Block 6: Causal research design – experimentation

(Activity solutions can be found at the end of the document.)

We extend our previous discussion of causal research design by identifying the necessary conditions for **causality**, examining the role of validity in experimentation and consider **extraneous variables** and procedures for controlling them. A classification of experimental designs is considered.

Learning Objectives

- explain the concept of causality as defined in market research and distinguish between the ordinary meaning and the scientific meaning of causality
- discuss the various extraneous variables that can affect the validity of results obtained through experimentation and explain how the researcher can control extraneous variables
- describe and evaluate experimental designs and the differences among pre-experimental, true experimental, quasi-experimental and statistical designs.

Reading List

Malhotra, N.K., D. Nunan and D.F. Birks. Marketing Research: An Applied Approach. (Pearson, 2017) 5th edition [ISBN 9781292103129] Chapter 11.

6.1 Causal research design - experimentation

For each section of *Causal research design - experimentation*, use the LSE ELearning resources to test your knowledge with the Key terms and concepts flip cards.

Causality and core concepts

Experimentation is commonly used to infer causal relationships. We begin by distinguishing between the ordinary meaning and scientific meaning of causality.

Ordinary meaning	**Scientific meaning**
X is the only cause of Y.	X is only one of a number of possible causes of Y.
X must always lead to Y.	The occurrence of X makes the occurrence of Y more probable (X is a probabilistic cause of Y).
It is possible to prove that X is a cause of Y.	We can never prove that X is a cause of Y. At best, we can infer that X is a cause of Y.

Concomitant variation is the extent to which a cause, X, and an effect, Y, occur together or vary together in the way predicted by the hypothesis under consideration. The time order of occurrence condition states that the causing event must occur either before or simultaneously with the effect; it cannot occur afterwards. The absence of other possible causal factors means that the factor or variable being investigated should be the only possible causal explanation.

Independent variables are variables or alternatives which are manipulated and whose effects are measured and compared - for example, price levels.

Test units are individuals, organisations or other entities whose response to the independent variables or treatments is being examined - for example, consumers or stores.

Dependent variables are the variables which measure the effect of the independent variables on the test units - for example, sales, profits and market shares.

Extraneous variables are all variables other than the independent variables which affect the response of the test units - for example, store size, store location and competitive effort.

An **experimental design** is a set of procedures specifying:

- the test units and how these units are to be divided into homogeneous subsamples
- which independent variables or treatments are to be manipulated
- which dependent variables are to be measured
- how the extraneous variables are to be controlled.

We define the following set of symbols commonly used in market research:

- X = the exposure of a group to an independent variable, treatment or event, the effects of which are to be determined
- O = the process of observation or measurement of the dependent variable on the test units or group of units
- R = the random assignment of participants or groups to separate treatments.

Internal validity refers to whether the manipulation of the independent variables or treatments actually caused the observed effects on the dependent variables. Control of extraneous variables is a necessary condition for establishing internal validity.

External validity refers to whether the cause-and-effect relationships found in the experiment can be generalised. To what populations, settings, times, independent variables and dependent variables can the results be projected?

Extraneous variables

Extraneous variables can be classified as follows.

- **History** (H) refers to specific events which are external to the experiment but occur at the same time as the experiment.
- **Maturation** (MA) refers to changes in the test units themselves which occur with the passage of time.
- **Testing effects** are caused by the process of experimentation. Typically, these are the effects on the experiment of taking a measure on the dependent variable before and after the presentation of the treatment. The **main testing effect** (MT) occurs when a prior observation affects a later observation. In the **interactive testing effect** (IT), a prior measurement affects the test unit's response to the independent variable.
- **Instrumentation** (I) refers to changes in the measuring instrument, in the observers or in the scores themselves.

- **Statistical regression** (SR) effects occur when test units with extreme scores move closer to the average score during the course of the experiment.
- **Selection bias** (SB) refers to the improper assignment of test units to treatment conditions.
- **Mortality** (MO) refers to the loss of test units while the experiment is in progress.

Extraneous variables can be controlled for in the following ways.

- **Randomisation** refers to the random assignment of test units to experimental groups by using random numbers. Treatment conditions are also randomly assigned to experimental groups.
- **Matching** involves comparing test units on a set of key background variables before assigning them to the treatment conditions.
- **Statistical control** involves measuring the extraneous variables and adjusting for their effects through statistical analysis.
- **Design control** involves the use of experiments designed to control specific extraneous variables.

