Evaluation of Interactive Systems

Anastasia Bezerianos
Caroline Appert
M1-HCID - (M2-HCID) - M2-Inter

http://www.lri.fr/~appert/eval
Eval of Interactive Systems

Housekeeping

This module: Introduction
mandatory for M1 HCID + M2 Interaction
general evaluation methodologies
50% class exercises + 50% exam
Volunteers: paper presentations (in groups of 3)

Advanced module:
mandatory for M1 HCID, elective for rest
mini-bootcamp
focus on experimental design (very important for research)
Presentations - Sep 21

make a group of three-four students volunteers for next week !!!

pick an article in the list available on the course’s website

https://www.lri.fr/~appert/eval/

register your group in google spreadsheet

only one group per article: if your choice is not available, choose another article
References and Inspirations


[4] Anastasia Bezerianos and Wendy Mackay classes
Evaluation of Interactive Systems

INTRODUCTION
Evaluating Interactive Systems
What: Measuring usability

Def#1: According to ISO 9241 standard *Ergonomics of Human-System Interaction*, a system usability depends on:

- system’s effectiveness -- the user can reach their goal
- users’ efficiency -- amount of human resources to reach a goal
- users’ satisfaction
Evaluating Interactive Systems
What: Measuring usability

Def#2: Usability is the distance between the user’s and designer’s mental models. The closer the distance, the higher the usability of the system.

“If we have trouble using something then it’s probably because that thing is badly designed, rather than us being stupid” [1]

Evaluating Interactive Systems
When: Iterative User Interface Design

Needs, wants, and limitations of end users of a product are given extensive attention at each stage of the design process.
Evaluating Interactive Systems
When: Iterative User Interface Design

Golden rule: Do not wait to get a refined design before testing, it as some usability problems can be structural and require a deep rethinking of the architecture

Early design stages

Low-fidelity (paper mock-ups)

Advanced design stages

High-fidelity (software and hardware implemented interactive system)
Evaluating Interactive Systems

How: Measures

Efficiency / Speed of performance

Once users have learned the design, how quickly can they perform tasks?

Errors

How many errors do users make, how severe are these errors, and how easily can they recover from the errors?

Satisfaction

How pleasant is it to use the design?
Evaluating Interactive Systems
How: Measures

Learnability

How easy is it for users to accomplish basic tasks the first time they encounter the design?

Memorability / Retention

When users return after a period of not using it, how easily can they reestablish proficiency?

Engagement and Affective Response

Do users feel engaged in the task? Do they have an emotional response (valence/arousal)
Evaluating Interactive Systems
How: Measures (aside)

Arousal/valence model developed by James Russell and Lisa Feldman Barrett
Forced tradeoffs -- almost impossible to optimize all usability measures at the same time

Example: keyboard shortcuts and macros are very efficient but take a long time to learn

Importance of each measure depends on

the type of system

the type of users
Evaluating Interactive Systems
How: Measures

Objective and subjective usability measures

Data collected on 298 designs by the Nielsen Norman Group

*Users generally prefer designs that are fast and easy to use, but satisfaction isn't 100% correlated with objective usability metrics*
Type of system

Different measures are important for different types of systems
Type of system

Life-critical systems
e.g., control air traffic, nuclear reactors, medical instruments...

- Importance
  - crucial
  - less of an issue
- Errors (or Efficiency)
- Learnability
- Satisfaction
An Evaluation of Interactive Map Comparison Techniques

María-Jesús Lobo  Emmanuel Pietriga  Caroline Appert

CHI 2015
Type of system

Industrial and commercial uses
(Goal = reduce costs)

e.g., banking, insurance, order entry, airline
and hotel reservations...

Importance

crucial

less of an issue

Speed of performance
Learnability
Satisfaction, Memorability
PHYSIKIT
Data Engagement Through Physical Ambient Visualizations in the Home

Steven Houben, Connie Golsteijn, Sarah Gallacher, Rose Johnson, Saskia Bakker*, Nicolai Marquardt, Licia Capra and Yvonne Rogers

University College London
*Eindhoven University of Technology

Contact: physikit.com or s.houben@ucl.ac.uk

Type of system

Toolkits for novices or designers and/or their final product for end-uses

e.g., development toolkits or platforms

Sometimes the user group may not be clear, you may want to evaluate both the support of the building process (users: builders); and of the end results if characteristics are enforced by the toolkit (users: final product users).

In some cases builders and final users can be the same people (diy)
Type of system

Toolkits/tools for building other tools/systems

e.g., programming toolkits, fabrication systems, game development platforms...

Importance

Speed of performance, errors (can be experts)
Satisfaction, Learnability (can be novices doing DIY)
Memorability

less of an issue

crucial
Type of system

Home and entertainment applications

e.g., e-mail clients, search engines, smartphones, music players, computer games, social media…

Importance

- crucial
- Satisfaction, Learnability, (Engagement)
- Errors
- Speed of performance

less of an issue
The DeepTree Exhibit:
Visualizing the Tree of Life to Facilitate Informal Learning

Florian Block 1, Michael S. Horn 2, Brenda Caldwell Phillips 4, Judy Diamond 3, E. Margaret Evans 4 and Chia Shen 1

1 Harvard University, 2 Northwestern University, 3, University of Nebraska, 4 University of Michigan
photo courtesy of Penelope Sanderson
Type of system

public walk-up and use interfaces

e.g., museum kiosks, ticket vending machines, …

Learnability

(Satisfaction, Engagement)

(Errors / Speed of performance)
Patrick Paczkowski*, Julie Dorsey*, Holly Rushmeier*, Min H. Kim+

*Yale University +KAIST

Paper3D: Bringing Casual 3D Modeling to a Multi-Touch Interface

Type of system

Exploratory, creative, and collaborative interfaces
e.g., architectural design, music-composition tools, and video-editing systems, collaborative editing...

Design and evaluation is very difficult to do because:
- Users’ motivation is often high, but so are their expectations.
- Usage can range from occasional and novice, to frequent and professional.
- Benchmark tasks are hard to describe because of the exploratory nature.

Speed of performance, Errors, Satisfaction, Learnability, Memorability …
Type of users

Experts vs. novices

Experts: speed of performance ++, learnability --

Novices: speed of performance --, learnability ++

Frequent vs. occasional

Frequent: memorability --

Occasional: memorability ++
Type of users

Cultural and international diversity

  e.g., Japanese or Chinese users scan a screen differently from French or American users

Disabilities

  e.g., vision-impaired, hearing-impaired, cognitively-impaired and mobility-impaired users require adapted input and output modalities (and adapted evaluation methods)

Older adults

Children
Evaluating Interactive Systems
How: Methods

Questionnaires
  oral interviews, paper surveys, ...

Inspection
  design walkthrough, task modelling, ...

Observation
  Think aloud, lab experiment, field study, ...