Quantitative Research Design: Descriptive and Correlational Research
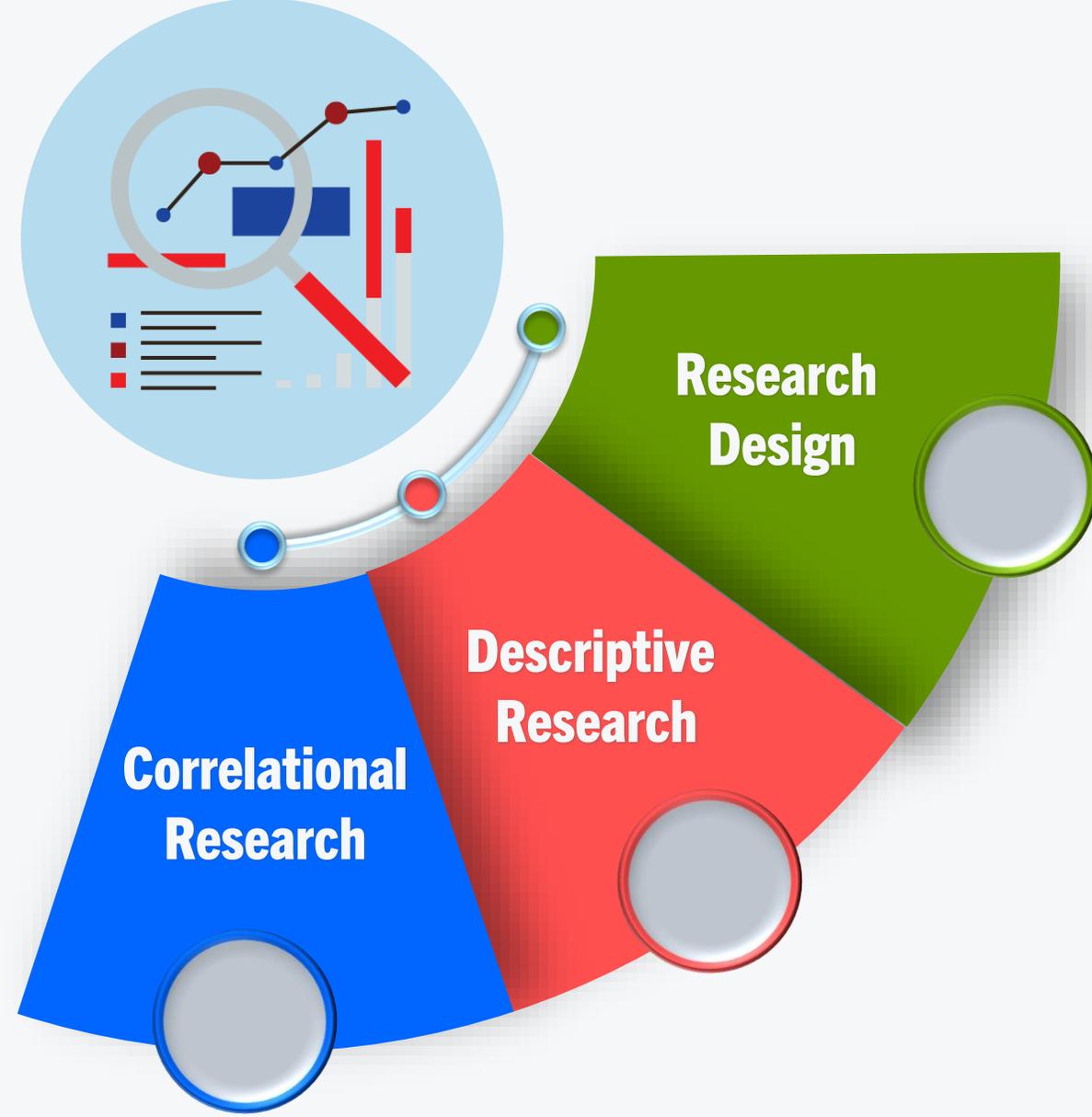
S
N
A

Siti Noormi Alias
Senior Lecturer
Department of Professional Development and
Continuing Education
Faculty of Educational Studies
Universiti Putra Malaysia

Research Design

Descriptive Research

Correlational Research

Research Design

What?

