

# **SAMPLING TECHNIQUES**

ASSOC. PROF. DR. MOHD ROSNI SULAIMAN  
FACULTY OF FOOD SCIENCE AND NUTRITION  
UNIVERSITI MALAYSIA SABAH

# INTRODUCTION

- **Why do scientists need to know about:**
  - **1) Experimental design?**
  - **2) Sampling techniques ?**
  - **3) Statistics?**

# PROBLEMS

- Imagine: to measure >>
  - the length of every anchovy in the South China Sea
  - the haemoglobin count of every adult in Malaysia
  - the diameter of every mangosteen tree in a plantation of 100 000
  - the individual protein content of 10 000 prawns in a large aquaculture pond.

# PROBLEMS

- Work on living things, impossible to get data from every individual of the group or species in question.
- If total number of individual in a population of study (e.g. anchovies) is too big, approx. 5 millions...
- How much money, time, effort, etc. etc.. needed to accomplish the study??
- MSc >> 2-3 years; PhD >> 3-5 years

# SOLUTIONS

- To be a scientist, each of us must know:
  - 1) How to properly design an experiment
  - 2) How to choose and carry out a correct **sampling method**
  - 3) How to choose and use a right statistical analysis
- By knowing all of these 3 main components of a research, we solved more than 70% of our problems.

# SAMPLING TECHNIQUES

- In this lecture, our focus only on sampling
- To really understand and mastery various techniques of sampling >> impossible to be achieved in just a lecture or in one semester course
- It is through a lifetime practice as a scientist
- But it is possible if just only one sampling technique
- Therefore, the aim of this lecture is to ensure each of us understand and mastery at least a type of sampling technique.

# SAMPLING TECHNIQUES

- Why doing sampling??
- As the reasons been mentioned before +
- Because we want to use sample to **represent** population.
- In other word, we use sample to **estimate** the population.

# SAMPLING TECHNIQUES

- We can say that there are three types of sampling:
  - 1) **Probability sampling**: it is the one in which each sample has the same probability of being chosen.
  - 2) **Non-probability sampling**: do not follow the theory of probability in the choice of elements from the sampling population
  - 3) **'Mixed' sampling**

# SAMPLING TECHNIQUES

- We will always make probability sampling
- Because it assures us that the sample is representative and
- We can estimate the errors for the sampling.
- There are different types of probability sampling.

# PROBABILITY SAMPLING

- Random sampling with and without replacement.
- Stratified sampling.
- Cluster sampling.
- Systematic sampling.
- Other types of sampling techniques.

# SIMPLE RANDOM SAMPLING

- Where we select a group of subjects (a sample) for study from a larger group (a population).
- Each individual is chosen randomly and each member of the population has an equal chance of being included in the sample.
- A lucky draw for six hampers in a UMS family day (e.g. 2500 staff attended) is a good example of simple random sampling.
- A sample of 6 numbers is randomly drew from a population of 2500, with each number having an equal chance of being selected.

# SIMPLE RANDOM SAMPLING

- Methods of drawing a random sample:
  - **1) The fishbowl draw** (total population is small)
  - **2) Computer programs**
  - **3) A table of random numbers**

# SIMPLE RANDOM SAMPLING

1	Albert D.	25	Monique Q.
2	Richard D.	26	Régine D.
3	Belle H.	27	Lucille L.
4	Raymond L.	28	Jérémy W.
5	Stéphane B.	29	Gilles D.
6	Albert T.	30	Renaud S.
7	Jean William V.	31	Pierre K.
8	André D.	32	Etienne M.
9	Jeremy W.	33	Marie M.
10	Anthony Q.	34	Gaétan Z.
11	James B.	35	Fidèle D.
12	Denis G.	36	Maria P.
13	Amanda L.	37	Anne Marie G.
14	Jennifer L.	38	Michel K.
15	Philippe K.	39	Gaston C.
16	Eve F	40	Alain M.
17	Priscilla O.	41	Olivier P.
18	Robert D	42	Geneviève M.
19	Brian F.	43	Berthe D.
20	Hellène H.	44	Jean Pierre P.
21	Isabelle R.	45	Jacques B.
22	Jean T.	46	François P.
23	Samanta D.	47	Dominique M.
24	Berthe L.	48	Antoine C.

# SIMPLE RANDOM SAMPLING

Table 12.3 Selecting a sample using a table for random numbers

	1	2	3	4	5	6	7	8	9	10
1	48461	14952	72619	73689	52059	37086	60050	86192	67 <del>049</del>	64739
2	76534	38 <del>749</del>	49692	31366	52093	15422	20498	33901	10 <del>379</del>	43397
3	70437	25 <del>861</del>	38504	14752	23757	29660	67844	78815	23 <del>758</del>	86814
4	59584	03370	42806	11393	71722	93804	09095	07856	55 <del>589</del>	50063
5	04285	58554	16085	51555	27501	73883	33427	33343	45 <del>507</del>	50 <del>063</del>
6	77340	10412	69189	85171	29802	44785	86368	02583	96 <del>483</del>	76553
7	59183	62687	91778	80354	23512	97219	65921	02035	59 <del>487</del>	91403
8	91800	04281	39979	03927	82564	28777	59049	97532	54 <del>540</del>	79472
9	12066	24817	81099	48940	69554	55925	48379	12866	41 <del>233</del>	21580
10	69907	91751	53512	23748	65906	91385	84983	27915	48 <del>491</del>	91 <del>068</del>
11	80467	04873	54053	25955	48518	13815	37707	68687	15 <del>570</del>	08890
12	78057	67835	28302	45048	56761	97725	58438	91529	24 <del>645</del>	18544
13	05648	39387	78191	88415	60269	94880	58812	42931	71 <del>898</del>	61534
14	22304	39 <del>246</del>	01350	99451	61862	78688	30339	60222	74 <del>052</del>	25740
15	61346	50269	67005	40442	33100	16742	61640	21046	31 <del>909</del>	72641
16	56793	37696	27965	30459	91011	51426	31006	77468	61 <del>029</del>	57 <del>008</del>
17	56411	48609	36698	42453	85061	43769	39948	87031	30767	13953
18	62098	12825	81744	28882	27369	88185	65846	92545	09 <del>063</del>	22653
19	68775	06261	54265	16203	23340	84750	16317	88686	86 <del>842</del>	00879
20	52679	19599	13687	74872	89181	01939	18447	10787	76 <del>246</del>	80 <del>072</del>
21	84096	87152	20719	25215	04349	54434	72344	93008	83 <del>282</del>	31670
22	83964	55937	21417	49944	38356	98404	14850	17994	17 <del>161</del>	98981
23	31191	75131	72386	11689	95727	05414	88727	45583	22 <del>568</del>	77700
24	30545	68523	29850	67833	05622	89975	79042	27142	99 <del>257</del>	32349
25	52573	91001	52315	26430	54175	30122	31796	98842	37 <del>600</del>	26 <del>023</del>
26	16586	81842	01076	99414	31574	94719	34656	80018	86988	79 <del>234</del>
27	81841	88481	61191	25013	30272	23388	22463	65774	10 <del>029</del>	58376
28	43563	66829	72838	08074	57080	15446	11034	98143	74 <del>989</del>	26885
29	19945	84193	57581	77252	85604	45412	43556	27518	90 <del>572</del>	00563
30	79374	23796	16919	99691	80276	32818	62953	78831	54 <del>395</del>	30705
31	48503	26615	43980	09810	38289	66679	73799	48418	12 <del>647</del>	40 <del>044</del>
32	32049	65541	37937	41105	70106	89706	40829	40789	59 <del>547</del>	40783
33	18547	71562	95493	34112	76895	46766	96395	31718	48 <del>302</del>	45893
34	03180	96742	61486	43305	84183	99605	67803	13491	09 <del>243</del>	29557
35	94822	24738	67749	83748	59799	25210	31093	62925	72 <del>061</del>	69991
36	4330	60599	85828	19152	68499	27977	35611	96240	62 <del>747</del>	89529
37	43770	81537	59527	95674	76692	86420	69930	10020	72 <del>881</del>	12532
38	56908	77192	50623	41215	14311	42834	80651	93750	59 <del>957</del>	31 <del>211</del>
39	32787	07189	80539	75927	75475	73965	11796	72140	48 <del>944</del>	74 <del>156</del>
40	52441	78392	11733	57703	29133	71164	55355	31006	25 <del>526</del>	55790
41	22377	54723	18227	28449	04570	18882	00023	67101	06895	08915
42	18376	73460	88841	39602	34049	20589	05701	08249	74 <del>213</del>	25 <del>220</del>
43	53201	28610	87957	21497	64729	64983	71551	99016	87 <del>903</del>	63875
44	34919	78801	59710	27396	02593	05665	11964	44134	00 <del>273</del>	76358
45	33617	92159	21971	16901	57383	34262	41744	60891	57 <del>624</del>	06962
46	70010	40964	98780	72418	52571	18415	64362	90637	38 <del>034</del>	04909
47	19282	68447	35665	31530	59838	49181	21914	65742	89 <del>875</del>	39 <del>231</del>
48	91429	73328	13266	54898	68795	40948	80808	63887	89 <del>939</del>	47938
49	97637	78393	33021	05867	86520	45363	43066	00988	64 <del>040</del>	09803
50	95150	07625	05255	83254	93943	52325	93230	62668	79 <del>529</del>	66964

Source: *Statistical Tables* by Rohlf and Sokal. Copyright © 1969 by W.H. Freeman and Company. Used with permission.

