Chapter 2
Personality Assessment, Measurement, and Research Design

Chapter Overview

This chapter provides students with an introduction to the sources of personality data, research designs in personality, and how personality measures are evaluated. The authors first address the four primary sources of data collected by personality psychologists. These are self-report data (S-data), observer-report data (O-data), test-data (T-data), and life-outcome data (L-data). The authors then address the conditions where there exists links among the different sources of data and also the conditions where there is no linkage among them. Because personality data are fallible, the authors recommend collecting data from more than one data source. Results that transcend data sources are more powerful. The authors then discuss how personality measures are evaluated. This section of the chapter includes discussions of a measure's reliability, validity, and generalizability. Next, the authors discuss the three key research methods used by personality psychologists. These methods are experimental designs, correlational designs, and case studies. Each research method has its own strengths and weaknesses. The strength of one design is a weakness of another, and the weakness of one design is a strength of another. The authors note that the type of design one uses will depend on the research question and the purpose of the investigation. The authors close by noting that no source of data is perfect and that no research method is perfect. Whether a data source or method is appropriate will depend on the research question and the purpose of the research.

Learning Objectives

1. Describe and provide examples of the four sources of data collected by personality psychologists: Self-report data (S-data), observer-report data (O-Data), test-data (T-data), and life-outcome data (L-data).

2. Identify the strengths and weaknesses of each source of personality data.

3. Discuss how each source of personality data can provide information not provided by other sources of data.

4. For O-data, discuss the problems of selecting observers (naturalistic and artificial

observations).

5. For T-data, discuss the strengths and weaknesses of mechanical recording devices and physiological recording devices, and provide examples of each type of device.

6. For T-data, provide and discuss examples of projective techniques, and identify the strengths and weaknesses of such sources of data.

7. Discuss the conditions under which one might expect links among different sources of data and how the presence or absence of these links can be interpreted.

8. Define reliability, and discuss test-retest reliability, inter-rater reliability, and internal consistency reliability.

9. Define validity, and discuss face validity, predictive or criterion validity, convergent validity, discriminative validity, and construct validity.

10. Define and discuss generalizability, and examine the different contexts to which a measure might be generalizable.

11. Provide and describe examples of the three types of research methods used by personality psychologists: experimental methods, correlational designs, and case studies.

12. Identify the strengths and weaknesses of each type of research method used by personality psychologists.

13. Identify and discuss when it might be appropriate to use only one of the three research methods.

14. Discuss how each type of research method provides information not provided by the other research methods.

Chapter Outline

I. Sources of Personality Data

- The most obvious source of information about a person is **self-report data (S-data)**—the information a person reveals.

A. Self-Report Data (S-Data)

- Self-report data can be obtained through a variety of means, including interviews that pose questions to a person, periodic reports by a person to record the events as they happen, and questionnaires.
- One good reasons for using self-report is that individuals have access to a wealth of information about themselves that is inaccessible to anyone else, such as their habitual level of anxiety (e.g., Vazire, 2010).
- Self-report can take a variety of forms, ranging from open-ended "fill in the blanks" to forced-choice true-or-false questions.
- Sometimes these are referred to as **unstructured** (open-ended, such as "Tell me about the parties you like the most") and **structured** ("I like loud and crowded parties"—answer "true" or "false") personality tests.
- **Likert rating scale** is a complex method, which involves requesting participants to indicate in numerical form the degree to which each trait term characterizes them, say on a 7-point rating scale of 1 (least characteristic) to 7 (most characteristic).
- For the self-report method to be effective, respondents must be both willing and able to answer the questions put to them.
    - People are not always honest, especially when asked about unconventional experiences, such as unusual desires, unconventional sex practices, and undesirable traits.
    - Some people may lack accurate self-knowledge.

B. Observer-Report Data (O-Data)

- **Observer-report data (O-data)** capitalizes on sources (friends, families, teachers, and casual acquaintances as observers) for gathering information about a person's personality.
- One advantage of using observer-report is that observers may have access to information not attainable through other sources.
- A second advantage of observer-reports is that multiple observers can be used to assess each individual, whereas in self-report only one person provides information (Connelly & Ones, 2010; Paunonen & O'Neill, 2010).
- The use of multiple observers allows investigators to evaluate the degree of agreement among observers—also known as **inter-rater reliability**.
- Personality researchers have developed the following strategies for selecting observers:
    - One strategy is to use professional personality assessors who do not know the participant in advance.
    - The other strategy is to use individuals who actually know the target participants.