Overall strategy that you choose to integrate the different components of the study in a coherent and logical way

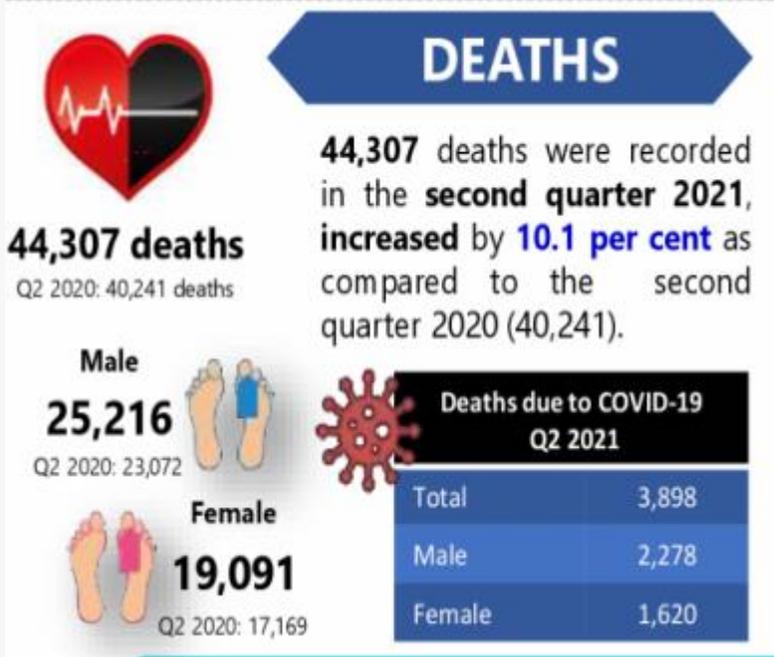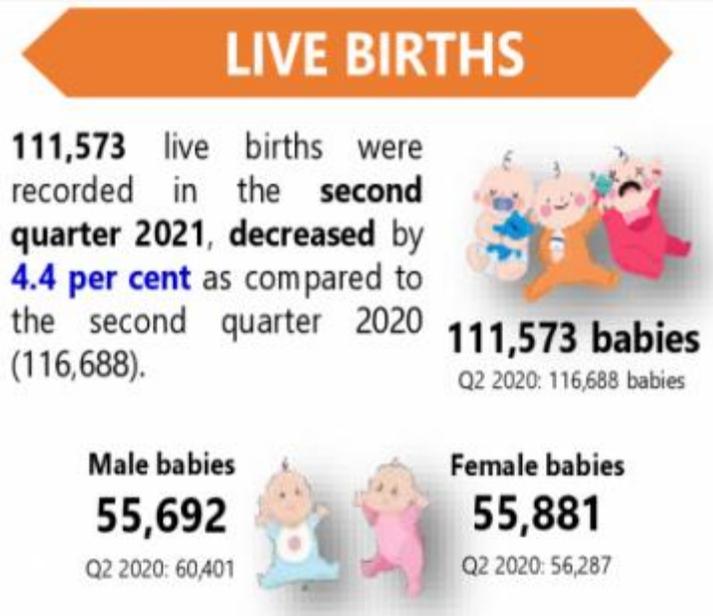**The Need for an Accurate Research Design**

Effectively address the research problem; it constitutes the blueprint for the collection, measurement, and analysis of data.

Note that the research problem determines the type of design you should use, not the other way around!

Descriptive Research

❑ The goal of descriptive research is to describe a single variable or to obtain separate descriptions for each variable when several are involved in as they exist.

❑ Does not concern with relationships between variables

❑ Focus on description of individual variables

❑ Useful as preliminary research (early stages of research) in its own right – to gain some idea of the variable of interest as it naturally exists

❑ Results from descriptive research help us to capture interesting, naturally occurring behavior

Examples of Descriptive Research

LIVE BIRTHS

111,573 live births were recorded in the **second quarter 2021**, **decreased** by **4.4 per cent** as compared to the second quarter 2020 (116,688).

111,573 babies
Q2 2020: 116,688 babies

Male babies
55,692
Q2 2020: 60,401

Female babies
55,881
Q2 2020: 56,287

DEATHS

44,307 deaths
Q2 2020: 40,241 deaths

44,307 deaths were recorded in the **second quarter 2021**, **increased** by **10.1 per cent** as compared to the second quarter 2020 (40,241).

