Design and Evaluation of Interactive Systems

Evaluating your system (Phase IV)

Anastasia Bezerianos
Wendy E. Mackay

in|situ| lab, INRIA & Univ Paris-Sud
29 January 2014

Homework due today
29 January 2014

1. Finished your video prototype

Some admin info

1. We meet again on 31 Jan (this Friday) @ 9am in E105
2. Exam (i.e. presentations) on 12 Fev (in 2 weeks) @ 9am in E203
3. (Short) report due on 14 Fev (the followin Friday) by email

Remember: Final project presentation (exam 12 Fev)

Oral presentation
- 15 minutes:
  - design problem
  - user profile
  - final design (concept and details)
  - video prototype (maximum 5 minutes)
  - justification and reasons for improvements (to discuss now)
- 5 minutes:
  - class discussion: every group asks at least one question

Also due: video prototype, transparencies, final storyboard, poster
Also due: any left-over exercises from the course

Remember: Final report (due on 14 Fev by email)

Executive summary of your project
10 pages maximum

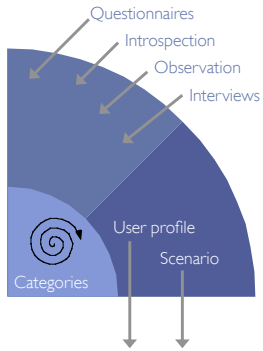
Potential users:
who are they? (refer to your data)
what do they need?
Design concept:
what is the design concept?
why is this a good solution? (avoid marketing!)
Next steps (future work)

Phase I Understanding Users

Finding out about users
Introspection
Observation
Interviews
Questionnaires

Analyze information
Grounded theory categories

Create resources for design
Scenario
User profile & Persona

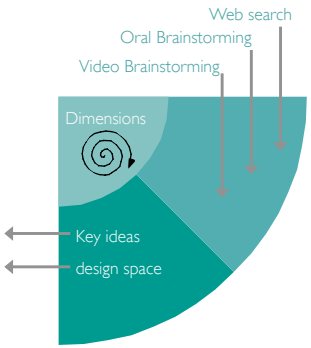


Phase II Invention

Collect or sample information
Web search
Oral Brainstorming
Video Brainstorming

Analyze information
Preference votes
Technology dimensions

Create resources for design
Key ideas
Design space (ideas to pursue)

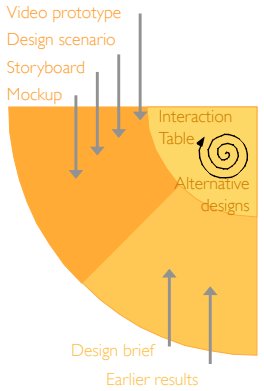


Design: What should it be?

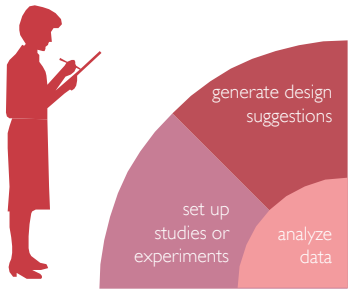
Collect or sample information
Design brief
+ results from earlier phases

Analyze information
Interaction table
Alternative designs

Create resources for design
Design scenario
Storyboard
Mockup
Video prototype



Evaluation: Does it work?



Evaluate the system

Homework of the previous week:
- Examine video prototypes

In class today:
- Phase IV: Evaluating the System

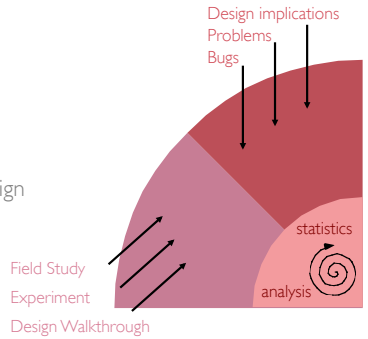
Exercises in class:
- Design Walkthrough
- Design Alternatives
- Experiment design

Evaluation: Does it work?

