



# DATA ANALYSIS, INTERPRETATION AND PRESENTATION

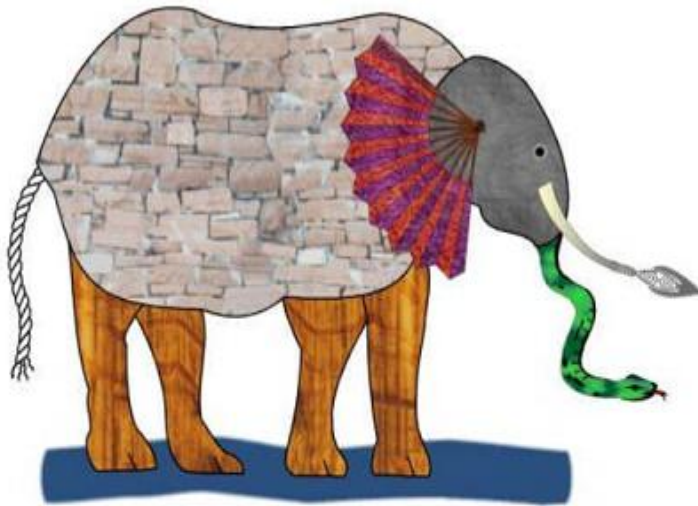
# OVERVIEW

- Qualitative and quantitative
- Simple quantitative analysis
- Simple qualitative analysis
- Tools to support data analysis
- Theoretical frameworks: grounded theory, distributed cognition, activity theory
- Presenting the findings: rigorous notations, stories, summaries

# WHY DO WE ANALYZE DATA

The purpose of analysing data is to obtain usable and useful information. The analysis, irrespective of whether the data is qualitative or quantitative, may:

- describe and summarise the data
- identify relationships between variables
- compare variables
- identify the difference between variables
- forecast outcomes



Blind men and an elephant  
- Indian fable

Things aren't always what we think!

Six blind men go to observe an elephant. One feels the side and thinks the elephant is like a wall. One feels the tusk and thinks the elephant is like a spear. One touches the squirming trunk and thinks the elephant is like a snake. One feels the knee and thinks the elephant is like a tree. One touches the ear, and thinks the elephant is like a fan. One grasps the tail and thinks it is like a rope. They argue long and loud and though each was partly in the right, all were in the wrong.

# SCALES OF MEASUREMENT

Many people are confused about what type of analysis to use on a set of data and the relevant forms of pictorial presentation or data display. The decision is based on the scale of measurement of the data. These scales are nominal, ordinal and numerical.

## Nominal scale

A nominal scale is where:

the data can be classified into a non-numerical or named categories, and the order in which these categories can be written or asked is arbitrary.

## Ordinal scale

An ordinal scale is where:

the data can be classified into non-numerical or named categories  
an inherent order exists among the response categories.

Ordinal scales are seen in questions that call for ratings of quality (for example, very good, good, fair, poor, very poor) and agreement (for example, strongly agree, agree, disagree, strongly disagree).

## Numerical scale

A numerical scale is:

where numbers represent the possible response categories

there is a natural ranking of the categories

zero on the scale has meaning

there is a quantifiable difference within categories and between consecutive categories.

## Common myths

- **Complex analysis and big words impress people.**
  - Most people appreciate practical and understandable analyses.
- **Analysis comes at the end after all the data are collected.**
  - We think about analysis upfront so that we HAVE the data we WANT to analyze.
- **Quantitative analysis is the most accurate type of data analysis.**
  - Some think numbers are more accurate than words but it is the quality of the analysis process that matters.

When using a quantitative methodology, you are normally testing theory through the testing of a hypothesis.

In qualitative research, you are either exploring the application of a theory or model in a different context or are hoping for a theory or a model to emerge from the data. In other words, although you may have some ideas about your topic, you are also looking for ideas, concepts and attitudes often from experts or practitioners in the field.

## Common myths cont...

- Data have their own meaning.
  - Data must be interpreted. Numbers do not speak for themselves.
- Stating **limitations** to the analysis weakens the evaluation.
  - All analyses have weaknesses; it is more honest and responsible to acknowledge them.
- Computer analysis is always easier and better.
  - It depends upon the size of the data set and personal competencies. For small sets of information, hand tabulation may be more efficient.