        - One advantage of using such strategy is that observers are in a better position to observe the target's natural behavior.
        - A second advantage of using intimate observers is that **multiple social personalities** can be assessed (Craik, 1986, 2008).
        - Because intimate observers have relationships with the target person, using such strategy may be biased in certain ways.
- In **naturalistic observation,** observers witness and record events that occur in the normal course of the lives of their participants.
- In contrast, observation can take place in contrived or artificial settings, such as occur at the Institute for Personality and Social Research (IPSR).
- Naturalistic observation offers researchers the ability to secure information in the realistic context of a person's everyday life, but at the cost of not being able to control the events and behavioral samples witnessed.
- Observation in experimenter-generated situations has the advantage of controlling conditions and eliciting the relevant behavior.
    - But this advantage comes at a cost—sacrificing the realism of everyday life.

C. Test-Data (T-Data)

- A third common source of personality-relevant information comes from standardized tests—**test data (T-data)**.
- Participants of this test are placed in a standardized testing situation.
- The idea is to see if different people react differently to an identical situation.
- The situation is designed to elicit behaviors that serve as indicators of personality variables (Block, 1977).
- The following are the limitations of T-data:
    - Some participants might try to guess what trait is being measured and then alter their responses to create a specific impression of themselves.
    - Verifying that research participants define the testing situation in the same way as the experimenter is difficult.
    - The situations are inherently *interpersonal,* and a researcher may inadvertently influence how the participants behave.
- Personality psychologists have been enterprising in adapting technological innovations.
    - The actometer is an example of one such technology, which is essentially a modified self-winding watch, which can be strapped to the arms or legs of participants (typically, children).
        - Movement activates the winding mechanism, registering the person's activity on the hands of the dial.
- Mechanical recording devices provide a mechanical means of assessing personality,

unhampered by the biases that might be introduced when a human observer is involved.
    - They can be obtained in relatively naturalistic settings—such as a children's playground.
    - Their primary disadvantage is that few personality dispositions lend themselves to being assessed by mechanical devices.
- Physiological measures can provide information about a person's level of arousal, reactivity to various stimuli, and the speed at which he or she takes in new information—all potential indicators of personality.
- A more recent physiological data source comes from **functional magnetic resonance imaging (fMRI),** a technique used to identify the areas of the brain that "light up" when performing certain tasks such as verbal problems or spatial navigation problems.
- One of the key benefits of physiological data is that it is difficult for participants to fake responses, particularly on measures of arousal or reflexive responses, such as the "eyeblink startle reflex."
- Physiological recording procedures, the recording is typically constrained by a relatively artificial laboratory situation.
- Another type of T-data is **projective techniques,** in which the person is given a standard stimulus and asked what he or she sees.
- The hallmark of any projective technique is that the person is presented with an ambiguous stimulus, such as an inkblot, and then asked to impose structure on this stimulus by describing what he or she sees—for example, what is in the inkblot.
- The idea behind projective techniques is that what the person sees in the stimulus reveals something about his or her personality. Presumably, the person "projects" his or her concerns, conflicts, traits, and ways of seeing or dealing with the world onto the ambiguous stimulus.
- Psychologists who advocate projective measures argue that they are useful for getting at wishes, desires, fantasies, and conflicts that the participants themselves may be unaware of and so could not report on a questionnaire.
    - Others are critical of projectives, questioning their validity and reliability as accurate measures of personality (Wood, Nezworski, & Stejskal, 1996).

D. Life-Outcome Data (L-Data)

- **Life-outcome data (L-data)** refers to information that can be gleaned from the events, activities, and outcomes in a person's life that are available to public scrutiny (for example, marriages and divorces are a matter of public record).
- L-data can serve as an important source of real-life information about personality.

E. Issues in Personality Assessment

- The issues can be classified into two broad categories:
  - The first issue involves using two or more data sources within a single personality study; what are the links among the various sources of personality data?
  - The second issue involves the fallibility of personality measurement and how the use of multiple data sources can correct some of the problems associated with single data sources.
- One powerful strategy of personality assessment is to examine results that transcend data sources—a procedure sometimes referred to as *triangulation*.

II. Evaluation of Personality Measures

A. Reliability

- **Reliability** can be defined as the degree to which an obtained measure represents the true level of the trait being measured.
- One way to estimate reliability is through **repeated measurement**.
- A common procedure is to repeat a measurement over time for the same people.
  - If the two tests are highly correlated, yielding similar scores for most people, the resulting measure is said to have high *test-retest reliability*.
- A second way to gauge reliability is to examine the relationships among the items themselves at a single point in time.
  - If the items within a test—viewed as a form of repeated measurement—all correlate well with each other, then the scale is said to have high *internal consistency reliability*.
- A third way to measure reliability—applicable only to the use of observer-based personality measures—is to obtain measurements from multiple observers.
  - When different observers agree with each other, the measure is said to have high *inter-rater reliability*.

