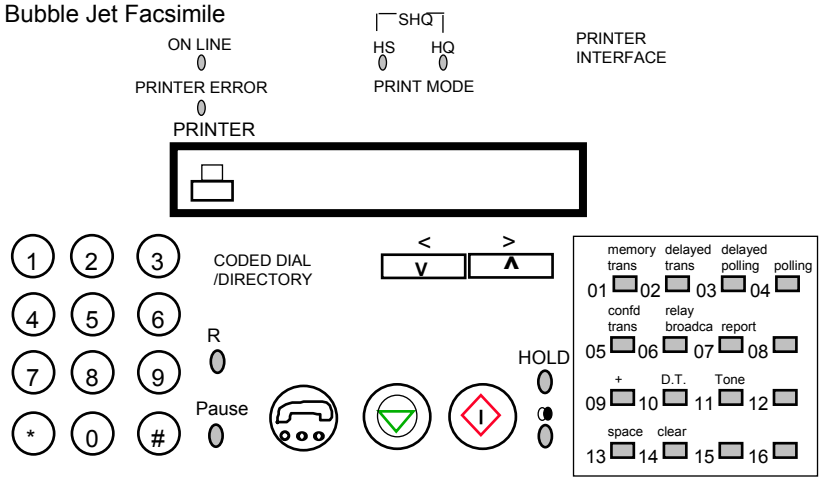
Qualitative Evaluation Techniques

Quickly debug and evaluate prototypes by observing people using them

Specific evaluation methods helps you discover people's thoughts and motivations as they are using your system

Canon
Fax-B320
Bubble Jet Facsimile



Discount Usability Evaluation

Low cost methods to gather usability problems

- Approximate: capture most large and many minor problems

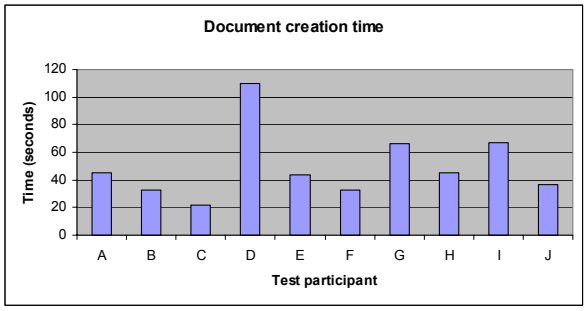
How?

- Quantitative
- Qualitative

Quantitative Approach For Usability Evaluation

Description of approach:

- Measure something of interest in user actions
- Count, log, speed, error rate



Qualitative Methods For Usability Evaluation

Description of approach:

- Data gathering
  - Observe the actions of the user
  - Gather opinions from the user
- Produces a description, usually in non-numeric terms
- May be subjective

Methods

- Inspection
- Extracting the conceptual model
- Direct observation
  - Simple observation
  - Think-aloud
  - Constructive interaction

Star Trek IV: The Voyage Home © Paramount Pictures

- Query via interviews and questionnaires
- Continuous evaluation via user feedback and field studies

The Inspection Method

Designer tries the system (or prototype) out

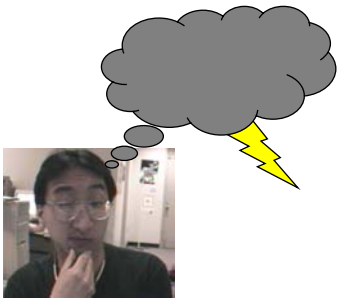
- Does the system "feel right"?

Benefits

- Can probably notice some major problems in early versions during every day use

Problems

- Not reliable as completely subjective
- Not valid as inspector is a non-typical user
- Intuitions and introspections are often wrong



Extracting The Conceptual Model

Show the user static images of:

- The paper prototype *or*
- Screen snapshots *or*
- Actual system screens during use

Have the user try to explain

- What all elements are
- What they would do to perform a particular task
- How they think that the system works

Initial vs. formative conceptual models

- **Initial:** How person perceives a screen the very first time it is viewed
- **Formative**: The same, except after the system has been used for a while

This approach is:

- Good for eliciting people's understanding before & after use
- Requires active intervention by evaluator, which can get in the way

James Tam

Direct Observation

Evaluator observes and records users interacting with design/system

- In lab:
  - User asked to complete a set of pre-determined tasks
  - A specially built and fully instrumented usability lab may be available
- In field:
  - User goes through normal duties

This approach is:

- Validity/reliability depends on how controlled/contrived the situation is
- Excellent at identifying gross design/interface problems

Three general approaches:

- Simple observation/Silent observer
- Think-aloud
- Constructive interaction

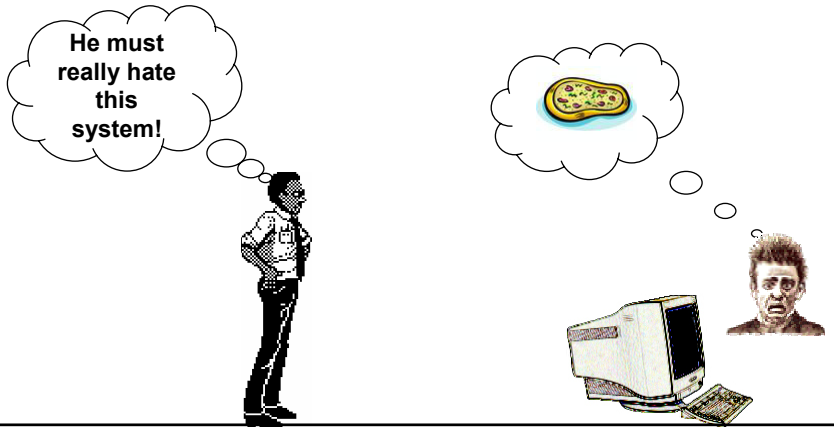
James Tam

Simple Observation Method

Person is given the task, and evaluator silently just watches while employing "The Silent Observer" technique.

Problem

- Does not give insight into the person's decision process or attitude

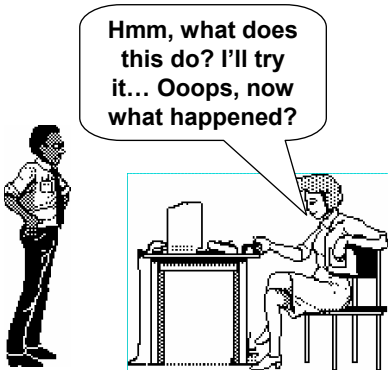


James Tam

The Think Aloud Method

Test participants are asked to say what they are thinking/doing

- Gives insight into what the person is thinking
  - What they believe is happening
  - What they are trying to do
  - Why they took an action



James Tam

The Think Aloud Method (2)

Problems

- Awkward/uncomfortable for person (thinking aloud is not normal!)
- "Thinking" about it may alter the way people perform their task
- Hard to talk when they are concentrating on problem

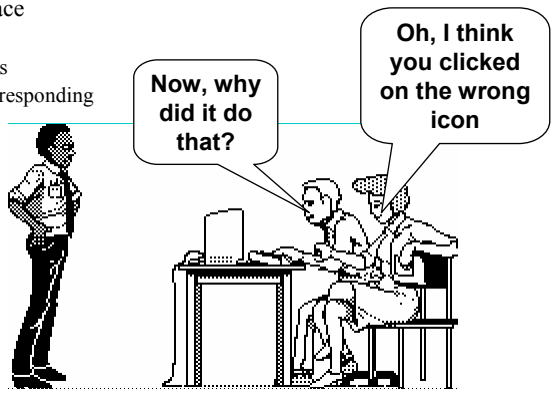
Most widely used "formal" evaluation method in industry

The Constructive Interaction Method

Two people work together on a task

- Normal conversation between the two users is monitored
  - Removes awkwardness of think-aloud
- *Variant*: Co-discovery learning
  - Use semi-knowledgeable "coach" and novice user together
  - Only novice uses the interface
  - Results in:

    Novice user asking questions
    Semi-knowledgeable coach responding
    Provides insights into thinking process of both user groups



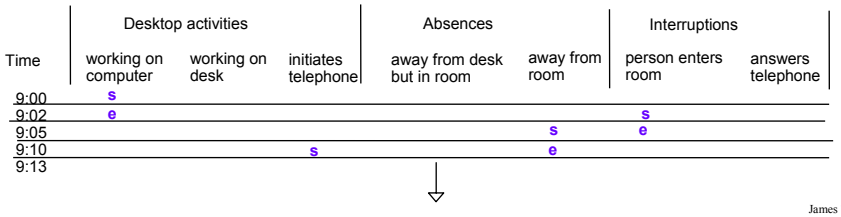
Recording Observations

How do we record user actions during observation for later analysis?