# SIMPLE RANDOM SAMPLING

Table 12.4 Selected elements using the table of random numbers

	<i>Elements selected</i>				
#	49	232	52	29	65
	246	161	243	61	213
	34	40			
*	63	68	108	72	25
	234	44	211	156	220
	231 (10th column)				
@	149	246			

# sample selected from the 9th column

\* sample selected from the 10th column

@ sample selected from the 2nd column

25 samples are selected from sampling population of 256 individuals

# STRATIFIED RANDOM SAMPLING

- Often factors which divide up the population into sub-populations (groups / strata)
- Measurement of interest may vary among the different sub-populations.
- This has to be accounted for when we select a sample from the population to ensure our sample is representative of the population.
- This is achieved by stratified sampling

# STRATIFIED RANDOM SAMPLING

- A stratified sample is obtained by taking samples from each stratum or sub-group of a population.
- Suppose a farmer wishes to work out the **average milk yield** of each **cow type** in his herd which consists of **Ayrshire, Friesian, Galloway** and **Jersey** cows.
- He could **divide** up his herd into the **four sub-groups** and **take samples** from these

# STRATIFIED RANDOM SAMPLING

- Divided into 2 types:
  - **1) Proportionate STRS**
  - **2) Disproportionate STRS**

# STRATIFIED RANDOM SAMPLING

- In the case of **Proportionate STRS**
  - Determine the proportion of each stratum in the study population
  - $p = \frac{\text{elements (\#) in each stratum}}{\text{total pop. size}}$
- Determine the **number of elements to be selected** from each stratum =  $(n) \times (p)$
- Select the required number of elements from each stratum with **SRS** technique.

# STRATIFIED RANDOM SAMPLING

- Say, sample size (n) required is 30% of N which equivalent to 266 cows.
- Ayr = 215 ; Fr = 223; Gal = 217; Jer = 230
- Total pop. size (N) = Ayr + Fr + Gal + Jer = **885**
- $p_{\text{Ayr}} = 215/885 = \underline{\mathbf{0.24}}$ ;  $p_{\text{Fr}} = 223/885 = \underline{\mathbf{0.25}}$ ;  
 $p_{\text{Gal}} = 217/885 = \underline{\mathbf{0.25}}$ ;  $p_{\text{Jer}} = 230/885 = \underline{\mathbf{0.26}}$
- Required number of each type of cow:
  - Ayr =  $266 \times 0.24 = \underline{\mathbf{64}}$ ; Fr =  $266 \times 0.25 = \underline{\mathbf{67}}$ ;
  - Gal =  $266 \times 0.25 = \underline{\mathbf{67}}$ ; Jer =  $266 \times 0.26 = \underline{\mathbf{69}}$

# STRATIFIED RANDOM SAMPLING

- In the case of **Disproportionate STRS**
  - Determine the **number of element to be selected** from each stratum = **Sample size (n)**  
**No. of strata (k)**  
**=  $266/4 = \underline{66 \text{ or } 67}$**
  - Select the required number of elements from each stratum with SRS technique i.e. a total number of 66 or 67 from each type of cow are taken at random in order to achieve  $n=266$ .

# SYSTEMATIC RANDOM SAMPLING

- Systematic sampling, sometimes called interval sampling, means that there is a gap, or interval, between each selection.
- **Often** used in **industry**, where an item is selected for testing from a production line (say, every fifteen minutes)
- To ensure that machines and equipment are working to specification.
- **Quality control (QC).**

# SYSTEMATIC RANDOM SAMPLING

- Alternatively, the manufacturer might decide to select every 20th item on a production line to test for defects and quality.
- This technique requires the first item to be selected at random as a starting point for testing and, thereafter, every 20th item is chosen.

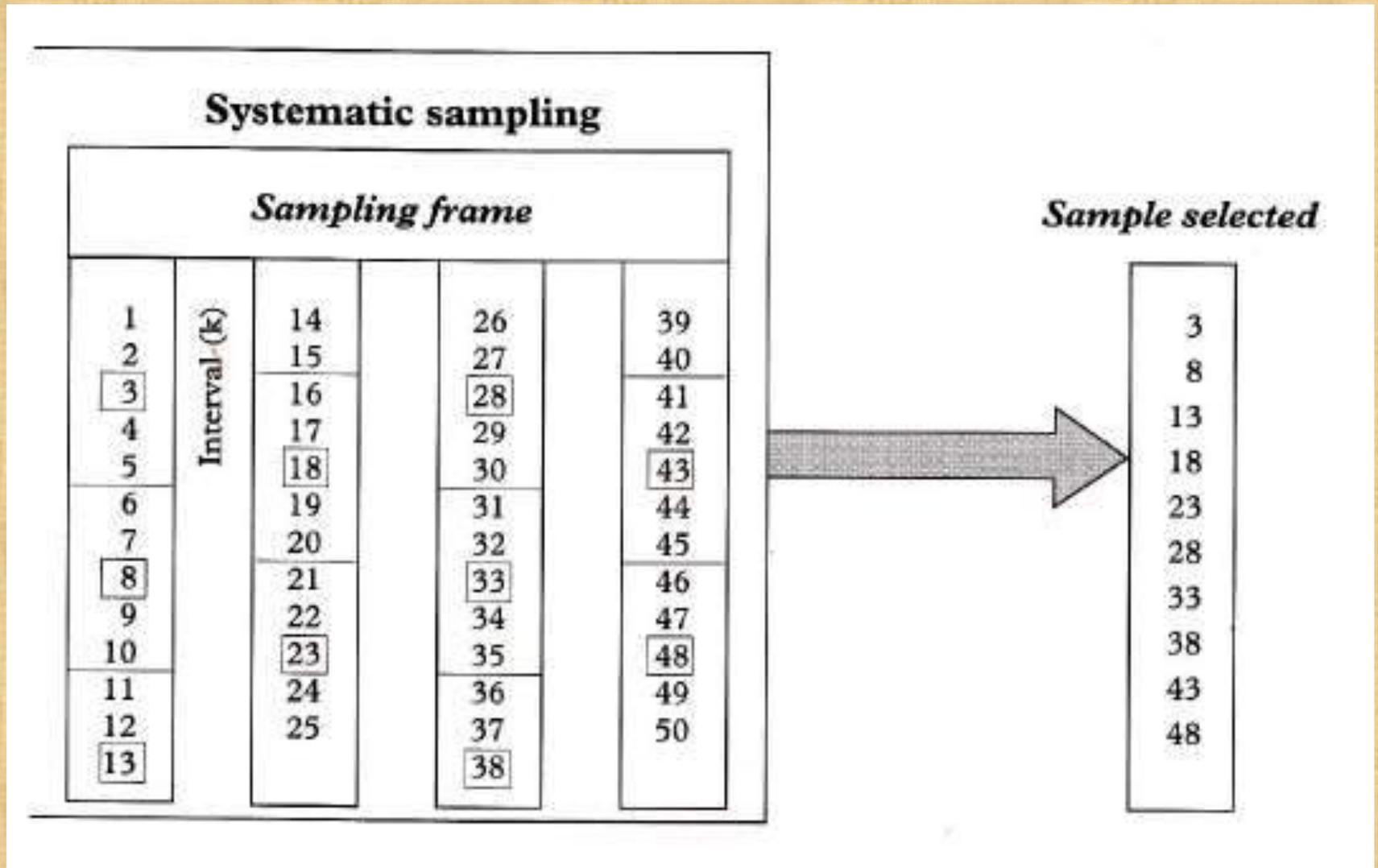
# SYSTEMATIC RANDOM SAMPLING

- If researcher wants to select a fixed size sample.
- In this case, it is first necessary to know the whole population size from which the sample is being selected.
- The appropriate *sampling interval*,  $l$ , is then calculated by dividing population size,  $N$ , by required sample size,  $n$ , as follows:

# SYSTEMATIC RANDOM SAMPLING

- If a systematic sample of 300 students were to be carried out in UMS with an enrolled population of 15,000, the sampling **interval** would be:
- $I = N/n = 15,000/300 = \underline{50}$
- This meaning that 1 element (student) will be selected in every 50 students from the list of 15,000 UMS students until the 300<sup>th</sup> student.
- This technique only can be used if a **complete list** of the **N elements** in a **population** is **available**.

