PSYCHOLOGICAL RESEARCH

MODULE 10

PSYCHOLOGICAL RESEARCH

The major research methods used in the field of Psychology are mentioned below.

10.1: ARCHIVAL RESEARCH

Research in which existing data, such as census documents, college records, and newspaper clippings, are examined to test a hypothesis.

In archival research, existing data, such as census documents, college records, and newspaper clippings, are examined to test a hypothesis. For example, college records may be used to determine if there are gender differences in academic performance (Sullivan, Riccio, & Reynolds, 2008).

Archival research is a relatively inexpensive means of testing a hypothesis because someone else has already collected the basic data. Of course, the use of existing data has several drawbacks. For one thing, the data may not be in a form that allows the researcher to test a hypothesis fully. The information could be incomplete, or it could have been collected haphazardly (Riniolo et al., 2003; Simonton, 2000a; Vega, 2006). Most attempts at archival research are hampered by the simple fact that records with the necessary information often do not exist.

10.2: NATURALISTIC OBSERVATION

Research in which an investigator simply observes some naturally occurring behavior and does not make a change in the situation.

In naturalistic observation, the investigator observes some naturally occurring behavior and does not make a change in the situation. The important point to remember about naturalistic observation is that the researcher simply records what occurs, making no modification in the situation that is being observed (Moore, 2002; Rustin, 2006; Schutt, 2001).

Although the advantage of naturalistic observation is obvious—we get a sample of what people do in their "natural habitat"—there is also an important drawback: the inability to control any of the factors of interest. Because naturalistic observation prevents researchers from making changes in a situation, they must wait until the appropriate conditions occur. Furthermore, if people know they are being watched, they may alter their reactions and produce behavior that is not truly representative.

10.3: SURVEY

Research in which people chosen to represent a larger population are asked a series of questions about their behavior, thoughts, or attitudes.

In survey research, a sample of people chosen to represent a larger group of interest (a population) is asked a series of questions about their behavior, thoughts, or attitudes. Survey methods have become so sophisticated that even with a very small sample researchers are able to infer with great accuracy how a larger group would respond.

However, survey research has several potential pitfalls. For one thing, if the sample of people who are surveyed is not representative of the broader population of interest, the results of the survey will have little meaning. Consequently, researchers using surveys strive to obtain a random sample of the population in question, in which every voter in the town has an equal chance of being included in the sample receiving the survey (Dale, 2006; Daley et al., 2003).

In addition, survey respondents may not want to admit to holding socially undesirable attitudes. And in some cases, people may not even be consciously aware of what their true attitudes are or why they hold them.

10.4: CASE STUDY

An in-depth, intensive investigation of an individual or small group of people.

A case study is an in-depth, intensive investigation of a single individual or a small group. Case studies often include psychological testing, a procedure in which a carefully designed set of questions is used to gain some insight into the personality of the individual or group (Addus, Chen, & Khan, 2007; Gass et al., 2000).

When case studies are used as a research technique, the goal is often not only to learn about the few individuals being examined but also to use the insights gained from the study to improve our understanding of people in general. Sigmund Freud developed his theories through case studies of individual patients. Similarly, case studies of terrorists might help identify others who are prone to violence.

Case study has some drawbacks too. If the individuals examined are unique in certain ways, it is impossible to make valid generalizations to a larger population. Still, they sometimes lead the way to new theories and treatments for psychological disorders.

10.5: CORRELATIONAL RESEARCH

Research in which the relationship between two sets of variables is examined to determine whether they are associated, or "correlated."

In correlational research, two sets of variables are examined to determine whether they are associated, or "correlated." The strength and direction of the relationship between the two variables are represented by a mathematical statistic known as a correlation (or, more formally, a correlation coefficient), which can range from +1.0 to –1.0.

10.5.1: Positive Correlation

A positive correlation indicates that as the value of one variable increases, we can predict that the value of the other variable will also increase. For example, if we predict that the more time students spend studying for a test, the higher their grades on the test will be, and that the less they study, the lower their test scores will be, we are expecting to find a positive correlation. The correlation, then, would be indicated by a positive number, and the stronger the association was between studying and test scores, the closer the number would be to +1.0.

