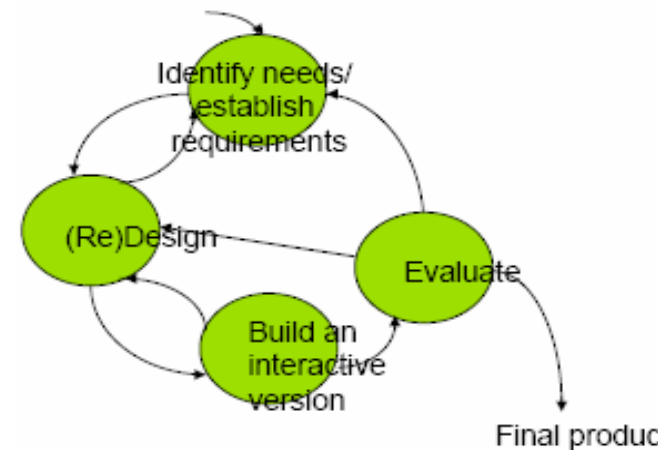# Usability evaluation

Unit 5

Learning outcomes

- Understand the main types of evaluation

- Know how to conduct a basic field study and how to analyse the data
    - Conceptual, practical and ethical issues of user evaluation in the field

- Understand how to conduct a basic user test and evaluate the results
    - Conceptual, practical and ethical issues of user evaluation

- Apply the DECIDE framework

When do you evaluate?

- **Formative evaluation**
  - During design and development process
    - Inform design
- **Summative evaluation**
  - After design is deployed
    - Measure effectiveness
    - Check standards
    - Guide adoption decisions
    - Collect requirements for future systems

Identify needs/ establish requirements

(Re)Design

Evaluate

Build an interactive version

Final produc

Evaluation

- Exploration – Conceptual design
    - Assess what it would take for a design to fulfills users' needs and likes – e.g., requirements
    - Based on scenarios, storyboard

- Development – Physical design
    - Evaluate alternatives
    - Anticipate breakdowns

- Deployment
    - Upgrade subsequent versions
    - Continual improvement
    - Collect user requirements for future systems

Overview of usability techniques

- observing users
- asking users' their opinions
- asking experts' their opinions
- testing users' performance
- modeling users' task performance

	Usability testing	Field studies	Analytical
Users	do task	natural	not involved
Location	controlled	natural	anywhere
When	prototype	early	prototype
Data	quantitative	qualitative	problems
Feed back	measures & errors	descriptions	solutions
Type	applied	naturalistic	expert

Method	Usability testing	Field studies	Analytical
Observing	X	X	
Asking users	X	X	
Asking experts		X	X
Testing	X		
Modeling			X

Usability testing

- controlled by the evaluator
- Record typical users' performance on typical tasks
- Users are monitored, recorded on video & their key presses are logged
- Output:
  - quantitative & (qualitative) data
  - performance times
  - errors
  - analysis of users' behaviour: what they did
- User satisfaction questionnaires & interviews are used to elicit opinions
  - Quantitative & qualitative data

Observation in usability testing

- More objective
- In usability lab equipment is in place
- Recording is continuous
- Analysis & observation almost simultaneous
- Care needed to avoid drowning in data
- Analysis can be coarse or fine grained
- Video clips can be powerful for reporting data

Simple observation

- User is given a task
- Evaluator just watches the user

- Problem
    - does not give insight into the user's decision process or attitude

Think aloud method

- Users speak their thoughts while doing the task
  - what they are trying to do
  - why they took an action
  - how they interpret what the system did

