Research Methods
for Business and Social Sciences

Pandelis Ipsilandis
Professor Emeritus, University of Thessaly

Contents (part II)

1. Formulating and clarifying the research topic
2. Critically reviewing the literature
3. Formulating the research design
4. Selecting Samples
5. Using secondary data
6. Primary Data thru Observation
7. Primary Data thru Interviews
8. Primary Data thru Surveys
9. Analysing Quantitative data
10. Analysing Qualitative data
11. Ethics

Part 6
Collecting primary data through observation

Observation as a data collection method

'Observation involves the systematic observation , recording, description analysis and interpretation of people's behaviour'

Saunders *et al*. (2009)

Types of observation

The two main types

- **Participant** observation – emphasises the discovery of meaning attached to actions (qualitative)

- **Structured** observation – is concerned with frequency of actions

  (quantitative)

Participant observation (1)

Definition

'Where the researcher attempts to participate fully in the lives and actions of subjects, enabling them to not merely observe what is happening but also feeling it'

Adapted from Gill and Johnson (2002)

Participant observation (2)

Points to consider

- Used both as the principle research method and in combination with other methods

- Researchers become immersed in the research setting

- Researchers try to understand the process by which individual identity is constructed and reconstructed (symbolic interactionism)

Examples of participant observation

- ## Street corner society - Whyte (1955)

Whyte lived in the North End of Boston, mostly inhabited by 1st and 2nd generation immigrants from Italy. Whyte, who came from a well-to-do family, considered the neighborhood a slum, and wanted to learn more about its "lower class" society.

Street Corner Society describes various groups and communities within the district. **Compaesani** – people originally from the same Italian town – are one example. How **local gangs were formed** and organized. Whyte differentiated between the lives of the "**corner boys**" revolved around particular street corners and the nearby shops and "**college boys**" who were more interested in good education and moving up the social ladder. Relations of social structure, politics, and racketeering in that district. It is also a testament to the importance of WPA jobs at the time

Through this work, Whyte became a pioneer in participant observation (which he called "participant observer research")

Examples of participant observation

- ## Machine shop piecework bonus system -

Roy (1952) worked as an employee in a machine shop for then months in order to understand how and why his "fellow workers" operated the piecework bonus system. Management destroyed trust by using trivial excuses to lower piece rates when employees found ways to be more productive

- ## Advertising agency – Rosen (1991)

The social drama played out during the breakfast at the advertising agency is one way in which social order is maintained and political control is sustained. A variety of symbolic techniques are used in a way to influence action, sharing experience, shaping ideas.. "commensality" – agency members break bread together.
Watch Netflix' "Mad Men

Researcher roles in participant observation

Typology of participant observation researcher roles
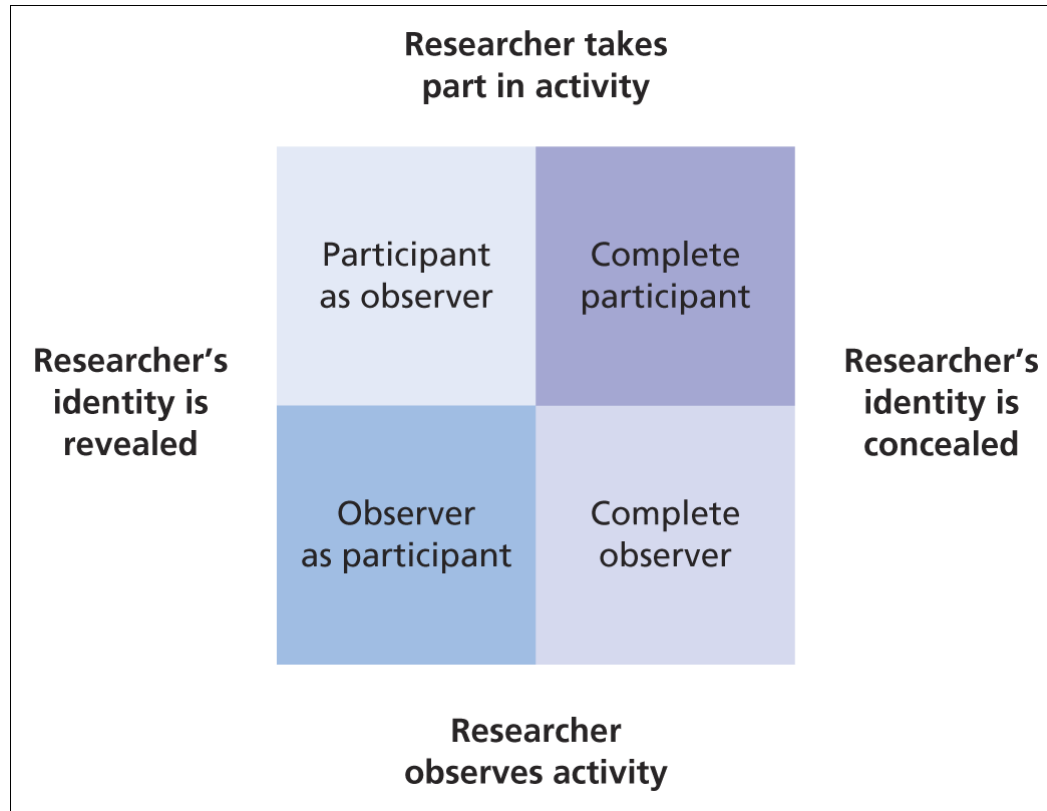
Source: Gill and Johnson (2002)

Choice of participant observer role

Determining factors

- Purpose of the research and time available

- Degree of suitability felt by the researcher
  - Not everybody is suitable

- Organisational access

- Ethical considerations

Data collection and analysis (1)

Types of data generated by participant observation

- Primary observations
  - Note what happened, what said. Keep diary
- Secondary observations
  - Observer's statements. (interpretations)
- Experiential data
  - Perception and feelings of the observer. Keeping a diary is useful
- Other
  - Organizational settings, changes, communications, etc.

Delbridge and Kirkpatrick (1994)

Data collection and analysis (2)

Points to consider

- Data may be classified as 'descriptive observation' and 'narrative account' Robson (2002)

- Data recording method(s) will depend on the role
  - If recording, that should take place on the same day as field work

- Data collection and analysis may be part of the same process

- Use of analytic induction leads to redefinition of the original hypothesis

Data collection and analysis (3)

Points to consider

- Threats to validity
  - Participant observation is very high on ecological validity; studying natural phenomena in their natural context

- The perspective of the subject - not the researcher
  - avoid observer's bias

Advantages	Disadvantages
• Good at explaining what is going on in particular social situations	• Time consuming
• Heightens researcher's awareness about the phenomenon	• Difficult ethical dilemmas
• Useful when working in own organization	• High levels of role conflict
• Opportunity for researcher to experience "for real" emotions of participants	• Closeness can lead to significant bias
• All data collected are useful	• Participant observer role is very demanding
	• Access to organizations may be difficult
	• Data recording is difficult

Saunders *et al*. (2009)

Structured observation (1)

Points to consider

- Thinking about **what you want to observe** rather than simply noting down whatever is visible and seems relevant.

- Structured observation is systematic and aims to establish straightforward facts

- Structured observation was an important part of Mintzberg's (1973) study of managerial work
    - Cast doubts on the theory that managerial work was a rational process of planning, controlling and directing

- Proliferation of the internet potentially widens the scope of participant observation in certain areas
    - Traces of behaviour

Saunders *et al*. (2009)

Structured observation (2)

Data collection and analysis

- Coding Scheme: Categorizing behavior so that you can code what you observe in terms of how often a type of behavior appears
  - Choosing an "off the shelf' coding schedule
  - Designing your own coding schedule
  - Combining both types of schedule

- Use of simple (manual) or complex (computer) methods of analysis

Structured observation (3)

Threats to validity and reliability

- Subject error
  - observations will be biased towards those behavior patterns and individuals which happen to be most conspicuous.

- Time error
  - Observing at different times / days

- Observer effects
  - conscious or unconscious prejudices.

Summary: Part 6

- Participant observation is used in a wide range of social settings

- Participant observation means adopting a number of potential roles

- Roles are differentiated according to the degree of concealed identity and participation in events adopted by the researcher

- The aim of participant observation is to **develop theory and avoid mere story telling**

- Structured observation is concerned with the frequency of events. It is characterised by high levels of predetermined structure and quantitative analysis

- The main threats to reliability and validity are subject error, time error and observer effects

Part 7

Collecting primary data using semi-structured, in-depth and group interviews

Research interviews

Definition

'An interview is a purposeful discussion between two or more people'

Kahn and Cannell (1957)

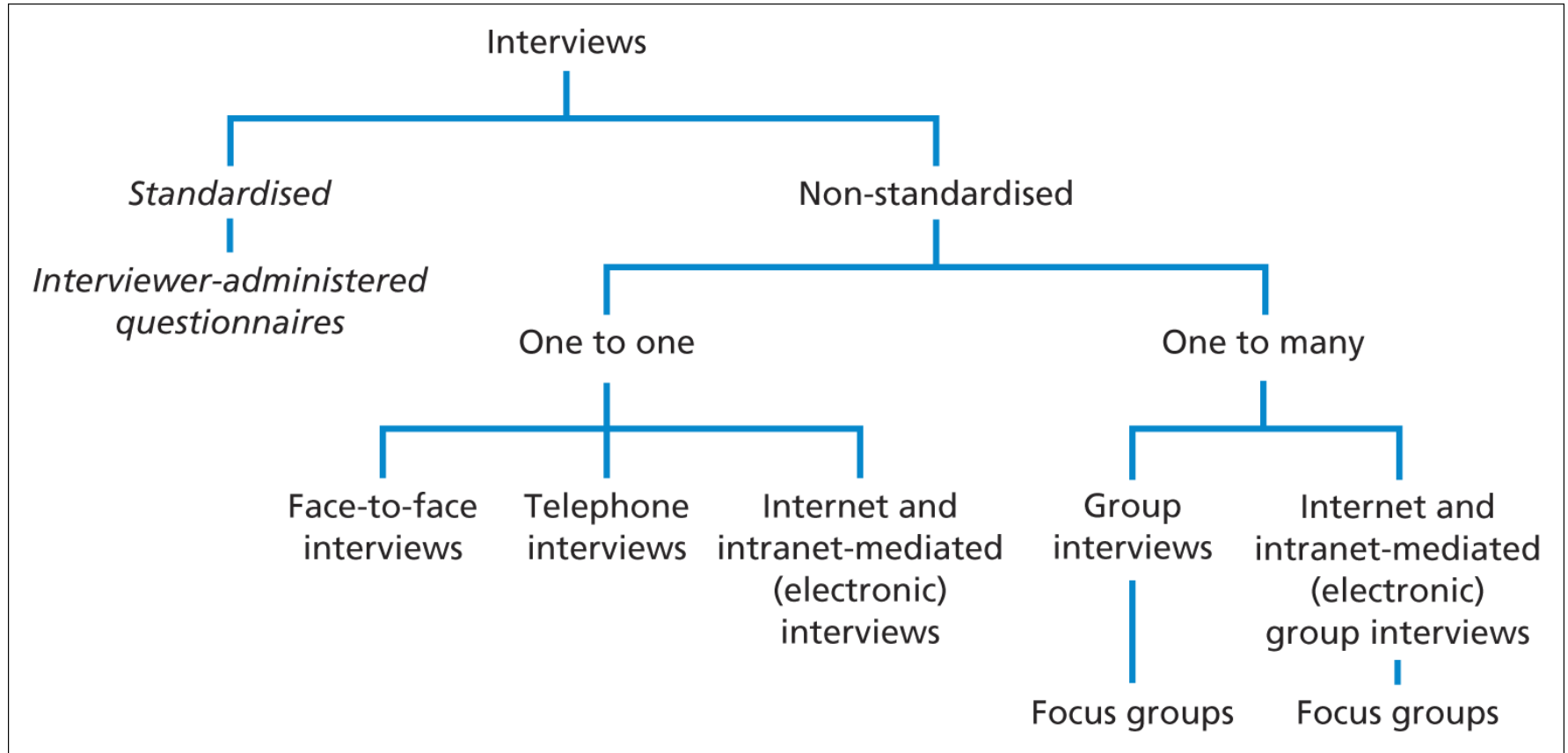## Types of interview used in research

Semi-structured
In-depth

Structured
Group

Saunders *et al*. (2009)

Research purpose and strategy (1)

Forms of interview

Saunders *et al*. (2009)

Research purpose and strategy (2)

Uses of different types of interview in each of the main research categories

	Exploratory	Descriptive	Explanatory
Structured		✓✓	✓
Semi-structured	✓		✓✓
Unstructured	✓✓		

✓✓ = more frequent, ✓ = less frequent.