## 1. Organizing the data

- Organize all forms/questionnaires in one place
- Check for completeness and accuracy
- Remove those that are incomplete or do not make sense; keep a record of your decisions
- Assign a unique identifier to each form/questionnaire





## Enter your data

- By hand
  - By computer
    - Excel (spreadsheet)
    - Microsoft Access (database mngt)
    - Quantitative analysis: SPSS (statistical software)
- Count (frequency)
  - Percentage
  - Mean
  - Mode
  - Median
  - Range
  - Standard deviation
  - Variance
  - Ranking
  - Cross tabulation

Which calculation do I use? It depends upon what you want to know.

Do you want to know how many individuals checked each answer?	Frequency
Do you want the proportion of people who answered in a certain way?	Percentage
Do you want the average number or average score?	Mean
Do you want the middle value in a range of values or scores?	Median
Do you want to show the range in answers or scores?	Range
Do you want to compare one group to another?	Cross tab
Do you want to report changes from pre to post?	Change score
Do you want to show the degree to which a response varies from the mean?	Standard deviation

### 3. Interpreting the information

Numbers do not speak for themselves.

*For example, what does it mean that 55 youth reported a change in behavior. Or, 25% of participants rated the program a 5 and 75% rated it a 4. What do these numbers mean?*

Interpretation is the process of attaching meaning to the data.

Interpretation demands fair and careful judgments. Often the same data can be interpreted in different ways. So, it is helpful to involve others or take time to hear how different people interpret the same information.

Think of ways you might do this...for example, hold a meeting with key stakeholders to discuss the data; ask individual participants what they think

## Part of interpreting information is identifying the lessons learned

What did you learn? – about the program, about the participants, about the evaluation.

- Are there any 'ah-has'? What is new? What was expected?
- Were there findings that surprised you?
- Are there things you don't understand very well – where further study is needed?

We often include recommendations or an *action plan*. This helps ensure that the results are used.

## 4. Discuss limitations

Written reports:

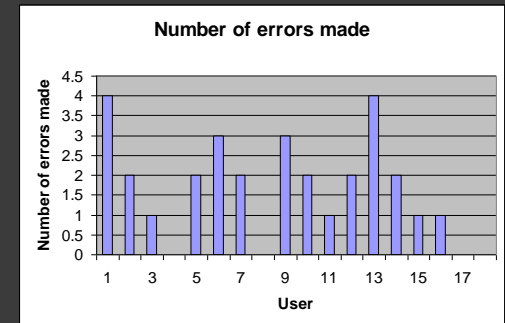
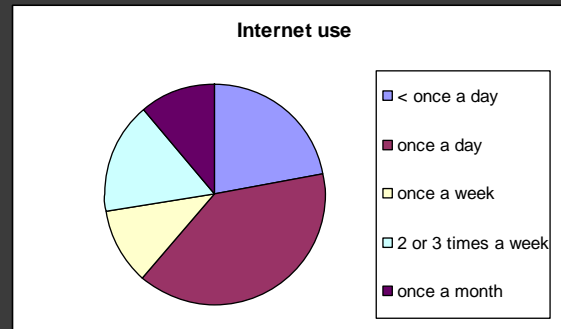
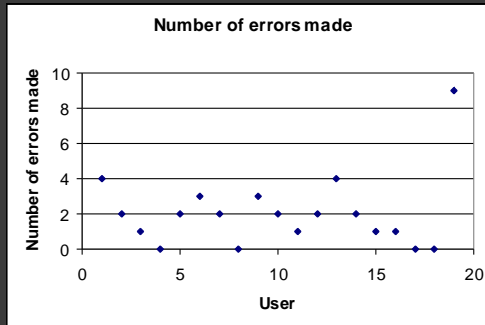
- Be explicit about your limitations

Oral reports:

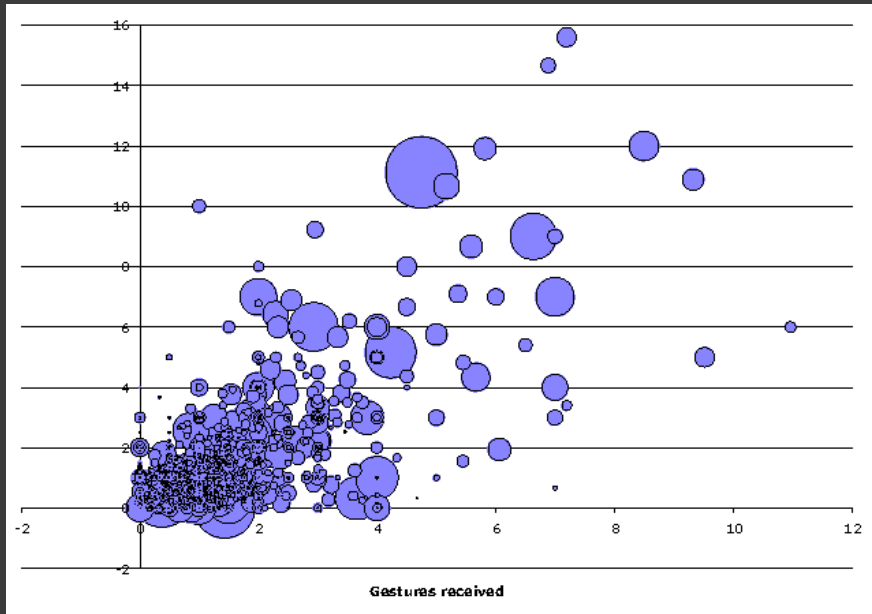
- Be prepared to discuss limitations
- Be honest about limitations
- Know the claims you cannot make
  - Do not claim causation without a true experimental design
  - Do not generalize to the population without random sample and quality administration (e.g., <60% response rate on a survey)

# GRAPHICAL REPRESENTATIONS

give overview of data

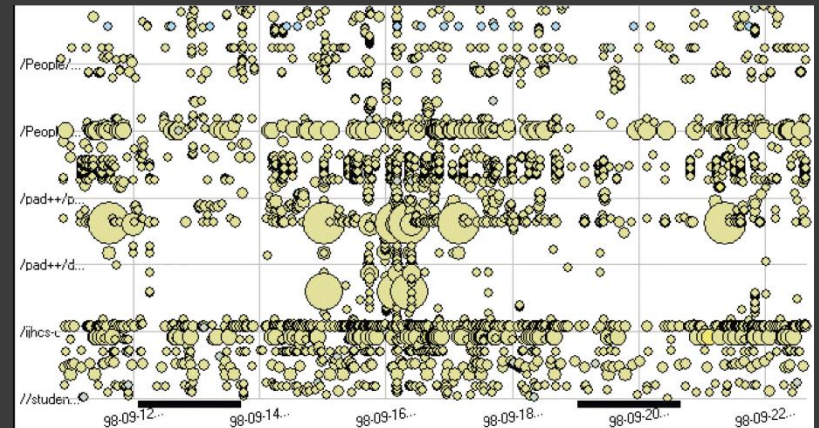


# Visualizing log data



Interaction profiles of players in online game

Log of web page activity





# QUALITATIVE ANALYSIS

"Data analysis is the process of bringing order, structure and meaning to the mass of collected data. It is a messy, ambiguous, time-consuming, creative, and fascinating process. It does not proceed in a linear fashion; it is not neat. Qualitative data analysis is a search for general statements about relationships among categories of data."

Marshall and Rossman, 1990:111

Hitchcock and Hughes take this one step further:

"...the ways in which the researcher moves from a description of what is the case to an explanation of why what is the case is the case."

