EXPERIMENTAL RESEARCH

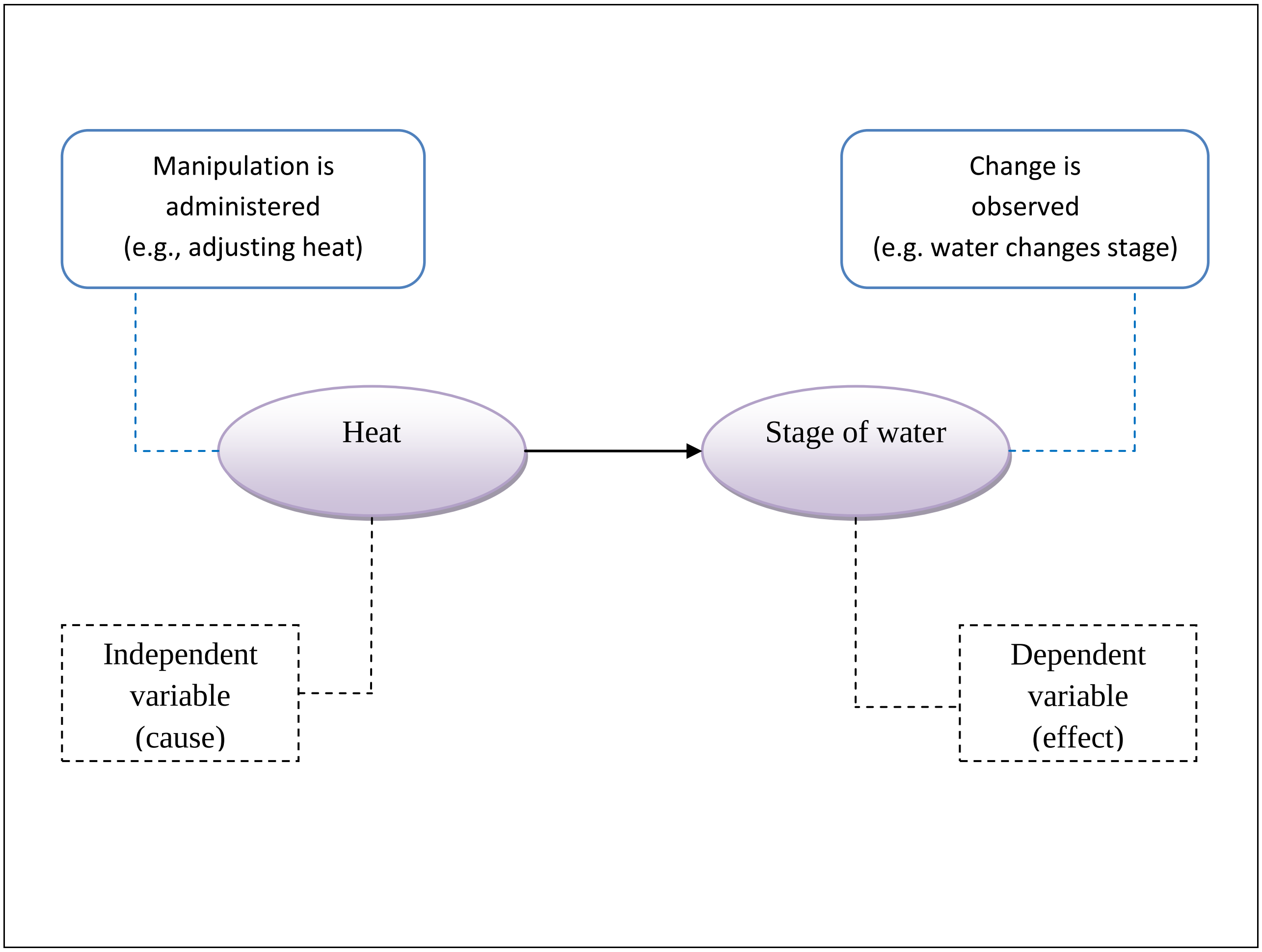
Experiment

In the previous chapter we already learned about a hypothesis, which is a statement that predicts the relationship between concepts. We also learned that a hypothesis is a statement that needs verification; we need to test it using observable data. In this chapter, we will focus on the method that researchers perform to test a hypothesis. This method is called an "*experiment*".

