

THEORIES OF PERSONALITY

II

SESSION 2: Personality Research and Assessment

[Pick the date]

SESSION 3

SESSION 2: Personality Research and Assessment

THEORIES OF PERSONALITY II

SESSION 2: Personality Research and Assessment

Influence of Research on Psychology

Psychology is not an absolute science and is often referred to as a 'Social Science' or a 'Soft Science.' This is because it deals with human thoughts, feelings, and behavior, and as we are all aware, humans are not always predictable and reliable. Instead, we interact with our environment in ways that alter how we behave, how we think, and how we feel. Change one thing and the domino effect can change everything else.

Nevertheless, research plays an extremely important role in psychology. Research helps us understand what makes people think, feel, and act in certain ways; allows us to categorize psychological disorders in order to understand the symptoms and impact on the individual and society; helps us to understand how intimate relationships, development, schools, family, peers, and religion affect us as individuals and as a society; and helps us to develop effective treatments to improve the quality of life of individuals and groups.

In this sense, psychological research is typically used for the following:

1. Study development and external factors and the role they play on individuals' mental health
2. Study people with specific psychological disorders, symptoms, or characteristics
3. Develop tests to measure specific psychological phenomenon
4. Develop treatment approaches to improve individuals' mental health

In the following sections, you will learn about how research is conducted and the different types of research methods used to gather information.

Experimental Methods

Starting from the general and moving to the more specific, the first concept we need to discuss is [Theory](#). A theory can be defined as a "general principle

proposed to explain how a number of separate facts are related." In other words, a theory is an "idea about a relationship." In order to test whether a theory is correct or not, we need to do research. Theories are stated in general terms, so we need to define more accurately what we will be doing in our experiment.

To do this, we need to define the variables in our theory so that they are testable, and every experiment has two types of variables:

- o Independent Variable (IV) – the variable that is manipulated by the experimenter (input variable)
- o Dependent Variable (DV) – the outcome variable (results of the experiment)

By defining our variables that we will use to test our theory we derive at our Hypothesis, which is a testable form of a theory.

As an example of this, let's say that we have a theory that people who drive sports cars are more aggressive in their interactions with others. Our independent variable would be the type of car you drive (sports, sedan, SUV, etc.). Our dependent variables, the outcome of our research, would be aggression. We would need to further define aggression so that it is something we can test such as speeding or cutting other people off in traffic. We now have the basics of our very simple experiment and can write our Hypothesis: People who drive sports cars drive over the speed limit more frequently than people who drive other types of cars.

Research Biases

Now we've got a hypothesis which is the first step in doing an experiment. Before we can continue, however, we need to be aware of some aspects of research that can contaminate our results. In other words, what could get in the way of our results in this study being accurate. These aspects are called research biases, and there are basically three main biases we need to be concerned with.

- Selection Bias – occurs when differences between groups are present at the beginning of the experiment.

- **Placebo Effect** – involves the influencing of performance due to the subject's belief about the results. In other words, if I believe the new medication will help me feel better, I may feel better even if the new medication is only a sugar pill. This demonstrates the power of the mind to change a person's perceptions of reality.

- **Experimenter Bias** – The same way a person's belief's can influence his or her perception, so can the belief of the experimenter. If I'm doing an experiment, and really believe my treatment works, or I really want the treatment to work because it will mean big bucks for me, I might behave in a manner that will influence the subject.

Controlling for Biases

After carefully reviewing our study and determining what might effect our results that are not part of the experiment, we need to control for these biases. To control for selection bias, most experiments use what's called **Random Assignment**, which means assigning the subjects to each group based on chance rather than human decision. To control for the placebo effect, subjects are often not informed of the purpose of the experiment. This is called a **Blind** study, because the subjects are blind to the expected results. To control for experimenter biases, we can utilize a **Double-Blind** study, which means that both the experimenter and the subjects are blind to the purpose and anticipated results of the study.

Standardization

We have our hypothesis, and we know what our subject pool is, the next thing we have to do is **standardize** the experiment. Standardization refers to a specific set of instructions. The reason we want the experiment to be standardized is twofold.

First, we want to make sure all subjects are given the same instructions, presented with the experiment in the same manner, and that all of the data is collected exactly the same or all subjects. Second, single experiments cannot typically stand on their own. To really show that are results are valid, experiments need to be replicated by other experimenters with different subjects. To do this, the experimenters need to know exactly what we did so they can replicate it.

Types of Research

What we've focused on is called Experimental Methods, the true experiment. It involves randomized assignment of subjects, standardized instructions, and at least one IV and one DV. There are several other types of research that are not as rigorous, but that you need to be aware of.

Perhaps the simplest form of research is [Naturalistic Observation](#).

Observing behavior in their natural environment

Often involves counting behaviors, such as number of aggressive acts, number of smiles, etc.

