HUMAN-COMPUTER INTERACTION THIRD EDITION

DIX FINLAY ABOWD BEALE

chapter 9

evaluation techniques

Evaluation Techniques

- Evaluation
  - tests usability and functionality of system
  - occurs in laboratory, field and/or in collaboration with users
  - evaluates both design and implementation
  - should be considered at all stages in the design life cycle

Goals of Evaluation

- assess extent of system functionality
- assess effect of interface on user
- identify specific problems

Evaluating Designs (expert based)

Cognitive Walkthrough
Heuristic Evaluation
Review-based evaluation

Cognitive Walkthrough

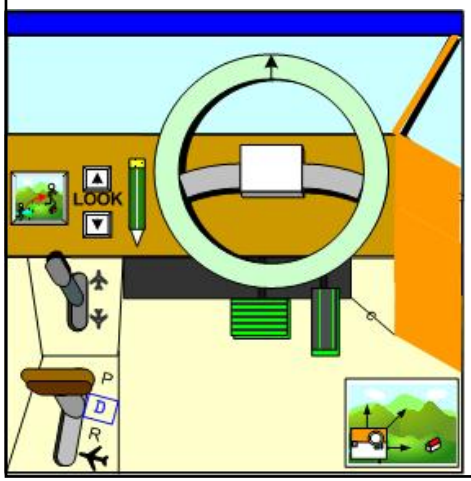
Proposed by Polson *et al. 1992*

- evaluates design on how well it supports user in learning task
- usually performed by expert in cognitive psychology
- expert 'walks through' design to identify potential problems using psychological principles
  - Based on the idea of a code walkthrough in conventional code testing
- forms used to guide analysis
- can be used to compare alternatives

Cognitive Walkthrough (ctd)

- For each task walkthrough considers
  - what impact will interaction have on user?
  - what cognitive processes are required?
  - what learning problems may occur?
- Analysis focuses on goals and knowledge: does the design lead the user to generate the correct goals?

Pen-based interface for LIDS



- UA1: Press look up button
- SD1: Scroll viewpoint up
- UA2: Press steering wheel to drive forwards
- SD2: Move viewpoint forwards
- UA3: Press look down button
- SD3: Scroll viewpoint down
-
-
-

Pen interface walkthrough

- UA 1: Press look up button
  1. Is the effect of the action the same as the user's goal at this point?
     Up button scrolls viewpoint upwards.
  2. Will users see that the action is available?
     The up button is visible in the UI panel.
  3. Once users have found the correct action, will they know it is the one they need?
     There is a lever with up/down looking symbols as well as the shape above and below the word look. The user will probably select the right action.
  4. After the action is taken, will users understand the feedback they get?
     The scrolled viewpoint mimics the effect of looking up inside the game environment.

Cognitive walkthrough results

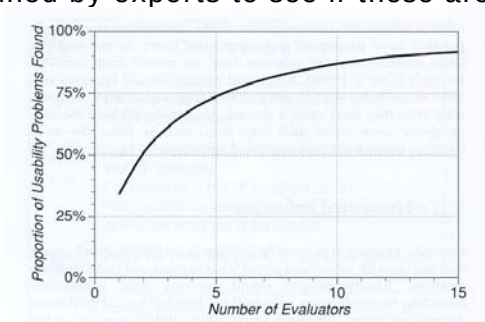
- Fill out a form
  - Track time/date of walkthrough, who the evaluators were
  - For each Action, answer the four proforma questions (as per prev slide, text pp. 321-322)
  - Any negative answer to any question should be documented on a separate Problem Sheet, indicating how severe the evaluators think the problem is, and whether they think it'll occur often

When to do a Cognitive Walkthrough

- Can be done at any stage in the development process once you have a 'prototype' or actual system implementation to work on
  - Can be done with paper prototype
  - Can be done with a shrink-wrapped product
- Focus on key tasks
  - Things that are done by most users
  - "Critical success factors" of the system
  - Consider something that matches the name of the product
    - If it's an email client, do a cognitive walkthrough of the task of writing and sending an email

Heuristic Evaluation

- Proposed by Nielsen and Molich.
- usability criteria (heuristics) are identified
- design examined by experts to see if these are violated



Heuristic Evaluation

- Rank by severity
  - 0=no usability problem
  - 1=cosmetic – fix if have extra time
  - 2=minor – fixing is low priority
  - 3=major – important to fix
  - 4=usability catastrophe – imperative to fix
- Heuristics such as 10 from Nielsen
  - Visibility of system status
  - Match between system and real world
  - User control and freedom, etc.