# SYSTEMATIC RANDOM SAMPLING



# CLUSTER OR MULTISTAGE SAMPLING

- SRS and STRS are based on researcher's ability to identify each element in a population.
- Practical for only total sampling population is small.
- In the case of large population e.g. city, state or country, it is impossible (difficult + expensive) to identify each sampling unit.
- Therefore, cluster sampling is more practical and appropriate.

# CLUSTER OR MULTISTAGE SAMPLING

- Cluster sampling is a sampling technique where the entire population is divided into groups, or clusters.
- Then a random sample of these clusters are selected using SRS.
- All observations in the selected clusters are included in the sample.

# CLUSTER OR MULTISTAGE SAMPLING

- Every element should have a specified (equal) chance of being selected into the final sample.
- Typically used when the researcher cannot get a complete list of the members of a population they wish to study.
- But can get a complete list of groups or 'clusters' of the population.
- Cheap, easy economical method of data collection.

# CLUSTER OR MULTISTAGE SAMPLING

- For example: a PhD student want to know the nutritional status of standard six students in Sabah (just before they left the school for their form one).
  - 1) He/she will cluster all the schools according to districts (e.g. Kota Kinabalu, Papar, Kota Belud, Penampang etc..)
  - 2) Under each districts, the schools will again be divided according to clusters (type of school i.e. SK, SJK ; category of school i.e. urban, sub-urban, remote; etc..)

# CLUSTER OR MULTISTAGE SAMPLING

- 3) Then, one school in each type and category of schools under each of district will be sampled using SRS.
- 4) Finally, standard six students will be selected proportionally from each of the selected school as in 3) according to the total sample size needed (as early calculated) through SRS.

# NON-PROBABILITY SAMPLING

- Convenience/ opportunity/accidental sampling.
- Purposive/ judgemental sampling
- Quota sampling
- Snowball sampling

# **CONVENIENCE/ OPPORTUNITY/ACCIDENTAL SAMPLING**

- Volunteer samples
- Sometimes access through contacts or gatekeepers
- 'Easy to reach' population.

# PURPOSIVE/JUDGEMENTAL SAMPLING

- Involves selecting a group of people because they have particular traits that the researcher wants to study
- e.g. consumers of a particular product or service in some types of market research

# QUOTA SAMPLING

- Widely used in opinion polls and market research.
- Interviewers given a quota of subjects of specified type to attempt to recruit.
- eg. an interviewer might be told to go out and select 20 male smokers and 20 female smokers so that they could interview them about their health and smoking behaviours .

# SNOWBALL SAMPLING

- Involves two main steps.
  1. Identify a few key individuals
  2. Ask these individuals to volunteer to distribute the questionnaire to people who know and fit the traits of the desired sample

# SAMPLE SIZE

- In general, the larger the sample size (selected with the use of probability techniques) the better.
- The more heterogeneous a population is on a variety of characteristics (e.g. race, age, sexual orientation, religion) then a larger sample is needed to reflect that diversity.

# SAMPLE SIZE

Sample Size Calculator - Confidence Level, Confidence Interval, Sample Size, Population Size, Relevant Population - Creative Research Systems - Mo...

File Edit View History Bookmarks Tools Help

http://www.surveysystem.com/sscalc.htm

Most Visited EPF : When Should E... Customize Links Free Hotmail Windows Marketplace Windows Media Windows ScienceDirect Error

Sample Size Calculator - Confidence ...

Gadgets powered by Google

have. Leave the Population box blank, if the population is very large or unknown.

### Determine Sample Size

Confidence Level:  95%  99%

Confidence Interval:

Population:

Sample size needed:

### Find Confidence Interval

Confidence Level:  95%  99%

Sample Size:

Population:

Percentage:

Confidence Interval:

Done

start SAMPLING TECHNIQ... Sampling and Basic D... QuantitativeAndSam... Sample Size Calculato... 7:13 AM

# REFERENCES

- Kumar, R. 1996. *Research Methodology: A Step-By-Step Guide For Beginners*. SAGE Publications, London.
- Pagano, M. and Gauvreau, K. 2000. *Principles of Biostatistics*. Duxbury Thomson Learning, USA.
- McKillup, S. 2005. *Statistics Explained: An Introductory Guide For Life Scientists*. Cambridge University Press, UK.
- Suphat Sukamolson. *Fundamentals of quantitative research*.
- PowerPoint Presentations By Leah Wild (*Sampling and Basic Descriptive Statistics. Basic concepts and Techniques*); David Arnott (*Experimental Research*); Moataza Mahmoud Abdel Wahab (*Sampling Techniques and Sample Size*)
- Most of the notes in this lecture are directly taken or slightly modified from the above mentioned references.

**THANK YOU**

# **FUNDAMENTALS OF QUANTITATIVE RESEARCH**

**ASSOC. PROF. DR. MOHD ROSNI SULAIMAN  
FACULTY OF FOOD SCIENCE AND NUTRITION  
UNIVERSITI MALAYSIA SABAH**

# INTRODUCTION

- **Definitions**

- Quantitative research is the numerical representation and manipulation of observations for the purpose of describing and explaining the phenomena that those observations reflect. It is used in a wide variety of natural and social sciences, including physics, biology, psychology, sociology and geology (Wikipedia Encyclopedia, 2005).

# INTRODUCTION

- **Definitions**

- Creswell (1994) has given a very concise definition of quantitative research as a type of research that is 'explaining phenomena by collecting numerical data that are analyzed using mathematically based methods (in particular statistics).'

# INTRODUCTION

- Let's study this definition step by step.
- The first element is *explaining phenomena*.
- This is a key element of all research, be it quantitative or qualitative.
- When we set out to do some research, we are always looking to explain something.
- The specificity of quantitative research lies in the next part of the definition.

# INTRODUCTION

- In quantitative research we collect *numerical data*.
- This is closely connected to the final part of the definition: analysis *using mathematically-based methods*.
- In order to be able to use mathematically based methods our data have to be in numerical form.
- This is not the case for qualitative research. Qualitative data are not necessarily or usually numerical, and therefore cannot be analyzed using statistics.

# INTRODUCTION

- The last part of the definition refers to the use of *mathematically based methods*, in particular *statistics*, to analyze the data.
- This is what people usually think about when they think of quantitative research.
- Is often seen as the most important part of quantitative studies.
- This is a bit of a misconception.

# INTRODUCTION

- It is important to use the right data analysis tools.
- It is even more important to use the right research design and data collection instruments.
- However, the use of statistics to analyze the data is the element that puts a lot of people off doing quantitative research
- Why does this happen?
- Because the mathematics underlying the methods seem complicated and frightening.

