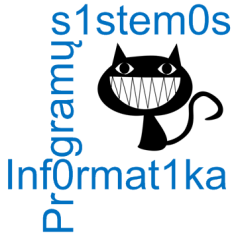


Data Gathering

Lecture 11

Slides adapted by
Dr. Kristina Lapin

Vilnius University

Overview

- Five key issues of data gathering
- Data recording
- Interviews
- Questionnaires
- Observation
- Choosing and combining techniques

Five key issues

1. Setting goals
   - Decide how to analyze data once collected
2. Identifying participants
   - Decide who to gather data from – the population
   - Probability sampling and non-probability sampling
   - Saturation sampling – access to all members
3. Relationship with participants
   - Clear and professional
   - Informed consent when appropriate
4. Triangulation
   - Look at data from more than one perspective
5. Pilot studies
   - Small trial of main study

Crowdsourcing Design for Citizen Science Organizations

SHORT VERSION OF CONSENT FORM for participants at the University of Maryland – 18 YEARS AND OLDER

You are invited to participate in a research project being conducted by the researchers listed on the bottom of the page. In order for us to be allowed to use any data you wish to provide, we must have your consent.

In simplest terms, we hope you will use the mobile phone, tabletop, and project website at the University of Maryland to

- Take pictures
- Share observations about the sights you see on campus
- Share ideas that you have to improve the design of the phone or tabletop application or website
- Comment on pictures, observations, and design ideas of others

The researchers and others using CampusNet will be able to look at your comments and pictures on the tabletop and/or website, and we may ask if you are willing to answer a few more questions (either on paper, by phone, or face-to-face) about your whole experience. You may stop participating at any time.

A long version of this consent form is available for your review and signature, or you may opt to sign this shorter one by *checking off all the boxes that reflect your wishes and signing and dating the form below.*

___I agree that any photos I take using the CampusNet application may be uploaded to the tabletop at the University of Maryland and/or a website now under development.

___I agree to allow any comments, observations, and profile information that I choose to share with others via the online application to be visible to others who use the application at the same time or after me.

___I agree to be videotaped/audiotaped during my participation in this study.

___I agree to complete a short questionnaire during or after my participation in this study.

NAME [Please print]	
SIGNATURE	
DATE	

[Contact information of Senior Researcher responsible for the project]

Data recording

- Notes, audio, video, photographs
  - Choice depends on the context, time available and the sensitivity of the situation.
- Notes plus photographs
  - Handwritten notes – flexible and less intrusive than typing,
  - Can be tiring to write, observe and listen at the same time
- Audio plus photographs
  - Audio – less intrusive than video
  - Attention to the interviewee rather than to taking notes
- Video
  - Requires additional planning

What to choose?

- Imagine you are employed to develop a new computerized garden system planning tool to be used by garden designers.
- Goal: to find out
  - how garden designer use an early prototype as they walk around their clients' gardens
  - sketching design ideas, taking notes and asking the clients about
  - what they like and how they use the gardens
- What are the advantages and disadvantages of the 3 approaches to data recording in this environment?

Criterion	Notes plus camera	Audio plus camera	Video
Equipment	Paper, pencil, and camera are easily available	Inexpensive, hand-held recorder with a good microphone	More expensive. Editing, mixing, and analysis equipmemnt needed
Flexibility of use	Very flexible. Unobtrusive.	Flexible. Relatively unobtrusive.	Needs positioning and focusing camera lens. Obtrusive.
Completeness of data	To get what note taker thinks is important and can record in the time available. Problem with unexperienced evaluators.	Complete audio recording but visual data is missing. Notes, photographs, sketches augment recording but need coordinating	Most complete data, especially if more than one camera is used, but coordination of video material is needed.
Disturbance of users	Very low	Low, but microphone needs to be positioned.	Medium. Camera needs to be positioned. Care needed to avoid Hawthorne effect.
Reliability of data	May be low. Relies on making a good record and knowing what to record.	High but ecternal moise, e.g. Fans in computers, can mffle what is sad	Can be high but depends on what camera is focused on
	Rich descriptions can be produced	Critical discussions can be identified	Critical incidents can be identified and tagged

