Observational Methods

Think Aloud
Cooperative evaluation
Protocol analysis
Automated analysis
Post-task walkthroughs

Think Aloud

- user observed performing task
- user asked to describe what he is doing and why, what he thinks is happening etc.

- Advantages
  - simplicity - requires little expertise
  - can provide useful insight
  - can show how system is actually used
- Disadvantages
  - subjective
  - selective
  - act of describing may alter task performance

Cooperative evaluation

- variation on think aloud
- user collaborates in evaluation
- both user and evaluator can ask each other questions throughout
- Advantages over individual think-aloud
  - less constrained and easier to use
  - user is encouraged to criticize system
  - clarification possible
- Disadvantage
  - Evaluator may inadvertently (or intentionally!) cover up the user getting stuck by being too helpful

---

Protocol analysis

- paper and pencil – cheap, limited to writing speed
- audio – good for think aloud, difficult to match with other protocols
- video – accurate and realistic, needs special equipment, obtrusive
- computer logging – automatic and unobtrusive, large amounts of data difficult to analyze
- user notebooks – coarse and subjective, useful insights, good for longitudinal studies

- Mixed use in practice
- Audio/video transcription difficult and requires skill
- Some automatic support tools available

Post-task walkthrough

• user reacts on action after the event
• used to fill in intention

• Advantages
  – analyst has time to focus on relevant incidents
  – avoid excessive interruption of task
• Disadvantages
  – lack of freshness
  – may be post-hoc interpretation of events (people often can rationalize action with no relation to what they were really thinking at the time)

5

Query Techniques

Interviews
Questionnaires

Interviews

• analyst questions user on one-to-one basis usually based on prepared questions
• informal, subjective and relatively cheap

• Advantages
  – can be varied to suit context
  – issues can be explored more fully
  – can elicit user views and identify unanticipated problems
• Disadvantages
  – very subjective
  – time consuming

7

Questionnaires

• Set of fixed questions given to users

• Advantages
  – quick and reaches large user group
  – can be analyzed more rigorously
• Disadvantages
  – less flexible
  – less probing

8

Questionnaires (contd.)

- Need careful design
  - what information is required?
  - how are answers to be analyzed?

- Styles of question
  - general
  - open-ended
  - scalar
  - multi-choice
  - ranked

- Likert scales are popular
  - see http://www.gifted.uconn.edu/siegle/research/instrument%20Reliability%20and%20Validity/Likert.html)
  - Commonly on a 4, 5 or 7-point scale (e.g., Strongly Disagree, Agree, Neutral, Agree, Strongly Agree)

9

Physiological methods

Eye tracking
Physiological measurement

---

eye tracking

- head or desk mounted equipment tracks the position of the eye
- eye movement reflects the amount of cognitive processing a display requires
- measurements include
  - fixations: eye maintains stable position. Number and duration indicate level of difficulty with display
  - saccades: rapid eye movement from one point of interest to another
  - scan paths: moving straight to a target with a short fixation at the target is optimal

11

---

physiological measurements

- emotional response linked to physical changes
- these may help determine a user's reaction to an interface
- measurements include:
  - heart activity, including blood pressure, volume and pulse.
  - activity of sweat glands: Galvanic Skin Response (GSR)
  - electrical activity in muscle: electromyogram (EMG)
  - electrical activity in brain: electroencephalogram (EEG)
- some difficulty in interpreting these physiological responses - more research needed
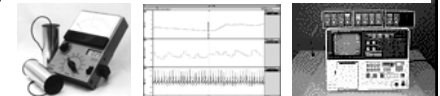
Choosing an Evaluation Method

when in process: design vs. implementation

style of evaluation: laboratory vs. field

how objective: subjective vs. objective

type of measures: qualitative vs. quantitative

level of information: high level vs. low level

level of interference: obtrusive vs. unobtrusive

resources available: time, subjects,
                       equipment, expertise