Vorlesung
Mensch-Maschine-Interaktion

Process, Methods & Tools

Ludwig-Maximilians-Universität München
LFE Medieninformatik
Andreas Butz & Albrecht Schmidt
WS2004/2005
http://www.medien.informatik.uni-muenchen.de/

Table of Content

Ethnographic Observation in HCI

- Traditional ethnographers immerse into other cultures over a extended period (weeks, month, years) and thereby study and understand the culture

- Ethnographic observations in HCI are a means of data collection
- Usually observing potential users (typical users) over a period of hours, days, or weeks. Include critical times (e.g. shift change)

- Goal
  - Acquire information that is required to create user interfaces and interaction mechanisms suitable
- Risk
  - Misinterpretation of observations (often due to a luck of insight)
  - Changing peoples behavior, disrupt processes
  - Overlooking / missing important facts

- Some problems occur infrequently – if you can not observe them conduct interviews

Guidelines for Ethnographic Observation in HCI (Shneiderman, chapter 3)

- Preparation
  - Understand the current system in the context of the organization and culture – don't be ignorant!
  - Describe the goals of the observation and prepare questions
  - Get permissions for observations and interviews
- Field Study
  - Establish contact, talk to people
  - Observe, interview, and collected data in situ
  - Document observations
- Analysis
  - Compile data, summaries and quantify
  - Provide interpretation of the data
  - Refine the goals and record issues about the process
- Reporting
  - Describe findings – possibly for different audiences

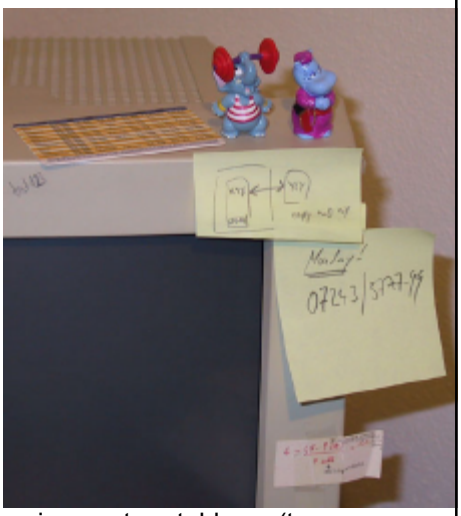
Ethnographic Observation in HCI Video Observation

- Capture work practices on video (consider legal and ethical issues)
- Different view points simultaneously
  - Camera overlooking the workplace
  - Camera looking from the screen to the user
  - Camera capturing what the user sees (e.g. camera mounted to glasses)
- User's view often provides significant insight
- Asking user's to talk (to describe) while doing a task provides generally a lot of useful information
- Raw material alone is of little value – need for analysis
- Analyzing video observations is hard and time consuming!
- Can be very useful
  - Multiple people interact (and observation of an individual and the whole group is of interest)
  - for tasks that are done very quickly or hard to observe
  - where observation is not possible (e.g. for safety or security reasons)
- Users may not like it! If they agree a person observing them they still may disagree to be videoed

Ethnographic Observation in HCI Contextual Enquires

- Learning about the way user's work in the users workplace
- Understanding the work practices and why certain tasks are performed
- Master – apprentice relationship
  - User (master) teaches the observer (apprentice) what they do and how they do it
  - master explains while working
  - Validate your observation by re-phrasing and discuss interpretations made
  - apprentice asks whenever it is not clear
- This method allows to understand how people work and WHY it is done in a certain way
- The observer must be prepared before the interview (understand the language)
- Limit the time of contextual interviews

Ethnographic Observation in HCI Interviews



- Prepare a set of questions beforehand (e.g. what do you want to know from the user)
- Tell people what are you doing
- Use capture (audio/video) if your communication partners agree
- If applicable capture (take photos/video) material they use in their work (e.g. a manual, a checklist, the post-its around the screen)
- Be nosy … ask for details
- If possible summaries what your interview partner told you (to minimize misunderstandings)

Vorlesung Mensch-Maschine-Interaktion

Evaluation

Ludwig-Maximilians-Universität München
LFE Medieninformatik
Andreas Butz & Albrecht Schmidt
WS2004/2005
http://www.medien.informatik.uni-muenchen.de/

What to evaluate?

