The user as wizard:

A method for early involvement in the design and evaluation of adaptive systems

Judith Masthoff

University of Aberdeen, Scotland

Introduction

- User Interface Design principle:
  Involve users early in the design process, not just in a final evaluation phase
- This also holds for adaptive systems
- However user involvement in the *design* of adaptive systems is rarely reported

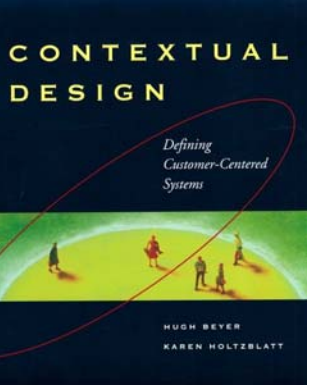
- Will present a method for doing this
  (this method is NOT new)

Early User Involvement in Design

- Many methods are available and may be applicable to adaptive systems:
  - Interviews and Questionnaires
  - Focus Groups
  - Contextual Design
  - Cultural Probes
  - Creative Brainstorming Sessions
  - …..

Early User Involvement in Design

- Many methods are available and may be applicable to adaptive systems:
  - Interviews and Questionnaires
  - Focus Groups
  - **Contextual Design**
  - Cultural Probes
  - Creative Brainstorming Sessions
  - …..

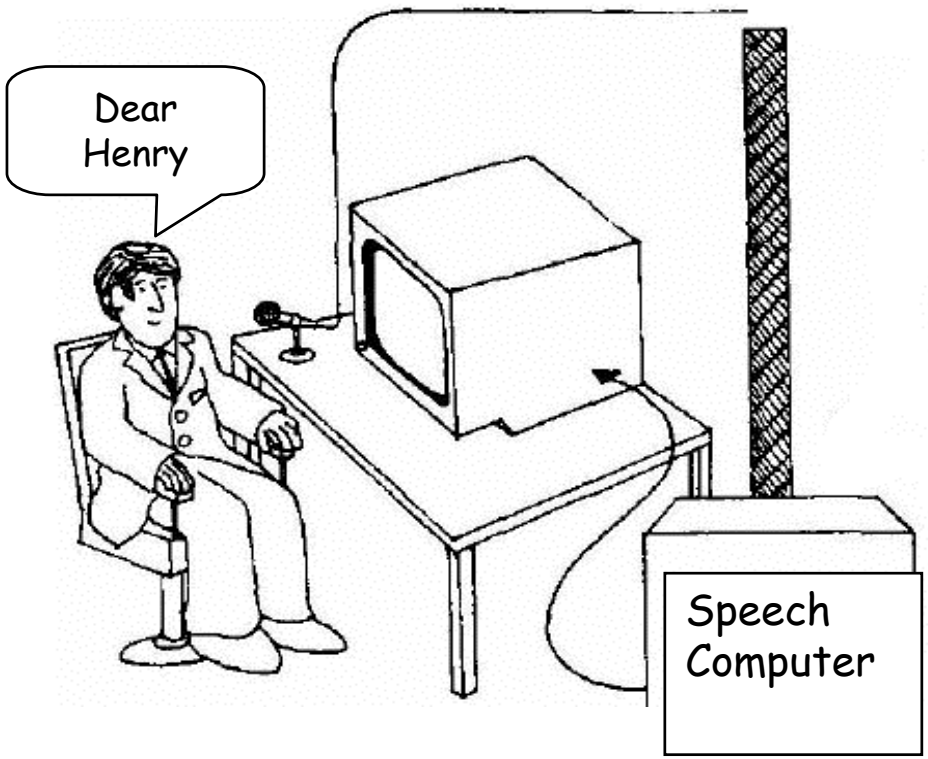


Contextual Design

- Ethnographic method
- Users are observed in their work place
  - how they go about their work
  - in what environment
  - using which artefacts
- Users are the experts in their tasks
- Observation will provide more detail