Male
25,216
Q2 2020: 23,072

Female
19,091
Q2 2020: 17,169

Deaths due to COVID-19 Q2 2021	
Total	3,898
Male	2,278
Female	1,620

Types of Descriptive Research

Observational Research

The researcher observes and systematically records the behavior of individuals for the purpose of describing behavior

Example, the mating behavior of birds, parent–child interactions on a playground, or the shopping behavior in a mall

Survey Research

A research study that uses a survey to obtain a description of a particular group of individuals

Example, Election Polling

Case Study

The in-depth study and detailed description of a single individual (or a very small group)

E.g., use in clinical psychology to determine rare phenomena

Observational Research Design

- As a measurement technique, behavioral observation can be used in a variety of research strategies including experimental and correlational designs

- A study using behavioral observation simply for descriptive purposes is classified as an observational research design.

Caution! Many researchers assume all research studies that use behavioral observation are observational research designs. However, observation can be used to measure variables in a variety of different designs. The defining element of an observational research design is that the results of the observations are used simply to describe the variable being studied.

Techniques to Quantify Observations

Frequency Method

Counting the instances of each specific behavior that occur during a fixed time observation period

E.g., The child committed three aggressive acts during the 30-minute period

Duration Method

Record how much time an individual spends engaged in a specific behavior during a fixed-time observation period

E.g., The child spent 18 minutes playing alone during the 30-minute period

Interval Method

Divide the observation period into a series of intervals and then recording whether a specific behavior occurs during each interval

E.g., The 30-minute observation period is divided into 30, 1-minute intervals

Strength and Weaknesses

Strength

- Researcher observes and records actual behavior; in contrast, survey research, for example, relies on the participants' *reports* of their behavior. Participants can distort or conceal the accuracy or truthfulness of their responses, and thus not reflect their actual behavior.

- Observational research have high external validity because most observational research is conducted in a field setting, and field research tends to have higher external validity.

- Flexibility. A researcher can complete a comprehensive observation of antecedents, behaviors, and consequences of the behaviors, whereas other studies examine a single, discrete behavior.

Weakness

- Ethical concern about spying on people. If participants are not aware that their behavior is being observed, the researcher may be violating a person's privacy and right to choose to participate in the study.

- Simply describe behavior and do not examine its causes.

Survey Research Design

Surveys and questionnaires are used extensively in the behavioral sciences as relatively efficient ways to gather large amounts of information

A researcher does not have to wait until a behavior or response occurs (e.g., it is not necessary to wait until after an election to discover people's attitudes about candidates or issues); we can ask at any time.

Obtain self-reported answers about attitudes, opinions, personal characteristics, and behaviors

Surveys can be used to obtain scores for a variety of different research designs, a survey often is conducted simply to obtain a description of a particular group of individuals. A study using the results from a survey simply for descriptive purposes is classified as a **survey research design.**

Rather than observe, we simply ask questions

…cont.

- Focuses on a specific characteristic such as eating behavior or political attitudes; or a more complex picture of a variety of behaviors and opinions(e.g., to investigate alcohol use at a local high school)

- Depending on the questions asked, the results could provide a description of how many students drink alcohol, how much they drink, and when and where.

- Other questions could yield a description of student attitudes toward alcohol use among their peers.

- A common application of survey research is by companies to obtain more accurate descriptions of their customers.
    - When you buy any electronic device, for example, a warranty registration card usually accompanies it. In addition to your name and address and the serial number of the product, other demographic questions are usually asked:
        - ✓ What is your age?
        - ✓ What is your occupation?
        - ✓ What is your income?
        - ✓ How did you hear about our product?
    - The purpose of these questions is to obtain the demographic characteristics of customers; to add description of the people who are likely to buy this product so that the company can do a better job of targeting its advertising.

Steps to Increase Accuracy in Survey Research

1. Survey questions must be developed.

2. The questions must be assembled and organized to produce a well-constructed survey.

3. A selection process must be developed to determine exactly who will participate in the survey and who will not; survey participants must be representative of the general group to be studied.

4. researchers must determine how the survey will be administered. Will participants receive printed surveys through the mail; will the survey questions be read to people over the telephone; or will participants complete the questions online in an Internet survey, or in person?

Types of Questions

- There are different ways to ask participants for self-report information.
    - You may be satisfied with a simple yes or no answer (Have you ever...)
    - You may want a quantitative answer (how much, how often).