# INTRODUCTION

- Quantitative research is essentially about collecting numerical data to explain a particular phenomenon.
- Therefore, particular questions seem immediately suited to being answered using quantitative methods.
- For example:
  - *How many species of snake that are still existing in Sabah?*
  - *What percentage of the primary school students in Kota Kinabalu has negative attitudes towards the Science subject?*
  - *On average, are there any significant differences between vegetables planted in Kundasang and Cameron Highland in terms of their copper content?*

# INTRODUCTION

- **These are all questions we can look at quantitatively, as the data we need to collect are already available to us in numerical form.**

# **ADVANTAGES OF QUANTITATIVE RESEARCH**

- 1. Provides estimates of populations at large.**
- 2. Indicates the extensiveness of attitudes held by people.**
- 3. Provides results which can be condensed to statistics.**
- 4. Allows for statistical comparison between various groups.**
- 5. Has precision, is definitive and standardized.**
- 6. Measures level of occurrence, actions, trends, etc.**
- 7. Can answer such questions as "How many?" and "How often?"**

# COMMON APPROACHES TO QUANTITATIVE RESEARCH

1. Surveys | 2. Custom surveys |
3. Mail/e-mail/Internet surveys | 4. Telephone surveys |
5. Self-administered questionnaire surveys | 6. Omnibus surveys |
7. Correlational research | 8. Trend analysis |
9. Exploratory research | 10. Descriptive research |
11. Experimental research |

# The research process



# SCIENTIFIC METHODS

**Three main classes of investigation:**

- **Descriptive studies**
  - variables or phenomena are described
- **Correlational studies**
  - relationships between variables are identified
- **Experiments**
  - manipulation and measurement of variables to infer causality

# RESEARCH QUESTIONS

- Research is the process of:
  - i) Asking important questions
  - ii) Answering them in a way that is **convincing** and **defensible**
- Any question that is capable of being confirmed or refuted is a potential target for experimentation.
- Methods should be guided by the questions.

# VARIABLES

- **Independent variables**
  - Variable whose effect we are interested in
  - Manipulated by the researcher
  - Levels - ways manipulated
  - Subject variables - selected not manipulated
- **Dependent variables**
  - The response or behaviour
  - Measures the influence of the independent variable

# VARIABLES

- **Intermediate variables**

- A variable in a causal pathway that causes variation in the dependent variable and is itself caused to vary by the independent variable
- Exercise vs High sugar foods intake vs Blood sugar (Diabetic)

- **Extraneous variables**

- A variable, other than the independent variable
- Capable of affecting the dependent variable
- Confounding variables or confounds

# GROUPS

- **Experimental group**
  - Treatment group
  - Group that receives the experimental treatment
- **Control group**
  - Does not receive treatment
- **Groups should be equivalent**
  - Control extraneous variables
  - Random assignment
  - Matched pairs

# HYPOTHESES

- Predictions about the effect of the independent var. on the dependent var.
- Research hypotheses:
  - **Alternative hypotheses ( $H_1$ )**
    - What the researcher expects
    - Two-tailed or one tailed
    - Direction is important
  - **Null hypotheses ( $H_0$ )**
    - What the researcher doesn't expect
    - No significant difference

# SIGNIFICANT DIFFERENCE

- Not sufficient to simply have a difference between the groups in an experiment to argue that the independent variable can affect the dependent variable in a causal way.
- The difference between two descriptive statistics that is of such magnitude that it is unlikely to have occurred by chance.
- **Significance level**
  - 95% or  $p \leq 0.05$  is acceptable
  - 99% or  $p \leq 0.01$  is a strong result

# THE NATURE OF THE INVESTIGATION

- Studies can be classified as:

- **Experimental**

*(researcher introducing the intervention that is assumed to be the “cause” of change and waiting until it has produced the change)*

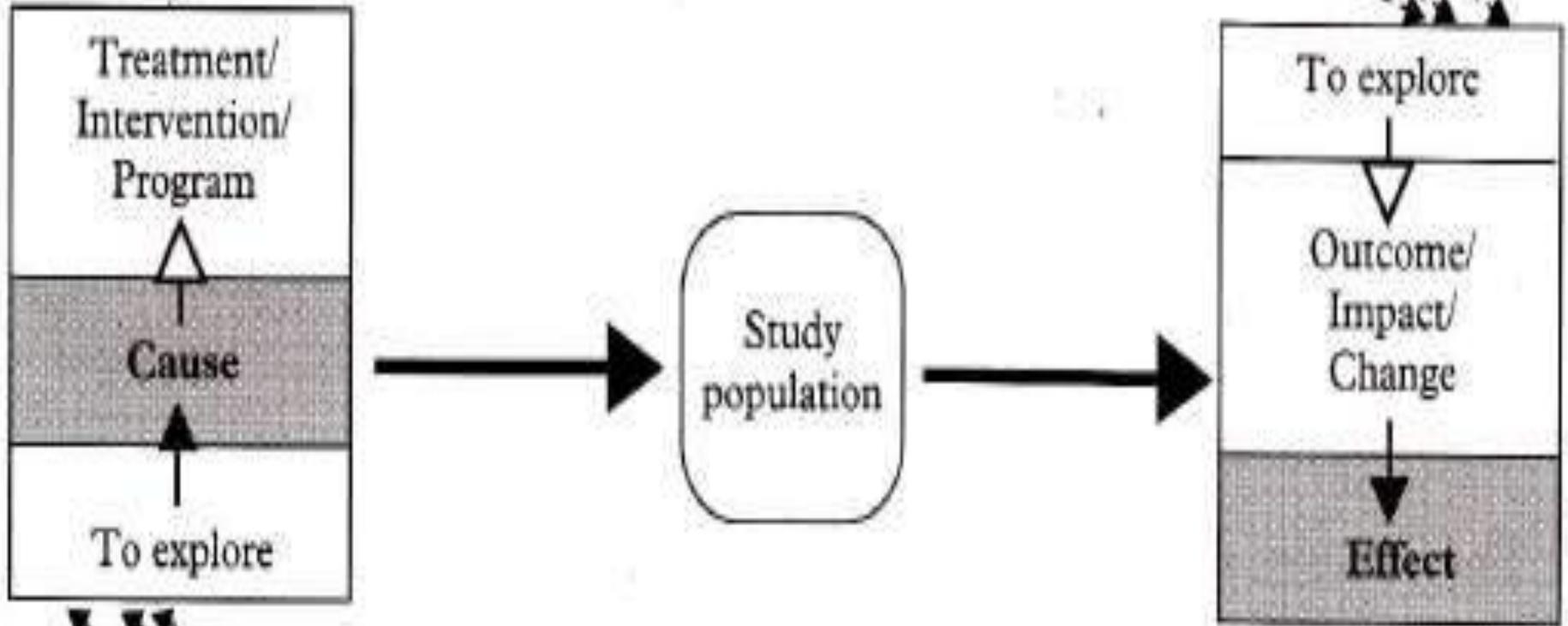
- **Non-experimental**

*(The researcher observing a phenomenon and attempting to establish what caused it)*