B. Response Sets

- The concept of **response sets** refers to the tendency of some people to respond to the questions on a basis that is unrelated to the question content.
  - Sometimes this is also referred to as **noncontent responding.**
- One example of the response set is **acquiescence,** or yea saying.
- **Extreme responding** is another response set, which refers to the tendency to give endpoint responses, such as “strongly agree” or “strongly disagree” and to avoid the middle part of response scales, such as “slightly agree” or “slightly disagree.”

- Another type of response set called **social desirability,** which is the tendency to answer items in such a way as to come across as socially attractive or likable.
- The following are the two views regarding the interpretation of social desirability.:
  - It represents distortion and should be eliminated or minimized.
  - It is a valid part of other desirable personality traits, such as happiness, conscientiousness, or agreeableness.
- The following are the different approaches one can take to solve the problems of socially desirable responses:
  - Assume that the responses are erroneous or deceptive, to measure this tendency, and to remove it statistically from the other questionnaire responses.
  - Develop questionnaires that are less susceptible to this type of responding.
  - Use a **forced-choice questionnaire** format—test takers are confronted with pairs of statements and are asked to indicate which statement in each pair is more true of them.
- Work on social desirability has attempted to disentangle self-deceptive optimism from impression management.

C. Validity

- **Validity** refers to the extent to which a test measures what it claims to measure (Cronbach & Meehl, 1955; Wiggins, 2003).
- The following are the five types of validity:
  - **Face validity**: It refers to whether the test, on the surface, appears to measure what it is supposed to measure.
  - **Predictive validity**: It refers to whether the test predicts criteria external to the test (thus it is sometimes called **criterion validity**).
  - **Convergent validity**: It refers to whether a test correlates with other measures that it should correlate with.
  - **Discriminant validity**: It is often evaluated simultaneously with convergent validity. Whereas *convergent validity* refers to what a measure *should* correlate with, *discriminant validity* refers to what a measure *should not* correlate with.
  - **Construct validity**: It is defined as a test that measures what it claims to measure, correlates with what it is supposed to correlate with, and does not correlate with what it is not supposed to correlate with.
    - This form of validity is called construct validity because it is based on the notion that personality variables are **theoretical constructs**.

D. Generalizability

- **Generalizability** is the degree to which the measure retains its validity across various contexts.
- One context of interest might be different groups of *persons*.
- Another facet of generalizability refers to *different conditions*.

III. Research Designs in Personality

A. Experimental Methods

- **Experimental methods** are typically used to determine causality—that is, to find out whether one variable *influences* another variable.
- In order to establish the influence of one variable on another, the following two key requirements of good experimental design must be met:
  - **Manipulation** of one or more variables
  - Ensuring that participants in each experimental condition *are equivalent to each other at the beginning of the study*
- If an experiment has manipulation among groups, then the **random assignment** of participants to experimental groups is a procedure that helps ensure that all groups are the same at the beginning of a study.
- Equivalence is obtained, in experiments where manipulation is within each single group, by **counterbalancing** the order of the conditions.

B. Correlational Studies

- In the **correlational method,** a statistical procedure is used for determining whether there is a relationship between two variables.
- Correlational designs typically try to determine what goes with what in nature.
- A major advantage of correlational studies is that they allow researchers to identify relationships among variables as they occur naturally.
- The most common statistical procedure for gauging relationships between variables is the **correlation coefficient**.
  - Correlation coefficients can range from +1.00 through 0.00 to −1.00.
- It is important to keep in mind that one cannot infer causation from correlations.
- There are at least two reasons correlations can never prove causality.
  - **Directionality problem**
  - **Third variable problem.**

C. Case Studies

- Sometimes a personality researcher is interested in examining the life of one person in-depth as a case study.
- The following are the advantages of using the **case study method**:
  - Researchers can find out about personality in great detail, which rarely can be achieved if the study includes large samples.
  - Case studies can give researchers insights into personality that can then be used to formulate a more general theory to be tested on a larger population.
  - The method can provide in-depth knowledge of particularly outstanding individuals, such as Mahatma Gandhi or Martin Luther King.
- Despite the strengths of the in-depth case study method, it has some critical limitations.
  - The most important one is that findings based on one individual cannot be generalized to other people.