- Different types of questions encourage different types of responses.
    - Different types of questions permit different degrees of freedom in the participants' answers.
    - E.g., a question may severely restrict response options (**Which of the following three flavors of ice cream do you prefer?**), or a question may give each participant complete freedom in choosing a response (**What is your favorite ice cream flavor?**).

- The wording of a question also can introduce bias into participants' answers (Are you one of those bland, unimaginative people who prefer vanilla ice cream?).

- Different types of questions permit different types of statistical analysis and interpretation.
    - If answers are limited to non-numerical categories on a nominal scale, for example, you cannot compute a group average.

Constructing a Survey

- Once the survey questions are determined, the next step is to organize the questions into a coherent survey that participants can easily understand and complete.

- There are a few general guidelines for creating a well-organized survey.
    1. Demographic questions (such as age, gender, level of education) should be placed at the end of the survey. These items are considered boring, and you do not want participants to quit because they are bored by the first few questions. In addition, identifying age, race, or gender first may influence how the participant answers survey questions that relate to these variables.
    2. Sensitive questions or items that may cause embarrassment or discomfort should be placed in the middle of the survey. By the time participants encounter these items, they are more likely to have warmed up to the topic and become committed to completing the survey.
    3. Questions dealing with the same general topic should be grouped together. Also, questions in the same format should be grouped together; e.g., all rating-scale questions should be grouped together. Grouping questions simplifies the survey so participants do not have to jump from one topic to another or switch from one type of question to another.
    4. If participants are going to read the survey, the format for each page should be relatively simple and uncluttered. Questions that are crammed together and seem to fill every square inch of the page create an overwhelming appearance that can intimidate participants.
    5. Vocabulary and language style should be easy for participants to understand. A survey with language appropriate for college students probably would not be appropriate for elementary school students.

Source Rea and Parker (2005) and Dillman, Smyth, and Christian (2009)

Selecting Relevant and Representative Individuals

- Quantitative research concern with generalization and representativeness

- The survey research design introduces a few additional concerns regarding sample selection.

- First, many surveys address a specific issue that is relevant to only a small subset of the general population.
    - For such a survey, care must be taken to select survey participants to whom the questions are relevant.
    - For a survey about childcare issues, for example, participants should be parents with small children.
    - A sensible strategy
      might be to hand out surveys to parents as they pick up their children at childcare centers around the city.
    - You might obtain mailing lists from the different childcare centers.
    - Participants for a shopping survey might be selected from the people in a shopping mall, and participants for an education survey could come from the parents of children in the local school district.

…cont.

- Second, there are some surveys seek to describe a broad cross-section of the general population.
  - In this case, the sample of survey participants must not be too restricted. E.g., administering surveys to the students in a psychology class would not result in an accurate description of the political attitudes of people in the community. A researcher should take some time to identify the group to be described, then make an effort to select individuals who accurately represent the group. This means that the individuals who participate in the survey are not necessarily the ones who are easiest or most convenient to obtain.
  - Seek for professional help preparing surveys and identifying participants. There are several research companies that design, administer, and analyze surveys. These companies usually have access to specialized mailing lists that can focus on a specific, well-defined population.

Statistical Procedure: Univariate Analysis

- The simplest form of quantitative analysis
- The analysis of a single variable, for purposes of description.
    - The major purpose is to describe
- Does not deal with causes or relationships
- Univariate data uses central tendency: mean, mode, media
- Its use dispersion method like range, variance, max, min, quartiles, standard deviation
- Frequency distributions
- Its result show in bar graph, histogram, pie chart, line graph, box-and-whisker plot

Correlational Research

- The goal of the correlational research strategy is to examine and describe the associations and relationships between variables.

- More specifically, the purpose of a correlational study is **to establish that a relationship exists** between variables and to describe the nature of the relationship.

- Notice that the correlational strategy **does not attempt** to explain the relationship and makes no attempt to manipulate, control, or interfere with the variables.

- The data for a correlational study consist of two or more measurements, one for each of the variables being examined. Usually, the scores are obtained from the same individual. For example, a researcher might record IQ and measure creativity for each person in a group of college students. Or a researcher could record food consumption and activity level for each animal in a colony of laboratory rats. Measurements can be made in natural surroundings or the individuals can be measured in a laboratory setting.

- The important factor is that the researcher simply measures the variables being studied. The measurements are then examined to determine whether they show any consistent pattern of relationship.

…cont.

