

Research Proposal and Ethics Workshop

Lecture 9 Data Collection

Dr Andre Samuel

andre.samuel@sam.edu.tt

<http://www.samuellearning.org>

Types of Research Design

POSITIVISTIC

PHENOMENOLOGICAL

Approach to Social Science



- CROSS SECTIONAL STUDIES
- EXPERIMENTAL STUDIES
- LONGITUDINAL STUDIES
- SURVEYS

- ACTION RESEARCH
- CASE STUDIES
- ETHNOGRAPHY
- GROUNDED THEORY
- HERMENEUTICS
- PARTICIPATIVE ENQUIRY

Source: Collis and Hussey (2003, pg.60)

Collecting Secondary Data

- This is data which already exists such books and documents
- One can think of their Literature Review as largely based on secondary data
- So there's a possibility of answering a part of your RQ or achieving one of your research objective through the use of Secondary Data
- Of course, it would involve the **reanalyzing** of the data that already exist, to relate to your research

- Secondary data include both quantitative and qualitative data
- The data you find may have been subjected to little or no processing, i.e. it is still largely raw data or
- The data you find may have been subjected to detailed and lengthy processing, i.e. compiled or published data
- It is important to remember that these data were previously collected for a different purpose

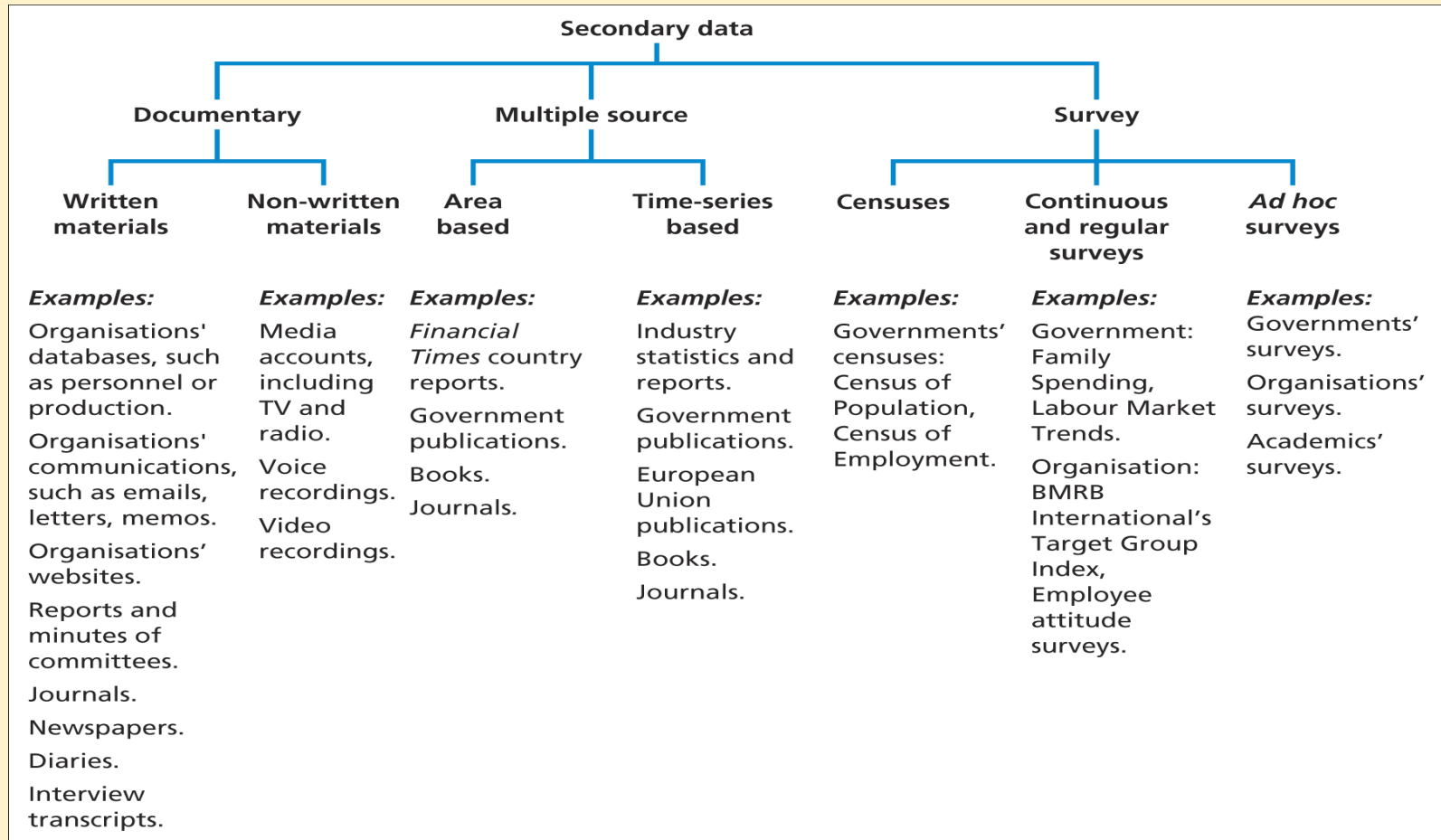
- So be careful in the use of secondary data, as its validity could be questionable
 - Is the data **applicable** or **suitable** to your research?
 - Is it a reliable source?
- Never take secondary data at face value!!!
- It should also be noted that where it is possible one should get to the '*primary or original source of the secondary data*'
- Do not rely on a secondary account of the secondary data