A classification of experimental designs

Pre-experimental designs do not employ randomisation procedures to control for extraneous factors: the one-shot case study, the one-group pre-test-post-test design, and the static group.

In **true experimental designs**, the researcher can randomly assign test units to experimental groups and treatments to experimental groups: the pre-test-post-test control group design, the post-test-only control group design, and the Solomon four-group design.

Quasi-experimental designs result when the researcher is unable to achieve full manipulation of scheduling or allocation of treatments to test units but can still apply part of the apparatus of true experimentation: time series and multiple time series designs.

Statistical designs are a series of basic experiments which allows for statistical control and analysis of external variables: randomised block designs, Latin square designs and factorial designs.

Figure 11.1 of the textbook provides a classification of experimental designs.

A **one-shot case study** can be represented symbolically as:

$$X_1 \quad O_1.$$

A single group of test units is exposed to a treatment X. A single measurement on the dependent variable is taken (O_1). There is no random assignment of test units. The one-shot case study is *more appropriate for exploratory research than for conclusive research*.

A **one-group pre-test-post-test design** can be represented symbolically as:

$$O_1 \quad \mathrm{X} \quad O_2$$

A group of test units is measured twice. There is no control group. The treatment effect is computed as O2−O1O2−O1. The validity of this conclusion is questionable since extraneous variables are largely uncontrolled.

In a **static group**, we have:

$$\text{EG:} \quad X \quad O_1 \quad \text{and} \quad \text{CG:} \quad O_2$$

A two-group experimental design. The experimental group (EG) is exposed to the treatment, and the control group (CG) is not. Measurements on both groups are made only after the treatment. Test units are not assigned at random. The treatment effect would be measured as $O_1 - O_2$.

In a **pre-test-post-test control group design**, we have:

$$\text{EG: } R \quad O_1 \quad X \quad O_2 \quad and \quad CG: \quad R \quad O_3 \quad O_4$$

Test units are randomly assigned to either the experimental or the control group. A pre-treatment measure is taken on each group. The treatment effect (TE) is measured as:

$$(O_2 - O_1) - (O_4 - O_3).$$

Selection bias is eliminated by randomisation. The other extraneous effects are controlled as follows:

$$O_2 - O_1 = \text{TE + H + MA + MT + IT +I + SR +MO}$$

and:

$$O_4 - O_3 = \text{H + MA + MT +I + SR +MO}$$

The experimental result is obtained by:

$$(O_2 - O_1) - (O_4 - O_3) = \text{TE+IT.}$$

Interactive testing effect is not controlled.

In a **post-test-only control group design**, we have:

$$\text{EG: } R \quad X \quad O_1 \quad \text{and} \quad CG: \quad R \quad O_2$$

The treatment effect is $O_2 - O_1 = TE$. Except for pre-measurement, the implementation of this design is very similar to that of the pre-test-post-test control group design.

The **Solomon four-group design** explicitly controls for interactive testing effects, in addition to controlling for all the other extraneous variables.

Quasi-experimental and statistical designs

A **time series design** involves a series of periodic measurements:

$$O_1 O_2 O_3 O_4 O_5 O_6 O_7 O_8 O_9 O_{10}$$

There is no randomisation of test units to treatments. The timing of treatment presentation, as well as which test units are exposed to the treatment, may not be within the researcher's control.

In a **multiple time series design**, we have:

$$\text{EG: } O_1 O_2 O_3 O_4 O_5 O_6 O_7 O_8 O_9 O_{10}$$

and:

$$\text{CG: } O_{11} O_{12} O_{13} O_{14} O_{15} O_{16} O_{17} O_{18} O_{19} O_{20}$$

If the control group is carefully selected, this design can be an improvement over the simple time series experiment. We can test the treatment effect twice: against the pre-treatment measurements in the experimental group and against the control group.

Statistical designs consist of a series of basic experiments which allow for statistical control and analysis of external variables and offer the following advantages.

- The effects of more than one independent variable can be measured.
- Specific extraneous variables can be statistically controlled.
- Economical designs can be formulated when each test unit is measured more than once.

The most common statistical designs are the **randomised block design**, the **Latin square design**, and the **factorial design**.