  - gives insight into the user's mental model of the system, but
    - may alter the way users do the task
    - unnatural (awkward and uncomfortable)
    - hard to talk if they are concentrating
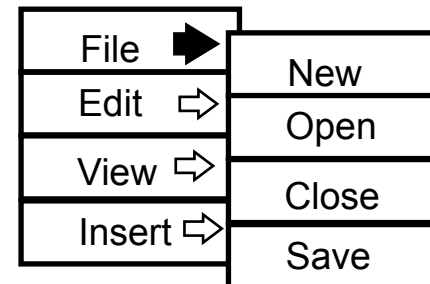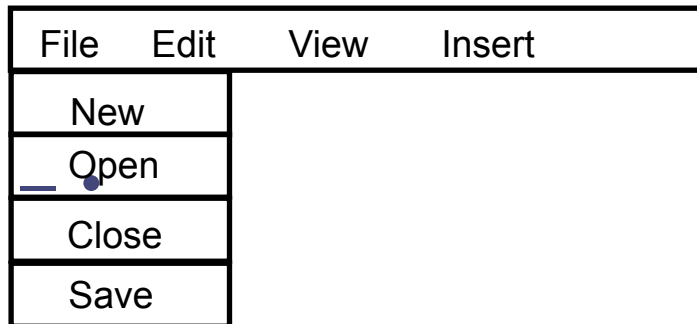
Constructive interaction method

- Two people work together on a task
  - monitor their normal conversation
  - removes awkwardness of think-aloud
  - Introduce group variability

- Co-discovery learning
  - use semi-knowledgeable "coach" and novice
  - only novice uses the interface
    - novice ask questions
    - coach responds
  - gives insights into two user groups

Testing users' performance

- Experimenter controls most environmental factors
  - Set forth a testable hypothesis
  - Manipulate one or more independent variables
  - Observe effect on one or more dependent variables
  - Nothing else changes
  - Can be reproduced by others

    *There is no difference in user performance (**time** and **error rate**) when selecting an item from a **pull down** or a **pull right** menu of 4 items*

File	Edit	View	Insert
New			
Open			
Close			
Save			

File ▶	New
Edit ⇨	Open
View ⇨	Close
Insert ⇨	Save

Field studies

- Field studies are done in natural settings

- The aim is to understand what users do naturally and how technology impacts them

- Can be used to:
    - identify opportunities for new technology
    - determine design requirements
    - decide how best to introduce new technology
    - evaluate technology in use

Field studies

- Are 'messy'
  - Activities often overlap and are constantly interrupted
  - It is often difficult to understand intention -
  - Observer immerse in the field – must have a very good knowledge of the context
  - Data is collected primarily by
    - observing or interviewing people
    - participants may also be required to fill out electronic or paper diary

Observation in natural settings

- Participant observation is key component of ethnography
- Must get co-operation of people observed
- Informants are useful
- Data analysis is continuous
- Interpretivist technique
- Questions get refined as understanding grows
- Reports usually contain examples

Internet based observations

- On-line communities, social-networking platforms, web 2.0

- The observer joins in the community and participate to it

- The observer only look at the community from the external

- Useful to understand how technology is used in real life

- http://www.wikimedia.org/

Diaries

- Critical incident diaries
  - write when something goes wrong
- Reflection diaries
  - post-usage
- Sampling activity
  - Timed
- Requires incentives
- Better if supported by different medias (e.g., cameras)

Decide

- *D*etermine the evaluation *goals*
- *E*xplore specific *evaluation questions*
- *C*hoose the *evaluation paradigm* and *techniques* to answer the questions.
- *I*dentify the *practical issues.*
- *D*ecide how to deal with the *ethical issues*.
- *E*valuate, interpret and present the *data.*

Determine the goals

- What are the high-level goals of the evaluation?

- Who wants it and why?

- The goals influence the paradigm for the study

- Some examples of goals:
  - Identify the best metaphor on which to base the design
  - Check to ensure that the final interface is consistent
  - Investigate how technology affects working practices
  - Improve the usability of an existing product

Explore the question

- *All* evaluations need goals & questions to guide them so time is not wasted on ill-defined studies

- For example, the goal of finding out why many customers prefer to purchase paper airline tickets rather than e-tickets can be broken down into sub-questions:
  - What are customers' attitudes toward e-tickets?
  - Are they concerned about security?
  - Is the interface for obtaining them poor?