When talking about an experiment, you may think back to the time when you took a sciences class in high school. Most of you must have a chance to conduct some experiments back then. Still, if you can't think of what is an experiment, try to think of some Sci-Fi movies; imagine the scientist was trying to mix some chemical together to see what would happen...and then…boom!! Well, an experiment in social science research seems not too different from the experiment in the Sci-Fi movies, except for that we don't actually deal with the explosive chemical.

In research, an *experiment* is defined as the orderly procedure conducted by the researchers with the main goal to verify or to refute the validity of a hypothesis. In other words, it is a procedure that the researchers perform to test a hypothesis; it is implemented for empirical testing. Generally, when researchers conduct an experiment, they aim to discover the cause-and-effect of phenomenon by determining what will happen when something is manipulated.

Manipulations in an experiment can be administered in many ways. For example, the scientists may introduce, increase, decrease or eliminate something in the environment to see what effect will happen to the phenomenon of interest. When something is manipulated, we can say that it receives *experimental treatment*. For example, if the scientists want to know whether the change in temperature will affect the form of water, they can manipulate the temperature by adding more heat to water to see what will happen. The experimental treatment in this case is heat. After the experimental treatment was administered, the researchers then observe the change that will occur to the phenomenon of interest. If nothing happens, they can conclude that the experimental treatment does not bring about an effect; thus, the cause-and-effect relationship does not exist. But when they observe some change occurs, they can confirm the existence of the cause-and-effect relationship.



For example, when the scientists increased heat to water and they observed that water began to evaporate, they can conclude that heat causes water to change form. But if they observed nothing happens to water no matter how much heat was increased, they conclude that there is no causal relationship between heat and states of water. Note that the factor that was manipulated in experiment can be called an "*independent variable*", whereas the factor that we expect change is called a "*dependent variable*".

PURE EXPERIMENT VS QUASI-EXPERIMENT

An experimental research can be classified into two types: (1) pure experimental research, and (2) quasi-experimental research. Basically, the main difference between these two types of experiment is how the experimental subjects are recruited. First, let's start from pure experimental research.

Pure experimental research

For an experiment in social sciences whether it is pure experimental research or quasi-experimental research, researchers have to work with "*subjects*". Subjects are basically the persons whom the researchers choose to participate in an experiment. By ethical standard, the subjects have to participate voluntarily; sometime they may also be compensated as well. After the subjects are obtained for an experiment, they will be divided into two groups. The first group is called the "*experimental group*"; and the second group is called the "*control group*". In particular, the experimental group is the group that will receive experimental treatment; but for the control group, no experimental treatment will be given to them.

Note that for pure experiment, subjects have to be *randomly* assigned to the experimental group and the control group. The *random assignment* is used in this case to guarantee that every subject will have equal chance to be assigned to both groups. This method also makes sure that the outcomes from the experiment will mainly come from experiment treatment, not from characteristics of individuals. Random assignment can be performed in many ways. For example, the researchers may wrote down the letter "E" and "C" on a small piece of paper separately and then ask each subject to draw one from the box. Those who pick the letter E will be assigned to the experimental group, whereas those who pick the letter C will be assigned to the control group.

Anyway, why do we need the control group? Well, the main reason is to make sure that the effect that we observe is actually caused by the experimental treatment. Because the experimental group is a group that receives experimental treatment, we expect something to happen to them. On the other hand, because the control group does not receive any treatment, so we expect nothing significant happen to them. When we compare the outcomes between the experimental group and the control group, the effect of the experimental treatment is simply the difference in the outcomes that we observe between these two groups.