Collect or sample information
- Design Walkthrough
- Experiment

Analyze information
- Qualitative
- Quantitative

Create resources for design
- List of problems found
- Implications for redesign



Evaluating our system

Informal and quick techniques:
- Heuristics
- Design Walkthroughs

Formal techniques:
- Usability Studies
- Controlled Experiments
- Quasi-experiments
- Use-case studies
- … and many many more

Quick and informal evaluation

Design Walkthrough
A group evaluates a specific aspect, step-by-step:

source of a program	to find the 'bugs'
design of a system	to understand the structure
graphic screens	to get feedback from users
text (scientific papers)	to verify the structure and comprehensibility
experiments	to verify the details of the method used

Exercise: Design Walkthrough

Purpose: Help identify problems informally and quickly, using some evaluation criteria

Procedure:
- Choose a small group with different roles and expertise
- Establish a duration time, not more than 1 hour
- Chose a presenter that explains the scenario of use, each action at a time
- Chose the level of critique (system, interface, specific component)
- The group identifies as many problems as possible
- Use rules to help find problems (e.g. the usability Heuristics)

Types of possible critiques (+/-) and comments

Specific
- It takes three steps to make a simple search

Missing functionality
- No help, need to search outside

Bugs
- The import functionality of X does not work

Suggestions
- An overview of all data created is needed

General (the least useful type)
- Difficult to use too many icons

some of the usability Heuristics (Norman 1983)

Make things visible
- Know the status of the interface by observing the system
- Know what actions that can be performed

Principle of "mapping"
- To understand the correspondence between
  - Actions and results
  - Controls and their effects

Principle of feedback
- Inform the user to system status
- (before - feedforward or after feedback)

Exercise 1
29 January 2014

1. Each group will evaluate other groups' video prototype
2. Two people per group stay (presenter/guardian & moderator/scribe).
   The rest move to another group as examiners/participants for another group.
   (one move to the left, one to the right)

In-class Exercise

Walkthrough
30 min

Evaluation: Formal and targeted

Usability Study or simple experiment
Test several alternatives for the system with users:

interaction techniques
layouts of screens
help
design alternatives ...

pop-up vs. drop down menus
hypertext vs. hierarchy
tutorials vs FAQ vs search

Usability Study

Purpose: To determine the best design choices by watching users try your prototype

Procedure:
- Describe the design objective
- Identify several alternatives
- Choose the dependent & independent variables
- Make a prediction and specify null hypothesis
- Prepare the environment for each test condition (imp. method)
- Use at least 3 subjects (5 better)
- Analyze the results. Are the differences significant?

Usability Study

Quality components (dependent variables/measures)
usually tested in Usability Studies:

- Learnability
- Efficiency
- Memorability
- Errors
- Satisfaction

Many useful usability links at the useit.com website
(e.g. http://www.useit.com/alertbox/20030825.html)

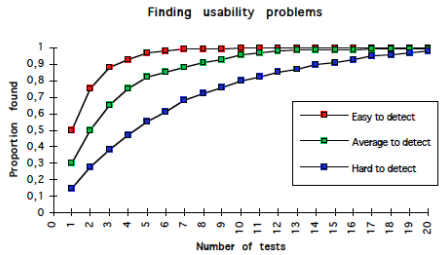
Always with real users

Problems

- How to define a step to evaluate? (one functionality vs the system)
- And if the user can not do a task?
- What if the system is not properly installed?
- What is the importance of prior user knowledge?
- Does it work differently in different contexts?

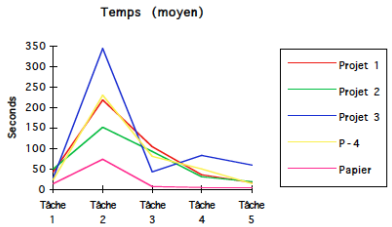
Other problems in Usability testing

How many testers are needed?



detectability d =
Probability that a test will identify a problem

Example of results: Time



++ usually accompanied by a report of the identified problems

Design a simple experiment (1)

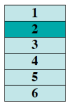
1. Specify the functionality offered to users
   What does the system do?
2. Specify alternative interaction techniques
   How does the user accomplish it?
3. Specify the hypothesis
   What do we compare and what do we predict?
4. Specify the independent variables
   What changes?
5. Specify the dependent variables
   What are we measuring?