Choose the evaluation paradigm

- The evaluation paradigm influences the techniques used, how data is analyzed and presented.
  - E.g. field studies do not involve testing or modeling
- Select the techniques (e.g., thinking aloud or interviews)

Identify practical issues

For example, how to:

- select users
- stay on budget
- stay on schedule
- find evaluators
- select equipment

Decide on ethical issues

- Testing is a distressing experience - Pressure to perform
- Develop an informed consent form

- Participants have a right to:
  - know the goals of the study
  - what will happen to the findings
  - privacy of personal information
  - not to be quoted without their agreement
  - leave when they wish
  - be treated politely

MANCHESTER BUSINESS SCHOOL

Consent form

Title of Project:

Fantasy Play in Real and Virtual Environments

The parent/guardian should complete the following section.

Please circle your

answer and provide

1. Have you read the Parent Information Sheet?	**YES/NO** ....
2. Have you received enough information about the study?	**YES/NO** ....
3. Do you consent to be audio taped/video taped/photographed as detailed in the Parent Information Sheet?	**YES/NO** ....
4. Do you understand that your child do not need to take part in the study and if your child do enter your child are free to withdraw:- * at any time * without having to give a reason for withdrawing * and without detriment to you	**YES/NO** ....
5. Do you/your child agree to take part in this study?	**YES/NO** ....

Name of child: ……….…..……..……. Signed: .... Date: ....
(parent/guardian)

Name of researcher: ....….. Signed: .... Date: ....

This project has been approved by the

Manchester Business School Research Ethics Committee

Project no

07149

MANCHESTER BUSINESS SCHOOL

Parent Information Sheet

Title of project:

Fantasy Play in Real and Virtual Environments

Introduction

Dear Parent(s) / Guardian(s),

My name is Evi Mansor and I am a PhD student working under the supervision of Dr Antonella De Angeli at Manchester Business School, University of Manchester where I am currently conducting a study on fantasy/pretend play in preschool children ages 3-4 years old.

Your child is invited to take part in this study. Please read the following information carefully and discuss it with your child. Feel free to contact me if there is anything that is not clear or if you would like more information.

What will my child have to do if he/she takes part?

The study will be conducted at the Webster Primary School from 3rd March 2008 – 19th March 2008. The class teacher will assign your child based on their friendships and each research session will involve two children. A guide of the exact procedure will be provided and explained to your child before the session. The session will be audio/video recorded with your consent to allow us to transcribe the children's responses. All research team members have completed the Criminal Records Bureau checks.

A) Part 1 - Observation

Children will be asked to "play" in pair alternately in two different conditions for 10 minutes in each condition and I will observe their behaviour. I'll take extreme care to use a language which is appropriated to your child age. Class teacher is invited to observe the study.

Condition 1 (Physical setting)

Children will be asked to play with physical objects in pairs on the provided small table with physical materials given (a wooden tree house, realistic and non realistic objects) to support their play.

Condition 2 (Virtual setting)

Children will be asked to play with virtual objects in pairs on the tabletop environment (60cm x 76cm) using the software that will be developed by the researcher. The software will contain a virtual tree house and virtual realistic and non realistic objects. Extra features will be included such as animations and sounds. Children can select, drag or organise the virtual objects on the table surface by using their fingers.

The *DiamondTouch* table will be used as interactive device. DiamondTouch is a front-projected table linked to a computer and a projector. Users can retrieve and interact with computer based resources by touching the surface of the table with their fingers. A short demonstration on how to use the application will be conducted by the researcher. Then, participants will be invited to "play" with the software with their partner and the researcher will observe their behaviour. A person (research assistant) will sit next to the table in case the children will need help while they are using the application.

At the end of the observation sessions, the researcher will briefly ask the children about their experience and preferences. Completion of the task will require approximately 30 minutes.

B) Part 2 - Interview

A week later, the children will be interviewed one by one by myself and complete two short language tasks after an initial chat. During the first part of the interview they will be asked to name their favourite story, favourite toy, favourite TV show, favourite game and they will also be asked about their imaginary companion/friends. After this initial chat they will be asked to complete two tests: a vocabulary test and a perspective taking test. During the vocabulary test the children will be given a word and a selection of four pictures. Their task is to select the picture that corresponds to the word they have just heard. During the second test they will take part in a short guessing game in which they have to find the location of a hidden object. The interview will last approximately 20 minutes.