- If no record is kept, evaluator may forget, miss, or mis-interpret events

- Paper and pencil
  - Primitive but cheap
  - Evaluators record events, interpretations, and extraneous observations
  - Hard to get detail (writing is slow)
  - Coding schemes or having a second observer may be helpful:

s = start of activity
e = end of activity



James Tam

Recording Observations (2)

- Audio recording
  - Good for recording talk produced by thinking aloud/constructive interaction
  - Hard to tie into user actions (i.e., what they are doing on the screen)

- Video recording
  - Can see and hear what a user is doing
  - One camera for screen, another for test user (picture in picture)
  - Can be intrusive during initial period of use

James Tam

Querying People Via Interviews

Use a set of pre-created questions

- Gets things started
- Focuses the interview
- Ensures a base of consistency
- Be sure to follow interesting leads rather than bulldozing through question list

Adding additional questions

- Could be based on results of user observations

Don't forget

- Balance each question
- Avoid bias
  - Try not to ask leading questions

THE INTERVIEWER DIDN'T EAT LUNCH.

Issues Associated With Interviews

Excellent for pursuing specific issues

- Flexible
  - You can vary questions to suit the context
- Provides a rich depth of data
  - Probe more deeply on interesting issues as they arise
  - Often leads to specific constructive suggestions

Problems:

- Accounts are subjective
- Time consuming
- Evaluator can easily bias the interview
- Prone to rationalization of events/thoughts by person
  - Reconstruction may be wrong
- Requires a skilled and/or experienced interviewer

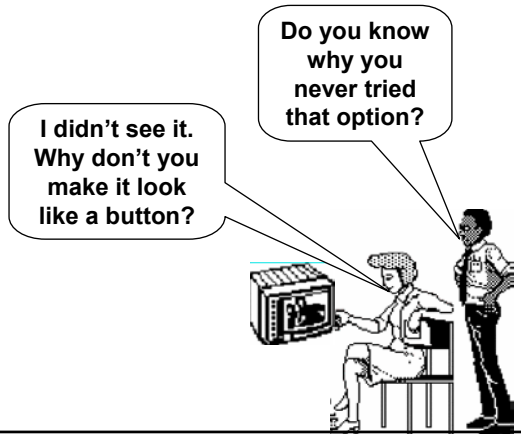
Group Discussions

- **Start with individual discussions to discover different perspectives, and then continue with group discussions**
- **Increasing group size may increase the universality of the comments**
- **May encourage cross discussions.**

Retrospective Testing

Post-observation interview to clarify events that occurred during system use

- Perform an observational test
- Create a video record of it
- Have users view the video and comment on what they did



Retrospective Testing (2)

Benefits

- Excellent for grounding a post-test interview
- Avoids erroneous reconstruction
- It can be used when thinking aloud is not possible
- Users often offer concrete suggestions

Drawbacks

- Time consuming

Querying People Via Questionnaires And Surveys

Questionnaires / Surveys

- Written queries for usability information

Benefits

- Administration cheap
  - Can reach a wide test group (e.g. mail)
- Administration requires little training
- Results can be quantified
- Anonymous

Drawbacks

- Preparation "expensive" – although this may balanced off by the administrative savings
- Inflexible

Querying People Via Questionnaires / Surveys (2)

Approach for all types

- Establish the purpose of the questionnaire
  - What information is sought?
  - How would you analyze the results?
  - What would you do with your analysis?

  http://www.cpsc.ucalgary.ca/~tamj/481/assignments/usability/questionnaire_tips.html

- Do not ask questions whose answers you will not use!
  - e.g. How old are you?

- Determine the audience you want to reach
  - Typical survey: random sample of between 50 and 1000 users of the product

- Determine how would you will deliver and collect the questionnaire
  - On-line for computer users
  - Web site with forms
  - Surface mail

    Be sure to include a pre-addressed reply envelope to get a far better response rate

Querying Users Via Questionnaires / Surveys (3)

- Determine the demographics
  - e.g., computer experience

Style Of Questions

Open-ended questions

- Asks for unprompted opinions
- Good for general subjective information but difficult to analyze rigorously

e.g., Can you suggest any improvements to the interfaces?

Style Of Questions

Closed-ended questions

- Restricts the respondent's responses by supplying alternative answers
- Data is more narrow (less rich but can be easily analyzed)
- But watch out for hard to interpret responses - alternative answers should be very specific
- Types: scalar, multiple choice, ranked
- Examples:

Do you use computers at work:

O Often O Sometimes O Rarely

vs.

