Introduction to HCI

Evaluation of Prototypes
Usability Testing

Prof. Narges Mahyar
UMass Amherst

nmahyar@cs.umass.edu
Courses, projects, papers, and more:
http://groups.cs.umass.edu/nmahyar/

© Mahyar with acknowledgements to Joanna McGrenere and Leila Aflatoony
Midterm evals [5 min]

- Course seems largely on track
- The attention to pre-readings have improved a lot
- Midterm results were good
- Working class seem to be effective
- In-class activities are well-received
- Workload has gone down
- In-class discussions/participation need improvement
- We will spend more time on milestone and instructions
Today

• Mid term eval [5 min]
• Discussion of readings [5 min]
• Usability testing lecture [20 min]
• In class activity [30 min]
  • Usability study
• Discussion [15 min]
Learning Goals

• Understand the role of usability testing in HCI
• Be able to define usability testing (nelson’s definition vs others)
• Understand how usability testing is different from other evaluation methods
• Explain when usability studies are typically conducted and why
  • Give examples of locations, tasks, metrics, evaluation methods that might be involved
• Explain how to plan and conduct a usability study
What is the role of usability in HCI?

• Usability: a primary focus of HCI
• **Evaluate** system usability
  • How **easy** it is for the user to **get** the system to do what s/he needs it to do
• **Design** for usability
• Establish/apply **metrics and standards** for usability
Observe, learn, iterate and learn (Don Norman)
What is the role of usability in HCI?

• HCI starts with understanding the problems that users are having

then designing a system that solves these problems
→ requirements, task examples specify what it should do
→ decide on conceptual/interface design for how system will do it

→ usability studies: see if we succeeded
Usability
(Nielsen’s definition)

• Learnability
  • easy to learn so a user can rapidly start to use it

• Efficiency
  • once the user has learned the system, a high degree of productivity is possible (better known as performance)

• Memorability
  • the user should be able to return to the system and not have to learn again

• Errors
  • users should make few errors and recover easily

• Satisfaction
  • the system should be pleasant to use

→ Usability study/test: evaluates an interactive system/prototype with respect to all/some of these elements, always involving real users
Elements of a usability test

1. Interactive system / prototype
2. Evaluation goals
3. Tasks
4. Measures/metrics
5. Data collection/recording methods
6. Participants
When designing a usability test:

- **Choice of methods:** *triangulate*
  - Typically: one instrument counts something, while another interprets what was counted

- **Choice of metrics:** driven by your requirements & eval goals
  - As well as basic usability principles

- **How many users:** should be representative of your user groups
  - E.g.: If you want to support both expert and novice users, should have good numbers of both!
  - Within a demographic, < 4-5 is dubious; often >10-12 is of marginal additional value.
  - Sometimes constraints dictate low numbers.
    - Examples?
    - If you have to generalize, consider who your test users are, and how representative they are?
Task

• Generally: user researcher specifies the task

• Can be:
  • At quite low level; e.g. The subtask that will take you from one screen to the next.
  • Or, at entire task level: see if someone can figure it out, start to finish, and watch / count / measure the challenges s/he has

• Not done with those task examples yet!

• Can use them as a basis for a stripped-down task description
  much as you did for cognitive walkthroughs

• (But don’t usually want to include the story)
Methods
examples of common ones

• Observational techniques:
  • silent
  • think aloud
  • constructive interaction
• Query techniques:
  • Interview
  • survey
  • questionnaire
Metrics
examples of common ones

• **Time:**
  • To **complete** a task (entire, or a portion)
  • **Learn** a task
  • **Resume** a task after interruption
  • **Find** something on a screen
  • **Attain** specified degree of proficiency

• **Errors:**
  • Number per task or unit of time
    • Different types: e.g., Navigation, selection, interpretation
  • Number of users making the error
  • Alternately: number of successes
Metrics

Examples of common ones

- **events of interest:**
  - page views or clicks
  - access of particular tools
  - timeouts
  - questions asked or help tools consulted
  - # users willing to recommend
- **subjective factors:**
  - task level satisfaction
  - perception of aesthetics
  - perceived ease of use
  - perceived preference
  - (all can be measured on a Likert or semantic rating scale)
Alternatives to usability testing

Usability testing requires users, relatively refined prototypes, and usually focusses on measuring something.

- “Discount” methods can also target prototypes at various stages and be done without users
  - Heuristic evaluation
  - Cognitive walkthrough

- Because you don’t need users . . .
  - Can do it first (before a usability study)
  - Possible to apply these methods yourself while iterating on a design (before it’s totally finished)
Biggest differences with alternatives

• Usability testing requires:

• A refined interface.

• This could be… your new medium fidelity prototype.

• Or it could be the bad old interface, which you plan to revise or replace i.e., Might be “evaluate for understanding the problem”

• Measured outcomes.

• Users (participants).
Note on terminology

• Not entirely standardized…

• **User Study** – very general. Any study that involves actual or prospective users. Can be anytime -- from before a system is built (Empathize / Pre-Design) right to a controlled experiment.

• **Usability Study** – more specific. Requires a system for which task performance can be measured (usually Mid / Late Design, but can be Pre-Designing for a system being re-designed)

• **Controlled Experiment** – a specific type of usability study with hypotheses and statistical testing, often comparing alternate designs (more on this later). (Test / Late Design)

• **Informal / Small User Study** – often used before a usability study, not ready to measure things yet, interested in higher-level feedback. (Early design)
Understand USERS:
- who they are
- their key tasks

Understand DESIGN:
- design space and risks
- choose design approach

REFINE Design:
- by element
  - considering task
  - varied contexts

CONFIRM & debug:
- performance in real use

Materials / Methods

Examine existing:
- user tasks & objectives
- contexts
- interfaces

Make use of:
- requirements
- task analysis
- real & virtualized users
- technology options
- company IP

Evaluate with:
- observation
- interview/quest
- participatory interaction
- task walk-throughs

Evaluate w/:
- usability testing - controlled, uncontrolled
- heuristic evaluation

PRODUCTS

PRE DESIGN
- user and task descriptions
- design requirements

EARLY DESIGN
- low fidelity prototyping methods
- throw-away prototypes
- design direction
- risk analysis

MID DESIGN
- testable medium-fidelity prototypes

LATE DESIGN
- alpha/beta systems or complete specification

K MacLean - derived from version by Saul Greenberg (U Calgary)
Usability testing
in your project – Fifth milestone: test

• Evaluation goals?
  • You will likely want to draw from your requirements and task examples; may need to prioritize;
  • Test how well your system supports what you intended it to
  • Metrics, evaluation methods, etc. Should follow

• Hi fidelity prototype scope?
  • Prototype should be a working system
  • It should do enough to test if your design will meet your goals (and be achievable in the time available)
On deck...

• Next class (Tuesday) ...
• Working Class and Prototype Review

• Forth project milestone: prototyping
  • due on Thursday Nov 14th (next week)
Activity [30 min]

• Read and discuss any questions about the next milestone [5min]
• Come up with evaluation elements for your projects [10 min]
• Document your metrics, methods, task, participants, data collection, procedure [15min]
• Develop unbiased questions [5 min]