A randomised block design is useful when there is only one major external variable, such as sales or store size, which might influence the dependent variable. The test units are blocked, or grouped, on the basis of the external variable. By blocking, the researcher ensures that the various experimental and control groups are matched closely on the external variable. Table 11.4 of the textbook shows an example of a randomised block design.

The Latin square design allows the researcher to statistically control two non-interacting external variables as well as to manipulate the independent variable. Each external or blocking variable is divided into an equal number of blocks, or levels. The independent variable is also divided into the same number of levels. A Latin square is conceptualised as a table with the rows and columns representing the blocks in the two external variables. The levels of the independent variable are assigned to the cells in the table. The assignment rule is that each level of the independent variable should appear only once in each row and each column. Table 11.5 of the textbook shows an example of a Latin square design.

A factorial design is used to measure the effects of two or more independent variables at various levels. A factorial design may also be conceptualised as a table. In a two-factor design, each level of one variable represents a row and each level of another variable represents a column.

Limitations of experimentation

Experiments can be time consuming, particularly if the researcher is interested in measuring the long-term effects.

Experiments are often expensive. The requirements of experimental group, control group and multiple measurements significantly add to the cost of research.

Experiments can be difficult to administer. It may be impossible to control for the effects of the extraneous variables, particularly in a field environment.

Competitors may deliberately contaminate the results of a field experiment.

Questions, solutions and case study

To access the solutions to these questions and case study, click here to access the printable Word document or click here to go to LSE's Elearning resources.

Questions on the block’s topics

1. What are the requirements for inferring a causal relationship between two variables?
2. Differentiate between internal and external validity.
3. List any six extraneous variables and give an example to show how each can reduce internal validity.
4. Describe the various methods for controlling extraneous sources of variation.
5. What is the key characteristic which distinguishes true experimental designs from pre-experimental designs?
6. List the steps involved in implementing the post-test-only control group design.
7. What is a time series experiment? When is it used?
8. How is a multiple time series design different from a basic time series design?
9. What advantages do statistical designs have over basic designs?
10. What are the limitations of the Latin square design?
11. Compare the characteristics of laboratory and field experimentation.
12. Should descriptive research be used for investigating causal relationships? Why or why not?
13. What is test marketing? What are the three types of test marketing?
14. What is the main difference between a standard test market and a controlled test market?
15. Describe how simulated test marketing works.

Activities on the block’s topics

Activity 1:

You are a research and insight manager for Louis Vuitton. The company would like to determine whether it should increase, decrease or maintain the current spend level of advertising. Design a field experiment to address this issue.

Activity 2:

What potential difficulties do you see in conducting the experiment above? To what extent could the Louis Vuitton management help you overcome these difficulties?

Activity 3:

Red Bull has developed three alternative package designs to replace its current can design. Design an online-based experiment to determine which, if any, of these new package designs is superior to the current one.

Discussion points

‘Is it possible to prove causality in any aspects of consumer behaviour?’

Issues which can be discussed include the role of inference in drawing conclusions, the difficulty of isolating all possible sources of variation in social science experimentation, the historic value of experimentation in deriving theories and inferring relationships and the nature or definition of science, (i.e. is science only limited to proofs, which account for all sources of variation?).

‘The potential to electronically observe consumer buying behaviour using the Internet has created great growth potential for the application of experimental techniques.’

This refers to the use of cookie technology to be able to observe the behaviour of computer users and how they relate to material presented on different websites. The discussion should develop around the nature of behaviour that could be observed (dependent variables) and the manipulation of independent variables.

Block 6: Causal research design – experimentation

Solutions to Questions on the block's topics

1. There are three requirements or conditions, which must be satisfied before a causal relationship can be inferred. The three conditions are as follows.

    - Concomitant variation of the two variables, i.e. the extent to which a cause, XX, and an effect, YY, occur together or vary together in the way predicted by the hypothesis under consideration.

    - Time order of occurrence of variables, i.e. the causing event occurs either before or simultaneously with the effect; it cannot occur afterwards.

    - Elimination of other possible causal factors, i.e. the absence of other possible causal factors means that the factor under investigation should be the only possible causal explanation.

2. Internal validity is concerned with the question of whether the observed effects on the dependent variables have been caused by variables other than the treatment. The control of extraneous variables is necessary in establishing internal validity. External validity, on the other hand, is concerned with whether the cause-effect relationship can be generalised. More often than not a trade-off is required between internal and external validity. Whilst the level of internal validity can be raised in an artificial environment, this reduces the generalisability of the results and, consequently, the external validity.