- In the **correlational research strategy,** two or more variables are measured to obtain a set of scores (usually two scores) for each individual. The measurements are then examined to identify any patterns of relationship that exist between the variables and to measure the strength of the relationship.

Major Articles

Health-Related Variables and Academic Performance Among First-Year College Students: Implications for Sleep and Other Behaviors

Mickey T. Trockel MS, Michael D. Barnes PhD & Dennis L. Egget PhD

Pages 125-131 | Published online: 24 Mar 2010

66 Download citation   https://doi.org/10.1080/07448480009596294

- A correlational study by Trockel, Barnes, and Egget (2000) examining the relationship between grade point average and sleep habits, specifically wake-up time, for college students.
  - The researchers measured the grade point average and wakeup time for each individual in a group of college students and found that earlier wake-up times were consistently related to higher grade point averages.
  - Although the study demonstrated a relationship between the two variables, it does not explain why the relationship exists. Specifically, the results do not justify a conclusion that waking earlier causes higher grades (or that higher grades cause students to wake earlier)
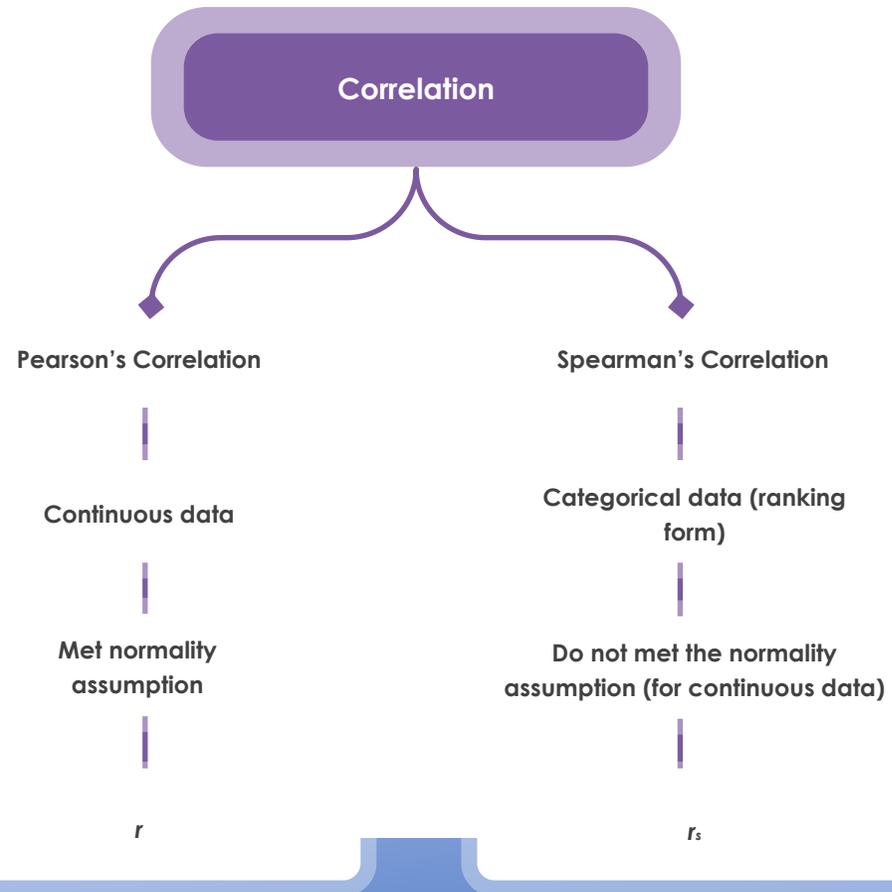
Data for a Correlational Study

- A correlational research study produces two or more scores for each individual. However, researchers are usually interested in the relationship between two variables at a time. Therefore, multiple scores are typically grouped into pairs for evaluation

- Traditionally, the scores in each pair are identified as *X* and *Y*. The data can be presented in a list showing the two scores for each individual or the scores can be shown in a graph known as a **scatter plot.**

- In the scatter plot, each individual is represented by a single point with a horizontal coordinate determined by the individual's *X* score and the vertical coordinate corresponding to the *Y* value.