- More than one type may be used
  - e.g. Unstructured, Semi structured to identify issues, questions followed by questionnaire administration in structured interviews

Saunders *et al*. (2009)

Non-standardised (qualitative) interviews

Four key aspects

- Purpose of the research
  - Exploratory element of research. Formulate research questions.
  - Discover reasons, add depth and significance to the data you obtained. Address research questions

- Significance of establishing personal contact
  - Managers are more keen to provide information through interviews than filling questionnaires. More trust.
  - Achieve higher response rate

- Nature of the data collection questions
  - Complex / large number / order varies depending on answers

- Time required and completeness of process
  - Inform interviewee about purpose, objectives, length of time, agree on a convenient time

Data quality (1)

Issues to consider

- **Reliability**
  - Reflect reality of time, circumstances

- **Forms of bias**
  - Reduction of willingness in lengthy interviews
  - Interviewees feeling uncomfortable / defensive

- **Validity and generalisability**
  - Access to interviewee's knowledge
  - Limited generalization of conclusions

Data quality (2)

The importance of preparation – the 5 Ps

'prior planning prevents poor performance'

Be trained and practice a lot your interview skills if you plan to gather data through interviews

Saunders *et al*. (2009)

Interview preparation (1)

Associated issues

- Interviewer's level of knowledge
  - Research topic and personal / organizational background of interviewees. Internet search.

- Level of information supplied to interviewees
  - Promote credibility by supplying information. Provide list of topics before the event, Access to company information

- Creating an interview guide
  - List of topics, initial questions, probes to follow up, comprehensive language

- Appropriateness of location
  - Participants may feel comfortable.

Interview preparation (2)

Associated issues

- Researcher's appearance – dress code
  - May affect the perception of the interviewee. Could result in biased information

- Shaping the interview - opening comments
  - Safeguard your reliability and interviewee's confidence
  - Interviewee may have uncertainties about sharing information

- Approach to questioning – clarity and reducing bias
  - Questions should be grounded on real experiences of participants not abstract concepts

- Use of critical incident technique
  - Asked to describe activity or event where consequences are so clear that interviewee has a definite idea about the effect

Interview preparation (3)

Associated issues

- Appropriate interviewer behaviour- verbal and non-verbal
  - Avoid comments or non-verbal behaviour, which indicate any bias
  - Posture and tone of voice could encourage or inhibit

- Attentive listening skills and testing understanding
  - Give respondent's time. Summarize

- Cultural differences and bias

Interview preparation (4)

Associated issues

- Approaches to data recording - notes and tape-recording
  - As well as audio recording make also some notes
  - Compile a full record right after the interview, including contextual data
    - Location
    - Date, time and duration
    - Setting of the interview
    - Background information about participants
    - Immediate impression of how well or badly did it go

Interview preparation (5)

Checklist (1/3)

- How might your level of preparation and knowledge in relation to the research context and your research question affect the willingness of the interviewee to share data?

- What will be the broad focus of your in-depth interview or what are the themes that you wish to explore or seek explanations for during a semi structured interview?

- What type of information, In any, would be useful to send to your interviewee prior to the interview?

- What did you agree to supply to your interviewee when your arrange the interview? Has this been supplied?

- How will you appearance during the interview affect the willingness of the interviewee to share data?

- Have you consider the impact that your interview location on participants responses and on your own personal safety?

Saunders *et al*. (2009)

Interview preparation (5)

Checklist (2/3)

- How will you prepare yourself to be able to commence the interview with confidence and   purpose?

- What will you tell your interviewee about yourself, the purpose of your research, its funding and your progress?

- What concerns or need for clarification, may your interviewee have?

- How will you seek to overcome these concerns or provide clarifications?

- What will you tell your interviewee about the right not to answer particular questions and to end the interview should they wish?

Saunders *et al*. (2009)

Interview preparation (5)

Checklist (3/3)

- In particular how do you intend to use the data to which you are given access, ensuring where appropriate it's confidentiality and your interviews anonymity?

- How would you like to record the data that are revealed to you during the interview? Will this involve audio recorder? Have you raised this as a request and provide a reason why it would help you to use this technique?

- How will you seek to overcome potential issues related to the reliability of the data you collect including forms of the interviewer bias (related to your role and contact) interviewee bias (the level of access that you gain to the data of those whom you interview(and sampling bias?

-

Interviewing competence (1)

Approaches to questioning

Open questions

- Why did the organization introduce its marketing strategy?
- What method have been used to make employees redundant?

Probing questions

That's interesting… Tell me more ….

- How would you evaluate the success of the marketing strategy?
- How did you choose a compulsory method to make redundancies?

Specific and closed questions

- How many people responded to customer survey?

Interviewing competence (2)

Advantages and disadvantages of audio-recording interviews

Advantages	Disadvantages
• Allows interviewer to concentrate on questioning and listening	• May adversely affect the relationship between interviewee and interviewer (possibility of 'focusing' on the audio-recorder)
• Allows questions formulated at an interview to be accurately recorded for use in later interviews where appropriate	• May inhibit some interviewee responses and reduce reliability
• Can re-listen to the interview	• Possibility of a technical problem
• Accurate and unbiased record provided	• Time required to transcribe the audio-recording (Section 13.3)
• Allows direct quotes to be used	
• Permanent record for others to use	

Saunders *et al*. (2009)

Interviewing competence (3)

Other issues to consider

- Managing resources – logistics and time
  - More costly than other methods
  - Give thought to time needed for interviews and compose notes / transcribe audio recordings
  - Schedule interviews well in advance

- Obtaining participants' permission for interview records (written and taped)

- Dealing with difficult participants

Interviewing competence (3)
Dealing with difficult participants

| Difficulty	Possible reasons	Suggestion
Provide monosyllabic answers little more than "yes" "no"	Time limitation. Worries about anonymity.	Opening the interview. Phrase questions more openly, make pauses
Repeatedly provide long answers which deviate from the focus of interview	Some should be tolerant Could lead to interesting point.	Cleverly ask to pause so you can note down responses, change the subject
Start interviewing you	Affinity, like-mindness	Emphasize that you are interested in their opinions and you can discussed more at the end
Proud of their status, intervening and criticizing your work	Want to show off their knowledge, superiority	Be confident and prepare to justify your research and the research design
Become noticeably upset during the interview.	Don't do anything to suggest you are feeling impatient. Tell them that they don't have to answer that question. Do not end the interview straight away. It will upset them more.	

Interviewing competence (4)

Additional forms of interviews:

- Group interviews

- Focus groups

  – Topic is defined clearly and precisely and there is a focus on enabling and recording interactive discussion among participants

- Telephone interviews

- Internet and intra-net mediated interviews

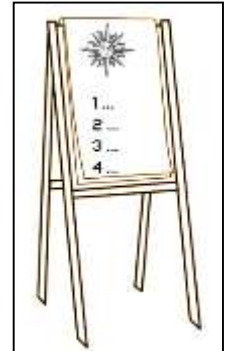# Interviewing competence (5)

Forms of electronic interviews
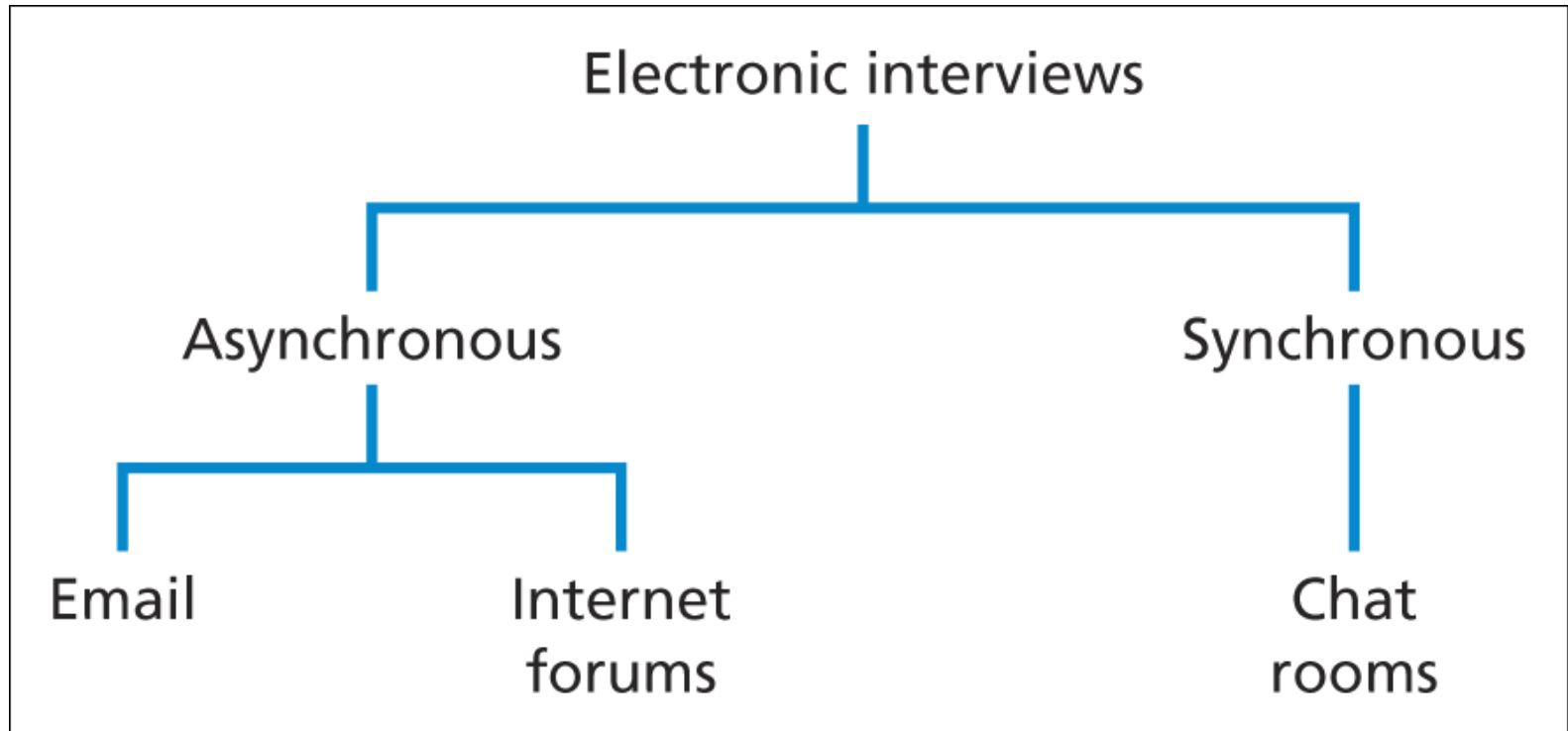
Saunders *et al*. (2009)

Types of Interviews

What kind of interview will you use in each of the following:

- – Market research project?

- – Research project seeking to understand whether trade unions attitudes have changed?

- – Following the analysis of a questionnaire?

