






Causal Research Design- Experimentation

In a social science (such as marketing) it is very important to understand that effects (e.g., consumers responding favorably to a new buzz marketing campaign) are caused by multiple variables. The relationships between these variables (e.g., the consumer's mood, the time of day, the campaign's novelty, etc.) tends to be probabilistic (= based on probability). As a result, it is not possible to 100% conclusively prove causality. In an experiment you usually only can *infer* a cause-effect relationship.

However, to attempt to infer causality, the following three conditions must be met:

(1) **Concomitant variation**: (Concomitant = means to exist or occur with something else). This type of 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 (e.g., correlation). Evidence pertaining to concomitant variation can be obtained in a qualitative or quantitative manner. Thus, we may hypothesize that increasing product price will lead to an increase in profits. It is important to remember that association (aka correlation) does not demonstrate causation (see image below). We can use stats to show us that X changes when Y changes, but this does not mean that we know the change in X caused the change in Y. The lesson here is that the stat method we use does not determine cause and effect. Instead, the design of our data collection is much more important in helping to infer causality.

Correlation is NOT Causation!		
		
All cats have four legs.	This animal has four legs.	Therefore, this animal is a cat. false!

(2) **Time order of occurrence of variables**: the causing event must occur either before or simultaneously *with* the effect; it cannot occur afterwards. Thus, our profits cannot increase until after we increase the price. This condition is also known as the “temporal ordering of variables.” Changes in the *causal* factor (independent variable) must influence changes in the *effect* (dependent variable), but not vice versa! This is also known as the direction of influence.

(3) **Absence of other possible causal factors**: This condition means that we can infer causality if all other factors affecting the subject are absent, i.e., these factors are held constant. Thus, other factors like store penetration, expanded consumer spending, or decreased costs must be held constant. The control of other causal factors is necessary. No *spurious* findings or



Causal Research Design- **Experimentation**

competing hypotheses. That's why random assignment (as is the case in experimental design) is helpful.

It is also important to understand the language we use in experimental research.

1. **Independent variable** (IV)—variables or alternatives that are manipulated and whose effects are measured and compared, such as price levels, different package designs, and advertising themes.
2. **Test units**—individuals (i.e., “consumers” “respondents” “participants”), organizations (i.e., “companies”), or other entities whose response to the independent variables or treatments is being examined.
3. **Dependent variables** (DV)—variables that measure the effect of the independent variables on the test units, such as sales, profits, and market shares. Examples of DVs are in the Marketing Handbook of Scales ([link](#)).
4. **Extraneous (extra) variables**—all the variables other than the independent variables or treatments that affect the response of the test units to the treatments, such as store size, store location, and competitive effort. These are the factors we (as experimenters) try to control.
5. **Experiment**—the process of manipulating one or more independent variables to determine their effect on the dependent variable, controlling for the effect of extraneous factors.
6. **Experimental design**—a set of procedures specifying (1) the test units and how these units are to be divided into homogeneous subsamples, (2) independent variables or treatments that are to be manipulated, (3) dependent variables are to be measured, and (4) how the extraneous variables are to be dealt with. An example of one of my own experimental designs is shown below:

The example below is known as a 2x2 design. (It's very simplistic, but drawing my experimental grids for each study is extremely helpful as I plan the design of each cell (i.e., what should each experimental block in Qualtrics look like?) *and* as I construct my hypotheses).

		Argument Strength	
		Strong	Weak
Message Source	Strong	1	2
	Weak	3	4

2 (Argument Strength: strong vs. weak) x **2** (Message Source: strong vs. weak) = 4 experimental conditions. IVs = “Argument Strength” and “Message Source”



Validity, in an experimental setting, is the ability to draw appropriate conclusions about the effects of independent variables and to make reasonable generalizations to a larger population of interest. With this definition, we can distinguish between the two primary types: (1) internal and (2) external validity.

Internal validity examines whether the manipulation of the independent variables actually caused the effects on the dependent variables. Your experiment has a high degree of internal validity if you can say that the only reason why your experimental groups are different is due to differences in your IV(s). You can ask yourself “Are there any other possible explanations for our findings?”

External validity examines whether the cause-and-effect relationships found in the experiment can be generalized. Your experiment has a high degree of external validity if you can say that your results extend to a (1) population, (2) situation, (3) time, and/or (4) environment that is different than your experiment. This is the concept of generalizability.

Generally speaking, as one type of validity increases, the other type decreases. For example, as you implement more controls in your experiment (i.e., increasing the internal validity), the external validity is decreased (i.e., more “artificial”).