D. When to Use Experimental, Correlational, and Case Study Designs

- Each of the three major types of research designs has strengths and weaknesses or, more precisely, questions that each is good at answering and questions that each is poor at answering.
- The experimental method is ideally suited for establishing causal relationships among variables.
- Correlational designs are ideally suited for establishing the relationships between two or more variables that occur in everyday life, such as between height and dominance, conscientiousness and grade point average, or anxiety and frequency of illness.
- Case studies are ideally suited for generating hypotheses that can be tested subsequently using correlational or experimental methods.

Summary and Evaluation

Personality assessment and measurement start with identifying the sources of personality data—the places from which researchers obtain information about personality. The four major sources of personality data are self-report (S-data), observer report (O-data), laboratory tests (T-data), and life history outcomes (L-data). Each of these data sources has strengths and weaknesses. In self-report, for example, participants might fake or lie. Observers in the O-data mode may lack access to the relevant information. Laboratory tests may be inadequate for identifying patterns that occur naturally in everyday life. Each source of personality data is extremely valuable, however, and each provides information not attainable through the other sources. Furthermore, new measurement techniques continue to be invented and explored; recent examples include assessments through electronic and internet technologies such as Smart Phones, Instagram, and Facebook.

Once sources of data have been selected for measuring personality, the researcher then evaluates their quality. Personality measures, ideally, should be reliable in the sense of attaining the same scores through repeated measurement. They should be valid, measuring what they are supposed to measure. And researchers should establish how generalizable their measures are—determining the people, settings, and cultures to which the measure is most applicable. Scales applicable only to college students in the United States, for example, are less generalizable than scales applicable to people of differing ages, economic brackets, ethnic groups, and cultures.

The next step in personality research involves selecting a particular research design within which to use the measures. There are three basic types of research designs. The first, the experimental research design, which involves controlling or manipulating the variables of interest, is best suited to determining causality between two variables. The second, correlational research design, is best for identifying relationships between naturally occurring variables but is poorly suited to determining causality. The third is the case study method, which is well suited to generating new hypotheses about personality and to understanding single individuals. Perhaps the most important principle of personality assessment and measurement is that the decisions about data source and research design depend heavily on the purpose of the investigation. There are no perfect methods; there are no perfect designs. But there are data sources and methods that are better suited for some purposes than for others.

Key Terms

Self-report data (S-data)
Structured
Unstructured
Likert rating scale
Experience sampling
Observer-report data (O-data)
Inter-rater reliability
Multiple social personalities
Naturalistic observation
Test data (T-data)
Functional magnetic resonance imaging (fMRI)
Projective techniques
Life-outcome data (L-data)
Reliability
Repeated measurement
Response sets
Noncontent responding
Acquiescence
Extreme responding
Social desirability
Forced-choice questionnaire
Validity
Face validity
Predictive validity
Criterion validity
Convergent validity
Discriminant validity
Construct validity
Theoretical constructs
Generalizability
Experimental methods

Manipulation
Random assignment
Counterbalancing
Statistically significant
Correlational method
Correlation coefficient
Directionality problem
Third variable problem
Case study method

Lecture Topics and Lecture Suggestions

1. *Personality and Mate Preferences: Five Factors in Mate Selection and Marital Satisfaction* (Botwin, Buss, & Shackelford, 1997)

   Students will appreciate the presentation of a research paper in personality psychology that employs multiple sources of data. In addition, the topics of mate preferences, mate selection, and relationship satisfaction are consistently well received. Instructors can use this study as a springboard for discussing the different sources of data, the issues and limitations of self-report and observer-report, as well as the relationship of personality to "real world" outcomes such as relationship satisfaction.

   - Personality characteristics play a prominent role in what people want in a mate (see, e.g., Buss, 2004, for a review).
   - Little is known, however, about the answers to the following questions:
     - Which personality characteristics are most important among mate preferences?
     - Do men and women differ in their personality preferences?
     - Do men and women, individually, differ in what they want in a mate, and do individuals actually get what they want in a mate?
   - To explore these issues, two parallel studies were conducted, one using a sample of dating couples (N = 118) and another with a sample of married couples (N = 216).
   - The five-factor model (FFM) of personality (which proposes that there are five major dimensions covering the range of personality variations: surgency or extraversion, agreeableness, emotional stability, conscientiousness, and openness/intellect) guided the investigation.
   - The FFM, operationalized in adjectival form, was used to assess personality characteristics from the following three sources of data:
     - Self-report (S-data)
     - Partner-report (O-data)
     - Independent interviewer-report (O-data)
   - Participants were evaluated on a parallel 40-item instrument to gain information on their preferences for the ideal personality characteristics of their mates.
   - Results were consistent across both the studies.