- **For example;** if I wanted to explore the experience of children in regard to absent fathers during Wars
- Then it might be best to accumulate hundreds perhaps thousands of letters written by the children, who are now adults ,to their fathers
- This would be more reliable and suitable than say using the account of a newspaper column during the War

- As another example, at this level we do not want you to use an authors work if you have not directly read it.
- That is, **do not** do this:
 - According to Smith (2003) as cited by King (2006, pg. 67)
- Get to the source of the work, do not rely on what King is saying about what Smith has said.
- You need to read Smith's work directly
- This must be noted when you are doing your Literature Review

Types of Secondary Data

Saunders et al (2009, pg. 259)

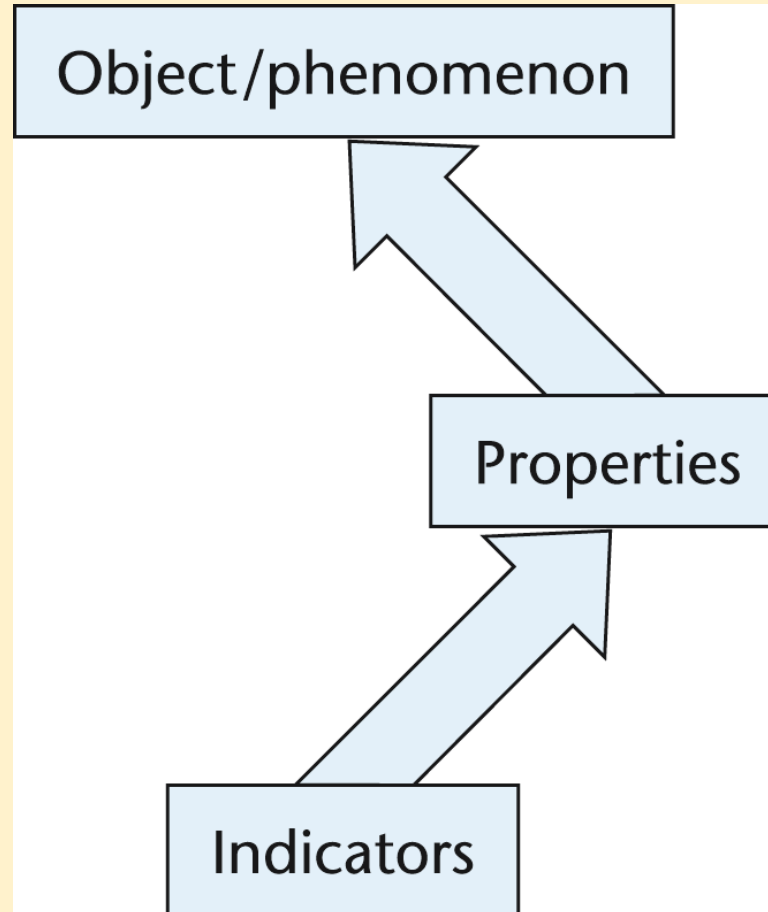


Collecting Primary Data

- Of course Primary data is original data, new data collected for the purpose of your research by **YOU**
- There are many methods:
 - Diaries
 - Focus groups
 - Interviews
 - Observation
 - Questionnaires

- In research we are interested about collecting data about **variables**
- A variable is an attribute of the entity which you have chosen as your unit of analysis
- For example, age and qualifications are variables of individuals.
- Number of employees and profit margin are attributes of an organization