Advantages: Behavior is naturally occurring and is not manipulated by a researcher and it can provide more qualitative data as opposed to merely quantitative information.

Limitations: Even the presence of someone observing can cause those being observed to alter their behavior. Researcher's beliefs can also alter their observations. And, it is very difficult to coordinate multiple observers since observed behaviors must be operationally defined (e.g. what constitutes an aggressive act)

[Case Study](#)

Following a single case, typically over an extended period of time

Can involve naturalistic observations, also can include psychological testing, interviews, interviews with others, and the application of a treatment or observation

Advantages: Can gather extensive information, both qualitative and quantitative and it can be helpful in better understanding rare cases or very specific interventions

Limitations: Only one case is involved, severely limiting the generalization to the rest of the population. Can be very time consuming and can involve other problems specific to the techniques used, including researcher bias.

Survey

Everyone has probably heard of this and many of you have been involved in research involving surveys. They are often used in the news, especially to gather viewer opinions such as during a race for president

Advantages: Can gather large amounts of information in a relatively short time, especially now with many surveys being conducted on the internet.

Limitations: Survey data is based solely on subjects' responses which can be inaccurate due to outright lying, misunderstanding of the question, placebo effect, and even the manner in which the question is asked

Correlational Studies

Correlation means relationship, so the purpose of a correlational study is to determine if a relationship exists, what direction the relationship is, and how strong it is.

Advantages: Can assess the strength of a relationship. Is popular with lay population because it is relatively easy to explain and understand.

Limitations: Can not make any assumptions of cause and effect (explain how third a variable can be involved, or how the variables can influence each other).

Psychological Testing

Utilizing testing to gather information about a group or an individual

Advantages: Most tests are normed and standardized, which means they have very reliable and valid results. Popular with businesses looking for data on employees and with difficult or specific therapy cases

Limitations: Tests which are not rigorously normed and standardized can easily result in inaccurate results.

What do all those Numbers Mean?

Unlike the often times subjective nature of psychology, research is a means to objectively measure psychological phenomenon. Research uses statistical measures to determine likelihood, probability, and relationships, and therefore, when reading a research paper you will often come across statistics that help you understand the results. Lets look at some of the statistics commonly used in psychological research, especially those related to the study of personality.

Averages

One of the simplest measures in a research study is that of average and variance. There are three types of averages: mean, median, and mode. The one most of us are aware of is [mean](#), referring to the total of the subjects scores divided by the total number of subjects. The [median](#), simply enough, is the score that falls at exactly the 50th percentile, or the mathematical middle. Finally, the [mode](#) refers to the score that occurs most often. Groups of scores that have more than one mode are considered bimodal. See the sample data set below.

SAMPLE DATA SET

<u>Subject</u>	<u>Score</u>
1	10
2	10
3	15
4	15
5	20
6	25
7	30
8	30
9	30
10	85

Sum of Scores = 270

Mean = TOTAL/SUBJECTS
= 270/10
= 27

Median = 50th PERCENTILE
= (20+25)/2 This represents the point
= 45/2 where the same percentage
= 22.5 of scores fall below and
above this score.

Mode = MOST FREQUENT SCORE
= 30

As you can see by the statistics above, the mean, median, and mode are not always the same. When all three are the same, the data set is said to be [normally distributed](#). In other words, the mean, mode, and median all fall at the 50th percentile so there are an equal number of scores on either side. A good

example of this is the intelligence quotient (IQ) which has a mean, median, and mode of 100 and 50% of scores fall above and 50% fall below this number.

Statistical Significance

Before you can understand statistical significance, you must first understand the role chance plays in any data set. If you flip a coin 100 times, for example, you would guess that 50% of the time the outcome would be heads, and 50% tails. The truth is, however, that you may actually get 47% heads and 53% tails, or perhaps even 60% heads and 40% tails. Does these mean the coin is flawed? No, it just means that chance dictates that you will average 50/50 or somewhere close to that.

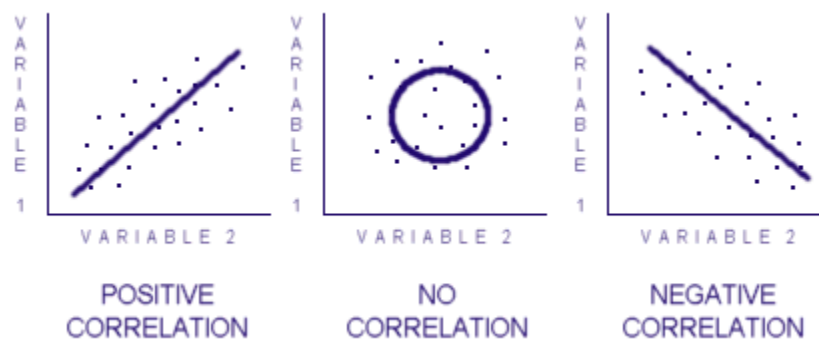
But what if we flipped this same coin 100 times and got 18 heads and 82 tails? What if we did it again and got 21 heads and 79 tails? Could we then say that there is something wrong with this coin? If we continue to see a pattern such as this that seems so far fetched, so different from each other, we could say that these coin toss results are statistically significant. In other words, the results are so different that they could not have been caused merely by chance.