Net effect of experimental treatment

= Outcome from the experimental group – Outcome from the control group

In order to have a clear understand about the experiment, let's have some basic example. A teacher wanted to know whether more homework assignments would improve class performance of her students. In fact, she personally believed that more homework assignments would help improve class performance because it allowed student to practice what they had learned from class. Based on her hypothesis that more homework assignments would lead to higher class performance, she set up an experiment to test this hypothesis. At the beginning of the semester, she recruited students who enrolled in her class to be the experimental subjects for her study. She randomly assigned these students to the experiment group and the control group. Her experimental treatment in this case is the load of homework assignments. Of course, those who were in the experimental group would receive the experimental treatment. In this case, students who were in the experimental group received more load of homework assignment. On the other hand, students who were in the control group received regular load of homework assignments. After the semester ended, she conducted an exam to measure the knowledge that students learned from the class. The test scores were found as the following:

	Experimental group (more homework assignments)	Control group (regular load of homework assignments)
Students' test scores after more homework assignments were given to an experimental group	80	60

From this observed data, the effect of the experimental treatment can be calculated as the following:

$$80 - 60 = 20$$

Evidently, there is a positive effect of experimental treatment. It shows that students who were in the experimental group (more homework assignments were provided) scored higher than students who were in the control group (regular load of homework assignments were provided). From this finding, she may conclude that more homework assignments actually helped students improve class performance. But still, before she decided to give more homework assignments to students in her next class, there is something that she was skeptical.

Pretest-posttest

From the example of the experiment shown earlier, although students were randomly assigned to the experimental group and the control group, there is still some bias that may present in the result that the teacher found. Remember that the random assignment is the method that minimizes the bias, but does not completely eliminate it. There is still a chance that the majority of smart students were assigned to the experimental group by chance. In order to avoid this issue, it is important to measure the initial outcome of the subjects in both groups before and after the experimental treatment is given. In this regard, the researchers must perform the *pretest/posttest* approach.

In the "*pretest*", we measure the initial outcome of the subjects in the experimental group and the control group before an experimental treatment is administered. After we obtain the initial outcome, then the experimental treatment can be administered to the experimental group. After the experimental treatments are administered, we measure the outcome of the subjects in both groups again. The outcome measurement after the experimental treatment is administered is called "*posttest*".

In order to obtain the net effect of the experimental treatment using pretest/posttest, first we have to subtract the pretest outcomes from the posttest outcomes of each group. To get the net effect of the experimental treatment, we then subtract the difference of the control group from the difference of the experimental group. The formula is shown in the box below.

Net effect of the experimental treatment using pretest/posttest

$$= (\text{Posttest}_E - \text{Pretest}_E) - (\text{Posttest of}_C - \text{Pretest}_C)$$

Where E=experimental group; C=control group

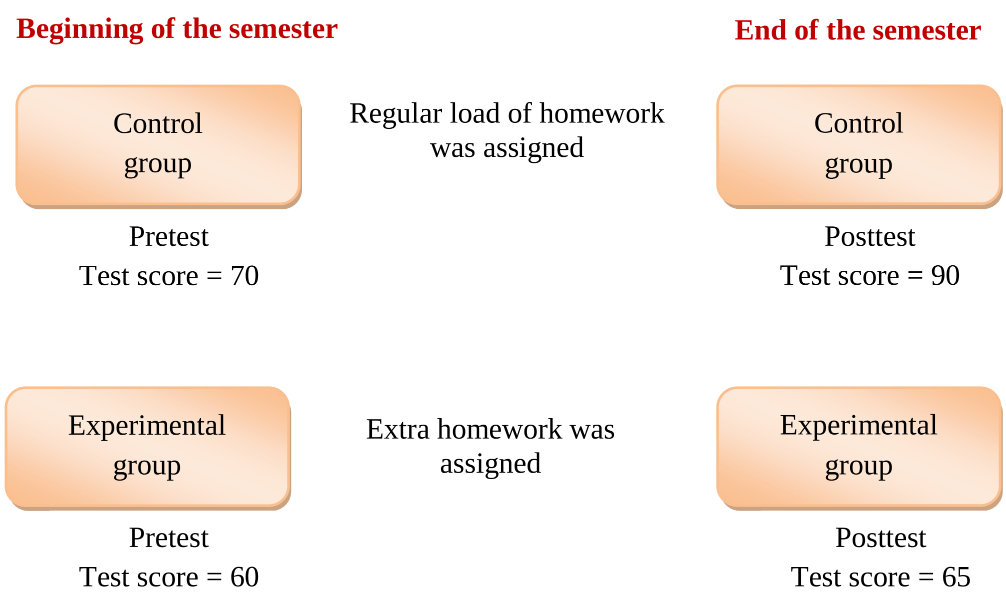
Referring to the previous example, a teacher decided to conduct the experiment again by using the pretest/posttest approach. This time she measured students' performance in both groups at the beginning of the semester as a pretest. Then more home assignments were given to students in the experimental group. After the semester had ended, she measured students' performance again as a posttest. The followings are the data that she obtained from evaluating students' performance before and after the experimental treatment was administered:

	Experimental group (more homework assignments)	Control group (regular load of homework assignments)
Before more homework assignments were given (Pretest)	70	60
After more homework assignments were given (Posttest)	90	65

The net effect of the experimental treatment using pretest/posttest approach can be calculated as the following:

$$
\begin{aligned}
&(90-70) - (65-60) \\
&= 20-5 \\
&= 15
\end{aligned}
$$

An example of pretest/posttest approach in an experiment



From this result, it shows a consistent result that the group that received extra homework assignment demonstrated higher improvement in class performance than the group that had regular load of homework assignment. Thus, her hypothesis was supported. This may be a bad news for the future students in the class because the teacher finally confirmed that more homework assignment have a positive effect on students' performance.

Quasi-experiment

Literally, the word "*quasi-*" means resembling or seeming. Thus, quasi-experiment research can be simply understood as the experiment that is quite similar to pure experiment research. In fact, the only difference between pure experiment research and quasi-experimental research is that, for quasi-experimental research, subjects are not randomly assigned to the experimental group like in the case of pure experimental research. Instead, subjects are selected and assigned to the experimental group based on some predefined criterion that makes them suitable to be the experimental subjects.

In particular, quasi-experiment is normally conducted when the experimental treatment cannot be randomly assigned to subjects. But what are the situations that make random assignment a difficult task in an experiment? Let's consider this example. If you want to study the effect of "amphetamine" on work productivity of employees, what will you do if you want to conduct an experiment to test it? Alright, if you decide to use pure experimental design like what we learned earlier, first you have setup the experimental group and the control group and then randomly assign the subjects to both groups. After the subjects are assigned to both groups, it is time to deliver the experimental treatment to people in the experimental group. In this case, the subjects who are in the experimental group will be given amphetamine to take, and then we observe what will happen to their work productivity, as compared to the subjects who are in a control group.

But wait!! Can you actually do that for real, giving amphetamine to the experimental subjects? Ethically and legally, the answer is definitely no. Amphetamine is considered the extremely addictive substance that can cause devastating impacts to those who abuse it. In this case, pure experiment design is not feasible. However, there is a solution for that. In the medical field, amphetamine is a medication used to treat narcolepsy and attention deficit disorder with hyperactivity (ADHD). It is a highly controlled drug that can still be prescribed to patients under close supervision of a doctor. By using quasi-experiment design, the researcher may recruit people who are currently prescribed amphetamine by a doctor to be the experimental subjects.

In this regard, quasi-experimental research can offer more advantage then pure experimental research mainly because it does not require randomization; thereby making it is easier to setup than pure experiment. Also, using quasi-experiment provides a chance for the researchers to tap on the issues that are susceptible to ethical and legal concerns (DeRue, 2012), like in the example of the amphetamine study discussed earlier. Nonetheless, one major disadvantage of quasi-experiment is that it tends to lack internal validity.

Internal validity and external validity in an experimental research

In experimental research, the main issue that the researchers concern is the validity of the results. In particular, there are two aspects of validity: internal validity and external validity.

<u>Internal validity</u>

Internal validity concern whether the result obtained from an experiment is actually caused by the experimental treatment. If the effect is actually caused by experimental treatment, we can warrant the internal validity of the results. Conversely, when it turns out that the effect is caused by other factor that is not the experimental treatment, the internal validity of the research finding is compromised.

