2. Variable: Independent, Dependent, Intervening and Confounding Variables

Variables: Whether we accept it or not, we all make value judgements constantly in our daily lives: 'this food is **excellent**'; 'I could not sleep **well** last night'; 'I do not **like** this'; and 'I think this is **wonderful**'. These are all judgements based upon our own preferences, indicators or assessment. Because these explain feelings or preferences, the basis on which they are made may vary markedly from person to person. There is no uniform yardstick with which to measure them. A particular food may be judged excellent by one person by awful by another, and something else could be wonderful to one person but ugly to another. When people express these feelings or preferences, they do so on the basis of certain criteria in their minds or in relation to their expectations. If you were to question them you will discover that their judgement is based upon indicators and/or expectations that lead them to conclude and express a particular opinion.

Definition of variable: An image, perception or concept that is capable of measurement – hence capable of taking on different values – is called a variable. In other words, a concept that can be measured is called a variable.

The difference between concept and a variable: Measurability is the main difference between a concept and a variable. Concepts are mental images or perceptions and therefore their meanings vary markedly from individual to individual, whereas variables are measurable, though, of course, with varying degrees of accuracy. A concept cannot be measured whereas a variable can be subjected to measurement by crude/refined or subjective/objective units of measurement. Concepts are subjective impressions which, if measured as such would cause problems in comparing responses obtained from different respondents.

TABLE 5.1 *Examples of concepts and variables*

Concepts	Variables
• Effectiveness • Satisfaction • Impact • Excellent • High achiever • Self-esteem • Rich • Domestic violence • Extent and pattern of alcohol consumption • etc.	• Gender (male/female) • Attitude • Age (*x* years, *y* months) • Income ($ ___ per year) • Weight (___ kg) • Height (___ cm) • Religion (Catholic, Protestant, Jew, Muslim) • etc.
– Subjective impression – No uniformity as to its understanding among different people – As such cannot be measured	– Measurable though the degree of precision varies from scale to scale and from variable to variable (e.g. attitude – subjective, income – objective)

Types of variables:

A. **On the basis of casual relationship:**
   1. **Independent variable:** The cause supposed to be responsible for bringing about changes in a phenomenon or situation.
   2. **Dependent variable:** The outcome or changes brought about by introduction of an independent variable.
   3. **Extraneous variable:** Several other factors operating in a real-life situation may affect changes in the dependent variable. These factors, not measured in the study, may increase or decrease the magnitude or strength of the relationship between independent and dependent variables.

4. **Intervening variable:** Sometimes called the confounding variable; it links the independent and dependent variables. In certain situations, the relationship between an independent and a dependent variable cannot be established without the intervention of another variable. The cause, or independent variable will have the assumed effect only in the presence of an intervening variable.