Interviews

- Conversation with a purpose
- Unstructured
  - are not directed by a script. Rich but not replicable.
- Structured
  - tightly scripted, often like a questionnaire.
  - Replicable but may lack richness.
- Semi-structured
  - guided by a script
    - but interesting issues can be explored in more depth.
    - Can provide a good balance between richness and replicability.
- Group interviews

Interview questions

- Two types:
  - 'closed questions' have a predetermined answer format, e.g., 'yes' or 'no'
  - 'open questions' do not have a predetermined format
- Closed questions are easier to analyze
- Avoid:
  - Long questions
  - Compound sentences - split them into two
  - Jargon and language that the interviewee may not understand
  - Leading questions that make assumptions e.g., why do you like …?
  - Unconscious biases e.g., gender stereotypes

Interview with children

- Childres think and react to situations different from adults
- Sitting 4-year-old down in a formal interview
  - is unlikely to result in anything other than a wall of silence
- Recording pose a problem:
  - Children have a tendency to perform in front of the camera
- Child-friendly methods
  - Images, e.g. smileys and chat.

Smilometer: Read ir kiti, 2002


HCIL

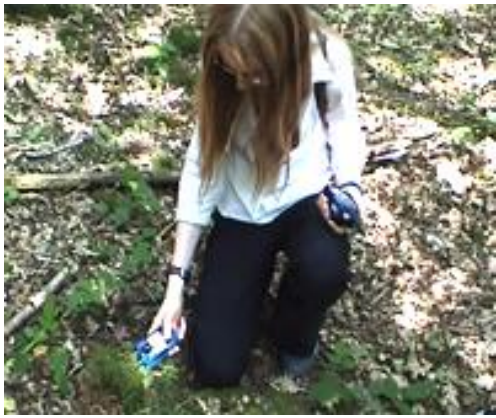
Example: a story-based interactive digital platform

- Duveskog et al. (2009) designed a platform to educate children about HIV and AIDS in Tanzania.
- Project group included:
  - secondary school pupils, university counseling students, HIV counseling experts and experts in ICT.
- Pupils were interviewed, students produced drawings to illustrate their stories, then tested the platform

Duveskog, M., Bednarik, R., Kemppainen, K., Sutinen. E. (2009) Designing a Story-Based Platform for HIV and AIDS Counseling with Tanzanian Children, IDC 2009, June 3–5, 2009, Como, Italy.

Example: Ambient Wood

- Investigates the use of ubiquitous computing and mobile technologies to support learning.
  - Children explored the inhabitants in a woodland area.
  - Each child was given PDA and mobile probing tool
  - The data collected dispayed on the PDA
  - Children's position was monitored and location-specific data was sent to their PDA, e.g. When they walked past a specific plant



Focus groups

- Frequently used in marketing and political campaigning.
- 3-10 people involved and discussion is led by trained facilitator
  - Participants are selected to provide a representative sample of the target population
- Benefit: diverse or sensitive issues can be raised investigating community issues
- Drawback: what they say is not always what they do

Focus group: problems

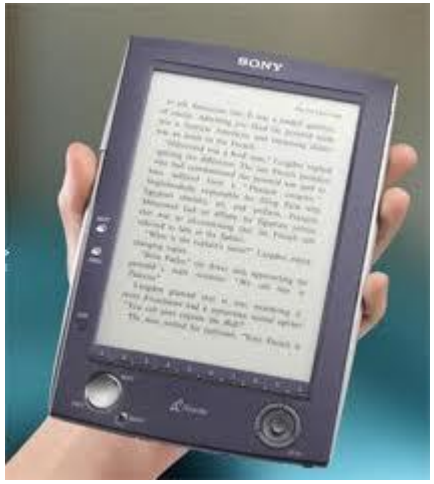
- Whar they say is not always what they do
  - People sometimes give the answers that they think show them in the best light
  - They may forget how long they spend on a particular activity
- So, can we believe all the responses we get?

Planning and conducting an interview

- Developing interview questions
  - Open – for exploratory sessions
  - Closed – when the possible answers are known in advance
  - structured interview usually consists of open questions,
  - while a structured – closed
  - Semistructured – a combination of both

Example

- How appealing are reading devices to people over 65?



Sony e-reader

Amazon Kindle

Apple iPad

Data gathering session

- The goal: to seek opinions whether e-readers would appealing to people over 65
- Suggest ways od recording the interview data.
- Suggest set of questions.