Early User Involvement in Evaluation

- Wizard-of-Oz studies:
  testing a non-existing system

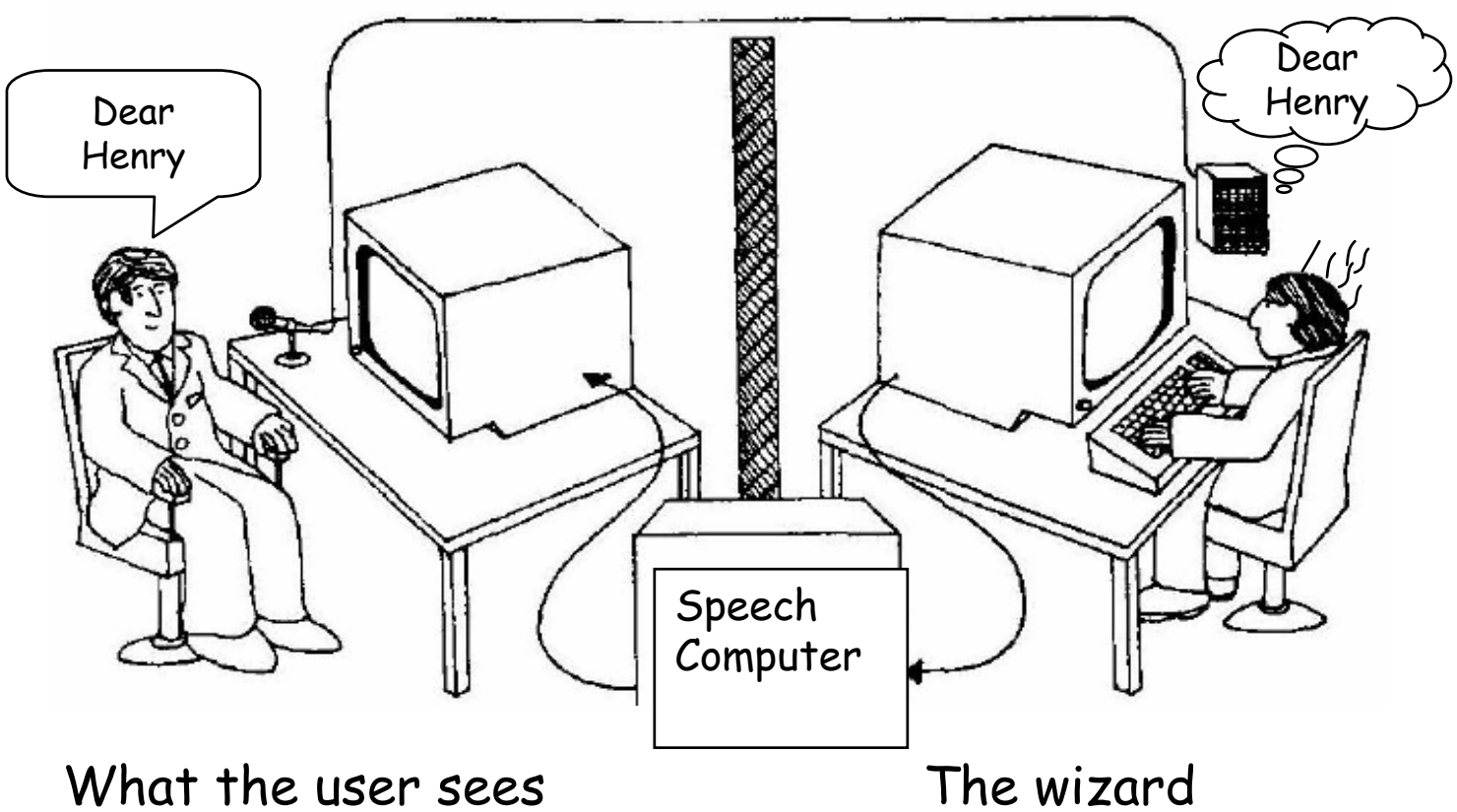


What the user sees

From Gould, Conti & Hovanvecz, Comm ACM 26(4) 1983.

Early User Involvement in Evaluation

- Wizard-of-Oz studies:
  testing a non-existing system



From Gould, Conti & Hovanvecz, Comm ACM 26(4) 1983.

Early User Involvement to Inspire Adaptation Algorithms

- Method inspired by Contextual Design and Wizard-of-Oz
- Humans tend to be good at adaptation, so, observing them in the role of the wizard may help to design the adaptation
- Different from observing experts (e.g. teachers) in the field

Users-as-Wizards

- Participants in the role of the wizard, no script
- Given the same info as the system would have
- Fictional users are used rather than real ones
- Two stages:
  - Exploration Stage:
    Participants take the role of the adaptive system (or a part of it: Layered Design!)
  - Consolidation Stage:
    Participants judge the performance of others

Case Study: Group Recommender System

- *System purpose*. To select a sequence of items (e.g. music clips) adapted to a group of users.
- *Problem*. How should ratings by individual users be aggregated into ratings for the group as a whole? Many different aggregation strategies existed, and we needed to know how suitable they might be, and how to judge them. There were far too many for an ordinary user test, and the domain was too complicated for normal user testing anyway

Exploration Stage (Step1)

- Give participants a scenario describing a fictional user (or group of users) and the user's intentions

"John, Mary, and Adam are going to watch video clips together. We know how interested they are in the topics of the clips. Each clip is rated from 1 - really hate this topic - to 10 - really like this topic. [A table shows each individual's liking for each of the clips]. They only have time to watch five clips."

Exploration Stage (Step 2)

- Give participants the task the adaptive system is supposed to perform

"Which clips should they watch?"

Exploration Stage (Step 3)

- Find out participants' reasons for their decisions and actions

"Why?"

Alternative: thinking-aloud or co-discovery

Exploration Stage (Step 4)

- Steps 1 to 3 can be repeated for a number of scenarios.

Consolidation Stage (Step 1)

- Give participants
  (a) a scenario involving a fictional user and the user's intentions, and
  (b) an associated task.

Normally the same as in Exploration phase.

Consolidation Stage (Step 2)

- Show the participants a performance on this task for this scenario (by participants in Exploration phase or System)

"The TV decides to show them the following sequence of clips [a sequence is given]." (Performance was that of various group aggregation strategies.)

Consolidation Stage (Step 3)

- Ask participants to judge the quality of task performance

"How satisfied do you believe John would be?
How satisfied do you believe Adam would be?
How satisfied do you believe Mary would be?"
(A seven point Likert-scale was used.)

Consolidation Stage (Step 4)

- Find out participants' reasons for their judgments

"Why?" for each judgement.

Consolidation Stage (Step 5-6)

- Normally, steps 2 to 4 are repeated for a number of task performances

  Participants were given 3 performances to judge

- Steps 1 to 5 can be repeated for a number of scenarios.

Conclusions

- Method can help to design adaptation algorithms
- Limitations:
  - Participants' judgments may not correspond with what would be best for users
  - Method not suitable for tasks that are inherently more difficult for humans than for computers
- Method is only initial step:
  User testing is still needed
- Making methods explicit may help in uptake