There are many factors that can cause the results from an experiment to lack internal validity. One issue is the *placebo effect*. The term placebo effect is first introduced to the clinical field by Henry K. Beecher (1955); since then, it has been widely used in the clinical studies on the effect of a drug or medical treatment on a particular condition of patients (Kienle & Kiene, 1997). A placebo is generally a fake drug that looks exactly like a real drug, except for that it does not contain any active ingredient meant to affect health conditions. In the clinical experiment on the effect of a drug, usually the subjects are divided into three groups: the first group is given real medication, the second group is given a placebo, and the third group receives nothing. In the end of the experiment, health conditions among these three groups are compared. If a drug actually affects health conditions, the results must reveal a significant improvement in health conditions of the subjects in the first group, but not the second group and a third group. More importantly, a placebo should not have any effect on health conditions of the subjects. Nonetheless, some clinical studies found that a placebo actually improved health conditions in patients. In particular, the rationale that explains the health improvement caused by the placebo effect is that some people may react to a medication treatment not because of the active ingredient in a drug, but the reaction simply comes from own psychological expectations that taking a drug will benefit health. Therefore, when the placebo effect is detected in an experiment, internal validity of the result is questionable.

Internal validity is confirmed

	Group 1 (real drug)	Group 2 (no drug)	Group 3 (placebo)
Health conditions after the treatment	**Improve**	Stable	Stable

Internal validity is questionable

	Group 1 (real drug)	Group 2 (no drug)	Group 3 (placebo)
Health conditions after the treatment	**Improve**	Stable	**Improve**

Another issue that can jeopardize the internal validity of the experimental research is known as the "*Hawthorne effect*" or "*the observer effect*" (Parsons, 1974; Wickström & Bendix, 2000). The term Hawthorne effect was first introduced to the academic field by a researcher named Henry A. Landsberger who was commissioned to conduct the experiments at the Hawthorne works electric company (located near Chicago, Illinois) during 1924 and 1932.

At that time, the team of researchers was interested to study whether productivity of workers was affected by work environment. One particular aspect of work environment is the amount of light. The experiment was administered by a researcher named Elton Mayo. He tried to manipulate the amount of light within the workplace by switching the light from low to high, and from high to low, to see the reactions of workers. At first, the amount of light within the workplace was increased for workers who were in the experimental group. The research team witnessed that workers increased the level of productivity as a result. Thus, they initially concluded that more light led to higher productivity. The, the opposite was performed in the experiment. This time, the amount of light within the workplace was reduced. Nonetheless, the level of productivity of workers still increased. This result surprised the research team. No matter the light was switched from low to high and from high to low, productivity still increased. In the end, the research team tried to come up with some logical reason to explain this perplexing finding. One possible explanation is that the reason why workers increased productivity is not because of the experimental treatment (which is the amount of light), but it occurred because workers were aware that the team of researchers were observing them. This circumstance caused them to alter their behaviors (increased productivity) just to impress researchers.

From what happened at the Hawthorne factory, the term "Hawthorne effect" is then used afterward in research to describe the situation when subjects tend to perform or act differently from what they should have done when they know that they participate in an experiment. Thus, when the Hawthorne effect presents in the experiment, the internal validity of the result is compromised.

<u>External validity</u>

External validity concern whether the results that are obtained from the experiment can be generalized to other situations or to other people. It concerns the extent to which the results obtained from a group of people in the experiment can be inferred to a larger population. One particular question to be asked to confirm external validity of the results is that if we conduct the same experiment in other group of people, will it still produce consistent results. Referring to the example of the experiment about the effect of homework assignments on class performance, if the experiment was conducted using students in one school as the experimental subjects, to what extent a researcher can ascertain that the results will be generalized to students in other schools? Is it possible that more homework assignments only increases class performance of students in one school, but will not improve class performance of students in other schools? In particular, when the result obtained from one setting can be generalized to other settings, we can conclude that the research exhibit external validity.

Generally, external validity problem can be a major concern when the sample used in the experiment is not a true representative of a population of interest. Some example of research that is susceptible to this problem is the study about the effects of workplace factors on job productivity that use experimental subjects who are students who don't actually work. Because the objective of this study is to investigate the work-related factors, using the student subjects who do not belong to the population of working people may cause the external validity of the results questionable. Therefore, it is important for the researchers to select the subjects who actually belong to the population of interest to avoid external validity problem. Also, using the random method to select the subjects can lessen the sampling bias that might affect external validity of the results. Finally, using a large amount of subjects also increases the chance that the results will be generalized to a larger population, thereby lowering the concern about external validity issue.

Laboratory study and field study

An experiment can be conducted in terms of laboratory research and field research. For *laboratory research* (or lab research), an experiment is performed in the artificial or close environment (or a laboratory) whereby all settings are fully controlled by the researchers. For a *field research*, on the other hand, an experiment is conducted in a natural setting or real-world environment, whereby the researchers have limit or no control over other factors in the environment.