- The usability of a system!
- … it depends on the stage of a project
  - Ideas and concepts
  - Designs
  - Prototypes
  - Implementations
  - Products in use
- … it also depends on the goals
- Approaches
  - Formative evaluation – throughout the design, helps to shape a product
  - Summative evaluation – quality assurance of the finished product.

Why Studies and Experiments?

- To measure more precisely the usability or other features
- Applicable mainly to
  - Functional prototypes
  - Testing an implementation
  - Quality monitoring of software products
- To compare solutions, e.g.
  - Users are quicker using version A than using version B
  - Users make 10% less errors when using version X than when using version Y
- To provide quantitative figures, e.g.
  - 90% of the users can complete the transaction using version Y in less than 3 minutes
  - On average users will be able to by a ticket using version A in less than 30 seconds

Designing the experiment

- Basic Scientific Method
  1. Form Hypothesis
  2. Collect data
  3. Analyze
  4. Accept/reject hypothesis
- Issues for user studies
  - System to test
  - Participants
  - Hypothesis
  - Variables
  - Experimental Methods
  - Statistical approach

Does computer science fit this traditional science approach?

Is it really possible to prove usability?

Procedure for user studies

- Set goals
- Design the experiment
- Schedule users
- For each user (typical example):
  - Inform the user and sign the consent form
  - Do a survey on demographics and questions of interest to the experiment
  - Give the participant instructions on the task – do not reveal the hypotheses
  - (optional) Make a training run - depends on the study
  - Perform the actual run and measure variables
  - (optional) do a survey on subjective measure
  - Be available for questions of participants or for their (informal) feedback
- Analyze the results

Participants (Subjects)

- How many participants do we need?
  - Depending on the project and the goals
  - Depending on the set-up
    - measuring the login-in time of remote users vs.
    - Doing a full video observation for a 1 hour task
  - Be pragmatic
  - Minimal size of about 10 participants
- Participants should be representative for the user group
  - Age, background, skills, experience, …
  - In most cases the other people on the team are NOT representative!
- How to recruit participants
  - Customer data base
  - Market research services
  - Volunteers (online, newspapers, etc) – this is risky because the people who will respond are often not representative

Selection of Participants

- Services offered that allow to get participants fitting a specific description
- Methods widely used in market research

- Example: Online Panel
  - For online questionnaires
  - Pool of users
  - Customer can specify the users that should take part

- How do companies get their subjects?
  - Incentive (money, prices, …)
  - Big set of questions when registering potential users, show examples from ComCult Online Panel

Variables

- Variables are manipulated and measured
  - Independent variables are manipulated
  - Dependent variables are measured
- The conditions of the experiment are set by independent variables
  - E.g. number of items in a list, text size, font, color
  - The number of different values used is called *level*
  - The number of experimental conditions is the product of the levels
  - E.g. font can be times or arial (2 levels), background can be blue, green, or white (3 levels). This results in 6 experimental conditions (times on blue, times, on green, …, arial on white)
- The dependent variables are the values that can be measured
  - Objective values: e.g. time to complete a task, number of errors, etc.
  - Subjective values: ease of use, preferred option
  - They should only be dependent on changes of the independent variables

Hypotheses

- Prediction of the result of an experiment
- Stating how a change in the independent variables will effect the measured dependent variables
- With the experiment it can be shown that the hypotheses is correct
- Usual approach
  - Stating a null-hypotheses (this predicts that there is not effect of the change in the independent variable on the measured variable)
  - Carrying out the experiment and using statistical measures to disprove the null-hypotheses
  - When a statistical test shows a significant difference it is probable that the effect is not random

Designing the experiment

- The experiment should be set up to be reproducible!
- Main factors
  - Participants
  - Independent variables
  - Hypotheses stated
- Approach
  - state the hypotheses – what do you want to proof
  - find the variables? Which are varied? which are measured?
  - Find participants – representative for the experiment
  - Fix the method to use (between-groups / within groups)

Experimental Method

- Within groups
  - Each user performs under all the different conditions
  - Important to randomize the order of the conditions for each participant
  - Problems
    - Learning may influence results
  - Advantages
    - The effect of differences between individuals are lessened
    - Fewer participants required
- Between groups (randomize)
  - One condition is selected for each participant
  - Each user performs only under one condition (avoids learning)
  - Careful selection of groups is essential
  - Drawback
    - Differences between individuals in different groups can play an important role (leads to large groups)
    - More user required
    - Usually harder to show significance