Person	X	Y
A	1	3
B	6	8
C	4	4
D	7	12
E	9	10

Statistical Procedure

Correlation

Pearson's Correlation

Continuous data

Met normality assumption

r

Spearman's Correlation

Categorical data (ranking form)

Do not met the normality assumption (for continuous data)

r_s

Measuring Relationships

Correlation Coefficient, *r*

Characteristic 1	Characteristic 2	Characteristic 3

Direction of Relationship

Form of Relationship

Strength/ Magnitude of Relationships

+VE /-VE Relationships

+ V E
- If X increase and Y Increase
- If X decrease and Y Decrease

- V E
- If X increase and Y decrease
- If X decrease and Y increase

Linear Relationship

- Data located around a straight line

- +1.00/ -1.00 indicates perfect relationship
- 0 indicates no relationship

Guildford Rule of Thumb

r	Strength of Relationship
< .2	Negligible Relationship
.2 - .4	Low relationship
.4 - .7	Moderate relationship
.7 - .9	High relationship
> .9	Very high relationship

Cohen Rule of Thumb

r	Strength of Relationship
.1	Weak
.3	Medium
.5	High

Cohen, S. (1988). Statistical Power Analysis for Behavioral Science (2nd ed.).
New Jersey: Lawrence Erlbaum Associates

- **Correlational studies provide the basic information needed to make predictions**
- **(e.g., research shows a good positive relationship between Scholastic Assessment Test scores and future grade point average in college (Camera & Echternacht, 2000; Geiseer & Studley, 2002)).**
- **Regression - goal is to find the equation that produces the most accurate predictions of *Y* (the criterion variable) for each value of *X* (the predictor variable).**

- **Reliability evaluates the consistency or stability of the measurements**
- **Validity evaluates the extent to which the measurement procedure actually measures what it claims to be measuring.**

Prediction

Evaluating Theories

Applications of Correlational Strategy

Reliability and Validity

Interpreting a Correlation

Evaluating and validating theories

Prediction (Example)

- Ng and Jeffery (2003) used regression to predict health behaviors for working adults using stress as the predictor variable. The results showed that higher levels of stress predicted a higher-fat diet and more cigarette smoking. However, stress was not a significant predictor of alcohol use.

- In situations in which a correlational study is not used for prediction, researchers still tend to refer to a predictor and a criterion variable. In these situations, the labels are usually determined by the purpose of the study.

- Typically, a correlational study begins with one of the two variables relatively known or understood, and the second variable is relatively unknown. Thus, the purpose of the study is to gain a better understanding of the unknown variable by demonstrating that it is related to an established, known variable.

- In this situation, the known variable is designated as the predictor and the unknown variable as the criterion. For example, researchers have found a positive relationship between IQ and processing speed in a variety of perceptual and cognitive tasks (Eysenck, 1999). In this research, IQ would be the predictor variable and speed would be the criterion variable.

- When a correlational study demonstrates a relationship between two variables, it allows researchers to use knowledge about one variable to help predict or explain the second variable. In this situation, the first variable is called the **predictor variable** and the second variable (being explained or predicted) is called the **criterion variable.**

Reliability and Validity

- Both reliability and validity are commonly defined by relationships that are established using the correlational
research design.

- For example, test-retest reliability is defined by the relationship between an original set of measurements and a follow-up set of measurements. If the same individuals are measured twice under the same conditions,
and there is a consistent relationship between the two measurements, then the measurement procedure is said to be reliable.

- The concurrent validity of a measurement procedure can also be defined in terms of a relationship. If a new test is developed to detect early-stage Alzheimer's disease, for example, the validity of the test can be established by demonstrating that the scores from the test are strongly related to scores from established tests.

- This is exactly what was done by Ijuin et al. (2008) to validate a relatively new 7-minute test that was developed as an alternative to other commonly used screening tests for Alzheimer's.

  - Correlations were computed to measure the relationship between the scores from the 7-Minute Screen and the scores from each of the three established cognitive tests for Alzheimer's.
  - The researchers obtained correlations around 0.70 for each test, indicating a strong positive relationship and high concurrent validity between the 7-Minute Screen and established screening tests.

Evaluating Theories

- Many theories generate research questions about the relationships between variables that can be addressed by the correlational research design.

- A good example comes from the age-old nature/nurture question as it applies to intelligence: "Is intelligence primarily an inherited characteristic, or is it primarily determined by environment?"

- A partial answer to this question comes from correlational studies examining the IQs of identical twins separated at birth and placed in different environments.

- Because these twins have identical heredity and different environments, they provide researchers with an opportunity to separate the two factors.