- **Quasi or semi experimental**

*(Has the properties of both experimental and non-experimental)*

Experimental studies



Non-experimental studies

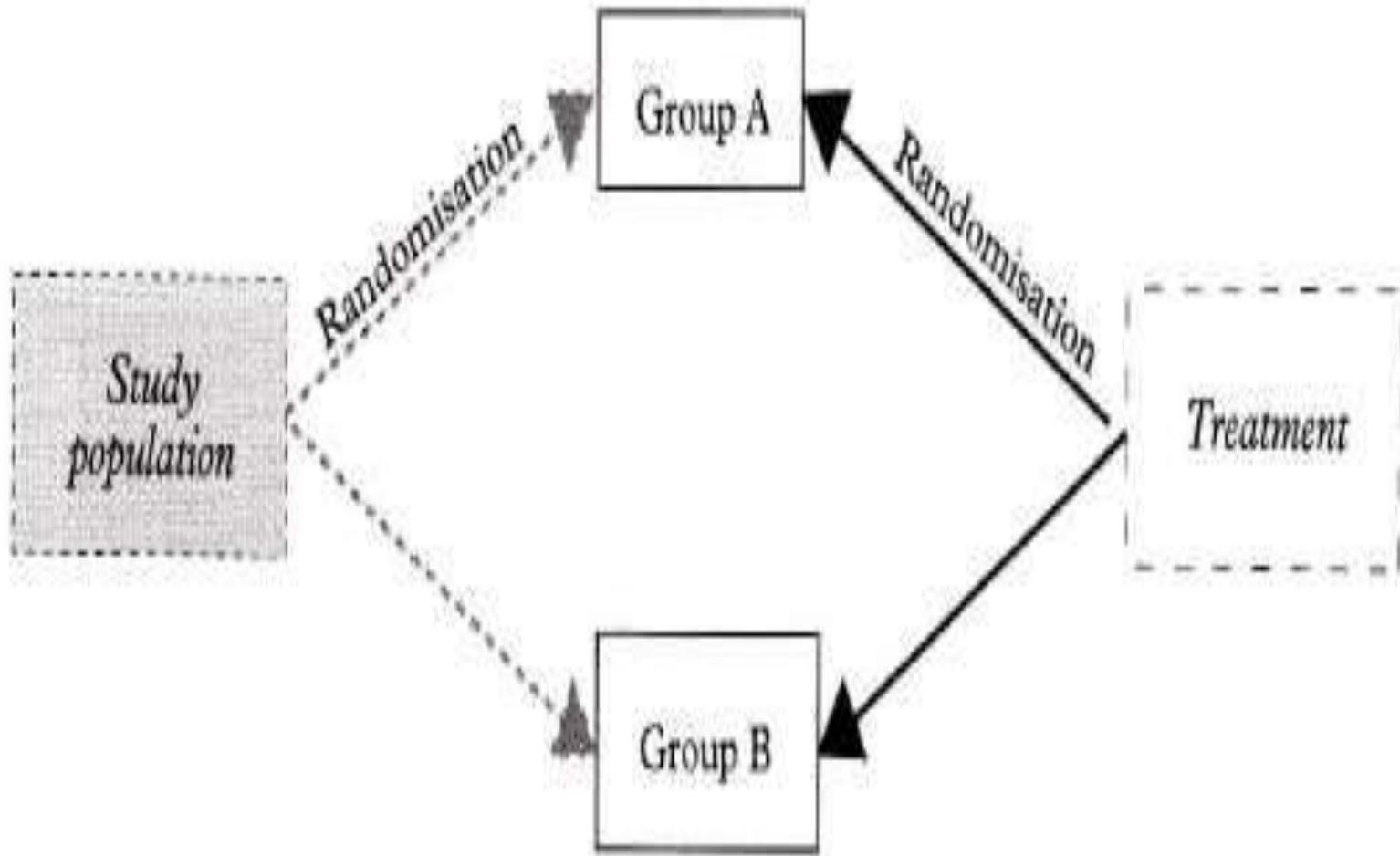
Source: Kumar (1996)

# EXPERIMENTS

- Investigations where groups are treated identically.
- Except for a manipulation of the independent variable.
- Changes in the dependent variable may be attributed to the difference in the independent variable.

# **RANDOMISATION IN EXPERIMENTS**

- Experimental studies can be further classified on the basis of whether or not the study population is randomly assigned to different treatment groups.
- One of the biggest problems in comparable designs is a lack of certainty that the different groups are in fact comparable in every respect except the treatment.
- The process of randomisation is designed to ensure that the groups are comparable.

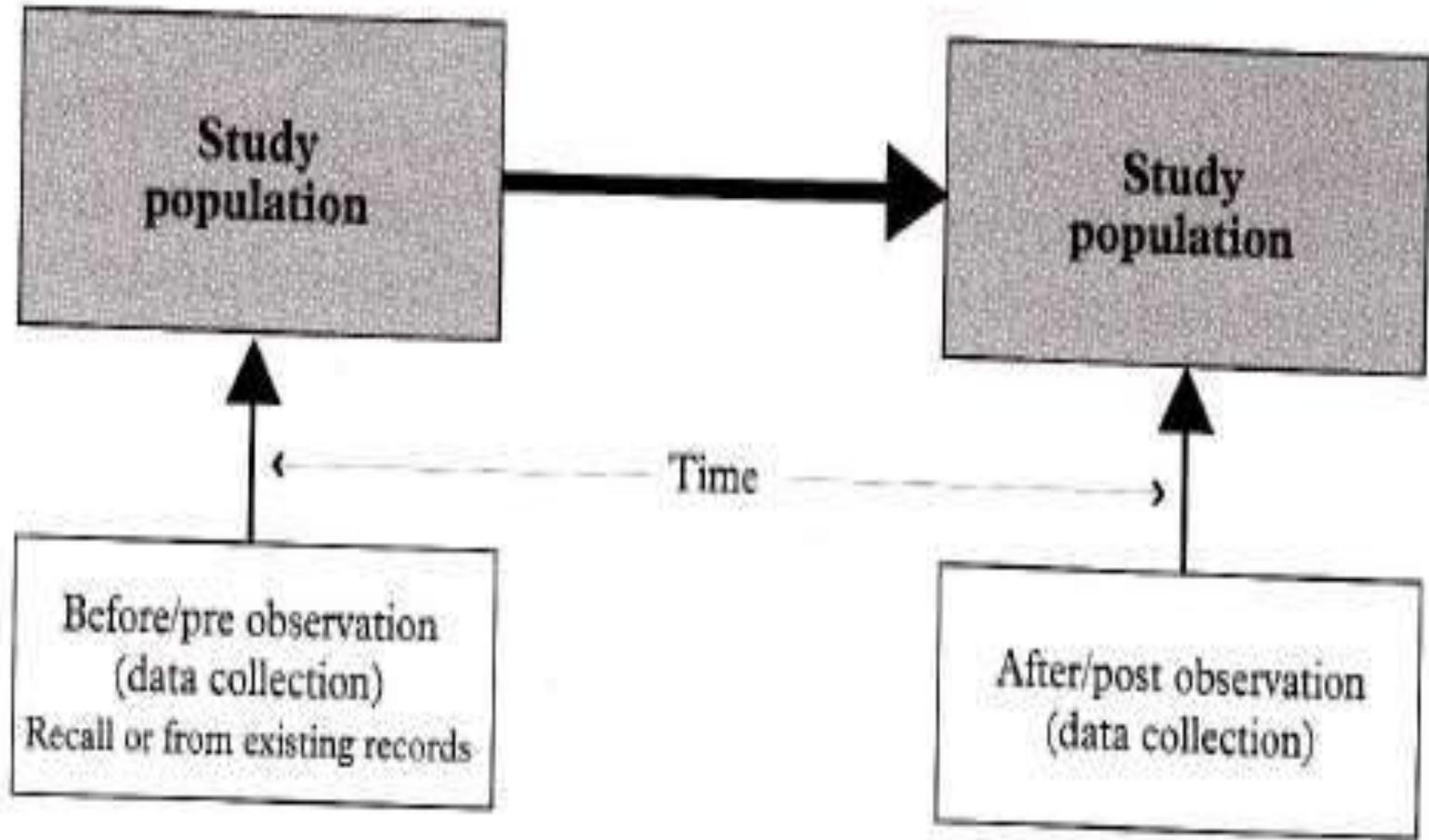


Source: Kumar (1996)