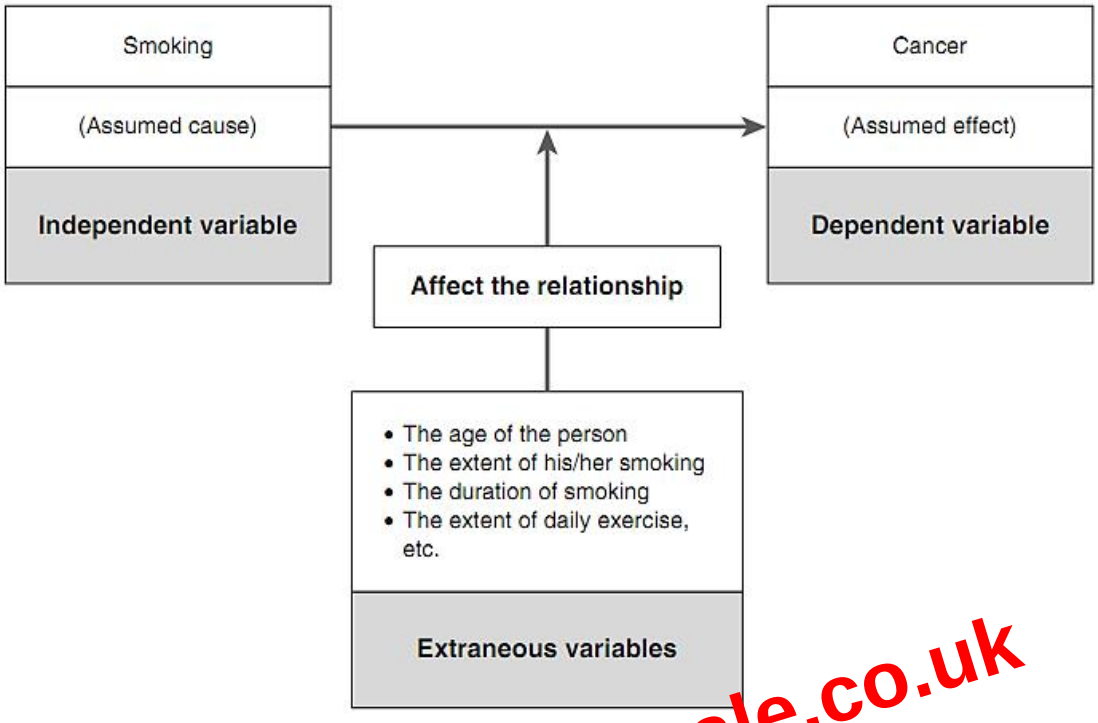

FIGURE 5.3 *Independent, dependent and extraneous variables in a causal relationship*

Explanation: Suppose one wants to study the relationship between fertility and mortality. His/her aim is to explore what happens to fertility when mortality declines. The history of demographic transition has shown that a reduction in the fertility level follows a decline in the mortality level, though the time taken to attain the same level of reduction in fertility varied markedly from country to country. As such, there is no direct relationship between fertility and mortality. With the reduction in mortality, fertility will decline if people attempt to limit their family size. History has shown that for a multiplicity of reasons, people have used one method or another to control their fertility, resulting in lower fertility levels. It is thus the intervention of contraceptive methods that completes the relationship: the greater the use of contraceptives, the greater the decline in the fertility level and the sooner the adoption of contraceptive methods by people, the sooner the decline. The extent of the use of contraceptives is also affected by a number of other factors, for example, attitudes towards contraception, level of education, socioeconomic status and age, religion, and provision and quality of health services. These are classified as extraneous variables.

In the above example, decline in mortality is assumed to be the cause of a reduction in fertility, hence the mortality level is the independent variable and fertility is the dependent variable. But this relationship will be completed only if another variable intervenes – that is, the use of contraceptives. A reduction in mortality (especially child mortality) increases family size and an increase in family size creates a number of social, economic and psychological pressures on families, which in turn create attitudes favourable to a smaller family size. This change in attitudes is eventually operationalized in behavior through the adoption of contraceptives. If people do not adopt methods of contraception, a change in mortality levels will not be reflected in fertility levels. The population explosion in developing countries is primarily due to lack of acceptance of contraceptives. The extent of the use of contraceptives determines the level of the decline in fertility. The extent of contraceptive adoption by a population is dependent upon a number of factors. As mentioned earlier, in this causal model, the fertility level is the dependent variable, the extent of contraceptive use is the intervening variable, the mortality level is the independent variable and the unmeasured variables such as attitudes, education, age, religion, the quality of services, and so on, are all extraneous variables. Without the intervening variable the relationship between the independent and dependent variables will not be complete.

Refereeing system

Reference: Reference is the detailed description of the document from which one has obtained his/her information. Referencing is a way of demonstrating that one has done that reading.

Citing: Acknowledging within one's text the document from which one has obtained one's information.

Reference vs bibliography: The term reference and bibliography are often used synonymously, but there is a difference in meaning between them.

References are items one has read and specifically referred to (or cited) in his/her work and his/her list of source at the end of the assignment will be headed.

Bibliography is a list of everything one reads whether or not he/she referred specifically to it. A bibliography can give a tutor an overview of which authors have influenced his/her ideas and arguments even if one does not specifically refer to them.

Importance of referencing:

1. To acknowledge others' works.
2. To allow others (readers) to find the original source easily (cited reference).
3. To get recognition and authentication of the work.
4. To make the work informative (quality).
5. To trace the intellectual development of the ideas one presents.
6. To avoid plagiarism.

Harvard style of referencing:

- Author's name followed by its initials
- Year of publication
- Article title with single quotation mark followed by full stop
- Name of journal in italic form
- Volume followed by a comma
- Issue no. in bracket
- Page no.

Example:

- Padda, J. (2003) 'creative writing in coventry'. *Journal of writing studies* 3 (2), 44 – 59.

Vancouver style

- Author surname followed by initials
- Title of article followed by double quotation
- Title of journal (abbreviated)
- Date of publication followed by double quotation
- Volume number
- Issue number in bracket
- Page number

Example:

1. Has AN, Susin C, Albandar JM, et al. Azithromycin as a adjunctive treatment of aggressive periodontitis: 12-months randomized clinical trial. N Engl J Med 2008 Aug; 35(8):696 – 704.