- Women expressed greater preference than men for a wide array of socially desirable personality traits.
- Individuals differed in the characteristics they desired, preferring mates who were similar to themselves.
  - They actually obtained mates who embodied what they desired.
- Personality characteristics of one's partner significantly predicted marital and sexual dissatisfaction, most notably when the partner was lower than desired on agreeableness, emotional stability, and openness/intellect.

References:

Botwin, M. D., Buss, D. M., & Shackelford, T. K. (1997). Personality and mate preferences: Five factors in mate selection and marital satisfaction. *Journal of Personality, 65,* 107–136.

Buss, D. M. (2004). *The Evolution of Desire* (rev. ed.). New York: Basic Books.

2. *Personality and Day-to-Day Physical Symptoms* (Larsen & Kasimatis, 1991)

One of the research methodologies used to study personality, which is not explicitly discussed in Larsen and Buss is what is often called the "daily diary design." This design is similar to an experience sampling design, in that data are collected on an ongoing basis from the same set of participants. In daily diary studies, data are collected on a daily basis about events such as physical symptoms, emotions, and self-esteem. In addition, personality researchers often collect personality data either before or after the daily diary phase. Students will likely enjoy hearing about this sort of research design, which highlights the critical role of participants in the personality research work. In addition, the topic of the relationships between personality and health is likely to capture the interest of a large portion of students enrolled in personality psychology courses.

- Larsen & Kasimatis (1991) explored the relationship between personality and ongoing health status for a sample consisting of 43 undergraduates.
- The students completed mood and symptom reports three times a day for eight weeks.
- A daily event approach was used to model the following three temporal parameters of day-to-day health:
  - Occurrence rate of symptoms
  - Duration of symptoms
  - Covariation of symptoms and moods over time
- The researchers then determined if these variables related to the following three personality variables:

    - Neuroticism (emotional instability)
    - Anger/hostility
    - Type A behavior (excessive achievement striving, competitiveness, impatience, hostility, and vigorous speech and motor mannerisms)
- The following are findings from the study:
    - The occurrence of an illness related most strongly to neuroticism.
    - The duration of an illness related most strongly to the trait of aggressive responding.
    - Type A behavior related to less unpleasant effects reported during episodes of respiratory infection, aches, and depressive symptoms.
- The researchers conclude with a discussion of how alternative models of health/illness are made possible by the daily event perspective.

Reference:

Larsen, R. J., & Kasimatis, M. (1991). Day-to-day physical symptoms: Individual differences in the occurrence, duration, and emotional concomitants of minor daily illnesses. *Journal of Personality, 59,* 387–423.

Classroom Activities and Demonstrations

1. Distribute Activity Handout 2-1 (“Twenty Statements Test,” or TST) to the class. Have students take about five minutes to complete the test during class. Ask for volunteers to share their responses. Use this discussion as a springboard to talk about the TST, in particular, and the value of self-report data, more generally. Highlight that the self-report information that needs to be filled in the TST cannot be obtained from any other person except the students themselves. Finally, ask students to discuss what they think this test reveals about themselves.

2. Distribute Activity Handout 2-2 (“How Accurately Can You Describe Yourself?”). This is a measure of standings on the five factors of personality, or the “Big Five.” The Big Five include the following factors: surgency, agreeableness, conscientiousness, emotional stability, and openness/intellect. Give students about 5 minutes to complete the inventory. Then allow students about 10 minutes to score their responses. Ask students to write down the scoring instructions because they will need them to complete a future exercise (see number 3 below). This measure is scored as follows: To get a score for each of the five factors, take the mean of the indicated items. Items with an asterisk (*) should be reverse coded before entering it into the mean. The reverse code is as follows: 1 = 7, 2 = 6, 3 = 5, 5 = 3, 6 = 2, and 7 = 1

Surgency: 1, *6, *11, 16, 21, *26, 31, *36
Agreeableness: 2, *7, 12, *17, *22, 27, *32, 37
Conscientiousness: 3, 8, 13, 18, *23, *28, 33, *38
Emotional Stability: *4, *9, *14, 19, *24, 29, *34, 39
Openness/Intellect: 5, 10, 15, *20, 25, 30, *35, *40

This is a valuable exercise, not only because students will learn where they stand based on the five major personality dimensions, but also because students will participate firsthand in taking and scoring a personality test. They will appreciate how item scores are aggregated to form scale scores, for example.