- The original work in this area, conducted by British psychologist Cyril Burt, showed a strong relationship between the twins' IQs, suggesting that hereditary factors overwhelmed environment (Burt, 1972).

- However, later evidence showed that Burt probably falsified much of his data (Kamin, 1974). Nonetheless, correlational results suggest a strong relationship between twins' IQs. Note that the correlational research design is being used to address a theoretical issue.

Interpreting a Correlation

- The numerical value of a correlation, ranging from 0.00 to 1.00, describes the consistency of the relationship with 1.00 (or −1.00) indicating a perfectly consistent relationship and 0.00 indicating a complete lack of consistency.

- However, there are two additional factors that must be considered when interpreting the strength of a relationship.

    i.   One is the coefficient of determination, which is obtained by squaring the correlation, and

    ii. The significance of the correlation.

- *The Strength of a Relationship*
  The most common technique for measuring the strength of the relationship between two variables is to compute the **coefficient of determination,** which is obtained by squaring the numerical value of the correlation.

- Because a correlation is typically identified by the letter r, the coefficient of determination is $r2$.

- This coefficient measures how much of the variability in one variable is predictable from its relationship with the other variable.

- For example, if two college students are randomly selected, they will almost certainly have different grade point averages. Although there are many explanations for different grades, one possibility is that the two students have different IQs.

- In general, there is a tendency for higher IQs to correlate with higher grades. If the correlation between IQ and grade point average is calculated and then squared, the result provides a measure of how much of the differences in grade point averages can be predicted by IQ scores.

- A correlation of r 0.80 would mean that $r2$ 0.64 (or 64%) of the differences in grade point average can be predicted by difference in IQ. A correlation of r 0.30 would mean that only 0.09 (9%) of the differences are predictable.

Guidelines for Interpreting the Strength of a Correlation

Degree of Relationship	Value of the Correlation Coefficient or Coefficient of Determination
Small	$r = 0.10$ or $r^2 = 0.01$ (1%)
Medium	$r = 0.30$ or $r^2 = 0.09$ (9%)
Large	$r = 0.50$ or $r^2 = 0.25$ (25%)

…cont.

- ***The Significance of a Relationship***

- Is an important factor for interpreting the strength of a correlation.

- Term *significant* means that a correlation found in the sample data is **very unlikely to have been produced by random variation**.
  - Whenever a sample correlation is found to be significant, you can reasonably conclude that it represents a real relationship that exists in the population.

- With a small sample, it is possible to obtain what appears to be a very strong correlation when, in fact, there is absolutely no relationship between the two variables being examined.
  - For example, with a sample of only two individuals, there are only two data points and they are guaranteed to fit perfectly on a straight line.
  - With a sample of two individuals, you will always obtain a perfect correlation of 1.00 (or −1.00) no matter what variables you are measuring.
  - As the sample size increases, it becomes increasingly more likely that the sample correlation accurately represents the real relationship that exists in the population.

- A correlation found in a relatively large sample is usually an indication of a real, meaningful relationship and is likely to be significant.

- A statistically significant correlation does not necessarily mean that the correlation is large or strong.

- With a very large sample, for example, it is possible for a correlation of r 0.10 or smaller to be statistically significant. Clearly, this is not a strong correlation.

Strength and Weaknesses

- Often used for the preliminary work in an area that has not received a lot of research attention.

- The correlational design can identify variables and describe relationships between variables that might suggest further investigation using the experimental strategy to determine cause-and-effect relationships.

- Allows researchers an opportunity to investigate variables that would be impossible or unethical to manipulate.

- For example, a correlational study could investigate how specific behaviors or skills are related to diet deficiencies or exposure to pollution.

- Although it is possible and ethical to record diet deficiencies and environmental pollution as they exist naturally, it would not be ethical to create these conditions in the laboratory.

- Countless other variables such as family size, personality, alcohol consumption, level of education, income, and color preferences can be interesting topics for behavioral research but cannot be manipulated and controlled in an experimental research study.

- However, these variables can be easily measured and described in correlational research.

…cont.

- One of the primary advantages of a correlational study is that the researcher simply records what exists naturally.

- Because the researcher does not manipulate, control, or otherwise interfere with the variables being examined or with the surrounding environment, there is good reason to expect that the measurements and the relationships accurately reflect the natural events being examined.

- In research terminology, correlational studies tend to have high external validity. In general, a correlational study can establish that a relationship exists and it can provide a good description of the relationship.