Using an interview approach to collect data

The following statement was made during the evaluation of a research proposal, where the candidate's proposed research methodology was based on data collection through interviews:
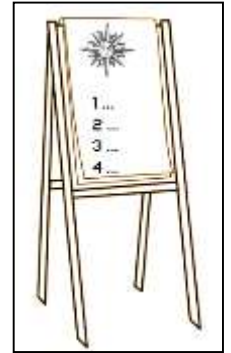
"It is not an easier alternative for those who want to avoid statistics"

Any comments?

Summary: Part 7

- Use of non-standard (qualitative) interviews should generate rich and detailed data

- Different types of interview are useful for different research purposes

- Qualitative interviews are generally categorised as in-depth (structured) and semi-structured

- Research design may incorporate more than one type of interview

- Using qualitative interviews is related to the research strategy and data collection questions

- Establishing personal contact with subjects and the length of time required are significant factors

- Data quality issues, interviewer competence and resource management are important considerations

- Face-to-face (individual, group and focus group) and electronic interviews can all be advantageous

Part 8
Collecting primary data using questionnaires

Use of questionnaires (1)

Definition of Questionnaires

Techniques of data collection in which each person is asked to respond to the same set of questions in a predetermined (*not always*) order
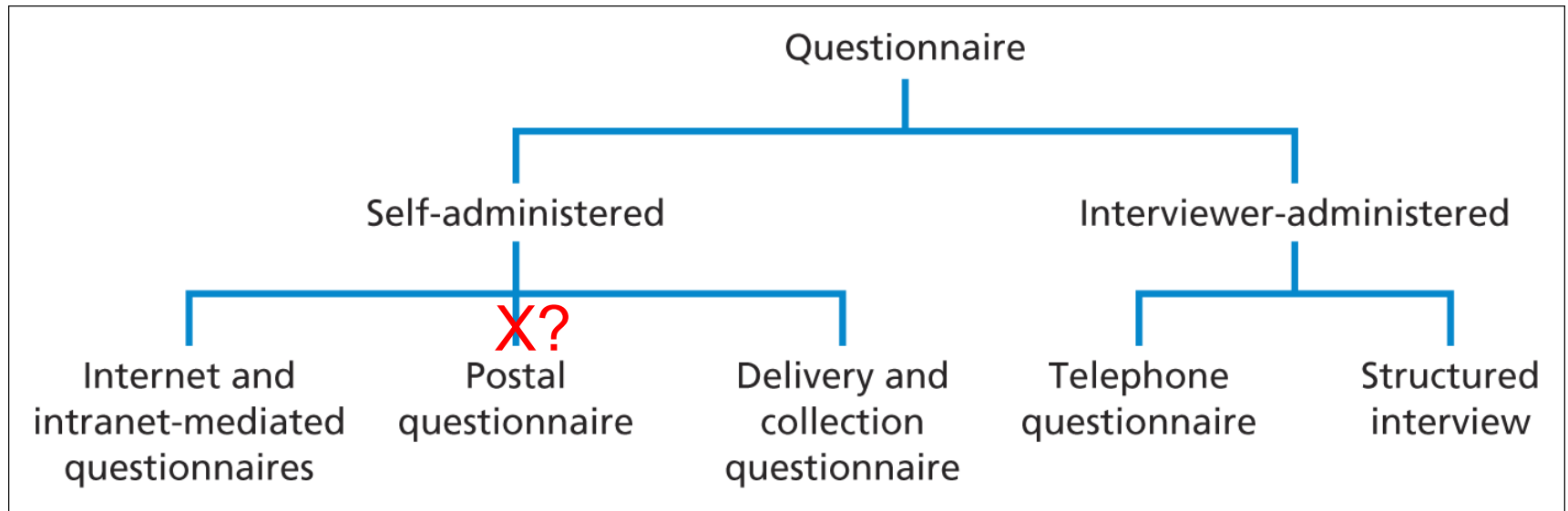
Adapted from deVaus (2002)

Use of questionnaires (2)

When to use questionnaires

- For explanatory or descriptive research

- Linked with other methods in a multiple-methods research design

- To collect responses from a large sample prior to qualitative analysis

Use of questionnaires (3)

Types of questionnaire

Choice of questionnaire

Related factors

- Characteristics of the respondents and access
  - if you want manager's view and not his/her assistance, it is better to interview
  - Delivery and collection allows to know who responded, assess the impact of non-respondents
- Respondents answers not being contaminated or distorted
  - Respondents not knowledgeable / discussing responses among them / answer to please you / providing socially desirable responses
- Size of sample required for analysis
- Type and number of questions required
  - Longer questionnaires and open questions are better presented in a structured interview
- Available resources including use of computer software

Data collection

Key factors

- Precisely defined questions
  - One chance. Not possible to prompt and explore issues further.
  - Time for planning is time is well spend
- Representative and accurate sampling
  - Especially if you plan to compare findings with previous researches in literature – off course it does not have to be the same.
- An understanding of the organisational context
  - Industry related, international aspects,..
  - Use the right terminology
- Relationships between variables
  - dependent, independent and extraneous
- Types of variable
  - Opinion, Behaviour, Attribute

Ensuring essential data are collected

- Subdivide Research questions or objectives into more specific investigative questions about which you need to gather data

- Repeat subdividing if you feel that questions are not sufficiently precise

- Identify variables about which you need to collect data

- Establish the level of detail required from the data for each variable

- Develop measurement questions to capture the data at the level required for each variable
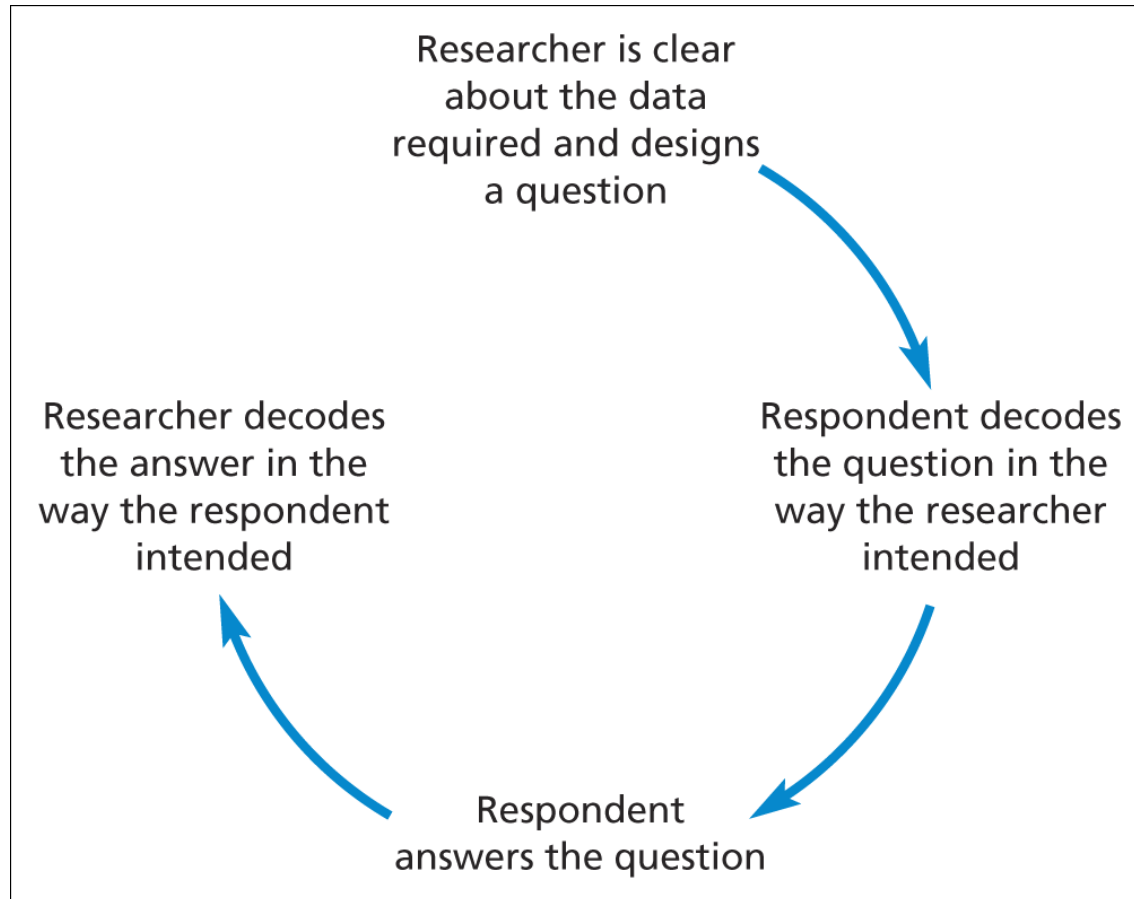
Data requirements table

Research Question: To establish customer's attitudes to the outside smoking area at restaurants and bars		
Type of research: Predominantly descriptive, although seeks to examine differences between restaurant and bars and between different customer groups		
Investigative Questions	**Variable(s) required**	**Detail in which data measured**
Do customers feel that they should have an outside smoking area at restaurants and bars as a right (opinion)	Opinion of customer on restaurants and bars providing an outside smoking area	Feel… Should be a right, should not be a right, no strong feelings (separate questions for restaurants and bars)
Do customers feel that restaurants and bars should provide an outside for smokers (opinion)	Opinion of customers to the provision of an outside smoking area	Feel…very strongly that it should, no strong opinion, quite strongly that it should not, very strongly that it should (separate for restaurants and bars)
Do customers' opinions differ depending on	-outlined above	Included above
Age? (attribute)	Age of customer	5 year band 16 to 65+
Whether or not a smoker (behaviour)	Smoker	Non-smoker, ex-smoker, smokes but not in own home, smokes in own home
How representative are the responses (attributes)	Age, Gender, Job, Education, Income	Male/female, High managerial Low Managerial, Technical, Self-employed, Technicians, …., Unemployed….. …Income brackets

Designing the questionnaire (1)

Stages that must occur if a question is to be valid and reliable

Researcher is clear about the data required and designs a question

Respondent decodes the question in the way the researcher intended

Respondent answers the question

Researcher decodes the answer in the way the respondent intended

Source: developed from Foddy (1994)

Designing the questionnaire (2)

Assessing validity

- Internal (measure what you intended to measure!!)
  - Relevant evidence, Own judgement

- Content
  - Adequate coverage of investigative issues
  - Panel reviews essential, useful, not essential

- Criterion – related (predictive)
  - Ability to predict, Correlation

- Construct
  - Attitude scales, aptitude and personality tests, etc. (Cooper & Schindler 2008)

Designing the questionnaire (3)

Testing for reliability- the 3 stage process
Consistency / Robustness of Questionnaire

- Test re-test
  - difficult to convince respondents,

- Internal consistency
  - Correlating responses to one question with others
  - Cronbach's alpha statistical test

- Alternative form
  - Include "attention" or "check" questions

Mitchell (1996)

Cronbach's alpha – What it is?

- Most commonly used measure of internal consistency

- Tests if multiple-questions in Likert scale surveys that measure latent variables — hidden or unobservable variables like: *satisfaction*, *openness, conscientiousness, neurosis* are reliable.

- These are very difficult to measure in real life. Cronbach's alpha will indicate if the construct you have designed is consistently measure the variable of interest.

- Can be used to check consistency of responses within a construct (group of related questions) or the entire questionnaire

- Lack of consistency could mean that: Respondents did not understood the question(s) / Paid little attention to think before they answer / They just pick an answer they "liked" / …

Interpreting scale reliability

- The higher the score, the more reliable the generated scale is

- A score of .70 or greater is generally considered to be acceptable
  - .90 or > = high reliability
  - .80-.89 = good reliability
  - .70-79 = acceptable reliability
  - .65-.69 = marginal reliability

- lower thresholds (> .50) are sometimes used in the literature.

"Good" or "Bad" questions in terms of alpha?

- "SPSS and SAS will report "alpha if item deleted", which shows how alpha would change if that one question was not on the test.

    - Low "alpha if item deleted" means a question is good because deleting that question would lower the overall alpha.

    - Questions with high "alpha if deleted" tend to have low inter-item correlations (Pearson's r).