3. After students have completed Activity Handout 2-2, distribute Activity Handout 2-3 ("How Accurately Can You Describe _____?"). Instruct students to have someone who knows them well rate them on the 40 items in the handout. Instruct students to bring the completed and scored measure to the next class session. Before the next class session, students should consider how their evaluations based on the five factors differ from their evaluations based on the observer-reports. Questions that could be raised for discussion include the following: How close were the results of the self-reported evaluation and your observer-reported evaluation on each of the factors? Which factors had the greatest discrepancy between self-report and observer-report? Which had the least? If there are discrepancies, which set of ratings is "correct?" Why?

Questions for In-Class Discussion

1. Self-report is a valuable tool for collecting personality data. However, self-report may not be appropriate for collecting certain classes of information. What might some of these classes of information be? Why might self-report be problematic for collecting these classes of information? Students often have lots of inputs when discussing these questions. If, however, students are sluggish to get started, instructors might provide a good example that can initiate a discussion. Criminal behavior, for example, may not be most appropriately assessed by self-report, because people may not be willing to report on how, when, and why they broke the law.

2. Larsen and Buss note that, if the same pattern of results is found with two or more data sources, then researchers can have greater confidence in the credibility of the findings. Ask students to discuss and elaborate on why this is the case. Relatedly, suggest and have students elaborate on the possibility that researchers should have greater confidence in a pattern of results if that pattern of results is documented using more than one research design.

3. Larsen and Buss discuss three basic research methods used by personality psychologists: Experiments, correlational studies, and case studies. The text focuses on different situations where each method is most appropriate. Have students discuss research questions that are not appropriately investigated by each of the three research methods. Students find it useful to discuss when each method is least appropriate. This discussion will further clarify the strengths and limitations of each method and will help students appreciate the fact that sometimes researchers simply cannot use a particular method, depending on the research question. The effects of child abuse on adult intelligence, for example, cannot ethically and legally be studied (at least not directly) using an experimental design. A correlational study or a case study would be more appropriate, ethically and legally.

Critical Thinking Essays

1. Larsen and Buss refer to Craik's (1987) proposal that people display "multiple social personalities." Discuss, in your own words, what it means to display multiple social personalities. Discuss how you might display multiple social personalities and briefly describe the key characteristics of each of these personalities. For example, you might present one personality when you are interacting with your mother, but a very different personality when you are interacting with your professor. Why do you think people display multiple social personalities?

2. According to Larsen and Buss, one of the issues that must be addressed by a researcher who wants to use observer-report data is the size of the observational unit. These units can be large, molar units, such as the global traits of intelligence, emotional stability, or conscientiousness. Or, they can be small, molecular units such as walking speed, number of miles per hour, or number of eye blinks. Develop a personality research question that is amenable to observational data, and describe how you might investigate this question using relatively molar units of observation. Specify the units of observation. Next, discuss how you might use relatively molecular units of observation. Again, clearly specify the units of observation. Given your research question, which observational unit that you proposed might be more appropriate and why?

3. The case study method is a valuable research method in personality psychology. A key limitation of this method, however, is that the results are based on a single individual, and therefore cannot be generalized to other people. Ask students why this is the case? Provide an example of a research question you might investigate using a case study, and discuss why it might be problematic to attempt to generalize the results of your investigation to

other people or a larger audience.

Research Papers

1. Larsen and Buss discuss four sources of data collected by personality psychologists. Conduct a search on the Internet or any other resources for psychological research literature, and locate four research articles published within the last five years. Select articles such that each of them uses only one of the four sources of data. For each article, first summarize the objectives, the process involved, and the findings of the research. Then, suggest how the researchers might have used each of the remaining three sources of data. Finally, address whether you think the results might have turned out differently if they had used different data sources; provide valid reasons for your answers.

2. Larsen and Buss note that there are three key issues that personality psychologists must address for a measure they have developed to assess a particular personality characteristic. These issues are reliability, validity, and generalizability. First, define, in your own words, what each of these concepts mean, and discuss the sub-types of reliability and validity. Next, conduct a search on the Internet or any other resources for psychological research literature. Identify an article that presents the development of a new measure of a personality trait or characteristic. Discuss how well the researchers address reliability, validity, and generalizability of the new measure. Did the researchers document these standards for the new measure? If you were a personality researcher, charged with ensuring that all aspects of the new measure's reliability, validity, and generalizability were well documented, what future research would you need to do on this new measure?