Statistical Tests

- See statistics text book (e.g. form psychology or medical tests)
- Software packages offer functions
- Test selected depends on
  - Distribution of the measured variables
  - The type of variables (continuous or discrete)
  - Experimental Method
- Example: Student's t-test
  - On the difference of means
  - Assumes a normal distribution
  - Functions available in spreadsheet software and statistics packages
- Example ANOVA
  - Analysis of Variance
- "significant difference"
  - Simplified: the probability that effect observed is random is less the 0.05

T-Test example in Excel

- TTEST(…)
- Parameters
  - Data row 1
  - Data row 2
  - Ends (1 or 2)
  - Type (paired, same variance, different variance)

User	Time M1	Time M2		
100	37	31		
101	44	38		
102	42	43		
103	56	37		
104	99	50		
105	33	30		
106	45	50		
107	49	36		
108	70	71		
109	63	56		
110	54	51		
111	61	46		
average	54,4167	44,9167		
t test (paired)			0,042	TTEST(B7:B18;C7:C18;2;1)
t test (un-paired)			0,137	TTEST(B7:B18;C7:C18;2;2)

Further Issues

- Consent form – get written consent from participants
  - Templates available
  - May be checked with the legal department / review board
- Let participants know what they are doing
  - What is the participant expected to do
  - Procedure
  - How long will it take, breaks
  - What is the study for in general – but do NOT tell about the specific purpose or your hypotheses
- Make sure they know
  - Quality of a UI / software is tested
  - They are NOT tested
- Ethical Issues

Participants Consent (Example)

Participants Consent Form

Study _____ **Institution** _____

Name: _____ Date of Birth: _____
Email: _____
Phone:_____

I have been informed on the procedure and purpose of the study and my questions have been answer to my satisfaction.

I have volunteered to take part in this study and agree that during the study information is recorded (audio and video as well as my interaction with the system). This information may only be used for research and teaching purpose. I understand that my participation in this study is confidential. All personal information and individual results will not be released to third parties without my written consent.

I understand that I can withdraw from participation in the study at any time.

Date: _____ Signature:_____

Example: Study on Text Input

- Is text input by keyboard really better than using T9 on a phone?
- Compare text input speed and errors made
  - Qwertz-keyboard on a notebook computer
  - T9 on a mobile phone
- Concentrate on test input only, ignore:
  - Time to setup / boot / initialize the device
  - Time to get into the application

Example: Study on Text Input (2)

- Participants
  - How many?
  - Skills
    - Computer user?
    - Phone/T9 users?
- Independent variables
  - Input method
  - Text to input
- Dependent variables
  - Time to input a text
  - Number of errors made

Example: Study on Text Input (3)

- Independent variables
  - Input method,
    - 2 levels: Keyboard and T9
  - Text to input
    - 1 level: text with about 10 words

- Experimental conditions
  - 2 conditions – T9 and Key
  - User 1,3,5,7,9 perform T9 than Key
  - User 2,4,6,8,10 perform Key than T9
  - Different texts in first and second run?
  - Particular phone model?
  - Completion time is measure (e.g. stop watch or application)
  - Number of error/corrections is observed

Example: Study on Text Input (4)

- Hypotheses
  - H-1: Input by keyboard is quicker than T9
  - H-2: fewer errors are made using keyboard input compared to T9

- Null-Hypotheses
  - Assumes no effect
  - H0-1: there is no difference in the input speed between keyboard and T9
  - H0-2: there is no difference in the number of errors made using a keyboard input compared to T9

- Experimental Method
  - Within groups
  - Randomized order of conditions

Example: Study on Text Input (5)

- Collect Data

User	Order	Time Cond1	Time Cond2	# Err Cond1	# Err Cond2
01	c1>c2	…	…	…	…
02	c2>c1	…	…	…	…
03	c1>c2	…	…	…	…

- Perform a statistical analysis
- … exercise on Friday.

Example: Study on Text Input (6)

- Fairness
  - Same conditions and procedure (e.g. light condition, interruptions, noise)
  - Specify procedure for exceptions (e.g. someone does not complete the test)
  - No bias
- Participants Consent
- Further Issues?
  - Ethical issues
  - Privacy