- If a question is "bad", this means it is not conforming with the rest of the test to measure the same basic factor (e.g., statistics knowledge).

    - The question is not "internally consistent" with the rest of the test.

Examples of question types (1)

Open questions

6   Please list up to three things you like about your job

1……………………………………………………

2……………………………………………………

3……………………………………………………

Saunders *et al*. (2009)

Examples of question types (2)

List questions

7 **What is your religion?**
Please tick ✓ the appropriate box

Buddhist ☐   None ☐
Christian ☐   Other ☐
Hindu ☐
Jewish ☐
Muslim ☐
Sikh ☐

Saunders *et al.* (2009)

Examples of question types (3)

Category questions

8   How often do you visit the shopping centre?

Interviewer: listen to the respondent's answer and tick ✔ as appropriate

(In general, use no more than 5 categories, unless you using flash cards)

Put them in a natural order

☐   First visit

☐   Once a week

☐   Less than fortnightly to once a month

☐   2 or more times a week

☐   Less than once a week to fortnightly

☐   Less often

Does this make sense

Saunders *et al*. (2009)

Examples of question types (4)

Ranking questions

9   **Please number each of the factors listed below in order of importance to you in choosing a new car. Number the most important 1, the next 2 and so on. If a factor has no importance at all, please leave blank.**

Factor	Importance
Carbon dioxide emissions	[ ]
Boot size	[ ]
Depreciation	[ ]
Price	[ ]

Adapted from Saunders *et al*. (2009)

Examples of question types (5)

Rating questions

10   **For the following statement please tick the box that matches your view most closely**

	Agree	Tend to agree	Tend to disagree	Disagree
I feel employees' views have influenced the decisions taken by management	☐	☐	☐	☐

Saunders *et al*. (2009)

Examples of question types (5)

Rating questions

- Four, Five. Six, Seven point rating scale
- The number of points depends on the level that respondents can accurately respond to. E.g. > 3 levels is very difficult for telephone surveys
- Even number of point force respondents to "take sides" – not always desirable
- For a series of statements keep the same order of response categories
- Use positive and negative statements to make sure that respondents read the question

Examples of question types (5)

Rating Scales

- Agreement of a respondent to a statement

- For the following statement please circle the number that matches your view point, most closely

Support provided through trainings is

Absolutely inadequate

Sufficient

1   2   3   4   5   6   7   8   9   10

)
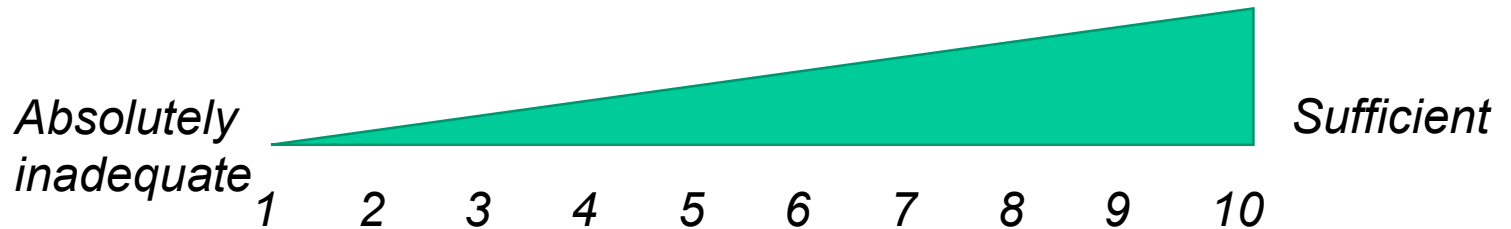
Examples of question types (5)

Matrix Questions

3. How satisfied or dissatisfied are you with each of the following?

	Very satisfied	Somewhat satisfied	Neither satisfied nor dissatisfied	Somewhat dissatisfied	Very dissatisfied
The interaction with the sales staff	○	○	○	○	○
Your experience at the register	○	○	○	○	○
The organization of the store	○	○	○	○	○

4. How satisfied or dissatisfied are you with each of the following?

	Very satisfied	Somewhat satisfied	Neither satisfied nor dissatisfied	Somewhat dissatisfied	Very dissatisfied
The products offered in the store	○	○	○	○	○
The price of the products	○	○	○	○	○

Examples of question types (6)

Quantity questions

14 What is your year of birth?

1	9		

(For example, for 1988 write:)

1	9	8	8

Saunders *et al*. (2009)

Designing individual questions

Other considerations

- Adopting or adapting existing questions – remember to check copyright

- Question wording

- Translating questions into other languages

- Question coding
  - Establish coding scheme before data collection

- Take into consideration Analysis Requirements
  - AHP
  - Multicriteria models

Checklist for designing individual questions (1)

- Will collected data be at the right level of detail?

- Will respondents have the right knowledge to answer the question?

- Does the question underestimate the respondents

- Is wording familiar to all so they understand it the same way?

- Any words that may cause offence?

- Can the question be shortened?

- Are more than one questions at the same time?

- Does the question include a negative or double negative? "*Would you rather not use a non-generic medicine…?*

Checklist for designing individual questions (2)

- Is the question unambiguous? *"When did you leave school? (year?, age?)" "Number of employers 0-100, 100-200 …"*

- Does the question imply that a certain answer is correct? *"Is this the first time you pretended to be sick?"*

- Is the question likely to embarrass the respondent? *Income question in brackets*

- Maximum number of categories

- Are closed questions written in a way that at least one choice will apply to every respondent and responses are mutually exclusive?

- Include "*prefer not to answer*" in your answer choices if at all possible.

  – *Many people will drop-out of a survey if they are uncomfortable with a particular question*

- Are instruction on how to record answers clear?

Examples of bad questions (2)

- Do you have any problems with your boss?

- Should concerned parents use infant car seats?

- Where do you enjoy drinking beer?

- How satisfied or dissatisfied are you with the pay and work benefits of your current job?

- How useful will this textbook be for students and young professionals in the field?

- Do you always eat breakfast? (Yes/No)

- Do you own a tablet PC?

Constructing the questionnaire

Main considerations

- Order and flow of questions
  - Filter questions assist the flow
  - Using survey s/w (CAPI, CATI) irrelevant questions do not appear

- Questionnaire layout

Explaining the purpose and testing
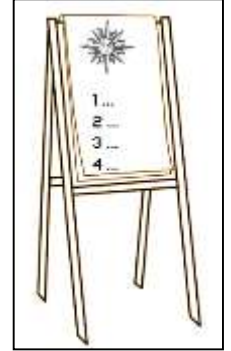
Key points

- The covering letter

- Introducing and closing the questionnaire

- Pilot testing and assessing validity

Administering the questionnaire

Points to consider

- Internet and intranet-mediated responses

- Postal questionnaires

- Delivery and Collection

- Telephone questionnaires

- Structured interviews

Using an interview approach to collect data

You are undertaking research on the use of children's book clubs by households within mainland Europe As part of this you have already undertaken interviews with households who belong and do not belong to children book clubs. This along with literature review has suggested a number of investigative questions from which you start to construct a table of data requirement.

- For each investigative question listed, decide whether you will need to collect data on opinions, behaviours or attributes.

Defining data requirements

Research Question: To establish mainland Europe's householder's opinion about children's book clubs		
Type of research: Predominantly descriptive, although wants to explain differences householders		
Investigative questions	Variable(s) required	Detail in which data measured
Do householders think that children's book clubs are a good or bad idea		
What things do householders like most about children's book clubs		
How much per year do households spend on children's books		
Do householders responses differ depending on: (i) number of children? (ii) whether already members of book club		

Summary: Part 8

- Questionnaires are often used to collect descriptive and explanatory data

- Five main types of questionnaire are Internet- or intra-net mediated, postal, delivery and collection, telephone and interview schedule

- Precise data that meet the research objectives can be produced by using a data requirements table

- Data validity and reliability and response rate depend on design, structure and rigorous pilot testing

- Wording and order of questions and question types are important considerations

- Closed questions should be pre-coded to facilitate data input and analysis

- Important design features are a clear layout, a logical order and flow of questions and easily completed responses

- Questionnaires should be carefully introduced and pilot tested prior to administration

- Administration needs to be appropriate to the type of questionnaire

Part 9
Analysing quantitative data

Quantitative data analysis (1)

Key points

- Data must be analysed to produce information
  - Data can be a product of all research strategies

  Data → Information → Knowledge

- Computer software analysis is normally used for this process
  - SPSS, SAS, Minitab, R-language, Statview, Excel, etc.
  - Specialized survey design and analysis packages
  - Transferable skills. Worth to invest time

- Data should be carefully prepared for analysis

Quantitative data analysis (2)

Key points

- Researchers need to know how to **select** and **use** different charting and statistical techniques

… a field where it is not at all difficult to carry out an analysis which is simply wrong or inappropriate for your purposes.
And the negative side of readily available analysis software is that it becomes that much easier to generate elegantly presented rubbish

Robson 2002

Quantitative data analysis (3)

Main concerns

- Preparing, inputting and checking data

- Choosing the most appropriate statistics to describe the data

- Choosing the most appropriate statistics to examine data relationships and trends

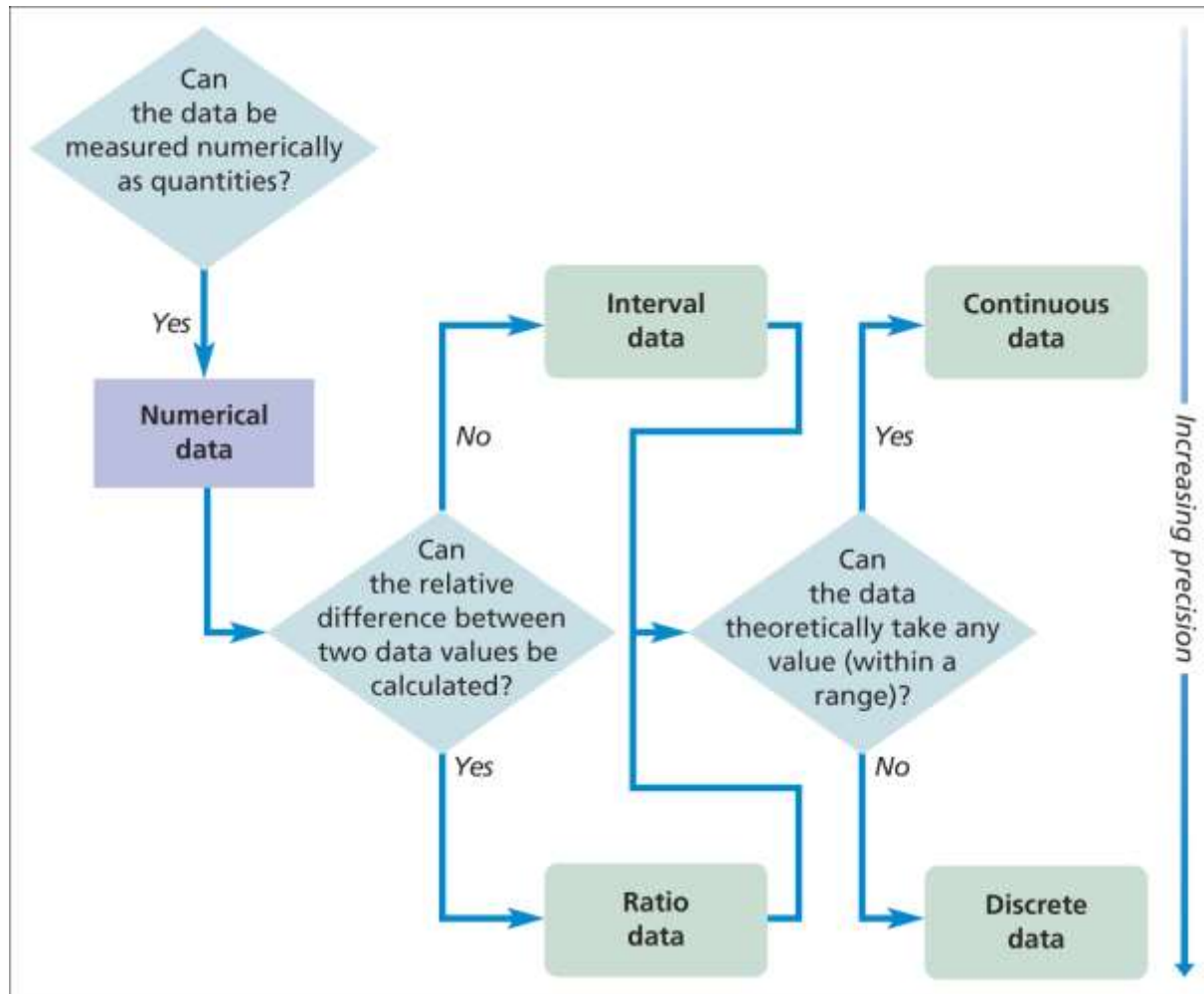# Preparing, inputting and checking data

Main considerations

- Type of data (scale of measurement)