3. Larsen and Buss discuss three types of research designs used by personality psychologists. Conduct a search of the psychological research literature on the Internet or any other resources, and locate three research articles published within the last five years. Select articles such that each of them uses only one of the four sources of data. For each article, first summarize the objectives, the process involved, and the findings of the research. Then, suggest how the researchers might have used each of the remaining two research designs. Finally, provide valid reasons, and address whether you think the results might have turned out differently if they had used different research designs.

Recent Research Articles and Other Scholarly Readings

Anastasi, A. (1986). Evolving concepts of test validation. *Annual Review of Psychology, 37,* 1–15.

Ben-Porath, Y. S., & Waller, N. G. (1992). “Normal” personality inventories in clinical assessment: General requirements and the potential for using the NEO Personality Inventory. *Psychological Assessment, 4,* 14–19.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56,* 81–105.

Cervone, D., Shadel, W. G., & Jencius, S. (2001). Social-cognitive theory of personality assessment. *Personality and Social Psychology Review, 5,* 33–51.

Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist, 49,* 997–1003.

Costa, P. T., Jr., & McCrae, R. R. (2005). Normal personality assessment in clinical practice: The NEO Personality Inventory. *Psychological Assessment, 4,* 5–13.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52,* 281–302.

Embretson, S. E. (1996). The new rules of measurement. *Psychological Assessment, 8,* 341–349.

Harkness, A. R., & Lilienfeld, S. O. (1997). Individual differences science for treatment planning: Personality traits. *Psychological Assessment, 9,* 349–360.

Hogan, R., DeSoto, C. B., & Solano, C. (1977). Traits, tests, and personality research. *American Psychologist, 32,* 255–264.

Jensen, A. R. (1980). Précis of Bias in mental testing. *Behavioral and Brain Sciences, 3,* 325–371.

Jones, L. V., & Appelbaum, M. I. (1989). Psychometric methods. *Annual Review of Psychology, 40,* 23–43.

Lubinski, D. (2000). Scientific and social significance of assessing individual differences: “Sinking shafts at a few critical points.” *Annual Review of Psychology, 51,* 405–444.

Matthews, G., Saklofske, D. H., Costa, P. T., Jr., Deary, I. J., & Zeidner, M. (1998). Dimensional models of personality: A framework for systematic clinical assessment. *European Journal of Psychological Assessment, 14,* 36–49.

McReynolds, P. (1989). Diagnosis and clinical assessment: Current status and major issues. *Annual Review of Psychology, 40,* 83–108.

Messick, S. (1981). Constructs and their vissitudes in educational and psychological measurement. *Psychological Bulletin, 89,* 575–588.

Michell, J. (1997). Quantitative science and the definition of measurement in psychology. *British Journal of Psychology, 88,* 355–383.

Ozer, D. J., & Reise, S. P. (1994). Personality assessment. *Annual Review of Psychology, 45,* 357–388.

Activity Handout 2-1: Twenty Statements Test

Instructions: Please complete the following 20 statements. There are no right or wrong answers. Please write the first thing that comes to mind, and try not to censor yourself.

1. I am _____
2. I am _____
3. I am _____
4. I am _____
5. I am _____
6. I am _____
7. I am _____
8. I am _____
9. I am _____
10. I am _____
11. I am _____
12. I am _____
13. I am _____
14. I am _____
15. I am _____
16. I am _____
17. I am _____
18. I am _____
19. I am _____
20. I am _____

Activity Handout 2-2: How Accurately Can You Describe Yourself?

Instructions: Please read the following pairs of characteristics and circle the number that best describes you, in general. For example, for number 1, if you see yourself as more passive than active, you should circle a number closer to "passive." If you see yourself as more active than passive, you should circle a number closer to "active."