- However, a correlational study usually does not produce a clear and unambiguous explanation for the relationship.

- In research terminology, correlational studies tend to have low internal validity.

- In particular, two limitations arise in explanations of results from a correlational study.
1. *The third-variable problem.* Although a correlational study may establish that two variables are related, it does not mean that there must be a direct relationship between the two variables. It is always possible that a third (unidentified) variable is controlling the two variables and is responsible for producing the observed relation, this is known as the **third-variable problem**.

- A recent television news program, for example, reported that higher participation in a company's fitness training program was associated with higher employee productivity and lower absenteeism. However, the company cannot conclude that their fitness program is causing benefits to the company; it may be that the employees who regularly participate were already healthier and had a higher level of fitness than those who rarely participate. Thus, a third variable (preexisting health) may be controlling both participation and productivity, resulting in the observed relationship

…cont.

- 2. *The directionality problem.* A correlational study can establish that two variables are related; that is, that changes in one variable tend to be accompanied by changes in the other variable.

- However, a correlational study does not determine which variable is the cause and which is the effect, is known as the **directionality problem**.

- For example, a recent study has found a relationship between exposure to sexual content on television and sexual behavior among adolescents (Collins, Elliott, Berry, Kanouse, Kunkel, Hunter, & Miu, 2004).

- Given this relationship, it is tempting to conclude that watching sex on television causes adolescents to engage in sexual behavior.

- However, it is possible that the true causal relationship is in the opposite direction. Adolescents who tend to be sexually active could simply choose to watch television programs that are consistent with their own behaviors. In this case, sexual behavior causes the teenager to prefer television programs with sexual content.

…cont.

A Summary of the Strengths and Weaknesses of the Correlational
Research Design

Strengths	Weaknesses
Describes relationships between variables	Cannot assess causality
Nonintrusive—natural behaviors	Third-variable problem
High external validity	Directionality problem
	Low internal validity

Relationships with more than Two Variables

- Most studies prefer a multitude of variables.

- For example, academic performance is probably related to IQ as well as to a number of other cognitive variables such as motivation, self-esteem, social competence, and a variety of other personal characteristics.

- One commonly used technique for studying multivariate relationships is a statistical procedure known as **multiple regression.**

- The underlying concept is that one criterion variable such as academic performance can be explained or predicted from a set of predictor variables such as IQ and motivation.

- IQ predicts part of academic performance, but you can get a better prediction if you use IQ and motivation together.

- For example, Collins and Ellickson (2004) evaluated the ability of four psychological theories to predict smoking behavior for adolescents in 10th grade. Although all four theories were good independent predictors, an integrated model using multiple regression to combine predictors from all four theories was more accurate than any of the individual models.

…cont.

- One interesting use of multiple regression is to examine the relationship between two specific variables while controlling the influence of other, potentially confounding variables.

- By adding predictor variables one at a time into the regression analysis, it is possible to see how each new variable adds to the prediction after the influence of the earlier predictors has already been considered.

- Earlier, we discussed a correlational study examining the relationship between adolescents' sexual behavior and the sexual content of the television programs they watch (Collins, Elliott, Berry, Kanouse, Kunkel, Hunter, & Miu, 2004). Because the age of the participants ranged from 12 to 17 years, the researchers were aware that participant age could create a third-variable problem.

- Specifically, the older the participants are, the more likely it is that they watch television programs with sexual content and that they engage in sexual behaviors. Thus, the participants' age can create an artificial relationship between sexual content and sexual behavior; individuals who watch less sexual content tend to engage in less sexual behavior (the younger participants), and individuals who watch more sexual content tend to engage in more sexual behavior (the older participants). However, the researchers were able to use multiple regression to eliminate this problem.

- Sexual content of the television programs was entered into the regression equation after the effects of age (and other variables) had been removed. The results indicated that sexual content still was a significant predictor of adolescent sexual behavior.

…cont.

- The language used to discuss and report the results from a multiple regression can be misleading. For example, you will occasionally see reports that the predictor variables *explained* the observed differences in the criterion variable. For example, a report might say that regression has demonstrated that variables such as intelligence, personality, and work drive *explain* differences in student grades. The truth is that the predictor variables only *predict* student grades; they do not really explain them. To get a cause-and-effect explanation, you must use the experimental research strategy. Unless a research study is using the experimental strategy (including manipulation and control), the best you can do is to describe relationships, not explain them.