- Data format for input to analysis software

- Impact of data coding on subsequent analyses

- Case weighting

- Methods for error checking

Ideally all of these should be considered
before obtaining the data

Defining the data type (1)
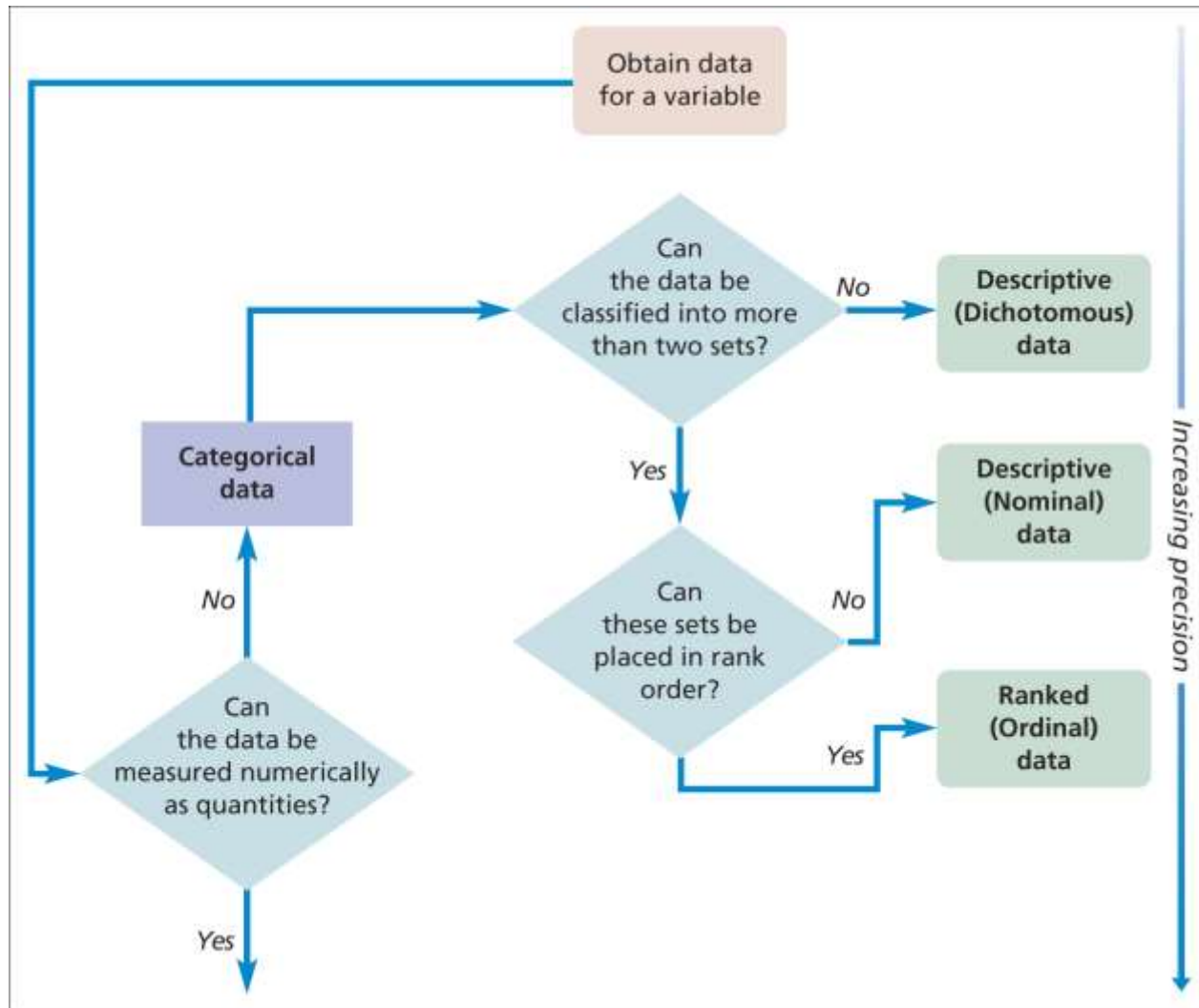
Saunders *et al*. (2009)

Defining the data type (2)

Preparing, inputting and checking data (4)

A simple data matrix

	Id	Variable 1	Variable 2	Variable 3	Variable 4
Case 1	1	27	1	2	1
Case 2	2	19	2	1	2
Case 3	3	24	2	3	1

Saunders *et al*. (2009)

Data Coding

- Coding general
  - All datatypes with few exceptions should be recorded using numerical codes. Why?
    - Enter the data quickly, fewer errors, simplify subsequent processing
  - Need to keep a list of codes. In statistical analysis s/w each code is automatically labelled

- Numerical data
  - Actual numbers
  - Can be combined to produce other variables
  - Can be recoded into groups

Data Coding

- Categorical data

  - Can keep the coding scheme used in data collection or modify it. Depends on level of precision vis-à-vis your data.

  - Design a coding scheme that will make subsequent analysis easier

  - Consider using existing coding when one exists. Save time / well-tested / allow comparisons of your results with others

  - Include codes in data collection form. Clear to interviewer or responder. Wait for the first 50-100 responses before you finalize the coding scheme.

  - Ensure that categories that may need to be aggregated receive adjacent codes

Data Coding

- Missing data
  - Each variable in each case should have a code even if no data are collected.
  - Missing data code could show why data are missing
    - Question skipped because of filter question
    - Respondent refused to answer
    - Respondent did not have an opinion
    - Respondent missed a question by mistake

- Weighting cases
  - Stratified sampling with different sampling factor for each strata
  - Different response rates among strata
    - Weight = highest proportion among all strata / proportion in this stratum
  - *Many authors question the validity of inferential statistics when weights are used*

Developing a Data Coding Scheme

- Survey includes an open question

  - List up to three things you like about this restaurant

- Let's collect some preliminary data
- Then develop a proper coding scheme
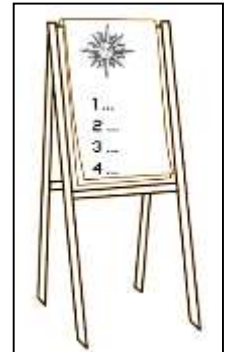
Preparing, inputting and checking data

Final stages of the process

- Entering data – rubbish in = rubbish out!
  - Make sure that labels correspond to exact words used in data collection to avoid problems later

- Always take time to check for errors
  - including illegitimate codes, illogical relationships and that rules were followed in filter questions
  - Discover when errors happened (coding or entry)

Checklist for exploring and presenting data

For both diagrams and tables

- Clear and descriptive title

- Are units of measurement used, stated clearly

- Are the sources of data used, stated clearly

- Are there notes to explain abbreviations and unusual terminology

- Stating the size of the sample on which the values are based

For diagrams

- Does it have clear axis labels

- Are bars and their components in the same logical sequence

- More dense shading used in smaller areas

- Avoided misrepresenting or distorting the data

- Is a key or Legend included

For tables

- Clear column and row headings

- Columns and rows in logical secrets

Exploring and presenting data (3)

Showing aspects of individual variables

- Specific values

- Highest and lowest values

- Trends

- Proportions

- Distribution of values

Graphical Excellence

"Complex ideas communicated with clarity, precision and efficiency"

- Shows the data

- Makes you think about substance rather than method, graphic design, or something else

- Many numbers in a small place

- Makes large data sets coherent

- Encourages the eye to compare different pieces of data

The Best and Worst of Statistical Graphics
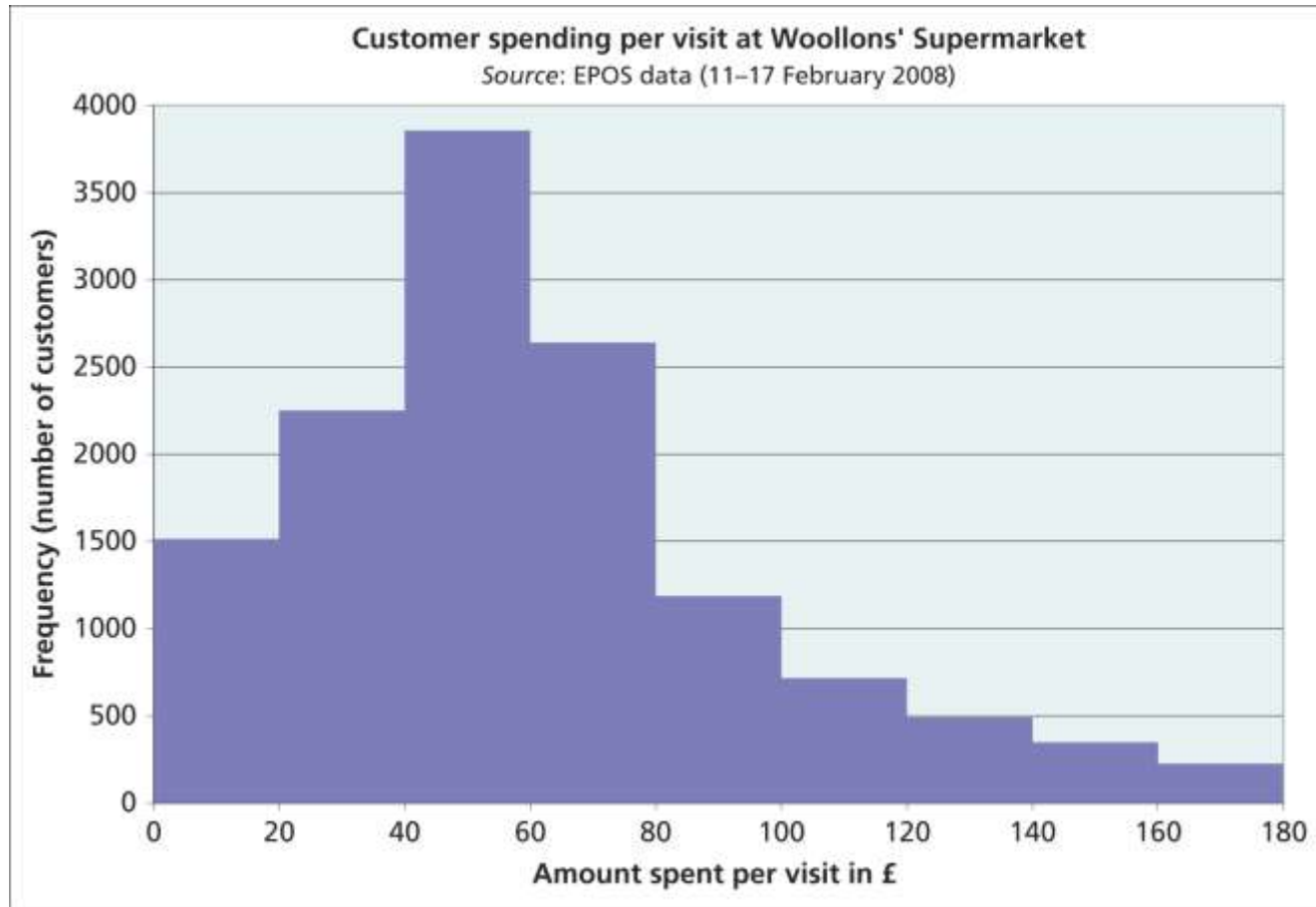http://www.datavis.ca/gallery/