In your typical work day, do you use computers:

O Over 4 hrs a day
O Between 2 and 4 hrs daily
O Between 1and 2 hrs daily
O Less than 1 hr a day

Closed-Ended Questions: Scalar

Scalar

- Ask user to judge a specific statement on a numeric scale
- Scale usually corresponds with agreement or disagreement with a statement

Characters on the computer screen are:

Hard to read Easy to read

1 2 3 4 5

Closed-Ended Questions: Multiple Choice

Multi-choice

- Respondent offered a choice of explicit responses

How do you most often get help with the system? (Check only one category)

O On-line manual
O Paper manual
O Ask a colleague

Which types of software have you used? (Check all that apply)

O Word processor
O Data base
O Spreadsheet
O Compiler

Closed-Ended Questions: Ranked

Ranked

- Respondent places an ordering on items in a list
- Useful to indicate a user's preferences
- Forces a choice

Rank the usefulness of these methods of issuing a command
(1 = Most useful, 2 = Next most useful..., 0 = Not used
__2__ Command line
__1__ Menu selection
__3__ Control key accelerator

Mixing Questionnaire Styles

Combining open-ended and closed-ended questions

- Get a specific response, but allows room for user's opinion

It is easy to recover from mistakes:

Disagree Agree
1 2 3 (4) 5 Comment: *The undo facility is really helpful*

Interviews Vs. Questionnaires: Summary Of The Pros And Cons

Preparation time

Unanticipated/unexpected events

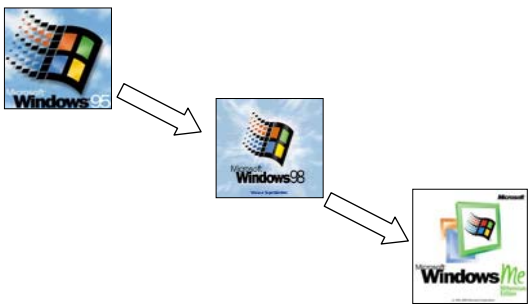
Depth of information

Analysis time

Continuous Evaluation

1) Developers monitor system while it's actually being used

- Usually done in later stages of development
  - i.e., Beta releases, delivered system
- Good for finding real-world problems
- Problems can be fixed in next release



Windows is the property of Microsoft Corporation

Continuous Evaluation (2)

2) Users can provide feedback

- Email
- Special built-in gripe facility (web site, bulletin board)
- Telephone hot line
- Help desks
- Suggestion boxes

Best combined with trouble-shooting facility

- Users always get a response (solution?) to their problem

Continuous Evaluation (3)

3) Case/field studies

- Careful study of "system usage" at the site
- Good for seeing "real life" use
- Can be informal or more rigorous qualitative approaches can be attempted

What You Now Know

Observing a range of users use your system for specific tasks reveals successes and problems

Qualitative observational tests are quick and easy to do

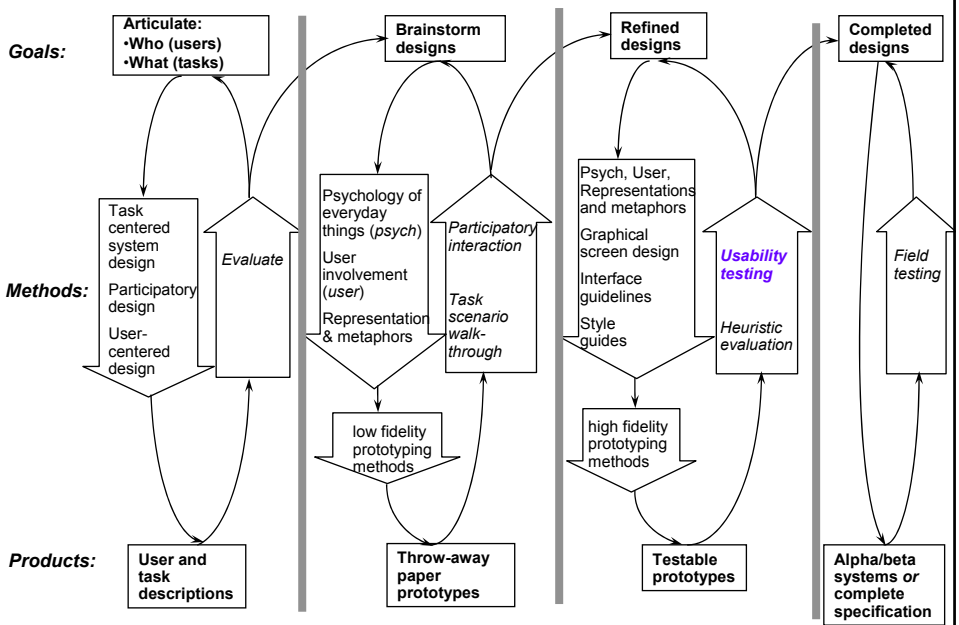
Several methods reveal what is in a person's head as they are doing the test

Particular methods include

- Conceptual model extraction
- Direct observation
  - Simple observation
  - Think-aloud
  - Constructive interaction (Co-discovery learning)
- Query via interviews, retrospective testing and questionnaires
- Continuous evaluation via user feedback and field studies

James Tam

Interface Design And Usability Engineering



This diagram is a variation of the one presented by Saul Greenberg

James Tam