3. Examples of extraneous variables which can reduce validity are as follows.
    - History: for example, a firm is measuring the effect of a new price level when a rival firm starts a promotional campaign.
    - Testing: for example, when measuring the effect of a celebrity in an advertisement, a pre-treatment questionnaire biases responses on the post-treatment questionnaire.
    - Selection bias: for example, while assessing the market size for a new soft drink, the sample was drawn only from households in the upper 50% income bracket.
    - Mortality: for example, 20% of participants failing to respond to the post-treatment questionnaire.
    - Instrumentation: for example, in the case of a pre-test-post-test control group design, using different formats (which could even be as simple as the size and font used as well as the nature and
    - Statistical regression: for example, in an experiment, suppose two measurements are made of a consumer's attitude toward a test product, one before tasting it and one afterwards. If extreme scores or ratings of the product move toward the average, regression has occurred.

4. The various methods for controlling extraneous sources of variation are as follows.
    - Randomisation. This involves randomly assigning test units to experimental groups by using random numbers. Treatment levels are also randomly assigned. Such random assignment usually results in an equal representation of extraneous variables in each treatment condition.
    - Matching. In this case, the test units are matched on a set of key characteristics before being assigned to the treatment conditions.

- Statistical control. This involves the use of statistical methods such as ANCOVA, which removes the effect of extraneous variables by adjusting the mean values of dependent variables within each treatment condition.
- Design control. Extraneous variables can also be controlled by using specific experimental designs.

5. The distinguishing feature of the true experimental design, as compared to the pre-experimental design, is randomisation. In true experimental designs, the researcher can randomly assign test units to experimental groups and also randomly assign treatments to experimental groups. Randomisation is the hallmark of objectivity.

6. Describe the design symbolically. The steps involved in implementing the post-test-only control group design are as follows.
   - Select a sample of participants at random.
   - Randomly assign the participants to two groups (i.e. experimental group and control group).
   - Participants in the experimental group would then be exposed to a treatment.
   - Post-treatment measurements are obtained from both groups using a measurement instrument like a questionnaire.

   The design is described symbolically as follows:

   - EG: $R\ X\ O_1$
   - CG: $R\ \ O_2$.

   The treatment effect is given by TE: $R\ \ O_1 - \ O_2$.

7. A time series experiment is a quasi-experimental design. It involves periodic measurement on the dependent variable for a group of test units. Next, the treatment is administered by the researcher or occurs naturally. After the treatment, periodic measurements are continued in order to determine the treatment effect. A time series experiment may be symbolically described as:

$$O_1 O_2 O_3 O_4 O_5 O_6 O_7 O_8 O_9 O_{10}$$

   This design is used in cases where the researcher lacks control over the scheduling of the treatment (for example, a sales promotion effort by a competitor) and the ability to randomly expose test units to the treatment but can control when measurements are taken and on whom they are taken.

8. The time series design involves a series of periodic measurements on the dependent variable for a group of test units. The treatment is then administered by the researcher or occurs naturally. After the treatment, periodic measurements are continued to determine the treatment effect. The multiple time series design is similar to the time series design except that another group of test units is added to serve as a control group. If the control group is carefully selected, this design can be an improvement over the simple time series experiment. The improvement lies in the ability to test the treatment effect twice: against the pre-treatment measurements in the experimental group and against the control group.

9. Statistical designs consist of a series of basic experiments which allow for statistical control and analysis of external variables. In other words, several basic experiments are conducted simultaneously. Therefore, statistical designs are influenced by the same sources of invalidity which affect the basic designs being used. The advantages statistical designs have over basic designs are as follows.

    - In the case of basic designs, the effects of only one independent variable or treatment can be measured, while in the case of a statistical design, the effects of more than one independent variable can be measured.
    - In the case of statistical designs, specific extraneous variables can be statistically controlled.
    - Economical designs can be formulated when each test unit is measured more than once.

10. A Latin square design allows the researcher to control statistically two non-interacting external variables as well as to manipulate the independent variable. Each external or blocking variable is divided into an equal number of blocks or levels. The independent variable is also divided into the same number of levels. The limitations of the Latin square design are as follows.

    - Such designs require an equal number of rows, columns and treatment levels.
    - Only two external variables can be controlled simultaneously.
    - They do not allow interactions of the external variables with each other or with the independent variable.