When we perform a statistical significance test on a data set, we are looking to determine how much of a difference in scores could be caused by chance and how much could be caused by what we are trying to measure. Researchers agree that if the odds are less than 5% (or 1% depending on the type of study and the researcher's goals) that chance caused the difference then the difference is said to be significant. If they show that chance played a role greater than 5%, the results are considered not significant.

Correlation

We talked about the research technique known as correlation earlier, but we didn't discuss how the correlation is determined. If you recall, a correlation represents a relationship between two variables and does not show cause and effect. This relationship is determined by a statistical analysis of the data that derives a correlation score ranging from +1.00 to -1.00.

A positive correlation, one that shows both sets of data moving in the same direction, is one that is greater than zero. The closer to +1.00 you get, the stronger the relationship. A negative correlation, where the two data sets respond in opposite directions, results in a correlation of less than zero. Again, the closer you get to -1.00, the stronger the relationship. A correlation of zero (or close to it) represents no relationship between the two data sets. If we were to graph the results of a correlation, the results may look something like the diagram below. Notice how a shape of the plotted data points emerges.



Assessment Basics

Many of us have taken some sort of psychological test, whether for mental health reasons, potential employment, job evaluations, or research. There are basically three reasons that assessment devices are developed. If you recall the five goals of psychology (describe, explain, predict, control, improve), you'll see how they are directed related to these goals.

Research

Psychological assessment is often developed to assist with research. Imagine that you wanted to determine if a difference existed between the motivation of first born children and last born children. You would first need to define the construct of motivation and then determine how you would measure that construct. A good way to do this is to develop a test, or assessment device, that measures motivation. Once the test is normed and standardized, you can then administer it to your group of first borns and your group of last borns and

then perform your statistical analysis on the results.

Diagnosis and Prediction

A second way psychological assessment is used is during treatment of disorders in order to help diagnosis or determine current level of functioning. In the non-mental health arena, such as in business and human relations, assessment techniques are used to help identify individual or group strengths and weaknesses. In either case, the results are used to make recommendations for improvement.

Treatment Progress

Some assessment, including those used for the above purpose, can be used to get a baseline of functioning. Once this is established, a therapist or medical provider can retest the individual at varying intervals to determine if changes are taking place. Imagine the person who presents with depression and scores a ten out of ten on some depression test. After treatment, they are given the same test and score a one out of ten. This would suggest (if of course the test was a good measure of depression) that treatment had been successful. Through the rest of this chapter we'll discuss different types of assessment and, like research, look at how to understand results.

What can Tests Measure?

Through the remainder of this text we'll discuss the major theories of personality. Along with each theory of personality development comes a different theory on how to measure it. Some see assessment as a means to uncover unconscious impulses and analyze dreams, others see it as a way to measure behavior, and still others use assessment to determine patterns of cognitions. No matter the theory, however, all personality assessment falls into three distinct categories: subjective, objective, and projective.

Subjective Tests

A good example of an subjective technique is a job interview. The purpose of an interview is to gather information, assess experience, abilities, and probably most importantly, to determine interpersonal skills and abilities to work with others. This type of technique can certainly help in gathering information and making decisions, but is not without its negatives.

The biggest strength as well as the biggest negative is its subjectivity. If the person interviewing has any biases, positive or negative, these can obviously effect the outcome of the assessment. The interviewer who believes brunettes make better employees may give blonds undeserving low scores and may not even be aware of this tendency. It is often thought best to combine subjective techniques with objective ones to help assure personal biases do not interfere.

Objective Tests

Objective tests do not have the biases that are common among subjective measures, but they also don't provide the breadth of information. An objective assessment is one that utilizes research to determine results. When you take an pre-employment pencil and paper test, chances are this is an objective measure. The results are given in the form of statistics and probabilities. Some of the most common objective tests include the Minnesota Multiphasic Personality Inventory, Second Edition (MMPI-2), the Myers-Briggs Type Indicator, and any of the major tests of intelligence.

Projective Tests

Projective techniques differ from the above in one major way. These tests are designed to uncover thoughts, emotions, and desires that may not be known to the test taker. In other words, unconscious impulses that may be driving current behaviors. The most common of this type include the Rorschach Inkblot Test, The Thematic Apperception Test (TAT) and the Incomplete Sentences Blank.