Will the data be anonymous?

Yes, the children's names will never be used during the publication of the results. For analysis purposes, each child's name will be replaced with a code. Only the codes will be used in presentations of the data.

Will the data be confidential?

Yes, only the research team will have access to the data. The data will be kept strictly confidential. Forms, written records, audio and video recordings will be kept in a secured unit in the Manchester Business School and electronic data will be kept in a password protected computer. The data and consent forms will be kept for a period of 5 years for the purpose of publishing the results.

Does my child have to take part?

Participation is entirely voluntary. If you and your child decide to take part please fill in the attached consent form. Please note that even if your child decides to take part, he or she is free to withdraw from the study at any time with no need to justify his/her decision. Similarly, you may decide to withdraw you child's participation in the study at any time if you wish so. I will take particularly care to interrupt the study if the child show any sign of distress.

Where can I obtain further information if I need it?

If you have any further questions or queries please contact:

Evi Mansor (PhD student): Evi.Mansor@postgrad.manchester.ac.uk

Dr Antonella De Angeli (Supervisor): Antonella.De-angeli@manchester.ac.uk

Evaluate, interpret & present data

- How data is analyzed & presented depends on the paradigm and techniques used

- The following points <u>MUST</u> be considered:
  - Reliability: can the study be replicated?
  - Validity: is it measuring what you thought?
  - Biases: is the process creating biases?
  - Scope: can the findings be generalized?
  - Ecological validity: is the environment of the study influencing it

Data collection

- Notes, paper and pencil
    - primitive but cheap
    - observer records events, comments, and interpretations
    - hard to get detail (writing is slow)
    - 2nd observer helps…

- audio recording
    - good for recording think aloud talk
    - hard to tie into on-screen user actions

- video recording
    - can see and hear what a user is doing
    - one camera for screen (screen capture software), rear view mirror useful…
    - initially intrusive

- Tracking users:
- diaries
- interaction logging

Usability specification table

Scenario task	Worst case	Planned Target	Best case (expert)	Observed
Find most expensive house for sale?	1 min.	10 sec.	3 sec.	??? sec
…				

Coding sheet example

tracking a person's use of an editor

	General actions			Graph editing			Errors	
Time	text editing	scrolling	image editing	new node	delete node	modify node	correct error	miss error
09:00	X							
09:02				X				
09:05							X	
09:10					X			
09:13								

Data analysis

- *Qualitative data - interpreted* & used to tell the 'story' about what was observed
  - *categorized* using techniques such as content analysis.

- *Quantitative data* - collected from interaction, questionnaires, & video logs.
  - Presented as values, tables, charts, graphs and treated statistically.

Quantitative data

- Data based on numbers
  - Number of students expressing satisfaction with a lecture
  - A company annual turnover
  - Time in milliseconds to activate a command
  - Number of people clicking a link more than three times per week

- Look for patterns in the data and draw conclusion
  - Table graphs or charts
  - Descriptive statistics
  - Inferential statistics

In practice

- Determine the TASK
  - Simple, clearly stated, understandable
- Determine the performance measures
- Develop the evaluation
- Ethical approval (?)
- Recruit participants
- Collect the data
- Inspect & analyze the data
- Draw conclusions to resolve design problems
- Redesign and implement the revised interface

Key points

- Field studies provide real life evidence
  - Data are obtained by asking or observing people
  - Qualitative analysis looks for patterns and common themes
- Usability studies provide controlled evidence
  - Different evaluation techniques: observation, think aloud
  - Questionnaire design: different question styles

- The *DECIDE* framework has six parts:
  - *D*etermine the overall goals
  - *E*xplore the questions that satisfy the goals
  - *C*hoose the paradigm and techniques
  - *I*dentify the practical issues
  - *D*ecide on the ethical issues
  - *E*valuate ways to analyze & present data