11. A laboratory experiment is one where the researcher creates an artificial environment for the purpose of experiment. Therefore, s/he has a great degree of control over extraneous variables. On the other hand, a field environment is one where the experiment is conducted under actual or natural conditions.

    Laboratory experiments have the following advantages over field experiments. The laboratory environment offers a high degree of control because it isolates the experiment in a carefully monitored environment. Therefore, the effects of history can be minimised. A laboratory experiment also tends to produce the same results if repeated with similar subjects, leading to high internal validity. Laboratory experiments tend to use a small number of test units, last for a shorter time, be more restricted geographically and are easier to conduct than field experiments. Hence they are generally less expensive as well.

    Compared with field experiments, laboratory experiments suffer from some main disadvantages. First, the artificiality of the environment may cause reactive error in that the participants react to the situation itself rather than to the independent variable. Also, the environment may cause demand artefacts, a phenomenon in which the participants attempt to guess the purpose of the experiment and respond accordingly. For example, while viewing the film clip, participants may recall pre-treatment questions about the brand and guess that the commercial is trying to change their attitudes toward the brand. Finally, laboratory experiments are likely to have lower external validity than field experiments. Because a laboratory experiment is conducted in an artificial environment, the ability to generalise the results to the real world may be diminished. It has been argued that artificiality or lack of realism in a laboratory experiment need not lead to lower external validity. One must be aware of the aspects of the laboratory experiment which differ from the situation to which generalisations are to be made.

12. Descriptive research should not be used for establishing causal relationships for the following reasons:
    - It is difficult to establish the prior equivalence of the groups with respect to both the independent and dependent variables in descriptive research.
    - It is difficult to establish the time order of occurrence.
    - It provides little control in eliminating other possible causes.

13. Test marketing is an application of a controlled experiment done in a limited but carefully selected part of the marketplace called test markets. It involves a replication of a planned national marketing programme for a product in the test markets. The three types of test markets are as follows.

    Standard test market. The product is sold through regular distribution channels and, typically, the company's own salesforce is responsible for distributing the product. It involves a one-shot case study.

    Controlled test market. The entire test-marketing programme is conducted by an outside research company, which also handles distribution and field sales operations in the test market.

    Simulated test market. Such markets are also called laboratory tests and yield mathematical estimates of market share based on initial reaction of consumers to the new product.

14. In a standard test market, test markets are selected and the product is sold through regular distribution channels. Typically, the company's own salesforce is responsible for distributing the product. Sales personnel stock the shelves, restock and take inventory at regular intervals. In a controlled test market, the entire test-marketing programme is conducted by an outside research company. The research company guarantees distribution of the product in retail outlets which represent a predetermined percentage of the market. It handles warehousing and field sales operations, such as stocking shelves, selling and stock control.

15. Also called a laboratory test or test market simulation, a simulated test market yields mathematical estimates of market share based on the initial reaction of consumers to a new product. In a simulated test market, participants are intercepted in high-traffic locations such as shopping centres and are pre-screened for product usage. The selected individuals are exposed to the proposed new product concept and given an opportunity to buy the new product in a real-life or laboratory environment. Those who purchase the new product are interviewed at a later date to determine their evaluation of the product and repeat purchase intentions. The trial and repeat purchase estimates so generated are combined with data on proposed promotion and distribution levels to obtain a projected share of the market.

Solutions to exercises on the block's topics

Exercise 1:

One of several designs can be used for this field experiment. A pre-test-post-test control group design is outlined below.

- Select three sets of test markets with similar demographic profiles, competitive environments and sales volume of Louis Vuitton products.
- Randomly increase advertising budgets in one set of markets, decrease them in a second set and maintain them in the third set of markets.
- Track sales volumes for three months in each set of markets.
- Analyse the sales results to determine differences in performance.

Exercise 2:

Potential difficulties include access to sales data, selecting three similar markets to test, changing competitive dynamics and objections from retailers in the test market where advertising expenditures will decrease. Management can assist by ensuring corporate compliance for gathering data, allocating the necessary funds, pacifying affected retailers.

Exercise 3:

The existing and the three new packages can be displayed, one at each specific location on the internet. Participants can be recruited and randomly assigned to each package location (URL), asked to view the package and answer several attitudinal and behavioural intent questions. The package with the most favourable attitude and/or the highest behavioural intent can then be identified.