An Evaluation Framework

The aims

- Explain key evaluation concepts & terms
- Describe the evaluation paradigms & techniques used in interaction design
- Discuss the conceptual, practical and ethical issues that must be considered when designing evaluations
- Introduce the DECIDE framework

Evaluation paradigm

- Any kind of evaluation is guided explicitely or implicitely by a set of beliefs, which are often under-pined by theory. These beliefs and the methods associated with them are known as an "**evaluation paradigm**"

User studies

- **User studies** involve looking at how people behave in their natural environments, or in the laboratory, both with old technologies and new ones

Evaluation paradigms

- "Quick and dirty"
- Usability testing
- Field studies
- Predictive evaluation

Quick and dirty

- "**Quick and dirty**" evaluation describes the common practice in which designers informally get feedback from users or consultants to confirm that their ideas are in-line with users' needs and are liked
- "Quick and dirty" evaluations are done any time
- The emphasis is on fast input to the design process rather than carefully documented findings

Usability testing

- Usability testing involves recording typical users' performance on typical tasks in controlled settings.
- As the users perform these tasks they are watched & recorded on video & their key presses and mouse clicks are logged
- This data is used to calculate performance times, identify errors & help explain why users did what they did
- User satisfaction questionnaires & interviews are used to elicit users' opinions

Field studies

- Field studies are done in natural settings
- The aim is to understand what users do naturally and how technology impacts them
- In design, field studies can be used to:
  - Identify opportunities for new technology
  - Determine design requirements
  - Decide how to best introduce new technology
  - Evaluate technology in use

Predictive evaluation

- Experts apply their knowledge of typical users often guided by heuristics, to predict usability problems
- Another approach involves theoretically-based models
- A key feature of predictive evaluation is that users need not be present
- Relatively quick and inexpensive

Overview of techniques

- Observing users
- Asking users their opinions
- Asking experts their opinions
- Testing users' performance
- Modeling users' task performance to predict the efficacy of a user interface
- IMPORTANT: some techniques are used in different ways in different evaluation paradigms

DECIDE: a framework to guide evaluation

- *D*etermine the goals the evaluation addresses
- *E*xplore the specific questions to be answered
- *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 to base the design
    - Check to ensure that the final interface is consistent
    - Investigate how technology affects working practices

Explore the questions

- All evaluations need goals & questions to guide them so time is not wasted on ill-defined studies
- For example, the goal of finding out how many customers prefer to purchase paper airline tickets rather than e-tickets can be broken down into subquestions:
  - What are the customers' attitudes to these new tickets ?
  - Are they concerned about security ?
  - Is the interface for obtaining them poor ?
- What questions might you ask about the design of a cell-phone ?

Choose the evaluation paradigm & techniques

- The evaluation paradigm strongly influences the techniques used, how data is analyzed and presented
- E.g field studies do not involve testing and modeling

Identify practical issues

- For example, how to:
    - Select users
    - Stay on budget
    - Staying on schedule
    - Find evaluators
    - Select equipment

Decide on ethical issues

- Develop an informed consent form
- Participants have a right to:
  – Know the goals of the study
  – What will happen with the findings
  – Privacy of personal information
  – Not to be quoted without their agreement
  – Leave when they wish
  – Be treated politely
- Uvic human ethics board
  – Review documentation, approve

Evaluate, interpret and present data

- How data is analyzed and presented depends on the paradigm and techniques used
- The following also need to be considered:
  - Reliability: can the study be replicated ?
  - Validity: is it measuring what you thought ?
  - Biases: it the process creating biases ?
  - Scope: can the findings be generalized
  - Ecological validity: is the environment of the study influencing it e.g Hawthorn effect (or easy midterm effect)

Pilot Studies

- A small trial run of the main study
- The aim is to make sure your plan is viable
- Pilot studies check:
  - That you can conduct the procedure
  - The interview scripts, questionnaires, experiments, etc. work appropriately
  - It's worth doing several to iron out problems before the main study
  - Ask colleagues if you can't spare "real" users

Key points

- An evaluation paradigm is an approach that is influenced by particular theories and philosophies
- Five categories of techniques were identified: observing users, asking users, asking experts, user testing, modeling users
- The DECIDE framework has six parts:
  - Determine the overall goals
  - Explore the questions that satisfy the goals
  - Choose the paradigm and techniques
  - Identify the practical issues
  - Decide on the ethical issues
  - Evaluate ways to analyze and present data
- DO A PILOT STUDY

Based on the slides available at www.id-book.com