Some of the extraneous factors that affect validity are as follows:

1. **History**—refers to specific events that are external to the experiment but occur at the same time as the experiment. These events may affect the dependent variable. E.g., if a natural disaster such as an earthquake happened at the same time as the experiment.
2. **Maturation**—is similar to history except that it is concerned with the changes in the participants themselves, where these changes are not due to the impact of the independent variables or treatments but occur with the passage of time. E.g., people getting tired of answering survey questions and/or experience mood changes. (See some of my tips re: survey design with crowdsourced data... no one wants to experience “bubble hell”!!)
3. **Testing**—is concerned with the effects that arise due to the process of experimentation itself affecting the dependent variables. There are two kinds of testing effects: (1) main testing effect, and (2) interactive testing effect.
4. **Instrumentation**—are the changes in the calibration of the measuring instrument. E.g., Are your DVs measuring what you *think* they’re measuring??
5. **Statistical Regression**—effects occur when test units with extreme scores move closer to the average score during the course of the experiment. The extreme test units may have been selected because of chance or by design.
6. **Selection Bias**—refers to the improper assignment of your participants to your experimental conditions where the bias occurs when selection or assignment of test units results in treatment groups that differ on the dependent variable before the exposure to the treatment condition. (This can be avoided by using *randomization* to ↓ sampling error!)
7. **Mortality**—refers to the loss of participants while the experiment is in progress. This could be caused due to the refusal of test units to continue in the experiment. On the other



hand, if you ask a lot of questions that appear very similar to the participant, he/she may tire of your survey and choose to drop out. (MTurkers ♥ when you include fun & *unique* attention checks. These ‘checks’ serve a functional purpose with the added benefit of reducing survey mortality! I strongly recommend regularly changing your attention checks since novelty is key.)

Some of the methods for controlling these factors (shown above) are as follows:

1. **Randomization**—involves randomly assigning test units to experimental groups. This should create representative samples free from bias. (i.e., reduce sampling error!)
2. **Matching**—test units are matched on a set of key background variables before being assigned to the treatment conditions. This ensures that a population has desirable traits that are to be studied and that the groups are homogeneous.
3. **Statistical Control**—it is possible to statistically control for the extraneous variables if these variables can be identified and measured, for example, in ANCOVA the effects of the extraneous variable on the dependent variable are removed by an adjustment of the dependent variable’s mean value within each treatment condition.
4. **Design Control**—using specific experimental designs that differ in the kinds of extraneous variables that can be used to control the extraneous variables.

There are two primary types of experiments: (1) *lab* experiments and (2) *field* experiments. Laboratory experiments (offline & online) make up the bulk of consumer research because of their ability to control extraneous variables and their relative efficiency in gathering data. However, for certain studies, field experiments are used & preferred. For example, Coca-Cola counts the shelf space it and its competitors receive in local grocery stores when promotional variables are manipulated. As a result, each type of experimentation has its role to play in marketing research. However, causality cannot be inferred from field experiments so laboratory experiments predominate in marketing.

Are there limits to experimentation?

Theoretically, experimentation is a very powerful way to uncover causal relationships. However, in practice, real world considerations may prevent experiments from achieving its optimal effectiveness. These considerations are the following three results:

1. **Time**: experiments can be time consuming; however, the experiments should be long enough in duration so that the DVs include most, or all the effects of the variables.
2. **Cost**: the requirements of experimental group, control group, and multiple measurements may significantly add to the cost of the experiment (although this limitation is reduced by using MTurk participants! However, the recommended *min.* pay = \$0.10/ minute).
3. **Administration**: experiments can be difficult to administer, and it is not always feasible to control for the effects of the extraneous variables. Administration difficulty ↑ as the number of experimental cells ↑.



Causal Research Design- **Experimentation**

Final thoughts: It is often necessary in experimentation to disguise the purpose of the research to produce valid results. I.e., you do *not* want your research participants to be able to guess what you are attempting to measure or manipulate. (Make sure your survey includes hypothesis-guessing questions!) Disguising the purpose of the research should not lead to deception, however. Although this seems like a paradox, one solution would be to disclose the possible existence of deception before the start of the experiment and allow the participants to have an opt-out option. The following four items should also be conveyed to participants: (1) inform participants that in an experiment of this nature a disguise of the purpose is often required for valid results, (2) inform them of the general nature of the experiment and what they will be asked to do, (3) make sure they know that they can leave the experiment at any time, and (4) inform them that the study will be fully explained after the data have been gathered and at that time they may request that their information be withdrawn. The procedure outlined in item (4) is called **debriefing**.

One further ethical concern in experimentation involves using the appropriate experimental design to control errors caused by extraneous variables. It is the responsibility of the researcher to use the most applicable experimental design for the problem. Determining the most appropriate experimental design for the problem requires not only an initial evaluation but also continuous monitoring.