Questions

1. Have you used an e-reader before? (Explore previous knowledge)
   Interviewer checks box ☐ *Yes* ☐ *No* ☐ *Don't remember/know*
2. Would you like to read a book using an e-reader? (Explore initial reaction, then explore the response)
   Interviewer checks box ☐ *Yes* ☐ *No* ☐ *Don't know*
3. Why?
   If response is 'Yes' or 'No,' interviewer says, 'Which of the following statements represents your feelings best?'
   For 'Yes,' interviewer checks the box
   ☐ *I don't like carrying heavy books*
   ☐ *This is fun/cool*
   ☐ *It's going to be the way of the future*
   ☐ *Another reason (interviewer notes the reason)*
   For 'No,' interviewer checks the box
   ☐ *I don't like using gadgets if I can avoid it*
   ☐ *I can't read the screen clearly*
   ☐ *I prefer the feel of paper*
   ☐ *Another reason (interviewer notes the reason)*
4. In your opinion, is an e-reader easy to handle or cumbersome?
   Interviewer checks box
   ☐ *Easy to handle*
   ☐ *Cumbersome*
   ☐ *Neither*

Data gathering session

- Based on results of unstructures interview developers have found that two important acceptance factors are:
  - whether the device can be handled easily;
  - whether the typeface and appearance can be altered.

Running the interviu

An introduction

- Interviewer introduces hweself, explains why the interview is being done, reassures interviewees regading any ethical issues, asks if they mind being recorded

A warm-up session

- Non-threatening questions, i.e. Demographic information

A main session

- Questions presented in logical sequence

A cool-off period

- Qa few easy questions to defuse tension if it has arisen

A closing session

- The interviewer than ks for the interviewee;
- Switches off the recorder or puts the notes away signaling that the interview has ended

Other forms of interview

- Telephone interviews
  - Much in common with face-to-face but it is not possible to see the interviewee's body language or facial expressons
- Online interviews
  - Emails, chats, video conferencing
- Retrospective interviews
  - Reflects on an activity or a data gathering session in the recent past
  - May be conducted to check that the interviewer has correctly understood what was happenning

Enriching the interview process

- Neutral meeting room
- Props - devices for prompting interviewee,
  - e.g., a prototype, scenario

Questionnaires

- Questions can be closed or open
- Closed questions are easier to analyze, and may be done by computer
- Can be administered to large populations
- Paper, email and the web used for dissemination
- Sampling can be a problem
  - when the size of a population is unknown as is common online

Questionnaire structure

1. Many start by asking for basic demographic information
   - Gender, age, place

PART 1

Name: _____ Contact #: _____

Email: _____ Mailstop: _____

Your current position: _____
Number of years in this position: ___

If you are a manager of people, how many people are in your group: ___

Questionnaire structure

- 2. Relevant experience

How much experience have you had with the following types of computers and computer devices?

Mac	___ Years	___ Months	
PC or Compatible	___ Years	___ Months	
Laptop	___ Years	___ Months	
Mainframe	___ Years	___ Months	Type: _____
Mouse	___ Years	___ Months	Manufacturer: _____
Trackball	___ Years	___ Months	Manufacturer: _____

About how many hours a week do you use a computer?

At home: _____ At work: _____

What type of computer do you use?

At home: _____ At work: _____

Do you use Microsoft Windows? Yes ___ No ___

What Windows applications have you used?:

Product name:		
1.	___ Years	___ Months
2.	___ Years	___ Months
3.	___ Years	___ Months

What do you typically use your computer for?

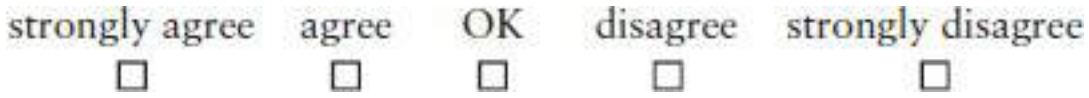
___	Games and Pleasure	___	Graphics
___	Accounting/Finance	___	Data storage (i.e., data bases)
___	Word Processing	___	Other _____
___	Decision Support	___	Other _____