Examples of diagrams (1)

Bar Chart
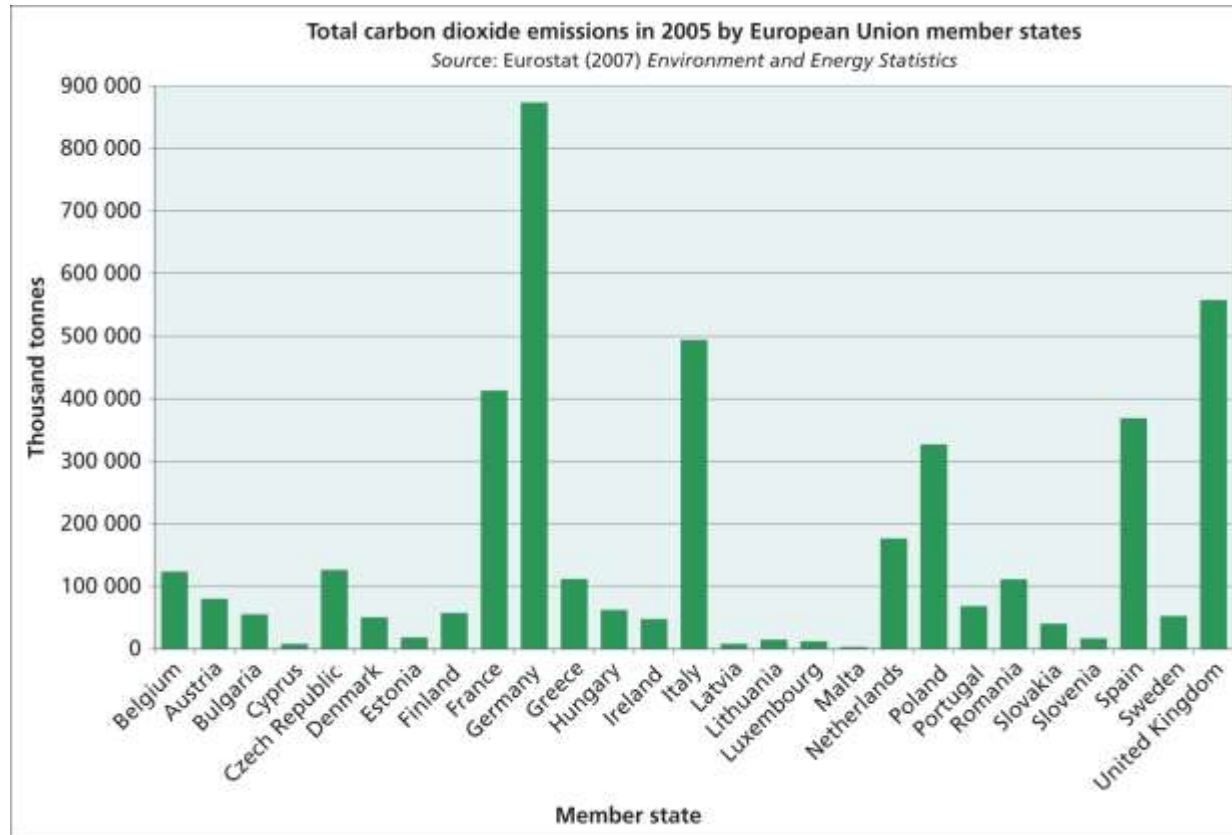
Source: adapted from Eurostat (2007) © European Communities, 2007
Reproduced with permission

Examples of diagrams (2)

Histogram

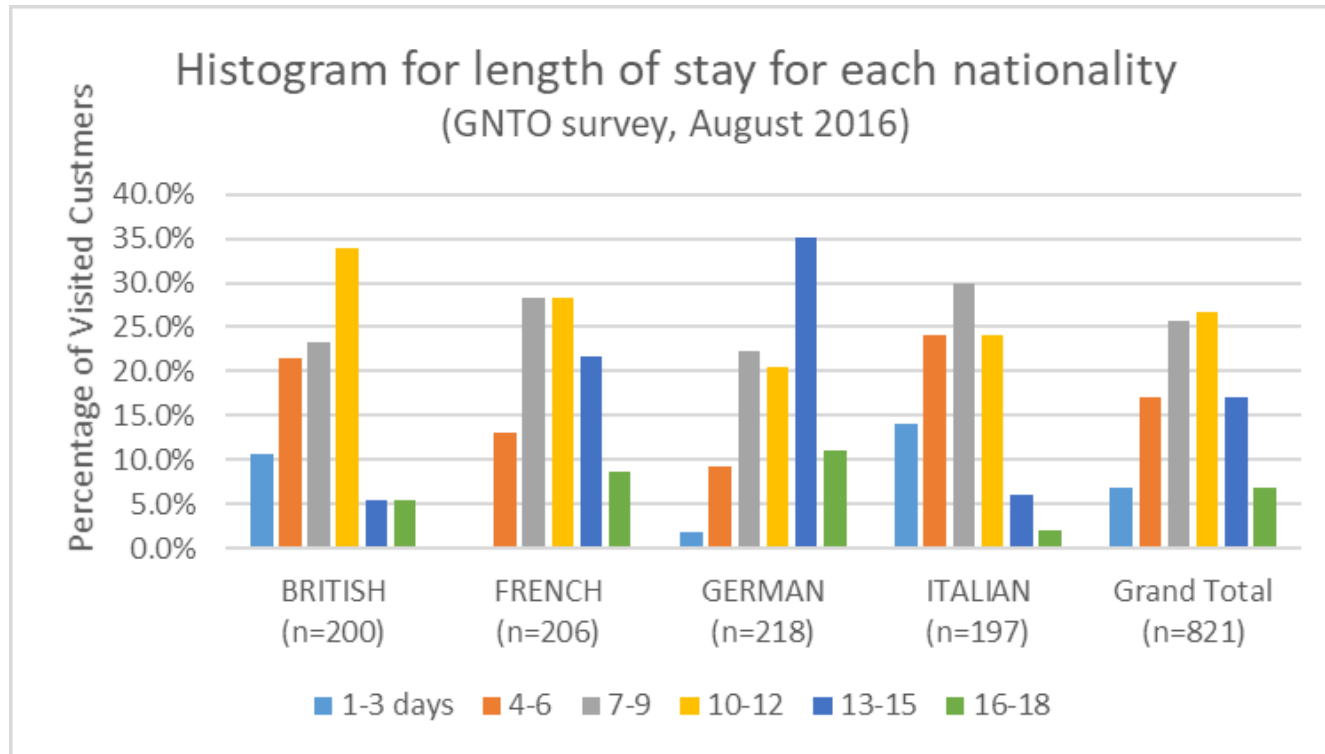

Customer spending per visit at Woollons' Supermarket
Source: EPOS data (11–17 February 2008)

Saunders *et al*. (2009)

Examples of diagrams (5)

Histograms

Histogram for length of stay for each nationality
(GNTO survey, August 2016)
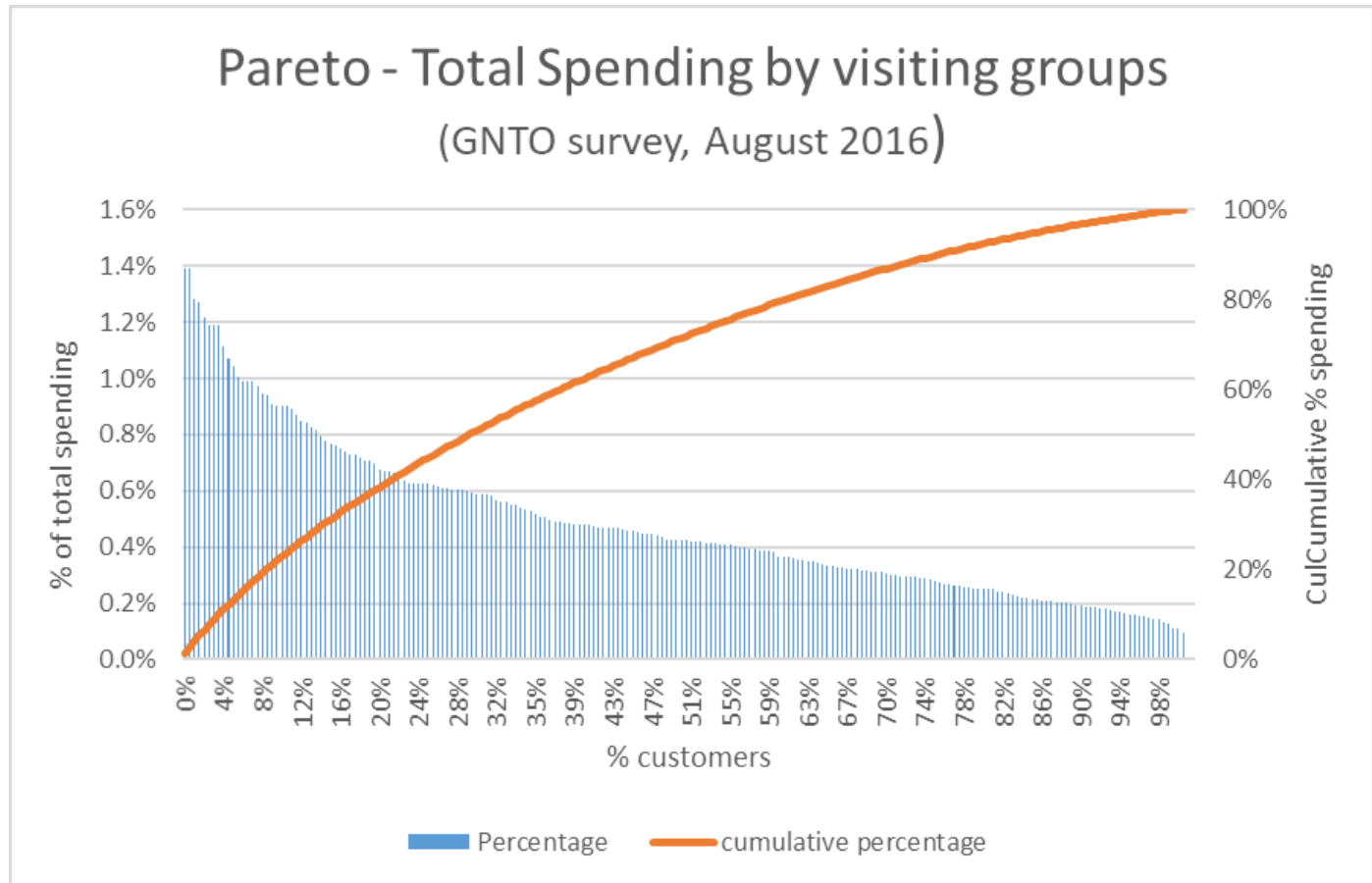
Examples of diagrams (5)