10.5.2: Negative Correlation

A negative correlation tells us that as the value of one variable increases, the value of the other decreases. For instance, we might predict that as the number of hours spent studying increases, the number of hours spent partying decreases. Here we are expecting a negative correlation, ranging between 0 and –1.0. More studying is associated with less partying, and less studying is associated with more partying. The stronger the association between studying and partying is, the closer the correlation will be to –1.0. For instance, a correlation of –.85 would indicate a strong negative association between partying and studying.

10.5.3: Zero Correlation

It's quite possible that little or no relationship exists between two variables. For instance, we would probably not expect to fi nd a relationship between number of study hours and height. Lack of a relationship would be indicated by a correlation close to 0. For example, if we found a correlation of –.02 or +.03, it would indicate that there is virtually no association between the two variables; knowing how much someone studies does not tell us anything about how tall he or she is.

It is a serious point to consider that finding that two variables are correlated does not mean that there is a causal relationship between them.

10.6: EXPERIMENTAL RESEARCH

The experiment is a research method in which the investigator manipulates a variable under carefully controlled conditions and observes whether any changes occur in a second variable as a result.

The experiment is a relatively powerful procedure that allows researchers to detect cause and effect relationships. Psychologists depend on this method more than any other. The purpose of an experiment is to find out whether changes in one variable cause changes in another variable. A well-designed experiment must take into account a number of factors that could affect the clarity of the results.

10.6.1: Independent Variable

An independent variable is a condition or event that an experimenter varies in order to see its impact on another variable.

The independent variable is the variable that the experimenter controls or manipulates. It is hypothesized to have some effect on the dependent variable, and the experiment is conducted to verify this effect.

10.6.2: Dependent Variable

The dependent variable is the variable that is thought to be affected by manipulation of the independent variable.

In psychology studies, the dependent variable is usually a measurement of some aspect of the participants' behavior. The independent variable is called independent because it is free to be varied by the experimenter. The dependent variable is called dependent because it is thought to depend (at least in part) on manipulations of the independent variable.

10.6.3: Extraneous Variable

Extraneous variables are any variables other than the independent variable that seem likely to influence the dependent variable in a specific study.

The logic of the experimental method rests on the assumption that the experimental and control groups are alike except for their treatment in regard to the independent variable. Any other differences between the two groups can cloud the situation and make it impossible to draw conclusions about how the independent variable affects the dependent variable.

Instead, experimenters concentrate on ensuring that the experimental and control groups are alike on a limited number of variables that could have a bearing on the results of the study. These variables are called extraneous, secondary, or nuisance variables.

10.6.4: Confounding Variable

A confounding of variables occurs when two variables are linked together in a way that makes it difficult to sort out their specific effects.

When an extraneous variable is confounded with an independent variable, a researcher cannot tell which is having what effect on the dependent variable. Unanticipated confounding of variables has wrecked innumerable experiments. That is why so much care, planning, and forethought must go into designing an experiment. One of the key qualities that separates a talented experimenter from a mediocre one is the ability to foresee troublesome extraneous variables and control them to avoid confounding.

10.6.5: Experimental Group and Control Group

In an experiment the investigator typically assembles two groups of subjects who are treated differently with regard to the independent variable. These two groups are referred to as the experimental group and the control group.

The experimental group consists of the subjects who receive some special treatment in regard to the independent variable. The control group consists of similar subjects who do not receive the special treatment given to the experimental group.

It is crucial that the experimental and control groups in a study be alike, except for the different treatment that they receive in regard to the independent variable. This stipulation brings us to the logic that underlies the experimental method. If the two groups are alike in all respects except for the variation created by the manipulation of the independent variable, any differences between the two groups on the dependent variable must be due to the manipulation of the independent variable. In this way researchers isolate the effect of the independent variable on the dependent variable.

Experimenters use a variety of safeguards to control for extraneous variables. For instance, subjects are usually assigned to the experimental and control groups randomly. **Random assignment** of subjects occurs when all subjects have an equal chance of being assigned to any group or condition in the study. When experimenters distribute subjects into groups through some random procedure, they can be reasonably confident that the groups will be similar in most ways.