Questionnaire design

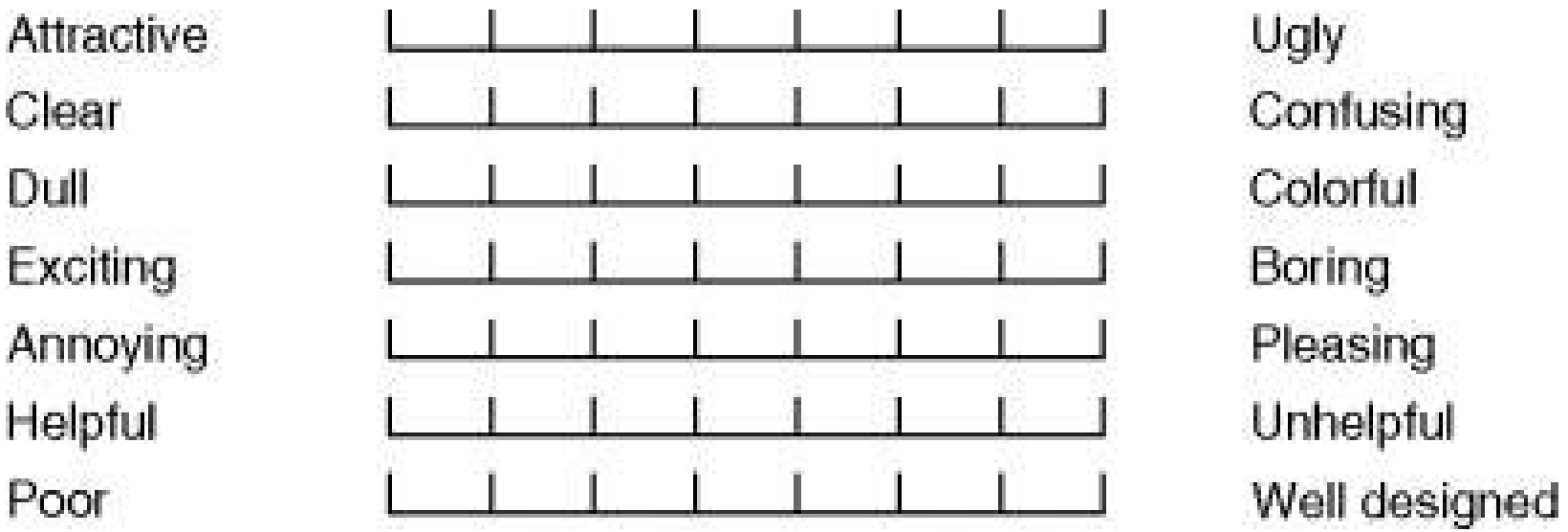
- The impact of a question can be influenced by question order.
- Do you need different versions of the questionnaire for different populations?
- Provide clear instructions on how to complete the questionnaire.
- Strike a balance between using white space and keeping the questionnaire compact.
- Decide on whether phrases will all be positive, all negative or mixed.

Question and response format

- 'Yes' and 'No' checkboxes
- Checkboxes that offer many options
- Rating scales
  - Likert scales
  - semantic scales
  - 3, 5, 7 or more points?
- Open-ended responses



Semantic differential scales



Encouraging a good response

- Make sure purpose of study is clear
- Promise anonymity
- Ensure questionnaire is well designed
- Offer a short version for those who do not have time to complete a long questionnaire
- If mailed, include a stamped addressed envelope
- Follow-up with emails, phone calls, letters
- Provide an incentive
- 40% response rate is high, 20% is often acceptable

Find poorly designed features

2. State your age in years ☐

3. How long have you used the Internet? (*check one only*)
   - ☐ <1 year
   - ☐ 1–3 years
   - ☐ 3–5 years
   - ☐ >5 years

4. Do you use the Web to:
   - purchase goods ☐
   - send e-mail ☐
   - visit chatrooms ☐
   - use bulletin boards ☐
   - find information ☐
   - read the news ☐