Note that, Vancouver style does not use the full journal name, only the commonly used abbreviation. For example, 'New England Journal of Medicine' is cited as 'N Engl J Med'.

6. Sampling Techniques

Sampling

Sampling: Sampling means selecting a given number of subjects from a defined population as representative of that population. A sample is a part of a target population, which is carefully selected to represent the population. Sampling frame is the list of elements from which the sample is actually drawn. Actually sampling frame is nothing but the correct list of population. Example: Telephone directory, Product finder, Yellow pages.

Sampling process: The sampling process comprises the following stage:

1. Defining the population of concern.
2. Specifying a sampling frame, a set of items or events possible to measure.
3. Specifying a sampling method for selecting items or events from the frame.
4. Determining the sampling size.
5. Implementing the sampling plan.
6. Sampling and data collection.
7. Reviewing the sampling process.

Survey methods: A survey may be conducted by either of the following two methods:

1. Census method or Parametric method.
2. Sampling method or Non-parametric method.

They are discussed below:

1. **Census method:** It deals with the investigation of the entire population. Here the data are collected for each and every unit of universe. This method provides more accurate and exact information as no unit is left out.
2. **Sampling method:** Here a small group is selected as representative of the whole universe. It works with the objective to obtain accurate and reliable information about the universe with minimum of cost, time and energy and to set out of limits of accuracy of such estimates. It makes exhaustive and intensive study possible with much less time, money and material. It is more popular in research work.

Population: Population or universe means the entire mass of observations, which is the parent group from which a sample is to be formed. The term population or universe conveys a different meaning than a traditional one. In census survey, the count of individuals (men, women and children) is known as population. But in Research Methodology, population means characteristics of a specific group. For example, secondary school teachers of, who have some specific features like teaching experience, teaching attitudes etc. one type of population distinguished by educational researchers is called the target population. By target population, also called universe, we mean all the members of a real hypothetical set of people, events or objects to which we wish to generalize the results of our research.

Advantages of sampling:

1. It has a greater adaptability.
2. It is an economical technique.
3. It has high speed for generalization.
4. According to W. G. Cocharan, 'It has greater precision and accuracy in the observation'.
5. This technique has great accuracy.
6. It has a greater speed in conducting a research work.
7. It has a greater scope in the field of research.
8. It reduces the cost of observation or data collection.

Disadvantages of sampling:

1. There is a scope of biasness (less accuracy).
2. Problem of representative sample-difficulty in selecting a truly representative sample.
3. Need of eligible researchers.
4. Instability of sample subjects or changeability of units i.e. in heterogeneous population.

Disadvantages:

i. It is a difficult and complex method of sampling.
ii. It involves errors when we consider the primary stages.
iii. It is again a subjective technique of sampling.

6. **Cluster sampling:** To select the intact group as a whole is known as a cluster sampling. In cluster sampling the sample units contain groups of element (cluster) instead of individual members or items in the population. Rather than listing all elementary school children in a given city and randomly selecting 15% of these students for the sample, a researcher lists all of the elementary schools in the city, selects at random 15% of these clusters of units, and uses all of the children in the selected schools as the sample.

Advantages:

i. It may be a good representative of the population.
ii. It is an easy method.
iii. It is an economical method.
iv. It is practicable and highly applicable in education.
v. Observations can be used for inferential purpose.

Disadvantages:

i. Cluster sampling is not free from errors.
ii. It is not comprehensive.

B. **Non probability sample:** Samples which are selected through non-random methods are called non probability samples. Depending upon the technique used it may be:

1. **Incidental or accidental sampling:** The term incidental or accidental applied to those samples that are taken because they are most frequently available i.e. this refers to the groups which are used as samples of a population because they are readily available or because the researcher is unable to employ more acceptable sampling methods.

Advantages:

i. It is very easy method of sampling.
ii. It is frequently used method in behavioral sciences.
iii. It reduces the time, money and energy i.e. it is an economical method.

Disadvantages:

i. It is not representative of the population.
ii. It is not free from errors.
iii. Parametric statistics cannot be used.

2. **Judgement sampling:** This involves the selection of a group from the population on the basis of available information assuming as if they are representative of the entire population. Here, group may also be selected on the basis of institution or on the basis of the criterion deemed to be self-evident. Generally, investigator should take the judgement sample of this sampling is highly risky.

Advantages:

i. Knowledge of investigator can be best used in this technique of sampling.
ii. This method of sampling is economical.

