Research in Psychology

Bill Altermatt, Ph.D.

As you read in chapter 1 of Mynatt and Doherty's text, science is a collection of tools for reducing error in conclusions. In this reading, my goal is to give you an overview of some of these tools.

Research Ethics

Psychological researchers are obligated to protect the dignity and welfare of the people who participate in the research. These "**participants**" must be informed about what they will do in a study, especially if there is any risk of harm, and they must give their consent to participate. This process is called **informed consent.** Some studies, especially those that investigate antisocial behavior, may mislead participants as to the true topic of the study. For example, a study on cheating may be advertised as a study on memory. These studies are said to involve **deception** because participants are deceived about the true nature of the study, but the deception can never be about the level of risk to a participant and the deception must be disclosed to participants at the end of the study.

Types of Variables

Many psychological studies can be represented by the formula "The effect of X on Y." For example, a researcher might study the effect of group size on conformity, or the effect of alcohol consumption on physical coordination. Researchers call the X in the formula above the **independent variable**, or **IV**. The IV is the alleged *cause* in a study. The Y in the formula above is called the **dependent variable**, or **DV**. The DV is the *effect*, or the behavioral response by participants that is observed.

Evaluating Research

In general, **validity** is the degree of confidence you can have in certain kinds of conclusions. The three kinds of confidence we will discuss are 1) confidence that a variable faithfully captures the concept it is representing; 2) confidence that a cause-and-effect relationship exists between the IV and DV; and 3) confidence that the results of a study can be generalized to people, times, or situations beyond the study.

1. Construct validity. Construct validity is the confidence that your variables are accurately capturing the ideas you wish to study. For example, a researcher wishing to investigate aggression as a DV must choose some behavioral response that indicates aggression. For ethical reasons, it would not be appropriate to measure actual physical aggression between participants in a laboratory, so researchers have developed a range of responses as substitutes: making participants believe they are administering an electric shock to another person, writing insulting feedback on an essay, etc. In every study you read about, ask yourself whether the researcher's choice of how to measure a variable has construct validity.

2. Internal validity. **Internal validity** is the degree of confidence in cause-and-effect conclusions. A cause-and-effect conclusion states that changes in the IV (e.g., exposure to television violence) directly cause changes in the DV (e.g., aggressiveness). A study has high internal validity when you can be confident that the IV is causing the DV. We will discuss below how one type of study, experiments, have higher internal validity than others.

3. External validity. **External validity** is the confidence you can have that the results in your study can be generalized beyond the study. In general, external validity is evaluated by the degree to which the participants in a study are representative of the population of interest. If you are studying voting patterns, your population of interest is "likely voters." This population will deliberately under-represent groups that typically do not vote (children, the very poor, etc.), but this should not be considered a weakness because you are only interested in people who *will* vote. To evaluate external validity, you must know both the population of interest and the characteristics of the sample used in the study. In particular, be skeptical when researchers attempt to generalize research from one group (e.g., males-only, Americans-only, etc.) to other groups.

Experiments

As mentioned above, experiments are a type of study that maximizes internal validity, or confidence in cause-and-effect conclusions. How can you be confident that the IV is causing the DV? The basic approach is to reduce the likelihood that any other variable could be causing the DV. Researchers call these other variables **confounding variables.** For example, let's say that you want to investigate whether exposure to television violence really does make people more aggressive. You want to know whether the IV (exposure to television violence) causes the DV (aggressiveness). Would it be reasonable to compare a group of children who watched 10 hours of television violence per week to a group of children who watched 0 hours of television violence per week? They differ in their exposure to television violence, which is your IV. The problem is that they will probably differ in other ways as well: age, level of parent involvement, or even the level of aggressiveness in their personality. Any of these are potential confounding variables. How can you rule them all out? The answer is **random assignment.** A researcher uses a coin toss or other *random* process to *assign* participants to groups (e.g., 10-hours of violence group, 0-hours of violence group). With enough participants, *the groups will tend to be nearly identical on any confounding variable.* The goal of random assignment is to make the groups in an experiment as similar as possible in every way except for one: the independent variable that the experimenter is investigating. Let's say you run this experiment and you find that the group who was randomly assigned to watch lots of violence later behaves more aggressively than the group who watched no violence. A critic would look for other ways that the two groups differ: age, level of parent involvement, etc. Because you used random assignment, it is very unlikely that there are *any* systematic differences between the groups other than the one difference that you created: the difference in the level of television violence they were exposed to. Random assignment increases internal validity because it decreases the likelihood that any other variable besides the IV could be responsible for changes in the DV.

One important limitation of experiments is that, although they may find one variable that is causally related to another variable, it is still possible that there are *other* causal factors that you have not identified. If you discover that watching violent television influences aggression, it is still possible that there are other, possibly more important, factors that would influence aggression.