5. How useful is the Internet to you?

_____

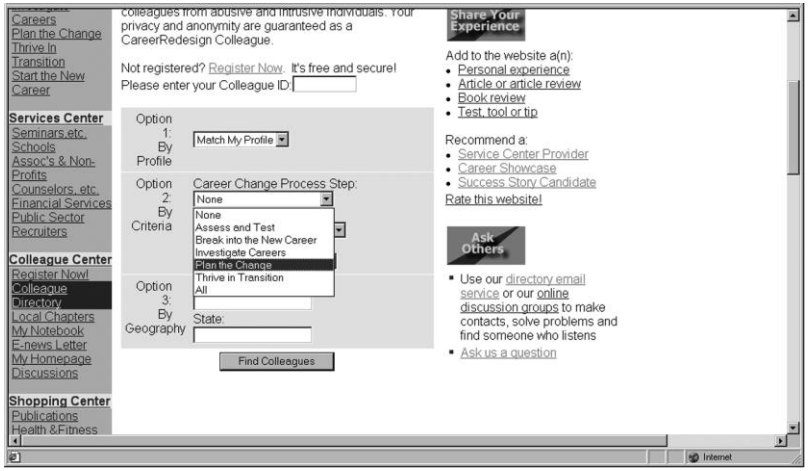
_____

Administering questionnaires

- Issues
  - Reaching a respresentative sample
  - Ensuring a resonable response rate
- For large surveys
  - Respondents are selected using sampling techniques
- Interaction designers commonly use small samples, less that 20 users
- 40% response rate is good
  - Much lower rates are common

Advantages of online questionnaires

- Responses are usually received quickly
- No copying and postage costs
- Data can be collected in database for analysis
- Time required for data analysis is reduced
- Errors can be corrected easily



Problems with online questionnaires

- Sampling is problematic if population size is unknown
- Preventing individuals from responding more than once
- Individuals have also been known to change questions in email questionnaires

Observation

- Direct observation in the field
  - Structuring frameworks
  - Degree of participation (insider or outsider)
  - Ethnography

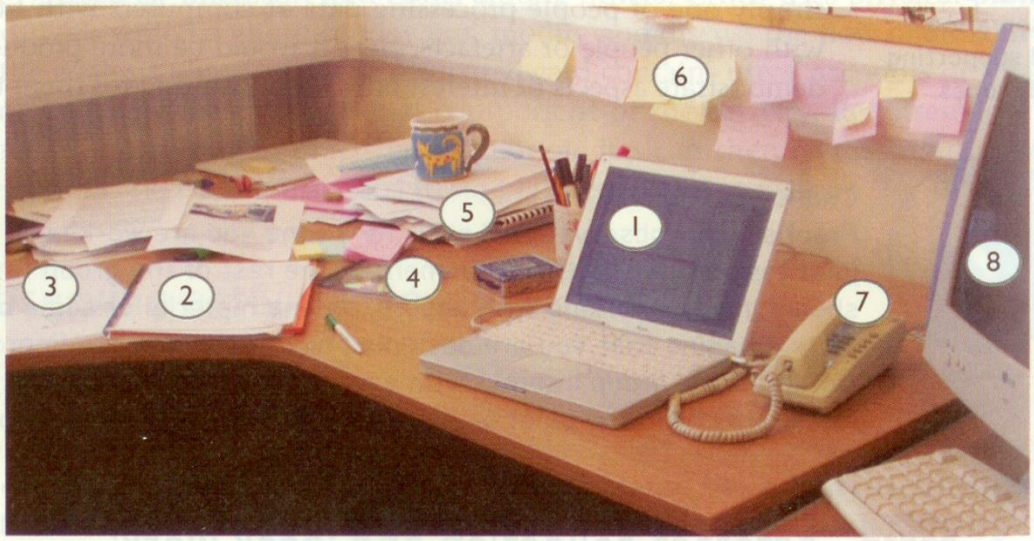


FIGURE 9-3 Artefacts on and around an office desk

Observation

- Indirect observation: tracking users' activities
  - Diaries
  - Interaction logging
  - Distant observation

Structuring frameworks to guide observation

- - *The person*. Who?
  - *The place.* Where?
  - *The thing.* What?

- The Goetz and LeCompte (1984) framework:
  - *Who* is present?
  - What is their role?
  - *What* is happening?
  - *When* does the activity occur?
  - *Where* is it happening?
  - *Why* is it happening?
  - *How* is the activity organized?