Getting your Idea for Action



- Variables can be classified as qualitative or quantitative
- A **qualitative variable** is non numerical attribute of an object
- For example variables like gender or colour are qualitative variables of an individual
- A **quantitative variable** is a numerical attribute of an object

- In order to obtain a quantitative measure you need to use a suitable **measuring tool**
- Some attributes like age, income, height or even weight have accepted and known measures i.e. years, \$, cm, kg
- But for some attributes there will be difficulty, such as honesty, loyalty or intelligence
- If there are no generally accepted measure, you must devise your own or find out what other research have used

Example

- Lets say I wanted to find out the level of employee commitment,
- How do I measure commitment to employment?
- I can devise a series of questions (indicators) and ask the respondent to specify their level of agreement or disagreement on a five point scale (Likert Scale)

To what extent, do you agree with the following statement.

I believe.....

	Strongly Agree	Agree	Neutral	Strongly Disagree	Disagree
Work is necessary					
Having a job is not important to me					
I would get bored if I don't go to work					
etc					

Levels of Measurement-

types of Variable

- **Nominal Scale-** describes variables that are categorical in nature and that differ in quality rather than quantity
 - So your observations can be placed into only one category
 - You are only able to either classify data, which allow you to make statements of equality or differences
 - For example, variable 'occupation' you can count how many directors, managers, administrators etc
 - Very few statistics can be applied

- **Ordinal Scale**- describes variables that can be ordered along some type of continuum
 - Not only can these values be placed in categories, but they can be **ordered** as well
 - Often referred to as ranking of various outcomes
 - It allows you therefore to make decisions on whether a score is greater than or less than
 - For example, your GPA might determine your ranking, you can be 1st of 30 or 15th of 30
 - Notice this does not tell us anything absolute about the GPA score from the ranking but only the position relative to others

- **Interval Scale-** describes variables that have equal intervals between them
 - Interval level variables allow us to determine the difference between point along some type of continuum
 - So it has the characteristics of both nominal and ordinal scales
 - Thus if you have an interval scale you can place each data item along the scale and determine exactly what the intervals are
 - Put simply you can tell the difference between points along a continuum

Sampling

- Probability Sampling
 - Random
 - Systematic
 - Stratified
 - Cluster
- A good sample must be:
 - Chosen at random, i.e. every member of the population has an equal chance of being selected
 - Not too small but not too large
 - unbiased

Random Sampling

- Each person or item has an equal opportunity of inclusion in the sample
- Provided that SAM had 1200 student and we decided to sample 300 students
- This means the probability of inclusion in the sample is $300/1200 = 0.25$ i.e 1 in 4
- This is called the **sampling fraction**

Key steps in Random Sampling

- Define the population
- Select sampling frame
- Decide your sample size
- List all students in the population and assign consecutive number 1 to N i.e. 1 to 1200
- Use a table of random numbers select the sample size from the list

Systematic Sampling

- A variation of random sampling
- Here you select directly from the sampling frame
- Use the sampling fraction, we know that we are to select 1 student in 4.
- Make a random start, thereafter take every fourth student on the list

Stratified Random Sampling

- Sampling to exhibit a proportional representation
- In the example, we might want to sample by different degrees for instance HR, Mkting, Mgt, CIS
- Thus if there are 400 HR students, using our sampling fraction of 1 in 4, the sample would be 100

Stratified Sample Solution

DEGREE	POPULATION	PROPORTION	STRATIFIED SAMPLE
HR	400	0.33*	100
MKTING	90	0.075	22
MGT	510	0.425	128
CIS	200	0.166*	50
TOTAL	1200		300

Cluster Sampling

- Imagine we wanted a national sample of students
- Clusters are groupings or aggregations of population to be sampled
- So say we want 5000 students
- We could sample from different tertiary institutes
- Now we could use random or systematic sampling to select say ten institute

- This would give us a list of ten clusters
- Then we could randomly, systematically or employ stratified sampling to select 500 students from each cluster

- **Another approach** is to divide the country into 2 regions
- Select 5 institute from each region
- Then sample 500 from each of the ten institute

Non-Probability Sampling

- Convenience Sampling
- Snowball Sampling
- Quota Sampling

Sample Size Calculator

- Qualtrics:

<https://www.qualtrics.com/blog/calculating-sample-size/>

Response rate

- The percentage of the sample that does in fact agree to participate and provides a usable response
- Response rate = $\frac{\text{number of usable questionnaires}}{\text{total sample} - \text{unsuitable sample}} \times 100$