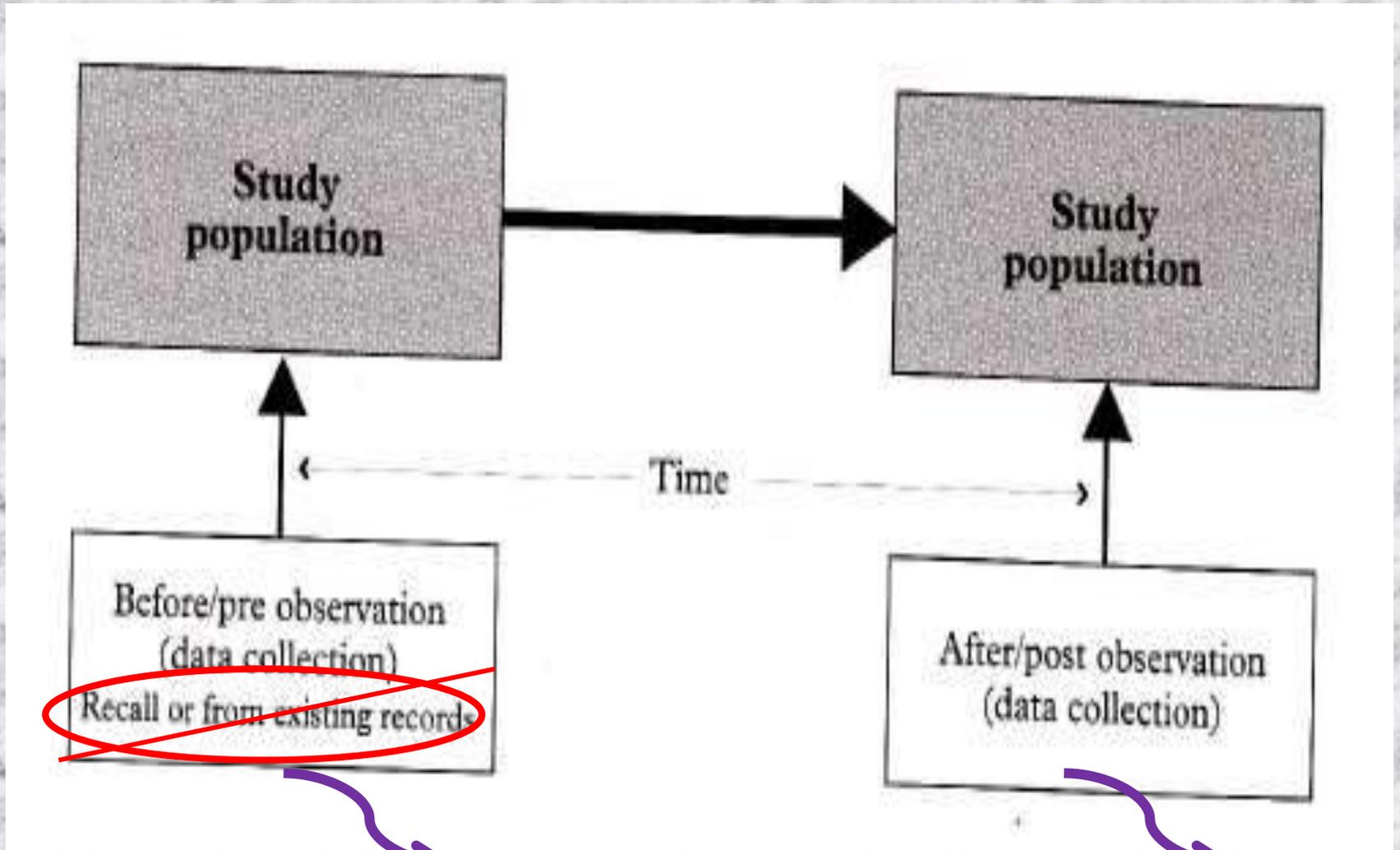
# THE EXPERIMENTAL STUDY DESIGN

- There are so many types of experimental design, some of the most commonly used are:
  - The after-only design
  - The before-and-after design
  - The control group design
  - The double control design
  - The comparative design
  - The matched control experimental design
  - The placebo design

# THE AFTER-ONLY DESIGN



# THE BEFORE-AND-AFTER DESIGN

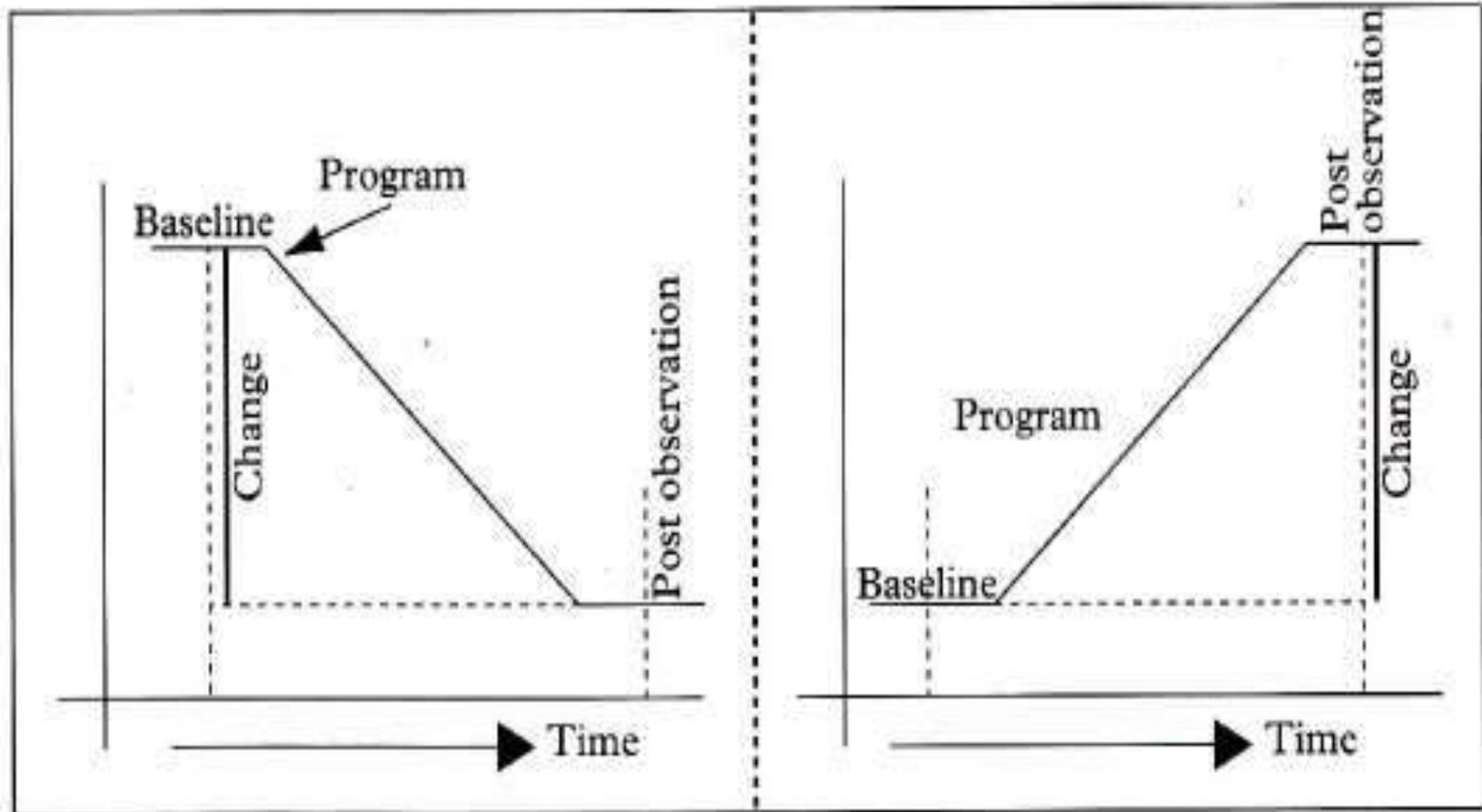


Source: Kumar (1996)