Research Method	Description	Advantages	Shortcomings
Descriptive and correlational research	Researcher observes a previously existing situation but does not make a change in the situation	Offers insight into relationships between variables	Cannot determine causality
Archival research	Examines existing data to confirm hypothesis	Ease of data collection because data already exist	Dependent on availability of data
Naturalistic observation	Observation of naturally occurring behavior, without making a change in the situation	Provides a sample of people in their natural environment	Cannot control the "natural habitat" being observed
Survey research	A sample is chosen to represent a larger population and asked a series of questions	A small sample can be used to infer attitudes and behavior of a larger population	Sample may not be representative of the larger population; participants may not provide accurate responses to survey questions
Case study	Intensive investigation of an individual or small group	Provides a thorough, in-depth understanding of participants	Results may not be generalizable beyond the sample
Experimental research	Investigator produces a change in one variable to observe the effects of that change on other variables	Experiments offer the only way to determine cause-and-effect relationships	To be valid, experiments require random assignment of participants to conditions, well-conceptualized independent and dependent variables, and other careful controls

10.7: RELIABILITY

Reliability can be defined as an indication of the consistency or stability of a measuring instrument.

One means of determining whether the measure using is effective is to assess its reliability. Reliability refers to the consistency or stability of a measuring instrument. In other words, the measuring instrument must measure exactly the same way every time it is used. This consistency means that individuals should receive a similar score each time they use the measuring instrument. For example, a weighing machine needs to be reliable, that is, it needs to measure the same way every time an individual uses it, or otherwise it is useless as a measuring instrument. Reliability is measured using correlation coefficients.

There are four types of reliability, namely; **test/retest reliability, alternate-forms reliability, split-half reliability, and interrater reliability**. Each type provides a measure of consistency, but they are used in different situations.

10.7.1: Test/Retest Reliability

A reliability coefficient determined by assessing the degree of relationship between scores on the same test administered on two different occasions.

One of the most often used and obvious ways of establishing reliability is to repeat the same test on a second occasion, a process known as test/retest reliability. The correlation coefficient obtained is between the two scores of an individual on the same test administered on two occasions. If the test is reliable, we expect the two scores for each individual to be similar, and thus the resulting correlation coefficient will be high (close to 1.00). This measure of reliability assesses the stability of a test over time.

A problem related to test/retest measures is that on many tests there are **practice effects**, that is, some people get better at the second testing, and this practice lowers the observed correlation. A second problem may occur if the interval between test times is short: Individuals may remember how they answered previously, both correctly and incorrectly. In this case we may be testing their memories and not the reliability of the testing instrument, and we may observe a spuriously high correlation.

10.7.2: Alternate-Forms Reliability

A reliability coefficient determined by assessing the degree of relationship between scores on two equivalent tests.

One means of controlling for test/retest problems is to use alternate-forms reliability, that is, using alternate forms of the testing instrument and correlating the performance of individuals on the two different forms. In this case the tests taken at times 1 and 2 are different but equivalent or parallel. As with test/retest reliability alternate-forms reliability establishes the stability of the test over time. In addition, it also establishes the equivalency of the items from one test to another. One problem with alternate-forms reliability is making sure that the tests are truly parallel. To help ensure equivalency, the tests should have the same number of items, the items should be of the same difficulty level, and instructions, time limits, examples, and format should all be equal often difficult, if not impossible, to accomplish. Further, if the

tests are truly equivalent, there is the potential for practice, although not to the same extent as when exactly the same test is administered twice.

10.7.3: Split-Half Reliability

A reliability coefficient determined by correlating scores on one half of a measure with scores on the other half of the measure.

A third means of establishing reliability is by splitting the items on the test into equivalent halves and correlating scores on one half of the items with scores on the other half. This split-half reliability gives a measure of the equivalence of the content of the test but not of its stability over time as test/retest and alternate-forms reliability do. The biggest problem with split-half reliability is determining how to divide the items so that the two halves are in fact equivalent. For example, it would not be advisable to correlate scores on multiple- choice questions with scores on short-answer or essay questions. What is typically recommended is to correlate scores on even-numbered items with scores on odd-numbered items. Thus if the items at the beginning of the test are easier or harder than those at the end of the test, the half scores are still equivalent.