Pie chart
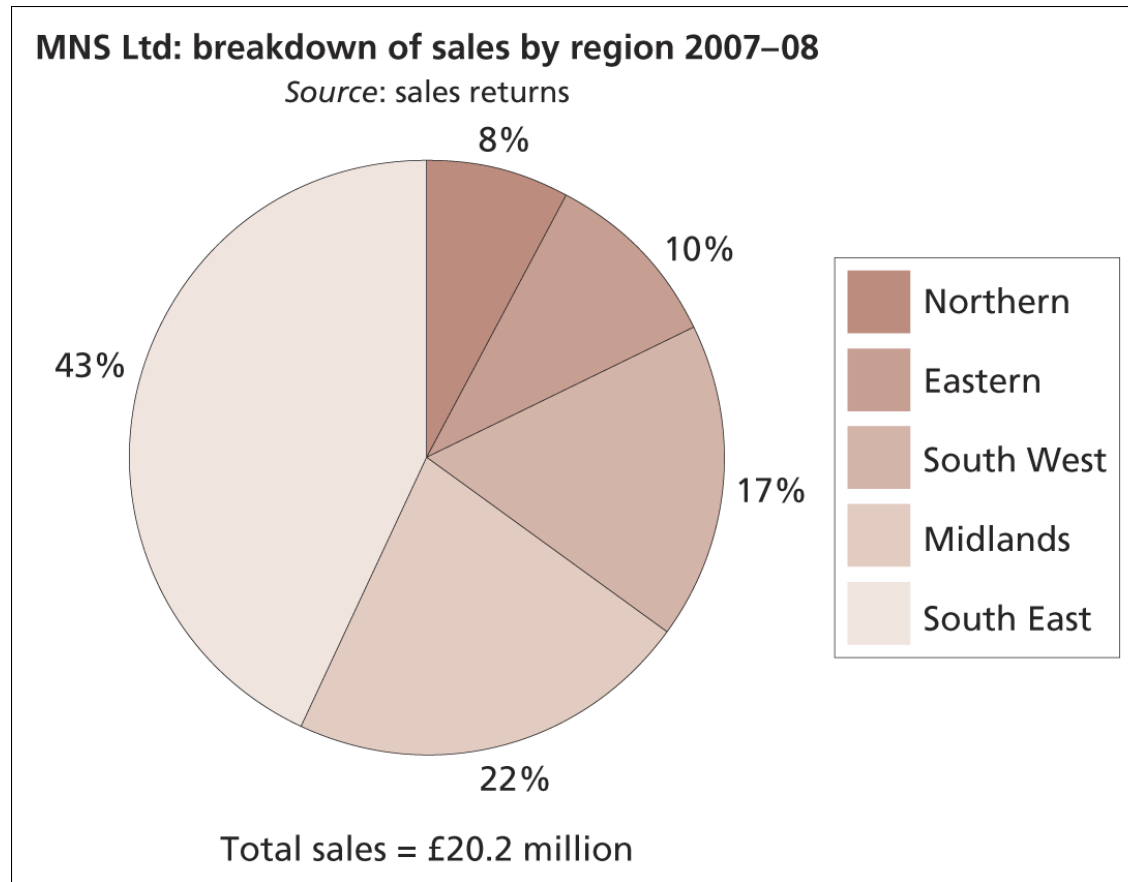
Saunders *et al*. (2009)

Examples of diagrams (5)

Pareto diagram

Examples of diagrams (5)

Box – Whisker plots
Identifying Skewness
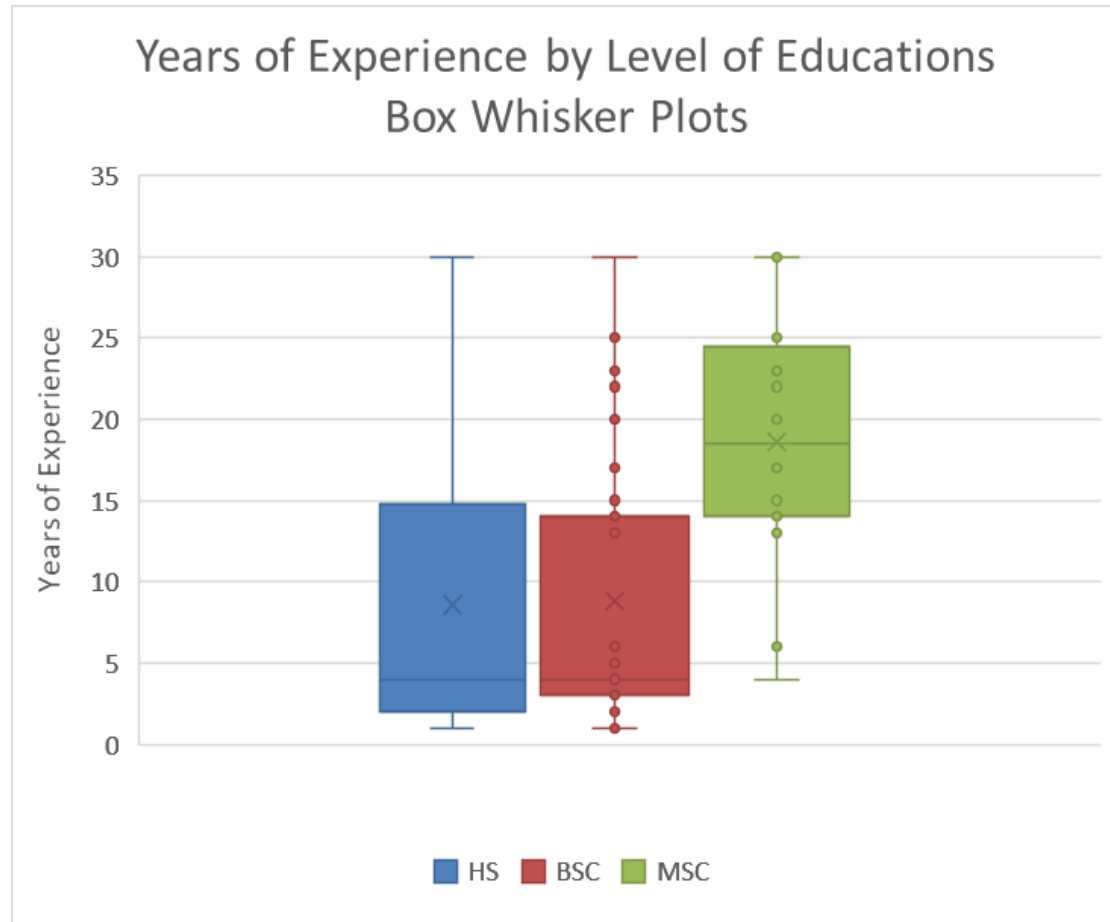
Exploring and presenting data (4)

Comparing variables to show

- Specific values and independence

- Highest and lowest values

- Proportions

- Trends and conjunctions

Exploring and presenting data (5)

Comparing variables to show

- Totals

- Proportions and totals

- Distribution of values

- Relationship between cases for variables

Describing data using statistics (1)

Statistics to describe a variable focus on two aspects

- The central tendency

- The dispersion

Describing data using statistics (2)

Describing the central tendency

- To represent the value occurring most frequently

  Mode: Proper for categorical data. Also for opinions on Likert scale

- To represent the middle value

  Median : Divides the sample in two equal parts (50%-50%). Not affected by extreme values. Personal income is a good example. Proper for numerical and ordinal data.

- To include all data values

  Mean: Average of sample values. Proper for numerical data

Describing data using statistics (3)

Describing the dispersion

- To state the difference between values
- To describe and compare the extent by which values differ from the mean

States the difference between highest and lowest	Range
States the difference between the middle 50% of values	Interquartile range
States the difference within another fraction of the values	Percentiles
Describes the extend to which the values differ from the mean	Variance, Standard deviation
Compares the extend to which data values differ from the mean between variables	Coefficient of variation
Allows the relative extend that different data values differ to be compared	Index numbers

Examining relationships, differences and trends

Using statistics to

- Test for significant relationships and differences

- Assess the strength of relationship

- Reduce the data

- Examine trends

Test for significant relationships and differences

- Comparing the data you selected with what you theoretically expect to happen (Hypothesis)
  - Helping to rule out the possibility that your results is just a matter of randomness in the data
  - Data must provide enough evidence, to reduce the level of error to an acceptable margin (usually 5%).
  - Test statistic, degrees of freedom, and p-value
  - Partly determined by sample size
  - If p-value > 0.05 then your results are not statistically significant. There still may be a relationship but you cannot be certain about it.

  Discuss significance of results based on the hypothesis tested the theoretical underpinnings and the research questions

Assess the strength of relationship

- The population correlation coefficient ρ (rho) measures the strength of the association between the variables

- The sample correlation coefficient r is an estimate of ρ and is used to measure the strength of the linear relationship in the sample observations

- Features
  - Unit free.   Range between -1 and 1
  - The closer to +1 or -1, the stronger the positive or negative linear relationship respectively
  - The closer to 0, the weaker the linear relationship

  It does not necessarily means causality

  Causality is established / explained based on theoretical underpinnings

What does correlation significance means?

Statical packages that compute correlation coefficients indicate their statistical significance at 5% and 1% level.

Values closer to ±1 tend to be significant.

Significance means that <u>95% of time the sample r is not zero</u>, therefore correlation does exist even if r is small..

It all depends on the value of r and the sample size

Smaller values of r can be significant as well.

Sample size (n)	Sample correlation coefficient (r)	t - statistic	t - critical	Significance of r
5	0.7	1.697749	3.182446	No
10	0.4	1.234427	2.306004	No
40	0.4	2.690371	2.024394	Yes
100	0.2	2.020726	1.984467	Yes

Data reduction Tools

Factor Analysis

- Removes redundancy or "duplication" from a set of correlated variables
  - Avoiding problems of collinearity

- Represents correlated variables with a smaller set of "derived" variables, called "factors".

- Factors are formed in a way that are relatively independent of one another.

- Two types of "variables":
  - latent variables: factors
  - observed variables

Factor Analysis. Why?

1. Identification of Underlying Factors:
    - clusters variables into homogeneous sets
    - creates new variables (i.e. factors)
    - allows us to gain insight to categories

2. Screening of Variables:
    - identifies groupings to allow us to select one variable to represent many
    - useful in regression (recall collinearity)

3. Summary:
    - Allows us to describe many variables using a few factors

4. Clustering of objects:
    - Helps us to put objects (people) into categories depending on their factor scores

Factor Analysis. Be cautious?

"Perhaps the most widely **used** (and **misused**) multivariate technique is factor analysis. Few statisticians are neutral about this technique. Proponents feel that factor analysis is the greatest invention since the double bed, while its detractors feel it is a useless procedure that can be used to support nearly any desired interpretation of the data. **The truth, as is usually the case, lies somewhere in between**. Used properly, factor analysis can yield much useful information; when applied blindly, without regard for its limitations, it is about as useful and informative as Tarot cards. **In particular, factor analysis can be used to explore the data for patterns, confirm our hypotheses, or reduce the Many variables to a more manageable number.**

-- Norman Streiner, *PDQ Statistics*

Statistics to examine relationships, differences and trends by data type

To test...	Categorical		Numerical Continuous & Discrete
	Descriptive	Ordinal	
whether two variables are associated	Chi Square * Gramer's V Phi coefficient (both variables dichotomous)		Chi Square (if variable grouped into discrete classes)
whether two groups (categories are different)		Kolmorogov-Smirnov * or Mahn-Whitney U test	Independent t-test or paired t-test or Mahn-Whitney U test (skewed data / small sample)
Whether three or more groups (categories are different)			Analysis of Variance (ANOVA)
Assess the strength of relationship between variables		Spearman's rank correlation coefficient or Kendall's rank order correlation coefficient	Pearson's product moment correlation coefficient

* data may need grouping

Statistics to examine relationships, differences and trends by data type

To test...	Numerical Continuous & Discrete
To assess the strength of a relationship between a dependent and one explanatory (independent) variable	Coefficient of determination (regression coefficient R-square)
To assess the strength of a relationship between one dependent and two or more explanatory (independent) variables For each explanatory variable	Coefficient of multiple determination (multiple regression coefficient, R-square) Regression F-test Significance of coefficient t-tests
To predict the value of a dependent from one or more explanatory variables	Regression equation
To examine and compare relative change (trend over time)	Index numbers
To determine the trend over time of a series of data	Time Series analysis: Moving averages, Seasonality indices, Regression analysis