1.	passive	1	2	3	4	5	6	7	active
2.	cold	1	2	3	4	5	6	7	warm
3.	undependable	1	2	3	4	5	6	7	reliable
4.	emotionally stable	1	2	3	4	5	6	7	emotionally unstable
5.	uncultured	1	2	3	4	5	6	7	cultured
6.	energetic	1	2	3	4	5	6	7	unenergetic
7.	agreeable	1	2	3	4	5	6	7	disagreeable
8.	negligent	1	2	3	4	5	6	7	conscientious
9.	secure	1	2	3	4	5	6	7	insecure
10.	ignorant	1	2	3	4	5	6	7	knowledgeable
11.	dominant	1	2	3	4	5	6	7	submissive
12.	critical	1	2	3	4	5	6	7	lenient
13.	careless	1	2	3	4	5	6	7	careful
14.	at ease	1	2	3	4	5	6	7	nervous
15.	stupid	1	2	3	4	5	6	7	intelligent
16.	timid	1	2	3	4	5	6	7	bold
17.	flexible	1	2	3	4	5	6	7	stubborn
18.	disorganized	1	2	3	4	5	6	7	well organized
19.	high-strung	1	2	3	4	5	6	7	relaxed
20.	perceptive	1	2	3	4	5	6	7	imperceptive
21.	conforming	1	2	3	4	5	6	7	independent
22.	trusting	1	2	3	4	5	6	7	suspicious
23.	hardworking	1	2	3	4	5	6	7	lazy
24.	even-tempered	1	2	3	4	5	6	7	temperamental
25.	uncreative	1	2	3	4	5	6	7	creative
26.	proud	1	2	3	4	5	6	7	humble
27.	unfair	1	2	3	4	5	6	7	fair
28.	traditional	1	2	3	4	5	6	7	untraditional
29.	emotional	1	2	3	4	5	6	7	unemotional
30.	simple	1	2	3	4	5	6	7	complex
31.	quiet	1	2	3	4	5	6	7	talkative
32.	selfless	1	2	3	4	5	6	7	selfish
33.	liberal	1	2	3	4	5	6	7	conservative
34.	not envious/ not jealous	1	2	3	4	5	6	7	envious/ jealous

35.	curious	1	2	3	4	5	6	7	uncurious
36.	sociable	1	2	3	4	5	6	7	retiring
37.	stingy	1	2	3	4	5	6	7	generous
38.	practical	1	2	3	4	5	6	7	impractical
39.	subjective	1	2	3	4	5	6	7	objective
40.	analytical	1	2	3	4	5	6	7	unanalytical

Activity Handout 2-3: How Accurately Can You Describe _____?

Instructions: Please read the following pairs of characteristics and circle the number that best describes _____, in general. For example, for number 1, if you see _____ as more passive than active, you should circle a number closer to "passive." If you see _____ as more active than passive, you should circle a number closer to "active."

1.	passive	1	2	3	4	5	6	7	active
2.	cold	1	2	3	4	5	6	7	warm
3.	undependable	1	2	3	4	5	6	7	reliable
4.	emotionally stable	1	2	3	4	5	6	7	emotionally unstable
5.	uncultured	1	2	3	4	5	6	7	cultured
6.	energetic	1	2	3	4	5	6	7	unenergetic
7.	agreeable	1	2	3	4	5	6	7	disagreeable
8.	negligent	1	2	3	4	5	6	7	conscientious
9.	secure	1	2	3	4	5	6	7	insecure
10.	ignorant	1	2	3	4	5	6	7	knowledgeable
11.	dominant	1	2	3	4	5	6	7	submissive
12.	critical	1	2	3	4	5	6	7	lenient
13.	careless	1	2	3	4	5	6	7	careful
14.	at ease	1	2	3	4	5	6	7	nervous
15.	stupid	1	2	3	4	5	6	7	intelligent
16.	timid	1	2	3	4	5	6	7	bold
17.	flexible	1	2	3	4	5	6	7	stubborn
18.	disorganized	1	2	3	4	5	6	7	well organized
19.	high-strung	1	2	3	4	5	6	7	relaxed
20.	perceptive	1	2	3	4	5	6	7	imperceptive
21.	conforming	1	2	3	4	5	6	7	independent
22.	trusting	1	2	3	4	5	6	7	suspicious
23.	hardworking	1	2	3	4	5	6	7	lazy
24.	even-tempered	1	2	3	4	5	6	7	temperamental
25.	uncreative	1	2	3	4	5	6	7	creative
26.	proud	1	2	3	4	5	6	7	humble
27.	unfair	1	2	3	4	5	6	7	fair
28.	traditional	1	2	3	4	5	6	7	untraditional
29.	emotional	1	2	3	4	5	6	7	unemotional
30.	simple	1	2	3	4	5	6	7	complex
31.	quiet	1	2	3	4	5	6	7	talkative
32.	selfless	1	2	3	4	5	6	7	selfish
33.	liberal	1	2	3	4	5	6	7	conservative
34.	not envious/ not jealous	1	2	3	4	5	6	7	envious/ jealous

35.	curious	1	2	3	4	5	6	7	uncurious
36.	sociable	1	2	3	4	5	6	7	retiring
37.	stingy	1	2	3	4	5	6	7	generous
38.	practical	1	2	3	4	5	6	7	impractical
39.	subjective	1	2	3	4	5	6	7	objective
40.	analytical	1	2	3	4	5	6	7	unanalytical