EVALUATION OF PROTOTYPES
USABILITY TESTING

CPSC 544 FUNDAMENTALS IN DESIGNING INTERACTIVE COMPUTATIONAL TECHNOLOGY FOR PEOPLE (HUMAN COMPUTER INTERACTION)
WEEK 9 – CLASS 17

© Joanna McGrenere
Includes slides from Leila Aflatoony, Karon MacLean and Jessica Dawson
TODAY

• Prototyping examples from ASSETS [10 min]
  • of a good poster and WoZ
  • Video – showing prototype
• Design sprint – method from Google [5 min]
• Usability testing lecture [20 min]
• In class activity [30 min]
  • Usability study
• Discussion [15 min]
PROTOTYPE EXAMPLES FROM ACM ASSETS 2017

• Poster example of WoZ in an AR system designed to help people with visual impairment acquire items (grocery shopping scenario shown)
  • Will post a PDF to Canvas

• Video of a hi-fidelity prototype to assist people with dementia to navigate and play songs
  • https://www.autodeskresearch.com/publications/AMI
GOOGLE
“DESIGN SPRINT”

https://designsprintkit.withgoogle.com/introduction/overview

• “The Design Sprint is a proven methodology for solving problems through designing, prototyping, and testing ideas with users.”

• all in 5 days!

• uses many of the methods you’ve been learning in this class.
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
  • for Prototype milestone you are doing an informal user test
  • for Test milestone you will do a usability study (experiment)
• 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
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 testing: “count” problems in refined prototype

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

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

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

Evaluate w/:
- observation
- many kinds
- ethnography
- interviews, questionnaires
- task analysis

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

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

Make use of:
- graphical design
- interface guidelines
- style guides
- real & virtualized users

Evaluate w/:
- observation
- many kinds
- ethnography
- interviews, questionnaires
- task analysis

Make use of:
- low fidelity prototyping methods

Make use of:
- throw-away prototypes
- design direction
- risk analysis

Make use of:
- med/ high fidelity prototyping methods

Make use of:
- testable medium-fidelity prototypes

Make use of:
- graphical design
- interface guidelines
- style guides
- real & virtualized users

K MACLEAN - DERIVED FROM VERSION BY SAUL GREENBERG (U CALGARY)
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

ersors: 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

- Interactive system / prototype
- Evaluation goals
- Tasks
- Measures/metrics
- Data collection/recording methods
- Participants
EXAMPLE OF A USABILITY TEST: HELP KIOSK

- Interactive system / prototype
  - Help Kiosk
- Evaluation goals
  - To see if older adults can learn to do basic tasks independently on their own smart phone after using HK for a short period of time (half hour or less)
  - To see if the concept of Help Kiosk is appealing and would be preferred to using a manual or getting help from a friend
- Tasks
  - Use HK to learn how to (1) a contact to your smartphone, (2) to send a text message, and (3) to call a contact
  - Repeat the same tasks without the support of HK, using printed manual
EXAMPLE (CONT’D)

• Measures/metrics
  • Time: from the moment user starts until they complete each task successfully
  • Errors: number of taps that deviate from correct task path
  • Satisfaction: Likert scale

• Data collection methods
  • Observation - video recording
  • Observation – software logging of all interactions
  • Survey – to capture satisfaction
  • Interview – probe on user experience and comparison to other learning methods

• Participants
  • 16 older adults (60+), mix of genders, range of computer expertise, all new to smartphones
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 (participant talks as they do the task)
- constructive interaction (participant is probed periodically)

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

events:
• 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 (LAST CLASS)

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-Desing 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

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

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

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

Make use of:
- graphical design
- interface guidelines
- style guides
- real & virtualized users

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

Evaluate w/:
- observation – many kinds
- ethnography
- interviews, questionnaires
- task analysis

Current milestone: Prototype

PROGRESS
- user and task descriptions
- design requirements

PERIOD
- pre design
- earl design
- mid design
- late design

MATERIALS / METHODS
- low fidelity prototyping methods
- throw-away prototypes
- design direction
- risk analysis

PRODUCTS
- med/ high fidelity prototypes
- testable medium-fidelity prototypes

CONSTRUCTION
- testable medium-fidelity prototypes
- risk analysis
- design direction
- throw-away prototypes

Current milestone: Prototype

usability testing: “count” problems in refined prototype

K MACLEAN - DERIVED FROM VERSION BY SAUL GREENBERG (U CALGARY)
USABILITY TESTING
IN YOUR PROJECT – TEST (NEXT) MILESTONE

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

medium fidelity prototype scope?
• prototype won’t be a complete working system
• it should do just enough to test if your design will meet your goals (and be achievable in the time available)
ACTIVITY
analyze a documented usability study
DISCUSSION ON READINGS [15 MIN]

- What surprised you? or
- What you disagreed with?
- Others?
ON DECK...

Next class (Tuesday) ...

• Readings and researcher journal
• Forth project milestone: prototyping  
  • due next Thursday