10.7.4: Interrater Reliability

A reliability coefficient that assesses the agreement of observations made by two or more raters or judges.

Finally, to measure the reliability of observers rather than tests, researchers can use interrater reliability, which is a measure of consistency that assesses the agreement of observations made by two or more raters or judges. Let's say that the panel of judges are observing play behavior in children. The observers all watch the children playing but independently count the number and types of play behaviors they observe. Once the data are collected, interrater reliability needs to be established by examining the percentage of agreement among the raters. If the raters' data are reliable, then the percentage of agreement should be high. If the raters are not paying close attention to what they are doing or if the measuring scale devised for the various play behaviors is unclear, the percentage of agreement among observers will not be high.

10.8: VALIDITY

Validity is an indication of whether the instrument measures what it claims to measure.

In addition to being reliable, measures must also be valid. Validity refers to whether a measuring instrument measures what it claims to measure. Validity is measured by the use of correlation coefficients. For instance, if researchers developed a new test to measure depression, they might establish the validity of the test by correlating scores on the new test with scores on an already established measure of depression, and as with reliability the researchers would expect the correlation to be positive.

There are four types of validity, namely; **content validity, face validity, criterion validity and construct validity**.

10.8.1: Content Validity

The extent to which a measuring instrument covers a representative sample of the domain of behaviors to be measured.

A systematic examination of the test content to determine whether it covers a representative sample of the domain of behaviors to be measured assesses content validity. In other words, a test with content validity has items that satisfactorily assess the content being examined. To determine whether a test has content validity, researchers consult experts in the area being tested.

10.8.2: Face Validity

The extent to which a measuring instrument appears valid on its surface.

Sometimes face validity is confused with content validity. Face validity simply addresses whether or not a test looks valid on its surface. Face validity is not really a type of validity in the technical sense because it not clearly refers what the test actually measures but to what it appears to measure. Face validity relates to whether the test looks valid to those who selected it and to those who take it.

10.8.3: Criterion Validity

The extent to which a measuring instrument accurately predicts behavior or ability in a given area.

Two types of criterion validity may be used, depending on whether the test is used to estimate present performance (**concurrent validity**) or to predict future performance (**predictive validity**).

The Scholastic Assessment Test (test used for the selection of graduate students) is an example of test that have predictive validity because performance on the tests correlates with

later performance in college and graduate school, respectively. The tests can be used with some degree of accuracy to predict future behavior.

A test also can use to determine whether someone qualifies as a pilot is a measure of concurrent validity. The test is estimating the person's ability at the present time, not attempting to predict future outcomes. Thus concurrent validation is used for the diagnosis of existing status rather than the prediction of future outcomes.

10.8.4: Construct Validity

The degree to which a measuring instrument accurately measures a theoretical construct or trait that it is designed to measure.

Construct validity is considered by many to be the most important type of validity. The construct validity of a test assesses the extent to which a measuring instrument accurately measures a theoretical construct or trait that it is designed to measure. Some examples of theoretical constructs or traits are verbal fluency, neuroticism, depression, anxiety, intelligence, and scholastic aptitude. One means of establishing construct validity is by correlating performance on the test with performance on a test for which construct validity has already been determined. Thus performance on a newly developed intelligence test might be correlated with performance on an existing intelligence test for which construct validity has been previously established.

10.9: NORMS

Norms can be defined as a summary of test results for a large and representative group of subjects.

Scores on psychological test are most commonly interpreted by reference to norm that represents the test performance on standardization sample. Norms always represent the best performance.

Basically there are two purposes of norms:

1. Norms indicate the individual's relative standing in the normative sample and thus permit evaluation of his/her performance in refer to other persons.
2. Norms provide compared measures that permitted a direct comparison of the individual performance on difference test.

10.9.1: Age Norms

An age norm depicts the level of test performance for each separate age group in the normative sample.