Summary: Part 9

- Data for quantitative analysis can be collected and then coded at different scales of measurement
- Data type constrains the presentation, summary and analysis techniques that can be used
- Data are entered for computer analysis as a matrix and recorded using numerical codes
- Codes should be entered for all data values
- Existing coding schemes enable comparisons
- Data must be checked for errors
- Initial analysis should use both tables and diagrams
- Subsequent analyses involve describing data and exploring relationships by using statistics
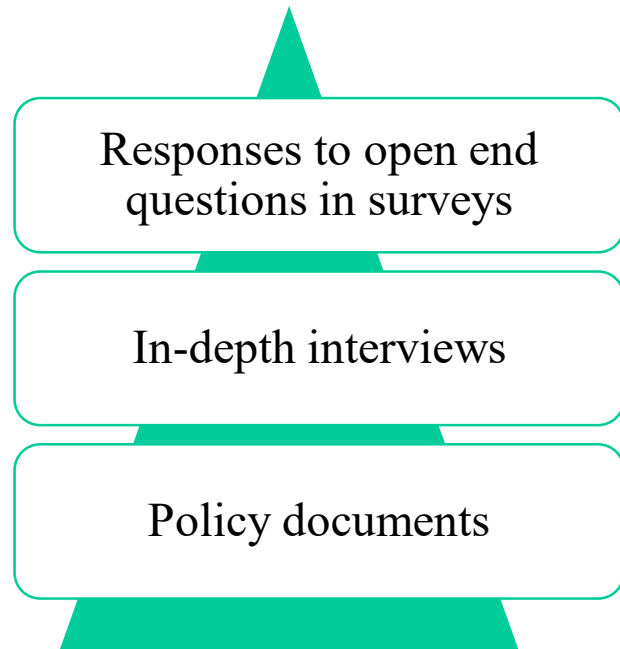- Longitudinal data may necessitate different statistical techniques

Part 10
Analysing qualitative data

Analysing qualitative data (1)

Definition

'Qualitative data refers to all non-numeric data or data that have not been quantified and can be a product of all research strategies'

Saunders *et al*. (2009)

Responses to open end questions in surveys

In-depth interviews

Policy documents

Analysing qualitative data (2)

Distinctions between quantitative and qualitative data

Quantitative data	Qualitative data
• Based on meanings derived from numbers	• Based on meanings expressed through words
• Collection results in numerical and standardised data	• Collection results in non-standardised data requiring classification into categories
• Analysis conducted through the use of diagrams and statistics	• Analysis conducted through the use of conceptualisation

Saunders *et al*. (2009)

Preparing data for analysis

Key issues (1)

- Transcribing qualitative data
    - Time consuming process. Record content, voice note and non-verbal communications
    - Could use voice recognition
    - Shortly after the interview
    - Sent transcript to interviewee ??
    - Keep each on separate files, Name properly.
    - Identify parts relevant to research
    - Group interviews – Need clear speaker identifiers
    - CAQDAS (Computer Aided Qualitative Data Analysis Software) requirements

Preparing data for analysis

Key issues (2)

- Using electronic textual data including scanned documents
    - Still need time to prepare them for analysis
    - Ensure anonymity
    - Appropriate storage for analysis
    - Free of typographical errors

- The interactive nature of the process
    - Recognize important themes, patterns, relationships as you go along. May re-categorize.
    - Could adjust future data collection

Approaches to qualitative analysis

Main approaches

- ## The deductive approach

  ❗ The prior specification of a theory tends to be disfavoured because of the possibilities of introducing a premature closure on the issues to be investigated, as well as the possibility of the theoretical constructs departing excessively from the views of participants in a social setting. *Bryman (1988)*

- ## The inductive approach

  – Difficult for inexperienced researchers

  – Analyse the data as you collect them to develop conceptual framework for subsequent work (grounded approach)

  – Do not start with a clearly defined theoretical framework

 – Theory emerges from the process of data collection and analysis

Types of qualitative analysis process

- Summarising (condensation) of meanings
  - Summary of key points of interview transcripts. Include other observation (contextual).
  - Summarize other documents (minutes, policy, planning, etc.) for triangulation purposes

- Categorising (grouping) of meanings
  - Terms emerged from the data
  - Actual terms used by participants ("in vivo")
  - Derived for existing theory,/ literature
  - "Unitising" data (attach units of data to appropriate category(ies). CAQDAS may help in that

- Structuring (ordering of meanings using narrative)

Saunders *et al*. (2009)

Categorising (grouping) of meanings - example

- How HRM professional manage downsizing processes
- *Categories derived from existing theory*

Saunders *et al*. (2009)
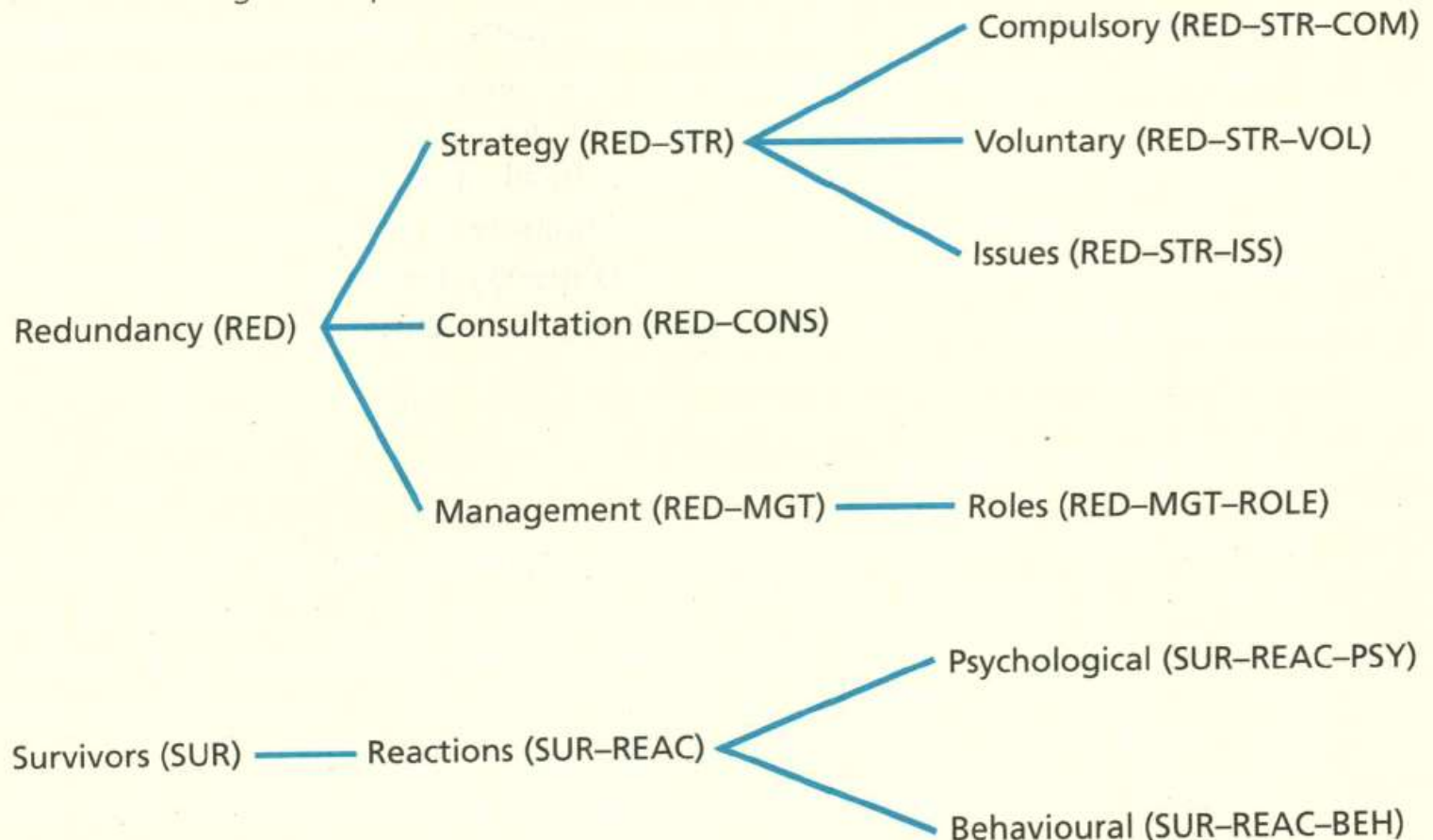
Categorising (grouping) of meanings - example

- *Categories attached to interview transcript*

RED–CONS	27MM The first stage is to find out what particular employees want for themselves and how they want this to happen. Staff are
RED–CONS	seen by their line manager and/or a member of personnel.
RED–MGT–ROLE	Employees might want to talk to someone from personnel rather than talk with their line manager – well, you know, for obvious reasons, at least as they see it – and this would be acceptable to the
RED–MGT–VOL	organisation. This meeting provides them with the opportunity to opt for voluntary redundancy. We do not categorise employees
RED–STR–ISS	into anything like core or non-core, although we will tell a group
RED–CONS	of employees something like 'there are four of you in this particular function and we only need two of you, so you think
RED–CONS	about what should happen'. Sometimes when we attempt to give employees a choice about who might leave, they actually ask us to make the choice. This is one such situation where a compulsory
RED–MGT–COM	selection will occur. We prefer to avoid this compulsory selection
SUR–REAC–PSY	because of the impact on those who survive – negative feelings, guilt and so on.

Saunders *et al.* (2009)

Categorising data

Points to consider

- Deriving categories

- 'Unitising' data

- Recognising relationships and developing categories

- Developing testable propositions

- Quantifying your qualitative data
  - Frequency of certain events / reasons / references to a phenomenon
  - Frequency of appearance of words with positive vs. negative meaning in the interview transcripts or texts used.

Analytical aids

Types of analytical aids

Contextual information

- Interim summaries: *what was found so far/ confidence in findings / conclusions / what to do next to improve quality of data*

- Self-memos: *During writing of observation notes, transcript production, categorizing data, report writing*

- Researcher's diary

Deductively based analytical procedures

Procedures applicable to qualitative analysis

- ## Pattern matching
  - Predict a pattern of outcomes based on theoretical propositions to explain what you expect to find.
  - Set of dependent variables associated with an independent one (*e.g. outcomes expected as a result of organizational change*)
  - Alternative explanations to explain the pattern of an outcome

- ## Explanation building
  - Test a theoretical proposition in an iterative way (similar to ground theory, but focuses in testing not generating)

Inductively based analytical procedures

Procedures applicable to qualitative analysis

- Data display and analysis

- Template analysis

- Analytic induction

- Grounded theory – open, axial and selected coding

- Discourse analysis *(how language is used to construct and change aspects of the world)*

- Narrative analysis

Using CAQDAS for qualitative analysis

Summary of functions

- Structure of work *store and provide connections between text files*

- Closeness to data and interactivity *instantaneous access to data*

- Explore the data *powerful text search*

- Code and retrieve *use inductive, deductive, combination schema to code, retrieve and organize data*

- Project management and data organisation *manage the research project as a whole, organize the data, could focus on subsets of data*

- Searching and interrogating *automatically coding data on basis of relationships between codes for different units of data*

- Writing memos, comments and notes

- Output *view material in hard copy, export to other apps*

Do you need CAQDAS ?

- How much data you have that need to be analysed quantitatively

- How important relative to other data

- Time you have to learn how to use the package

- Training support from the University

- Specifications of your computer (OS, memory, etc.)

- Time to undertake your analysis

- Do you need a s/w that

  - Allow inductive and deductive approach

  - Help you manage your thinking and develop codes

  - Explore the way language is used in the data

  - Display relationships diagrammatically

  - Allow quantification of your data

Summary: Part 10

- Qualitative data result from the collection of non-standardised data that require classification and are analysed through use of conceptualisation

- Qualitative analysis can involve summarising, categorising and structuring data

- The process of data analysis and collection are necessarily interactive

- Aids to analysis include interim summaries, self-memos and maintaining a researcher diary

- Qualitative analysis procedures can be related to using either a deductive or inductive approach

- Computer-assisted qualitative data analysis software (CAQDAS) can help with project management and data organisation