Actual measurement 1

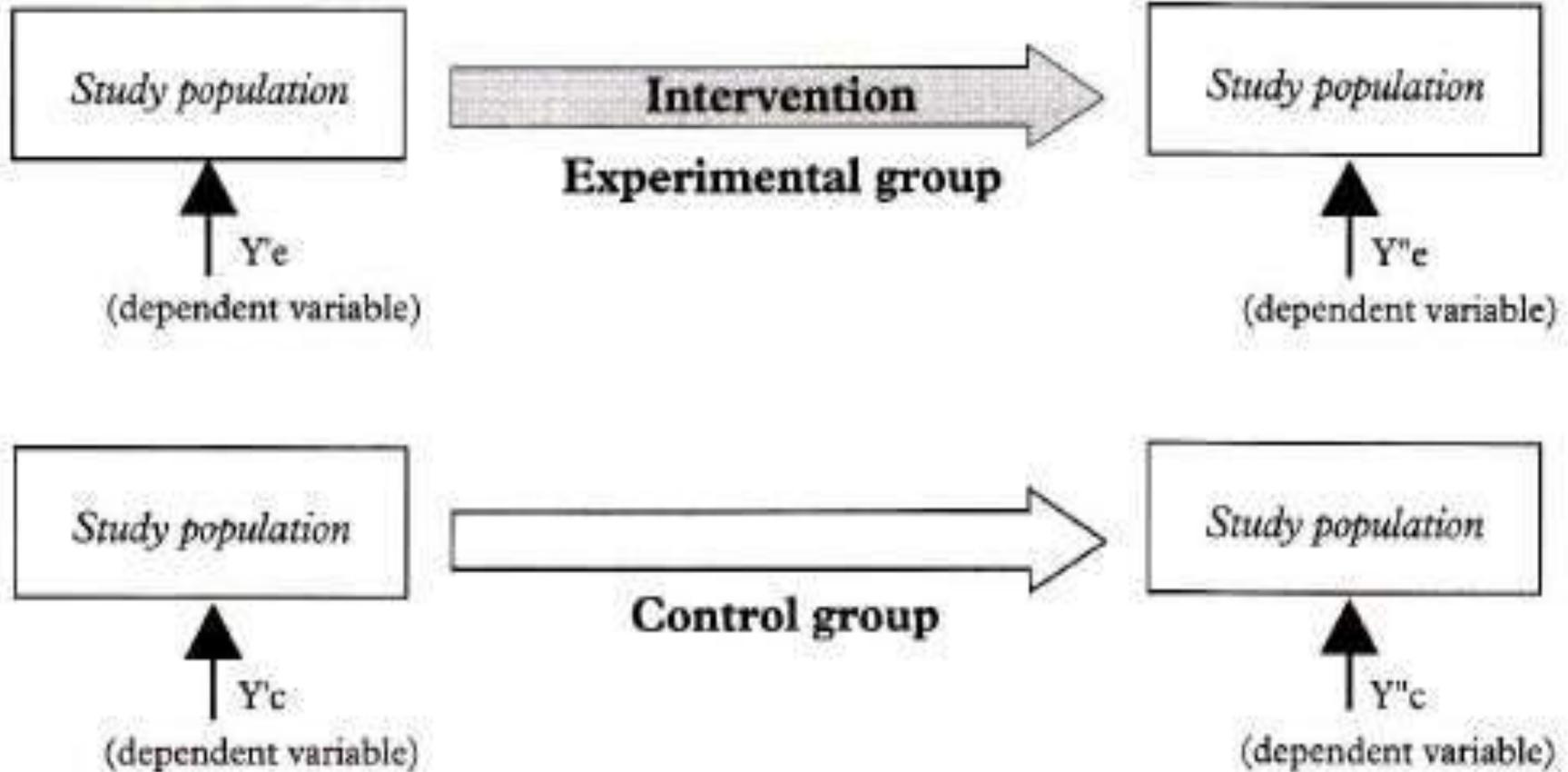
Actual measurement 2

# THE BEFORE-AND-AFTER DESIGN



Measurement of change

# THE CONTROL GROUP DESIGN



# THE CONTROL GROUP DESIGN

- In the experimental group, total change in the dependent variable ( $Y_e$ ):

$$(Y_e) = (Y''_e - Y'_e), \text{ where}$$

$Y''_e$  = 'after' observation on the experimental group

$Y'_e$  = 'before' observation on the experimental group.

- In other words:

$$(Y''_e - Y'_e) = (\text{Impact of program intervention}) \pm (\text{Impact of extraneous variables}) \pm (\text{Impact of chance variables})$$

# THE CONTROL GROUP DESIGN

- In the control group, total change in the dependent variable ( $Y_c$ ):

$$(Y_c) = (Y''_c - Y'_c), \text{ where}$$

$Y''_c$  = post-test observation on the control group

$Y'_c$  = pre-test observation on the control group.

- In other words:

$$(Y''_c - Y'_c) = (\text{Impact of extraneous variables}) \pm (\text{Impact of chance variables})$$

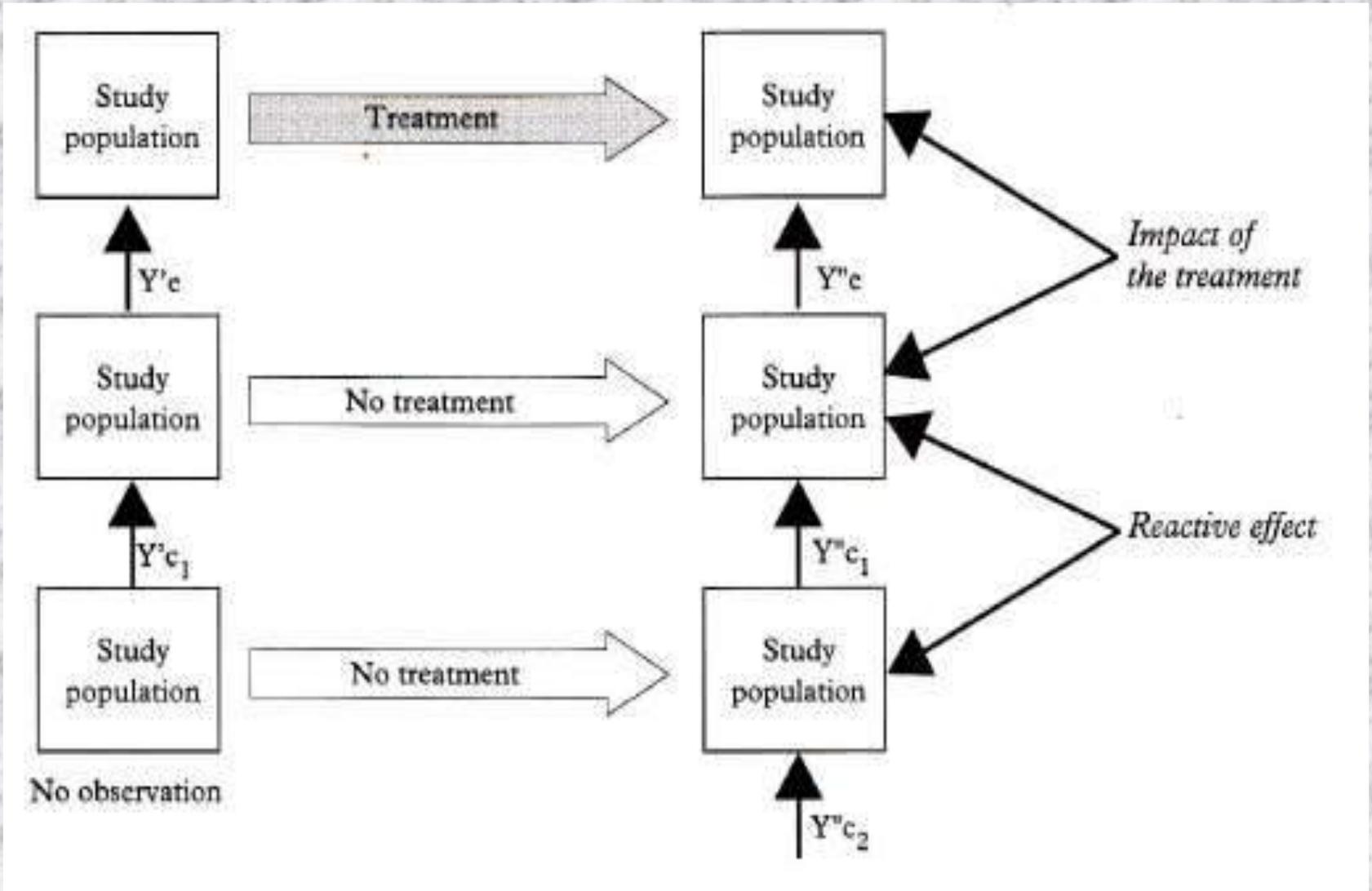
# THE CONTROL GROUP DESIGN

- The difference between control and experimental groups =  $(Y''e - Y'e) - (Y''c - Y'c)$ , which is