The purpose of age norms is to facilitate same-aged comparison. With age norms, the performance of an examinee is interpreted in relation to standardization subjects of the same age. The age span for a normative age group can vary from a month to a decade or more, depending on the degree to which test performance is age dependent. For characteristics that change quickly with age such as intellectual abilities in childhood, test developers might report separate test norms for narrowly defined age brackets.

10.9.2: Grade Norms

A grade norm depicts the level of test performance for each separate grade in the normative sample.

Grade norms are especially useful in school settings when reporting the achievement level of school children. Since academic achievements in many content areas are heavily dependent on grade - based curricular exposure, comparing a student against a normative sample from the same grade is more appropriate than using an age-based comparison.

10.9.3: Percentile Norms

A percentile expresses the percentage of a person in the standardization sample, who scored below a specific raw scores. The higher percentiles indicate higher scores. A percentile indicates only how an examinee compares to the standardization sample and doesn't convey the percentage of questions answered correctly.

Percentiles can also be viewed as ranks in a group of 100 representative subjects, with 1 being the lowest rank and 100 the highest. Percentile ranks are the complete reverse of usual ranking procedures. A percentile rank of 1 is at the bottom of the sample, while a percentile rank of 99 is near the top.

10.10: ETHICS IN RESEARCH

Nonetheless, because research has the potential to violate the rights of participants, psychologists are expected to adhere to a strict set of ethical guidelines aimed at protecting participants (American Psychological Association, 2002). Those guidelines involve the following safeguards:

- Protection of participants from physical and mental harm.
- The right of participants to privacy regarding their behavior.
- The assurance that participation in research is completely voluntary.
- The necessity of informing participants about the nature of procedures before their participation in the experiment.

One of psychologists' key ethical principles is **informed consent *(A document signed by participants affirming that they have been told the basic outlines of the study and are aware of what their participation will involve).*** Before participating in an experiment, the participants must sign a document affirming that they have been told the basic outlines of the study and are aware of what their participation will involve, what risks the experiment may hold, and the fact that their participation is purely voluntary and they may terminate it at any time.

Furthermore, after participation in a study, they must be given a **debriefing** in which they receive an explanation of the study and the procedures that were involved. The only time informed consent and a debriefing can be eliminated is in experiments in which the risks are minimal, as in a purely observational study in a public place (Barnett, Wise, & Johnson-Greene, 2007; Fallon, 2006; Koocher, Norcross, & Hill, 2005). Other important ethical principles are:

10.10.1: Deception

- Psychologists do not conduct a study involving deception unless they have determined that the use of deceptive techniques is justified by the study's significant prospective scientific, educational, or applied value and that effective nondeceptive alternative procedures are not feasible.
- Psychologists do not deceive prospective participants about research that is reasonably expected to cause physical pain or severe emotional distress.
- Psychologists explain any deception that is an integral feature of the design and conduct of an experiment to participants as early as is feasible, preferably at the conclusion of their participation, but no later than at the conclusion of the data collection, and permit participants to withdraw their data.

10.10.2: Invasion of Privacy

- Psychologists include in written and oral reports and consultations, only information relevant to the purpose for which the communication is made.

- Psychologists discuss confidential information obtained in their work only for appropriate scientific or professional purposes and only with persons clearly concerned with such matters.

10.10.3: Lasting Harm

Psychologists take reasonable steps to avoid harming their clients/patients, students, supervisees, research participants, organizational clients, and others with whom they work, and to minimize harm where it is foreseeable and unavoidable.

10.11: REFERENCES

Baron, R.A. (2004). Psychology, (5th ed.). New Delhi: Pearson Education.

Coon, D., & Mitterer, J.O. (2007). *Introduction to psychology: Gateways to mind and behavior,* (11th ed.). New Delhi: Thomson Wadsworth.

Feldman, R. S. (2011) *Understanding psychology.* New York: Tata McGraw Hill.

Myers, D. G., & DeWall, C. N. (2015). *Psychology*. New York, NY: Worth Publishers.

Weiten, W. (2016). *Psychology: Themes and Variations*. Boston, MA: Cengage Learning.