Data Collection Methods

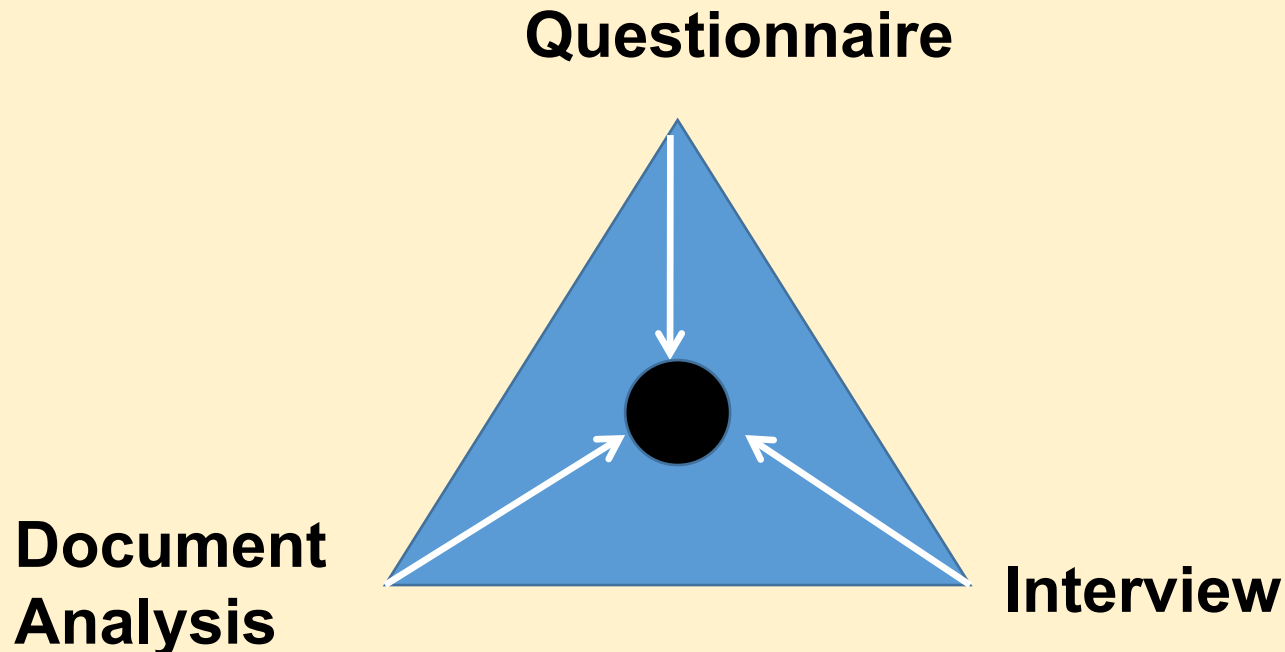
- Whether you are following a broadly Positivist or Phenomenological Paradigm,
- There will always be a combination of Quantitative and Qualitative inputs into your data generating activities
- Each approach presents a mixture of advantages and disadvantages

- Quantitative approach allows analytical and predictive power through the use of Statistical Analysis. However suffers from reductionist tendencies
- Qualitative provides a more 'real' basis for analysis and interpretation of phenomena, but expensive and time consuming
- It must be noted that the methods you utilize to collect your data can be either quantitative or qualitative
- **So you can use Mixed Methods!!**

Concept of Triangulation

- Triangulation is a way of assuring the validity of research results through the use of a variety of research methods or techniques
- It is a means of overcoming the weaknesses and biases which can arise from the use of only one of the methods such as observation, questionnaires etc.
- For example, a researcher might choose to begin their research with an unstructured interview

- This will allow them to identify key issues and which they can then use as a basis for more formal interviews and questionnaires
- Think of it as a Surveyor trying to find a point, they will locate the point based on three views



- Triangulation also allows researchers to collect both quantitative and qualitative data from both primary and secondary sources
- For Example:
 - in a study of the long term effects on victims of crime,
 - it would be possible to use both questionnaires and interviews to assess the effects of crime on victims
 - as well as investigating eye witness accounts in newspapers or reports of trials

Triangulation Defined

- “the use of more than one method or source of data in the study of a phenomenon so that findings may be cross-checked”

Bryman (2008, pg. 700)

Questionnaires

- Questionnaires are associated with both positivistic and phenomenological methodologies
- It is a list of carefully listed questions, with the view of eliciting reliable responses from a chosen sample
- The aim is to find out what a selected group of participants do, think or feel