Some laboratory research was conducted in a classroom setting. Some example of this type of laboratory research is the study by Lewis (2000) that investigated the effect of emotions that a leader displays on the perception of people about leadership effectiveness. In particular, the researcher aimed to explore whether different emotional tones that a leader express would influenced how they are evaluated by other people or not. The subjects for this experiment are undergraduate students. To simulate a situation for the subject to evaluate a leader, the researcher recruited the professional actors to take the role of a CEO who came to report the bad news that recently happen to the company; each of them reported the bad news with different emotional tones (e.g., anger, sadness). After each report ended, the researcher then asked the undergraduate students to evaluate the effectiveness of the actor. By using laboratory research, every factor in the study can be simulated and be conducted in a close setting like in a classroom. It provide convenient for the researcher have full control over the environment. If the study like this was conducted using field research, it would require the real organizational setting where the subjects can evaluate a real CEO.

Both laboratory research and field research have their own advantages and disadvantages. Although laboratory research allows researchers to have full control over the environment, the fact that the study is conducted in an artificial setting can make the validity of the results questionable. Critics may argue that the results from laboratory research may not be applied to the real-world situation. On the other hand, although the results from field research derive from the real-world situation, many factors in the environment are beyond the researchers' control tends to make it difficult to rule out the possibility that the results may be caused by other uncontrolled factors rather than the factor that the researchers are interested. Therefore, it is important for the researchers to understand the pros and cons of each method so that they can select the method that match their research objectives the most.

Cross-sectional design and longitudinal design in experimental research

An experiment can also be conducted using cross-sectional deign and longitudinal deign. For *cross-sectional design*, the researchers collect the data from multiple subjects over a single point in time. On the other hand, when *longitudinal design* is used, the data are collected from the same subjects over a different period of time.

In practice, using longitudinal design provides more advantages over cross-sectional design for several reasons. By collecting the data from the same subjects across time, researchers are able to track changes in the phenomenon of interest more accurately. This benefit is especially crucial in the situation when it takes some time for the cause to bring about the effect. Importantly, using longitudinal design also allows the researchers to confirm the causal relationship between phenomena. Referring to the rule of temporal sequence that was discussed in the early chapter, causality can be established when a factor regards as a cause occurs before a factor regards as an effect. Therefore, by using longitudinal design, the researchers can observe the cause in the current period and then observe the effect in the next period. This method allows researcher to rule out the possibility of the reverse-causality issue that might bias the interpretation of the results.

Although longitudinal design provides more advantages than cross-sectional design, it also has some weaknesses. In particular, one major limitation of longitudinal design is that it is more expensive and more time-consuming than cross-sectional design. Some longitudinal research took many years to collect the data from the same subject over time. Another difficulty of using longitudinal design is that sometime the same subjects may not be available to participate in the future data collection.

References

Beecher, H. K. (1955). The powerful placebo. *Journal of the American Medical Association, 159*(17), 1602-1606.

DeRue, S. (2012). A Quasi Experimental Study of After-Event Reviews. *Journal of Applied Psychology, 97*(5), 681–689.

Kienle, G. S., & Kiene, H. (1997). The Powerful Placebo Effect: Fact or Fiction? *Journal of Clinical Epidemiology, 50*(12), 1311-1318. doi: http://dx.doi.org/10.1016/S0895-4356(97)00203-5

Lewis, K. M. (2000). When leaders display emotion: how followers respond to negative emotional expression of male and female leaders. *Journal of Organizational Behavior, 21*(2), 221-234. doi: 10.1002/(SICI)1099-1379(200003)21:2<221::AID-JOB36>3.0.CO;2-0

Parsons, H. M. (1974). What Happened at Hawthorne?: New evidence suggests the Hawthorne effect resulted from operant reinforcement contingencies. *Science, 183*(4128), 922-932. doi: 10.1126/science.183.4128.922

Wickström, G., & Bendix, T. (2000). The "Hawthorne effect" — what did the original Hawthorne studies actually show? *Scandinavian Journal of Work, Environment & Health, 26*(4), 363-367.