporting the user's goal for using the app
 - ▣ other functionality getting in the way
- Context of Use
 - ▣ environment in which the app is being used
 - physical
 - social

Attributes of Usability

- Effectiveness
 - ▣ ability of a user to complete a task in specific context
- Efficiency
 - ▣ speed at which the user can accurately complete the task in the specific context
- Satisfaction
 - ▣ perception of comfort and pleasantness while and after the user completes the task in a specific context
- Learnability
 - ▣ ease by which the user gains proficiency with the app for a given task

Attributes of Usability

- Memorability

- ▣ ability of the user to retain and recall how to use the app for a given task

- Errors

- ▣ ability of the user to complete a task in a specific context without making errors
 - ▣ number and type of errors

- Cognitive Load

- ▣ impact of using the app to perform a task in a specific context on the other things the user is doing in this context (walking, watching TV, etc.)

Applying PACMAD via Inspection

- Specify the types of users (u_1, \dots, u_x), tasks (t_1, \dots, t_y) and contexts (c_1, \dots, c_z)
- Put yourself in the role of the user
- Inspect the app for each task/context pair and document
 - ▣ effectiveness
 - ▣ efficiency
 - ▣ satisfaction
 - ▣ learnability
 - ▣ memorability
 - ▣ errors
 - ▣ cognitive load
- Explain how the app could be improved to address these problems

Applying PACMAD in a User Study

- This model can be used to design a user study
 - ▣ controlled laboratory study
 - ▣ field trials
 - ▣ longitudinal study
- It provides guidance regarding:
 - ▣ participant recruitment (user)
 - ▣ situated work tasks (task, context)
 - ▣ task order and rest time (learnability, memorability)
 - ▣ data collection (effectiveness, efficiency, satisfaction, errors, cognitive load)

Final Words on Evaluation

- Do not treat the evaluation of your mobile software as an after-thought
 - plan for evaluations early
 - incremental inspections & testing
 - expert reviews
 - usability tests
 - careful evaluation of alternatives (if necessary)
 - controlled laboratory tests
 - evaluation of complete prototypes
 - field trials
 - longitudinal studies