  (p. 325-326) [remember – these will be used to assess your assignment 1 prototype!]
- Heuristic evaluation `debugs' design

Nielsen's 10

- Visibility of system status
- Match between system and real world
- User control and freedom
- Consistency and standards
- Error prevention
- Recognition rather than recall
- Flexibility and efficiency of use
- Aesthetic and minimalist design
- Help users recognize, diagnose and recover from errors
- Help and documentation

When to use Heuristic Evaluation

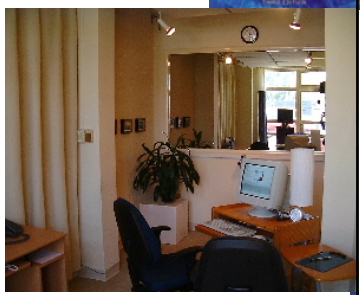
- Particular advantage that it can be used very early
  - From first sketches and outline descriptions
  - May head off a mistake rather than having to fix it
- Called a 'discount usability' method, because it's relatively cheap (doesn't require a lot of time and effort)

Review-based evaluation

- Results from the literature used to support or refute parts of design
  - Care needed to ensure results are transferable to new design
- Model-based evaluation (e.g., GOMS, keystroke)
  - Cognitive models used to filter design options
    e.g. GOMS prediction of user performance
  (we look at these later in the semester)
- Design rationale can also provide useful evaluation information

Evaluating through user Participation

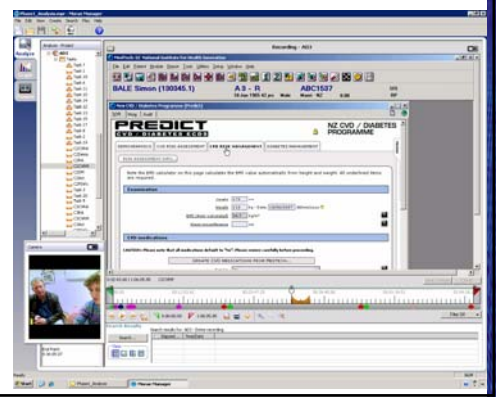
Laboratory studies



- User taken out of their normal work environment for a controlled test
- Advantages:
  - specialist equipment available
  - uninterrupted environment
- Disadvantages:
  - lack of context
  - difficult to observe several users cooperating
- Appropriate
  - if system location is dangerous or impractical for constrained single user systems to allow controlled manipulation of use

Laboratory study with Morae

- We've been analyzing a clinical decision support tool
  - PREDICT CVD/Diabetes estimates risk of a 'cardiovascular disease event' (e.g., stroke or heart attack) and gives recommendations for action, as well as information for patients
- We're interested in just how GPs and nurses are using the system – how much time they spend on what tasks, what they convey to patient.
- Difficult to coordinate detailed study in a real General Practice (consent of patient to be videotaped, placing instrumentation in the practice – e.g., Morae)
  - So, we use 'medical actors' and bring the real GP to the Tamaki clinic



Field Studies

- Advantages:
  - natural environment
  - context retained (though observation may alter it – *see next slide*)
  - longitudinal studies possible (i.e., over longer periods of time than are feasible for bringing people into a lab)
- Disadvantages:
  - distractions
  - noise
- Appropriate
  - where context is crucial and for longitudinal studies

Unwanted biases in studies

- You can't always take a study result at face value… must be attentive to what subjects are feeling
- Hawthorne effect
  - Worker is more productive when observed
- John Henry effect
  - Worker is [stubbornly] more productive when using his old tools (see http://www.ibiblio.org/john_henry/)
- Placebo effect
  - [Patient usually] gets some benefit just because they expect a benefit
- Pygmalion effect
  - Student performs better simply because they are expected to do so