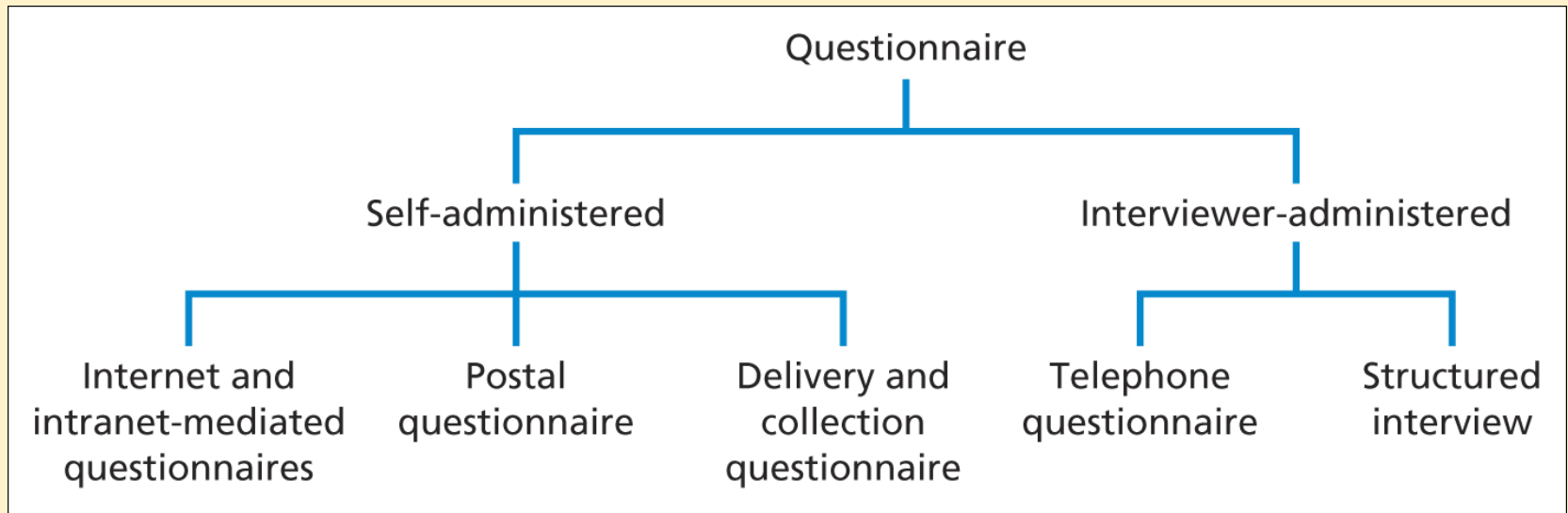
Main Decisions when using a Questionnaire

- Sample size
- Type of questions
- Wording questions and how to ensure that they are intelligible and unambiguous
- Design of questionnaire, including instructions

- Wording of any accompanying Participant Information Sheet (PIS) OR cover letter
- Method of distribution and return of completed questionnaires
- Tests for validity and reliability
- Methods for collating and analyzing the data

Types of Questionnaires

Saunders et al (2009, pg. 363)



Type of Questions

Open questions

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

1.....

2.....

3.....

List questions

7 What is your religion?

Please tick ✓ the appropriate box

Buddhist

None

Christian

Other

Hindu

Jewish

Muslim

Sikh

Category questions

8 How often do you visit the shopping centre?

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

- 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

Saunders *et al.* (2009)

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.

<i>Factor</i>	<i>Importance</i>
<i>Carbon dioxide emissions</i>	[]
<i>Boot size</i>	[]
<i>Depreciation</i>	[]
<i>Price</i>	[]

Adapted from Saunders *et al.* (2009)

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)

Method of Distribution

- By post
- By telephone
- Face to Face
- Group Distribution
- Individual Distribution
- Online Distribution
 - Emails
 - Facebook

Interviews

- It is a method of collecting data in which selected participants are asked questions
- Interviews make it easy to compare answers
- Issues to consider:
 - Confidentiality
 - Bias of interviewer
 - Recording mechanism
 - Organizing the interview
 - Developing interview themes

- A positivistic approach suggest closed questions, which have been prepared beforehand
- A phenomenological approach suggest unstructured questions, where the questions have not been prepared beforehand