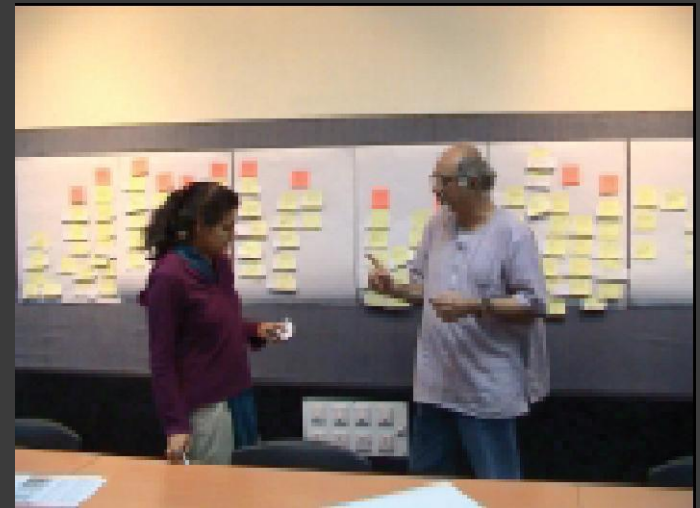
Hitchcock and Hughes 1995:295

# Simple qualitative analysis

- Unstructured - are not directed by a script. Rich but not replicable.
- Structured - are 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.

# Simple qualitative analysis

- Recurring patterns or themes
  - Emergent from data, dependent on observation framework if used
- Categorizing data
  - Categorization scheme may be emergent or pre-specified
- Looking for critical incidents
  - Helps to focus in on key events



# TOOLS TO SUPPORT DATA ANALYSIS

- Spreadsheet – simple to use, basic graphs
- Statistical packages, e.g. SPSS
- Qualitative data analysis tools
  - Categorization and theme-based analysis, e.g. N6
  - Quantitative analysis of text-based data
- CAQDAS Networking Project, based at the University of Surrey (<http://caqdas.soc.surrey.ac.uk/>)

# Theoretical frameworks for qualitative analysis

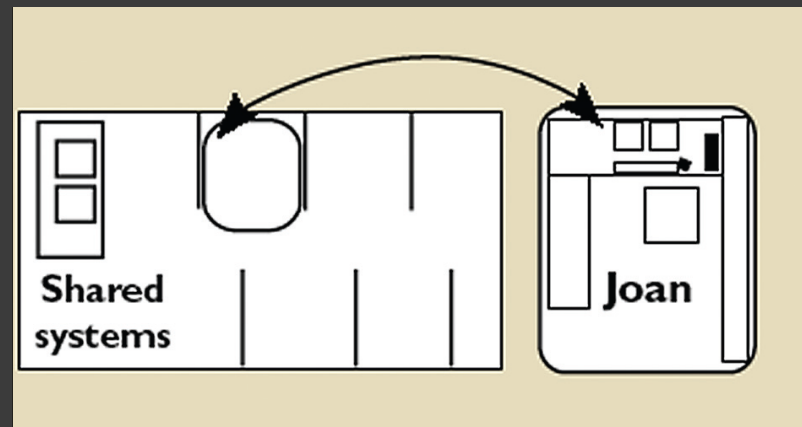
- Basing data analysis around theoretical frameworks provides further insight
- Three such frameworks are:
  - Grounded Theory
  - Distributed Cognition
  - Activity Theory

# Grounded Theory

- Aims to derive theory from systematic analysis of data
- Based on categorization approach (called here 'coding')
- Three levels of 'coding'
  - Open: identify categories
  - Axial: flesh out and link to subcategories
  - Selective: form theoretical scheme
- Researchers are encouraged to draw on own theoretical backgrounds to inform analysis

# Distributed Cognition

- The people, environment & artefacts are regarded as one cognitive system
- Used for analyzing collaborative work
- Focuses on information propagation & transformation



# Activity Theory

- Explains human behavior in terms of our practical activity with the world
- Provides a framework that focuses analysis around the concept of an 'activity' and helps to identify tensions between the different elements of the system
- Two key models: one outlines what constitutes an 'activity'; one models the mediating role of artifacts



