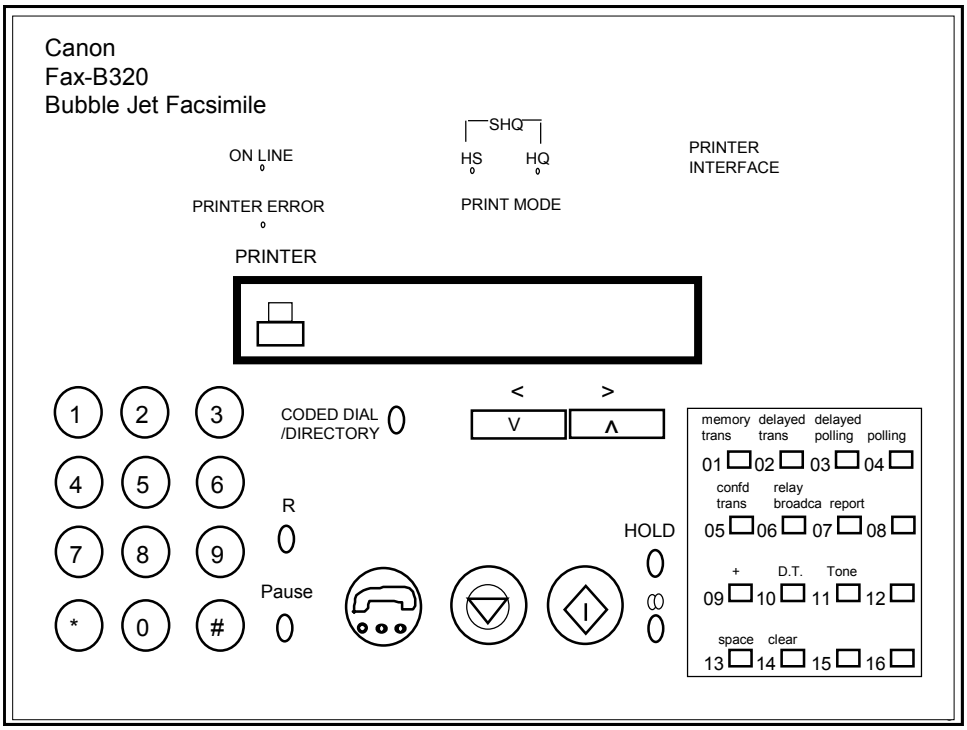
Qualitative Evaluation Techniques

Quickly debug and evaluate prototypes by observing people using them

Specific evaluation methods helps you discover what a person is thinking about as they are using your system

Saul Greenberg



Qualitative methods for usability evaluation

Qualitative:
- produces a description, usually in non-numeric terms
- may be subjective

Methods
- Introspection
- Extracting the conceptual model
- Direct observation
  - simple observation
  - think-aloud
  - constructive interaction
- Query via interviews and questionnaires
- Continuous evaluation via user feedback and field studies

Saul Greenberg

The Introspection Method

Designer tries the system (or prototype) out
- does the system "feel right"?
- most common evaluation method

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

Problems
- not reliable as completely subjective
- not valid as introspector is a non-typical user

Intuitions and introspection are often wrong

Saul Greenberg

Conceptual Model Extraction

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

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:
- Excellent for extracting a novice's understanding of system
- Poor for examining system exploration and learning
- Requires active intervention by evaluator, which can get in the way

Saul Greenberg

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

Excellent at identifying gross design/interface problems

Validity/reliability depends on how controlled/contrived the situation is

Three general approaches:
- simple observation
- think-aloud
- constructive interaction

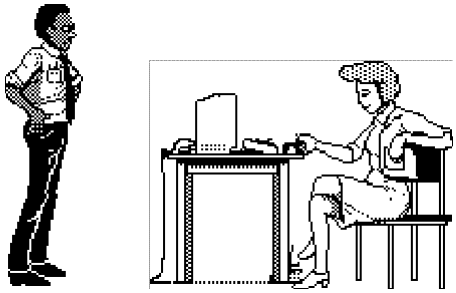
Saul Greenberg

Simple Observation Method

User is given the task, and evaluator just watches the user

Problem

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



Saul Greenberg

The Think Aloud Method

Subjects are asked to say what they are thinking/doing

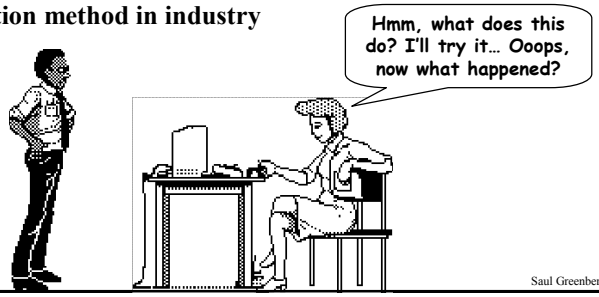
- what they believe is happening
- what they are trying to do
- why they took an action

- Gives insight into what the user is thinking

Problems

- awkward/uncomfortable for subject (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 evaluation method in industry