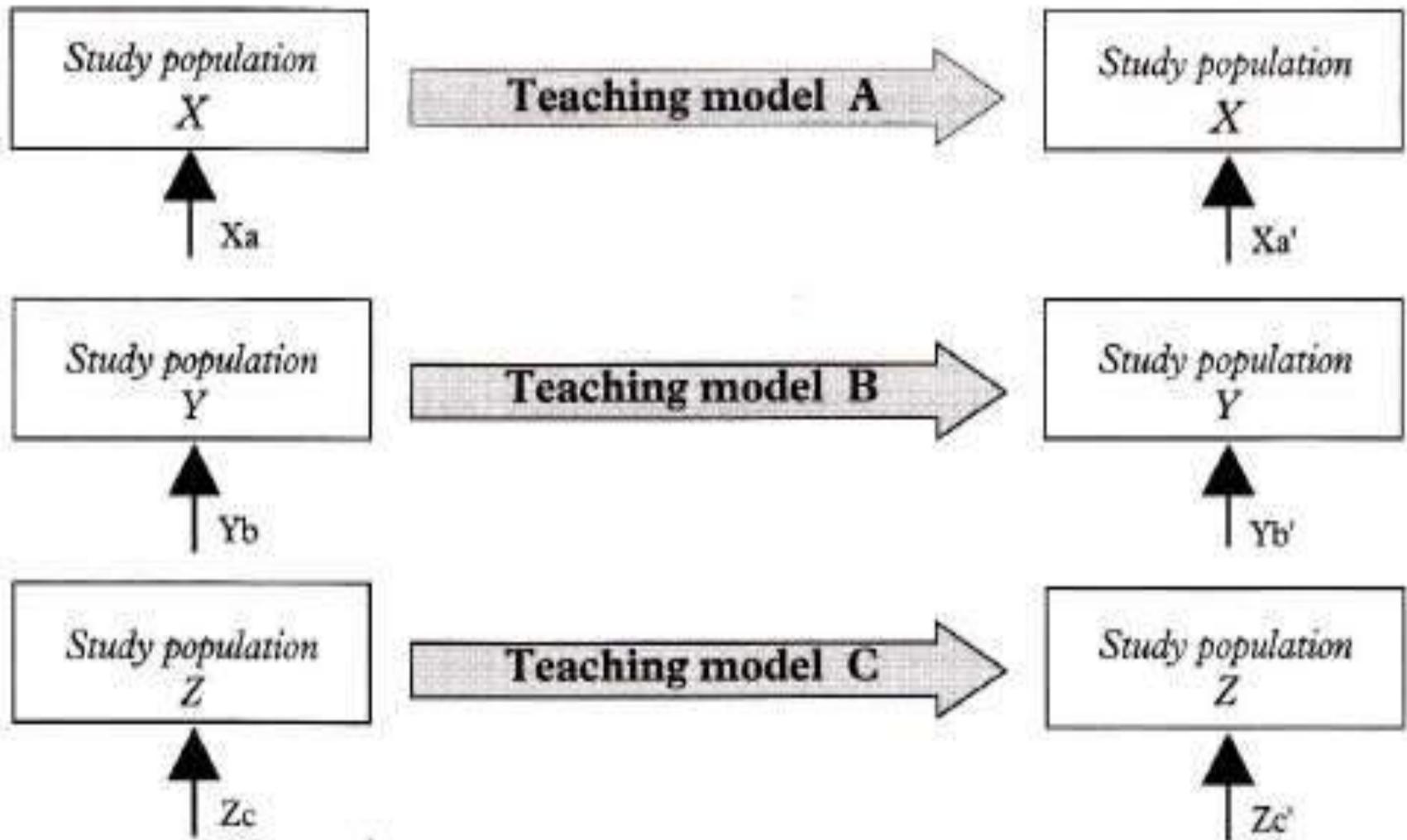
{(Impact of program intervention)  $\pm$  (Impact of extraneous variables\*)  $\pm$  (Impact of chance variables#)} - {(Impact of extraneous variables\*)  $\pm$  (Impact of chance variables#)}.

- Using a simple arithmetic operation this = **Impact of the intervention**

# THE DOUBLE CONTROL DESIGN



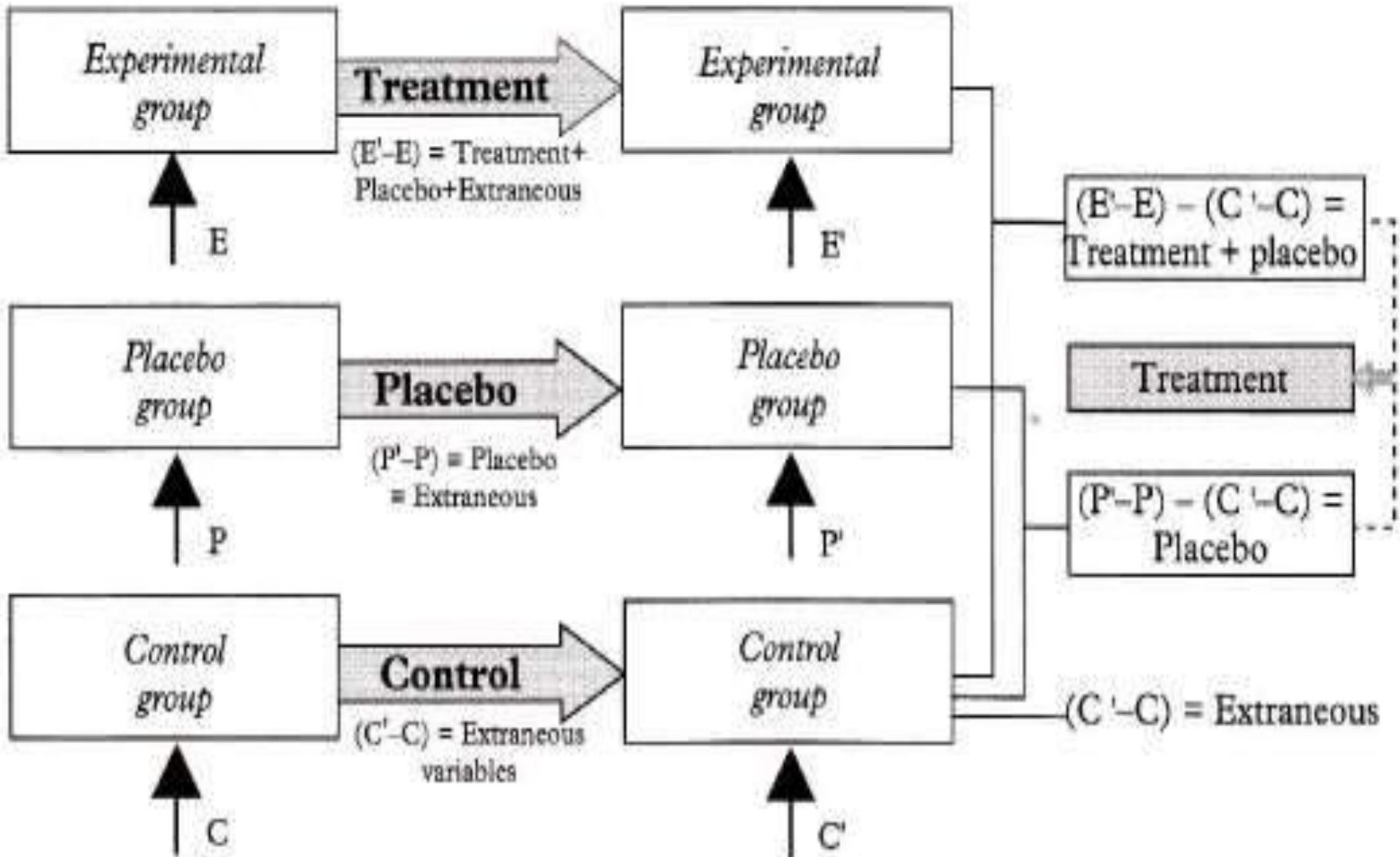
# THE COMPARATIVE DESIGN



# THE MATCHED CONTROL EXPERIMENTAL DESIGN

- Comparability is determined on an individual-by-individual basis.
- Two individuals from the study population who are almost identical (age, gender, type of illness etc..), are matched.
- Then each is allocated to a different group.
- Once two groups are formed, the researcher decides (randomisation) which group as control and which group as experimental.
- E.g. testing of new drugs.

# THE PLACEBO DESIGN



Source: Kumar (1996)

# THE PLACEBO DESIGN

- A patient's belief that he or she is receiving treatment can play an important role in his/her recovery from an illness even if treatment is ineffective.
- This psychological effect is known as the placebo effect.
- A placebo design attempts to determine the extent of this effect.
- A placebo design involves two or three groups, depending on whether or not the researcher wants to have a control group.

# REFERENCES

- Kumar, R. 1996. *Research Methodology: A Step-By-Step Guide For Beginners*. SAGE Publications, London.
- Pagano, M. and Gauvreau, K. 2000. *Principles of Biostatistics*. Duxbury Thomson Learning, USA.
- McKillup, S. 2005. *Statistics Explained: An Introductory Guide For Life Scientists*. Cambridge University Press, UK.
- Suphat Sukamolson. *Fundamentals of quantitative research*.
- PowerPoint Presentations By Leah Wild (*Sampling and Basic Descriptive Statistics. Basic concepts and Techniques*); David Arnott (*Experimental Research*); Moataza Mahmoud Abdel Wahab (*Sampling Techniques and Sample Size*)
- Most of the notes in this lecture are directly taken or slightly modified from the above mentioned references.

**THANK YOU**