# Individual model

Activity – Motive

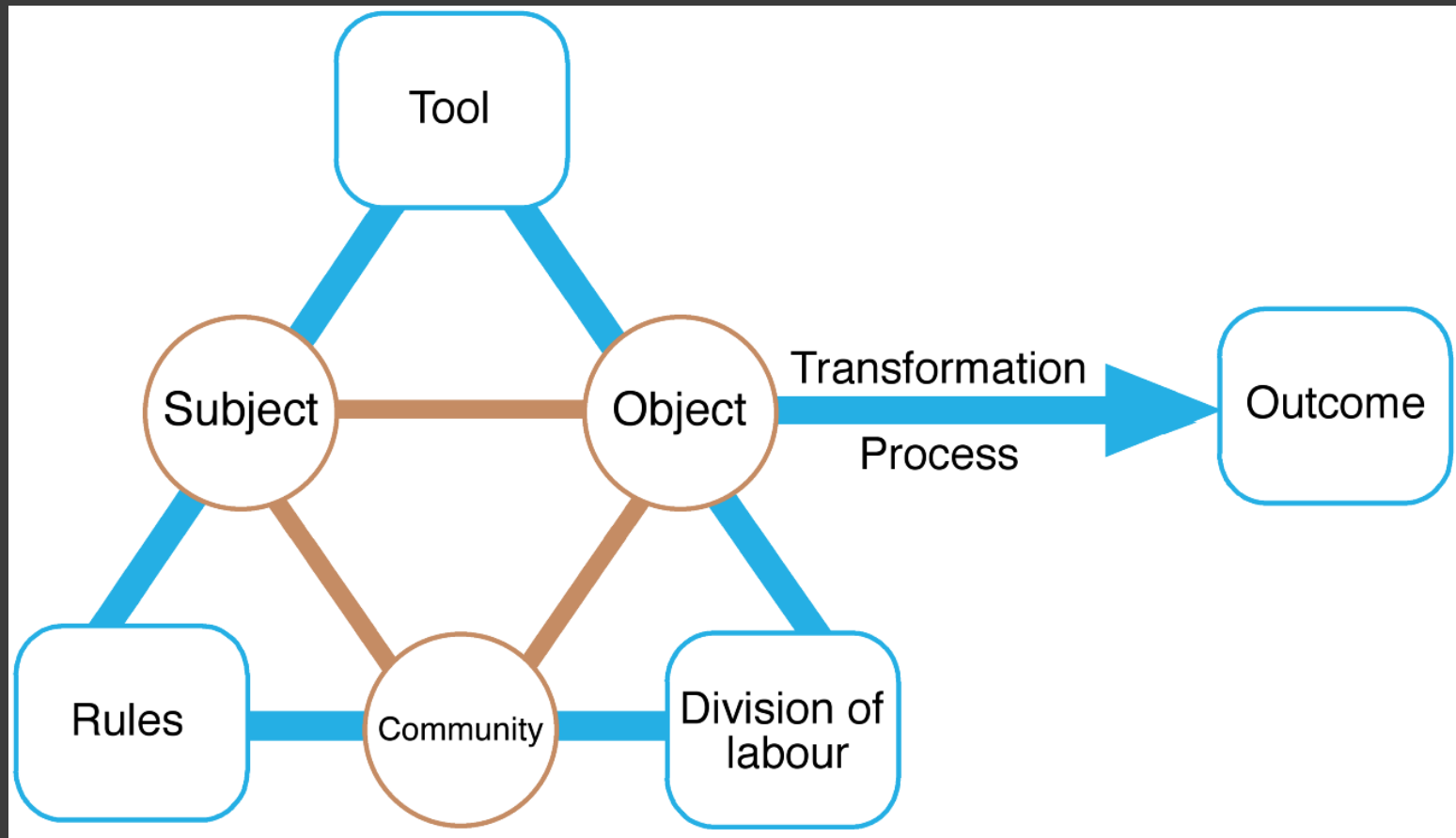


Action – Goal



Operation – Conditions

# *Engeström's (1999) activity system model*



# Presenting the findings

- Only make claims that your data can support
- The best way to present your findings depends on the audience, the purpose, and the data gathering and analysis undertaken
- Graphical representations (as discussed above) may be appropriate for presentation
- Other techniques are:
  - Rigorous notations, e.g. UML
  - Using stories, e.g. to create scenarios
  - Summarizing the findings

# SUMMARY

- The data analysis that can be done depends on the data gathering that was done
- Qualitative and quantitative data may be gathered from any of the three main data gathering approaches
- Percentages and averages are commonly used in Interaction Design
- Mean, median and mode are different kinds of 'average' and can have very different answers for the same set of data
- Grounded Theory, Distributed Cognition and Activity Theory are theoretical frameworks to support data analysis
- Presentation of the findings should not overstate the evidence