Disadvantages:

i. This technique is objective.
ii. It is not free from errors.
iii. It includes uncontrolled variation.

- When our population is very small.
- When we have extensive resources.
- When we don't expect a very high response.

Steps in sampling design: Sampling process consists of seven steps. They are-

1. Defining the population.
2. Identifying the sampling frame.
3. Specifying the sampling unit.
4. Selection of sampling method.
5. Determination of sample size.
6. Specify sampling plan.
7. Selection of sample.

construct, distribute and manage results. E mail questionnaire is totally e mail based. It works with an existing e mail system making online questionnaire surveys available to anyone with an Internet Connection.

Advantages:

1. Speed: An e mail questionnaire can gather several thousand responses within a day or two.
2. There is practically no cost involved once the setup has been completed.
3. Pictures and sound files can be attached.
4. The novelty element of an e mail survey often stimulates higher response levels than ordinary mail surveys.

Disadvantages:

1. Researcher must possess or purchase a list of email address.
2. Some people will respond several times or pass questionnaires along to friends to answer.
3. Many people dislike unsolicited e mail even more than unsolicited regular mail.
4. Findings cannot be generalized with e mail surveys. People who have e mail are different from those who do not, even when matched on demographic characteristics, such as age and gender.
5. E mail surveys cannot be automatically skip questions or randomize question.

- **Internet or intranet (web page) survey:** Web surveys are rapidly gaining popularity. They have major speed, cost and flexibility advantages, but also significant sampling limitations. These limitations restrict the groups that can be studied using this technique.

Advantages:

1. Web page surveys are extremely fast. A questionnaire posted on a popular website can gather several thousand responses within a few hours. Many people who will respond to an e mail invitation to take a web survey will do so the first day, and most will do so within a few days.
2. There is practically no cost involved once the setup has been completed.
3. Pictures can be shown. Some web survey software can also show video and play sound.
4. Web page questionnaires can use complex question skipping logic, randomizations and other features which is not possible with paper questionnaires. These features can assure better data.
5. Web page questionnaires can use colors, fonts and other formatting options not possible in most email surveys.
6. A significant number of people will give more honest answers to questions about sensitive topics, such as drug use, sex, when giving their answers to a computer, instead of to a person or on paper.
7. On an average, people give longer answers to open-ended questions on web page questionnaires than they do on other kinds of self-administered surveys.

Disadvantages:

1. Current use of the internet is far from universal. Internet surveys do not reflect the population as a whole. This is true even if a sample of internet users is selected to match the general population in terms of age, gender and other demographics.
2. People can easily quit in the middle of a questionnaire. They are not as likely to complete a long questionnaire on the web as they would be if taking with a good interviewer.
3. Depending on your software, there is often no control over people responding multiple times to bias the results.

- **Mailed questionnaire:** Mailed questionnaire is a paper questionnaire, which is sent to select respondents to fill and post filled questionnaire back to the researcher.

Advantages:

1. Easier to reach a larger number of respondents throughout the country.
2. Since the interviewer is not present face to face, the influence of interviewer on the respondent is eliminated.
3. This is the only kind of survey one can do if he/she has the names and addresses of the target population, but not their telephone numbers.

identified. The data collection methods will depend on the exposure, outcome and study setting, but include questionnaire and interviews as well as medical examinations. Routine data sources may also be used.

4. **Analysis.**
5. **Calculation of prevalence rate and odds ratio.**

Advantages:

1. It is relatively inexpensive.
2. It can be conducted within a short timescale.
3. As the information is collected about disease state and exposures currently, the problem of recalling past events is less than if subjects were asked about exposures about exposures and disease state in the medium or distant past.

Disadvantage:

1. There is lack of information on temporality.
2. Recalls bias susceptibility.
3. Confounders may be unequally distributed.
4. Neyman bias.
5. Group sizes may be unequal.
6. Establishes association at most, not causality.

Uses of cross-sectional study design:

1. Cross sectional studies are used to estimate the prevalence of disease or the prevalence of exposure to risk factors or both.
2. Often used to study conditions that are relatively frequent with long duration of expression (nonfatal, chronic conditions).
3. It measures prevalence, not incidence of disease.
4. It is not suitable for studying rare or highly fatal diseases or a disease with short duration of expression.
5. It is useful in studying disease etiology.
6. It is used to plan for healthcare utilization and resource allocation.
7. It is used to describe distribution pattern of the disease in population.

Design of a cross sectional study