Ethnography (1)

- Ethnography is a philosophy with a set of techniques
  - that include participant observation and interviews
- Debate about differences between participant observation and ethnography
  - Ethnographers immerse themselves in the culture that they study
- A researcher's degree of participation can vary along a scale from 'outside' to 'inside'
- Analyzing video and data logs can be time-consuming
- Collections of comments, incidents, and artifacts are made

Ethnography (2)

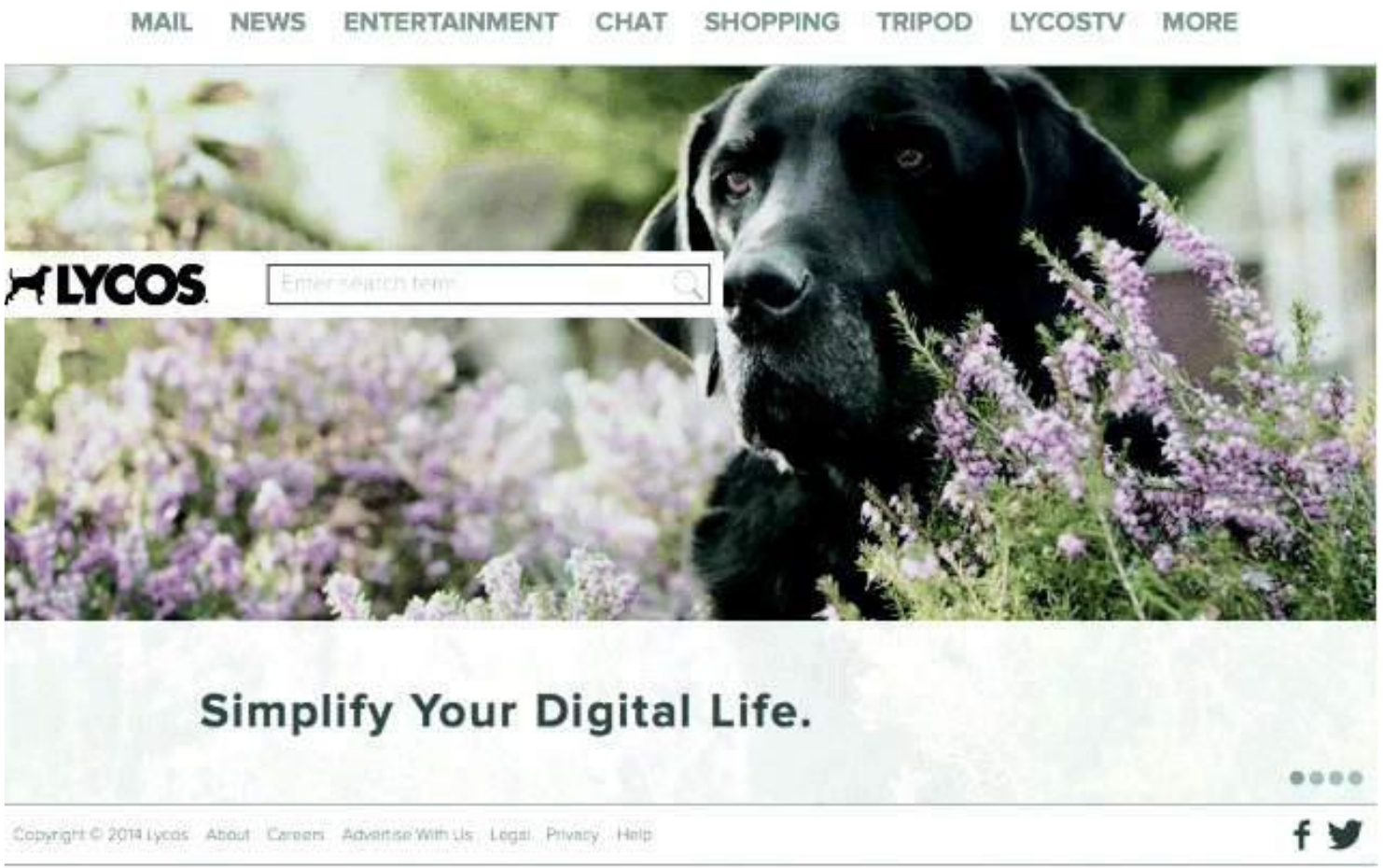
- Co-operation of people being observed is required
- Informants are useful
- Data analysis is continuous
- Interpretivist technique