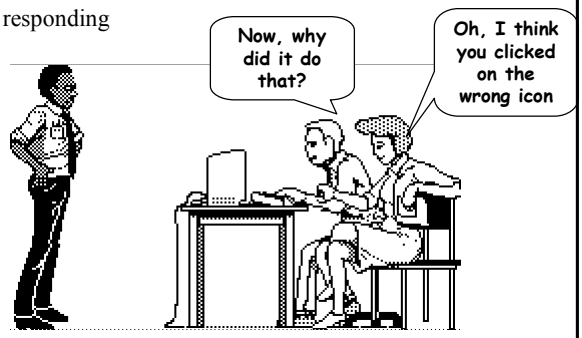


Saul Greenberg

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 naive subject together
  - make naive subject use the interface
- results in
  - naive subject asking questions
  - semi-knowledgeable coach responding
  - provides insights into thinking process of both beginner and intermediate users



Saul Greenberg

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 help…

- audio recording
  - good for recording talk produced by thinking aloud/constructive interaction
  - hard to tie into user actions (ie what they are doing on the screen)

- video recording
  - can see and hear what a user is doing
  - one camera for screen, another for subject (picture in picture)
  - can be intrusive during initial period of use

Saul Greenberg

Coding scheme example...

tracking a person's activity in the office

s = start of activity
e = end of activity

	Desktop activities			Absences		Interruptions	
Time	working on computer	working on desk	initiates telephone	away from desk but in room	away from room	person enters room	answers telephone
9:00	s						
9:02	e					s	
9:05					s	e	
9:10			s		e		
9:13							

Saul Greenberg

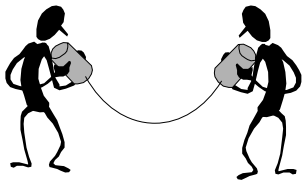
Querying Users via Interviews

Excellent for pursuing specific issues

- vary questions to suit the context
- probe more deeply on interesting issues as they arise
- good for exploratory studies via open-ended questioning
- 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 user
  - user's reconstruction may be wrong



Saul Greenberg

How to Interview

Plan a set of central questions

- could be based on results of user observations
- gets things started
- focuses the interview
- ensures a base of consistency

Try not to ask leading questions

Start with individual discussions to discover different perspectives, and continue with group discussions

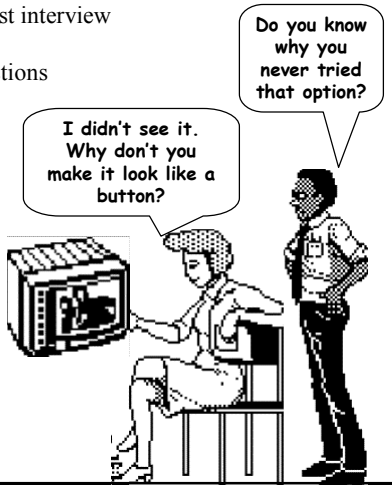
- the larger the group, the more the universality of comments can be ascertained
- also encourages discussion between users

Saul Greenberg

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
  - excellent for grounding a post-test interview
  - avoids erroneous reconstruction
  - users often offer concrete suggestions



Saul Greenberg

Querying users via Questionnaires and Surveys

Questionnaires / Surveys

- preparation "expensive," but administration cheap
  - can reach a wide subject group (e.g. mail)
- does not require presence of evaluator
- results can be quantified
- only as good as the questions asked

Saul Greenberg

Querying Users via Questionnaires / Surveys

How

- establish the purpose of the questionnaire
  - what information is sought?
  - how would you analyze the results?
  - what would you do with your analysis?
- 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
    including a pre-addressed reply envelope gives far better response
- determine the demographics
  - e.g. computer experience

Saul Greenberg

Styles of Questions

Open-ended questions

- asks for unprompted opinions
- good for general subjective information
  - but difficult to analyze rigorously

Can you suggest any improvements to the interfaces?

Saul Greenberg

Styles of Questions

Closed questions

- restricts the respondent's responses by supplying alternative answers
- makes questionnaires a chore for respondent to fill in
- can be easily analyzed
- but watch out for hard to interpret responses!
  - alternative answers should be very specific

Do you use computers at work:

✓ 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
✓ less than 1 hr a day

Saul Greenberg

Styles of Questions

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

Saul Greenberg

Styles of Questions

Multi-choice

- respondent offered a choice of explicit responses

How do you most often get help with the system? (tick one)

O on-line manual
O paper manual
✓ ask a colleague

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

✓ word processor
O data base
✓ spreadsheet
O compiler

Saul Greenberg

Styles of Questions

Ranked

- respondent places an ordering on items in a list
- useful to indicate a user's preferences
- forced choice

Rank the usefulness of these methods of issuing a command
(1 most useful, 2 next most useful..., 0 if not used
__2__ command line
__1__ menu selection
__3__ control key accelerator

Saul Greenberg

Styles of Questions

Combining open-ended and closed questions

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

It is easy to recover from mistakes:

disagree agree comment: *the undo facility is really helpful*
1 2 3 (4) 5

Saul Greenberg

Continuous Evaluation

Usually done in later stages of development
- (ie beta releases, delivered system)

Good for monitoring problems of system in actual use

Problems can be fixed in next release

a) User feedback via gripe lines
- users can provide feedback to designers while using the system
  - email
  - special built-in gripe facility
  - telephone hot line
  - help desks
  - suggestion box
  - bulletin board
- best combined with trouble-shooting facility
  - users always get a response (solution?) to their gripes

Saul Greenberg

Continuous evaluation...

b) Case/field studies
- careful study of "system usage" at the site
- good for seeing "real life" use
- external observer monitors behaviour or gets feedback via methods described above

Saul Greenberg

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

Saul Greenberg