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

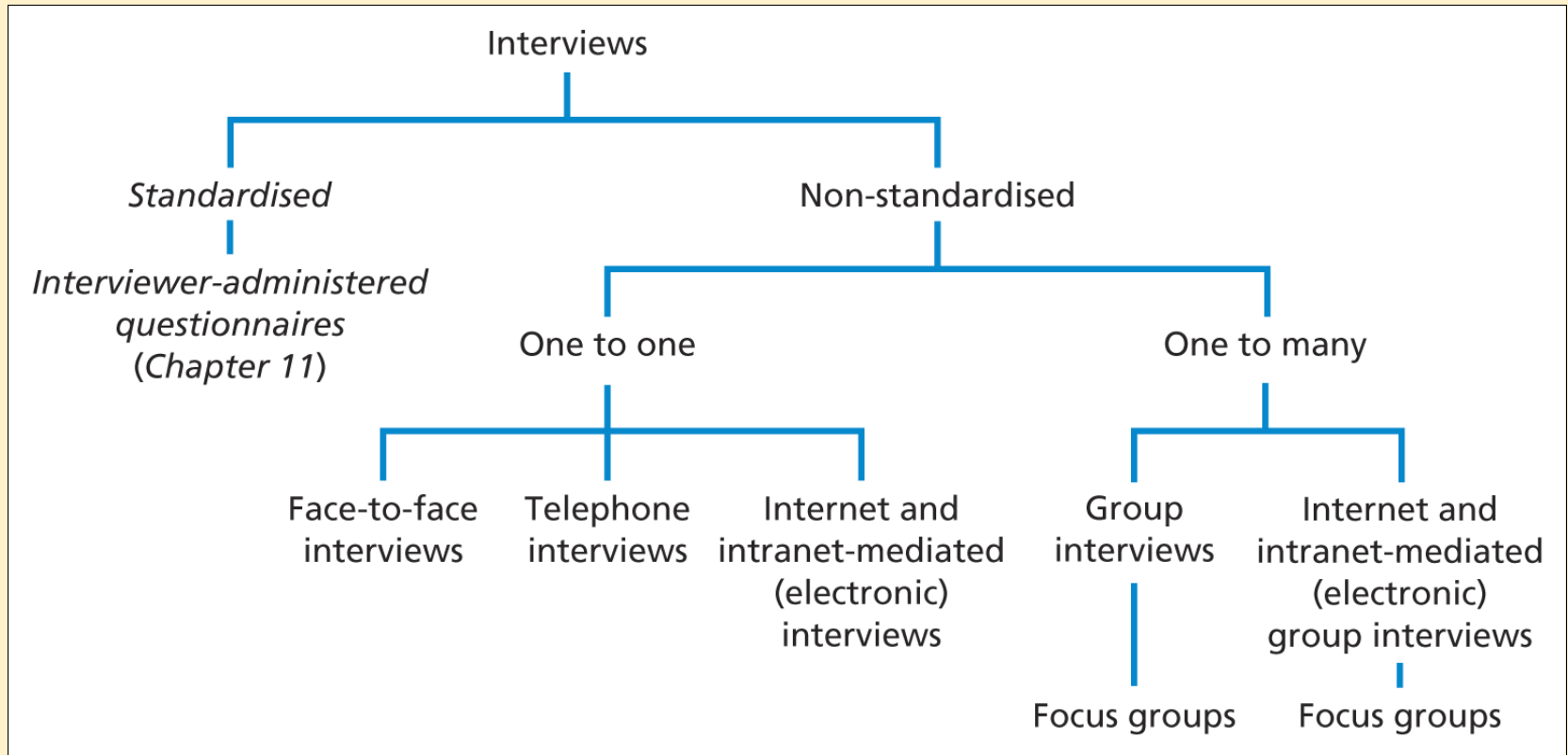
Saunders et (2009, pg. 323)

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

✓✓ = more frequent, ✓ = less frequent.

Forms of Interviews

Saunders et al (2009, pg. 321)



Bibliography

- Collis. J., Hussey R (2003) Business Research 2nd edition, Palgarve Macmillan
- Saunders, M., Lewis, P., Thornhill, A. (2009) Research Methods for Business Students, 5th edition, Prentice Hall
- Creswell. J.W., (2009) Research Design, Qualitative, Quantitative and Mixed Method Approaches, 3rd Edition, Sage Publications Inc.
- Salkind, N. (2003) Exploring Research, 5th Edition, Prentice Hall
- Cooper, D.R., Schindler, P.S. (2006) Business Research Methods, McGraw-Hill