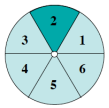
Design a simple experiment (1)

6. Operationalize the behavior (remove biases and noise)
   What are we studying?
7. Specify procedures
   What are the experimental and control groups?
8. Identify the appropriate statistical tests
   How to analyze results? Is there a difference?

Example of a hypothesis

Compare linear to pie menus





Hypothesis: pie menus are faster

Null hypothesis (that we try to disprove):
There is no difference in user performance in terms of time and error rate for the selection of an item in a linear and in a pie menu, regardless of previous user experience of using a mouse or other types of menus.

Specify the independent variables

The Independent variables (= factors) are
those we want to verify or that we want to control, independently of each other
The combinations of variables define the conditions

Independent Variables :
2 Types of menus : linear, pie
5 Number of menu items : 3, 6, 9, 12, 15
3 levels of experience : expert, novice, intermediate
=> 2 x 5 x 3 = 30 unique conditions

Note : we can treat the user experience as a factor

Specify the dependent variables

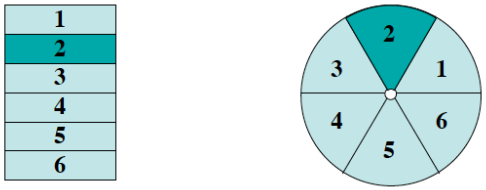
The Dependent variables (= measures) are
those we measure: they depend on the behavior of the subject and (hopefully) the independent variables

To make a reliable statistical analysis we must have adequate measures (user data) for each condition

Typical dependent variables in HCI:
- Time to select an item
- Number of errors
- Others?

Operationalize the behavior

Define the specific menu selection task
Ensure that the conditions are as similar as possible:
- Same labels for menu items
- Same menu position (center of screen)
- Highlight the item to select (to avoid searching)



Run the experiment

Prediction:
- Always write your subjective predictions
  before you discover the results
- Another example of looking for surprises

Control any factors that might bias the results:
- All subjects receive the same instructions
- All subjects perform tasks under the same conditions
- All instructions are simple and clear
  Informal contact kept to a minimum

Run the experiment

Ask subjects to sign an informed consent

Identify subjects and ensure their anonymity
- Assign a number to each subject
- Choose conditions based on that number

Collect experimental data
- Make sure they are reliable and valid (no system crashes)
- Minimize treatment when collecting
  collect raw data, you can sort them out later

Prediction ≠ Null Hypothesis

For our experiment:
- I think that whatever the expertise and size of the menu, the circular menu will be faster than the linear menu

Other predictions:
- For linear menus, performance decreases more items
- For pie menus, performance decreases with more items

Collect data (log)

Save a file that is easy to analyze by man and machine

```
Start My S1 E CL 3-12-15-9-6 November 21, 2005 3:45:54 p.m.
Condition S1 E C 3 My November 21, 2005 3:46:35 p.m.

# Subject expertise item type size hit / miss tps (ms)
Condition S1 E 12 C 3 November 21, 2005 Mon 3:54:22 p.m.
Trial S1 E C 3 2 Hit 1254
Trial S1 E C 3 1 Miss 885
...
End S1 E 12 C 3 November 21, 2005 Mon 4:23:55 p.m.
```

Exercise: Alternatives

Purpose:
- To consider more design options

Procedure:
- Choose a specific function from your functional table
- Imagine three alternative interactions (which we will operationalize to test in an experiment)

What are the advantages and disadvantages of each alternative?
- (Helps you to form predictions later on)

In-class Exercise

Define Alternatives
30 min

Exercise: Experiment and Hypothesis testing

Purpose:
- To determine the best design choices between alternatives

Procedure
- Describe your hypothesis and the null hypothesis
- Identify the independent and dependent variables
- Operationalize the behavior (interactions)
- Prepare the environment for each test condition
- ~~Use at least 3 subjects~~
- ~~Analyze the results: are the differences significant?~~

In-class Exercise

Hypothesis testing
30 min