- Questions get refined as understanding grows
- Reports usually contain examples

Online Ethnography

- Virtual, Online, Netnography
- Online and offline activity
- Interaction online differs from face-to-face
- Virtual worlds have a persistence that physical worlds do not have
- Ethical considerations and presentation issues are different

Direct observation in a controlled environment



Indirect observation

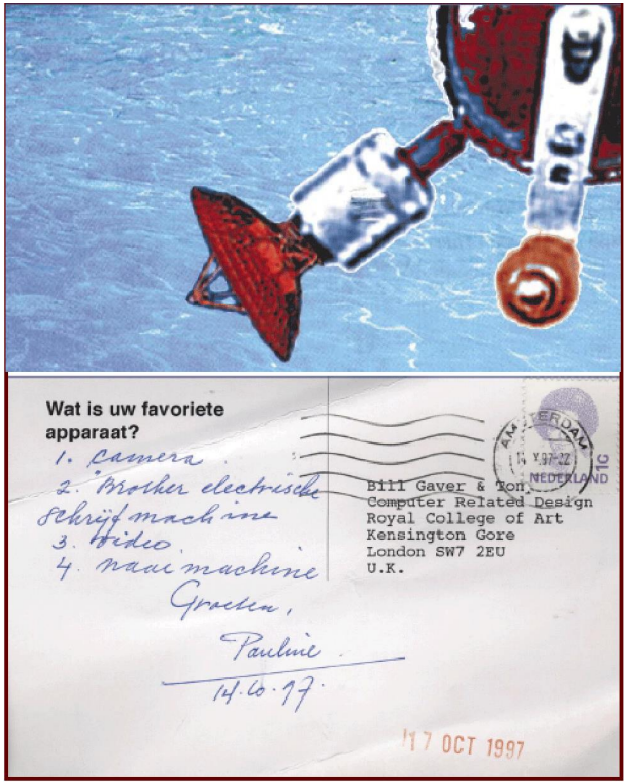
- Diaries
- Interaction logs
- Web analytics

Indirect observation

- Web analytics

(a)

Cultural probe



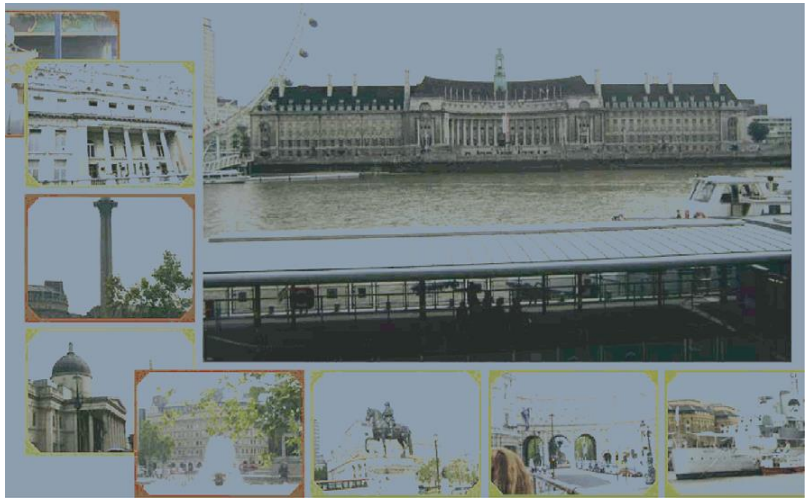
Postcard used as a ‘cultural probe’.

Source: Gaver, W.W., Dunne, T. and Pacenti, E. (1999) Cultural Probes, Interactions*, 6(1), pp. 21–29. © 1999 ACM, Inc.*

From Benyon, Designing interactive systems, Pearson Education, 2005.

Choosing and combining techniques

- Depends on
  - The focus of the study
  - The participants involved
  - The nature of the technique
  - The resources available



Summary

- Three main data gathering methods:
  - interviews, questionnaires, observation
- Five key issues of data gathering:
  - goals, choosing participants, triangulation, participant relationship, pilot
- Interviews may be
  - structured, semi-structured or unstructured
- Questionnaires may be
  - on paper, online or telephone
- Observation may be
  - direct or indirect, in the field or in controlled setting
- Techniques can be
  - combined depending on study focus, participants, nature of technique and available resources