Would you be nervous about signing your child up to participate in research on the effect of exposure to television violence? If you are pretty sure that television violence makes people more aggressive (and you assume that children should not be made more aggressive), you may think that the study described above is unethical. Despite their ability to increase internal validity, experiments cannot always be used. You could not, in good conscience, randomly assign expectant mothers to consume large amounts of alcohol or randomly assign people to gender using surgery and hormones. *In general, a limitation of experiments is that there are many topics that cannot be studied with random assignment because of ethical or practical considerations.* One technique to use when participants cannot be randomly assigned to the IV is correlational research.

Correlational Research

When used to make a contrast with experimental research, correlational research refers to research that does not use random assignment. Thus, correlational research tends to have lower internal validity than experimental research, meaning that researchers should be more cautious about making cause-and-effect inferences.

In addition to using it to mean "non-experimental" research, correlation is also a technique for statistical analysis. It is a way to measure how strongly high values on one variable (e.g., outdoor temperature) tend to co-occur with high values on another variable (e.g., sales of soft drinks). The correlation statistic is expressed as a number between -1 and +1 and it is often referred to using the lower-case letter "r", usually in italics: *r*. The number contains two pieces of information: its sign (whether it is positive or negative), and its absolute value (how far away it is from zero). If high values on one variable tend to occur with high values on the other variable, as would be the case with outdoor temperature and sales of soft drinks, the correlation would be positive. If high values on one variable tend to occur with *low* values on the other variable, as would be the case with outdoor temperature and sales of *hot* beverages, the correlation would be negative. If high values on

one variable are unrelated to values on the other variable, as would be the case with outdoor temperature and sales of vacuum cleaners, the correlation would be close to zero. This is summarized below in Table 1.

Table 1.

Variable 1		Variable 2	Correlation
high values	found with	high values	positive
high values	found with	low values	negative
high values	unrelated to	values	zero

The farther *r* is from zero (or the closer it gets to either -1 or +1) the stronger the relation between the two variables. A handy way to gauge the size of *r* is to square it (multiply it by itself). Squaring *r* gets you the percentage of one variable that can be predicted by another variable. For example, Anderson (1987) found that the correlation between the number of hot (above 90°) days in a city was correlated *r* = +0.52 with the rate of violent crime in that city. Cities with more hot days were more violent than cities with colder days. If you square 0.52, you get 0.27, or 27%. 27% of the differences between cities in their violent crime rate can be explained simply by the number of hot days. 27% is not 100%, so there are a number of other factors contributing to violent crime besides the number of hot days, but 27% is far above 0%. One goal of psychological research is to identify the most important predictors of behavior so that we can try to explain the greatest percentage of behavior possible. Because human behavior is usually caused by multiple factors, it is unusual for any single factor to explain more than 10% of behavior.

Correlation and causation. In the previous example, we saw that there was a positive correlation between the number of hot days and violent crime. Although it would be tempting to conclude that hot days *cause* violent crime, we cannot make that conclusion because we did not randomly assign cities to their number of hot days. Because we did not use random assignment, it is possible that some factor other than number of hot days could be responsible for causing violent crime. For example, hotter cities tend to be in the southern United States, so perhaps the causal factor is not temperature but rather geographical region. *In general, avoid inferring causation from correlation.* To make a cause-and-effect conclusion, you would need to conduct an experiment.

Sampling

As mentioned above, external validity is the confidence we can have that the results of our study can be generalized beyond the people who actively participated in our study. A major determinant of external validity is **sampling**: the process of obtaining participants for research. If you want your results to generalize to a target population, you must make your sample representative of that population. The best way to do this is to conduct a **random sample** of your target population. This means that every member of the target population has an equal chance of being included in your study. To do this for the target population of "Hanover students", you would begin with a list of all students and then use a computer program or other random process to select 100 or so people (the exact number depends on how precise you want your final estimates to be). Random samples are tremendously powerful. With a random sample of approximately **1,100** people, you can make generalizations of election preferences (where people are given 2 to 3 options to choose among) to millions of people. This is why most national election polls report their results as based on samples of approximately 1,100 people.

Remember that random *assignment* of participants to groups is related to internal validity (confidence in making conclusions about cause and effect), whereas random *sampling* is related to external validity (confidence in making conclusions about generalizing results to people outside of the study). These relationships are summarized in Table 2.

Table 2.

random	**validity**	**conclusion**
assignment	internal	cause-effect
sampling	external	generalizability

Other Types of Research

A lot of research falls between purely experimental and purely correlational research. Two of the most common methods are field

research and case studies.

Field research. In field research, researchers conduct their investigation in a naturalistic setting but still make an effort to keep their measures as objective and unbiased as possible. For example, researchers might study cheating in an actual classroom setting rather than in a laboratory. An advantage with field studies can be greater realism, but a disadvantage can be less control over the environment.

Case study. In a case study, researchers examine a single person, group, or event in great detail, typically summarizing their observations in a written report without using statistics. Case studies are commonly used in situations that are unusual or dangerous (e.g., brain damage from unusual trauma).