The strength of this technique lies in its ability to uncover issues that the other two techniques are likely to miss. The negatives include less objectivity in that these techniques typically require human interpretation rather than statistical. There

are also many who argue that these tests are not as effective as the other two, and may even be harmful if interpretation is incorrect.

Assessment Batteries

The best way to gain the positives of each technique while minimizing the negative is to use a combination of all three. When different approaches of assessment are used, it is called an assessment battery. If you want to assess depression and functioning in an individual, the best way might be to use the following approach:

1. Interview the person and ask subjective questions about his or her feelings, thoughts and behaviors. This will gather information in the person's own words, allowing the interviewer to better understand the person. This could also alert the interviewer to associated issues such as anxiety. (Subjective)
2. Utilize one or more of the many depression tests to compare the person's symptoms to others. Knowing now that anxiety may also be an issue, an anxiety test could also be given. (Objective)
3. Administer one or more projective tests to look at issues that unconscious issues or those that were not looked at by the other measures. (Projective)

How Accurate are Personality Tests?

The accuracy of any assessment measure can be quite complicated, with some tests requiring hundreds of statistical analyses just to give you a few simple numbers. There are two statistics, however, that every professionally published assessment technique must provide in order to be accepted as 'good' tests: reliability and validity.

Reliability

[Reliability](#) refers to a test's ability to yield similar results each time it is taken. It is best to see reliability as synonymous with consistency. When measuring personality traits we would expect results to be similar each time the test is taken due to the relative stability of personality. For example, if you scored high on a

test of extroversion today, you would expect to score high on the same test next week or even next year.

Suppose, however, that you scored high today and scored low next week. How would you know your true score? A test measuring a stable trait must yield stable results in order to be reliable. As you can see from this example, an unreliable test is worthless as a measuring device. For a reliable assessment, on the other hand, you are more likely to get similar results each time you take it.

There are two major ways to determine the reliability of a test. The first is called *test-retest reliability*. To determine this statistic, the developers of an assessment technique would administer it to a group of individuals and then administer it again to the same people under the same circumstances some time in the future. A correlation would then be determined and knowing what you know now about this technique, you would expect the two scores to be positively correlated. In other words, a test with high test-retest reliability is one where the scores are strongly related in a positive manner.

Another technique to determine reliability is called *internal consistency*. Basically, a new assessment technique would be divided in two; the first half of the test versus the second half or odd questions versus even questions for example. The results of each score should be positively correlated if the test is truly a reliable technique. The benefits of splitting the test are in the test items themselves.

A 100 item test designed to measure assertiveness may have high test-retest reliability, but what if the first 50 questions are not correlated with the last 50 items? This test would have low internal consistency that suggests some of the questions are not measuring what they are intended to measure. This brings us to the next assessment statistic.

Validity

Simply put, a [valid](#) assessment is one that measures what it is intended to measure. Imagine taking your first test on the material you are learning here. As you sit down and the test is handed out, you look down and see only one

question: $4 + 6 = \underline{\quad}$. While this test may be very reliable since you are likely to answer "10" every time you take the test, it is not a valid measurement of your knowledge of personality theory. There are basically four different types of validity that we will discuss: face validity, predictive validity, congruent validity, and discriminant validity.

Face Validity. The easiest type of validity to determine is face validity because it basically asks 'does the test look like it measures what it is intended to measure?' The example above would have very low face validity because the question $4 + 6 = \underline{\quad}$ obviously has little to do with psychology. However, a test of extroversion that asks questions such as "Do you enjoy group activities?" would have high face validity.

Predictive Validity. If you recall the five goals of psychology, you'll remember that making predictions is an important aspect of reaching the ultimate goal of improving lives. Predictive validity refers to an assessment's ability to do this. A valid test of relationship skills, for example, might predict an individual's ease of making friends, comfort in group settings, or ability to effectively communicate.

Congruent Validity. Suppose you want to get an idea of a person's intelligence but do not have the time to administer the more commonly used assessment techniques. You may want to use a less expensive or quicker measurement. If the test has high congruent validity, it would be a valid substitution. Congruent validity refers to a test's congruency or relationship with a known valid and reliable measure of the same construct. In other words, a test that is positively correlated with a previously validated test is said to have high congruent validity with that test.

Discriminant Validity. Discriminant validity is just the opposite of congruent validity. If we want to validate our measurement of extroversion and we know of a valid test of introversion, we could give both tests to a group and expect the results to be opposite. Those who score high on the introversion test should score low on the extroversion test; they should be negatively correlated.

Specific Tests of Personality

As we progress through the text, we will discuss specific tests related to each theory. They will vary in terms of their validity and reliability as well as their approach, as no test has been shown to be perfect. In general, the higher the validity and the higher the reliability, the better the test. Understanding these concepts, the different types of assessment, as well as the basics of research will help you analyze the theories and assessment approaches that will be discussed throughout the rest of the text.