

- Freud S [1916] 1961 Introductory lectures on psychoanalysis. In: J. Strachey (ed.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*. Hogarth, London, Vols. 15–16
- Hall C S, Lindzey G 1957 *Theories of Personality*. Wiley, New York
- Hogan R, Johnson J, Briggs S 1997 *Handbook of Personality Psychology*. Academic Press, San Diego, CA
- Loevinger J 1976 *Ego Development: Conceptions and Theories*. Jossey-Bass, San Francisco
- McAdams D P 2001 *The Person: An Integrated Introduction to Personality Psychology*, 3rd edn. Harcourt College Publishers, Fort Worth, TX
- McClelland D C 1961 *The Achieving Society*. D. Van Nostrand, Princeton, NJ
- McCrae R R, Costa Jr. P T 1990 *Personality and Adulthood*. Guilford Press, New York
- Mischel W 1968 *Personality and Assessment*. Wiley, New York
- Murray H A 1938 *Explorations in Personality: A Clinical and Experimental Study of Fifty Men of College Age*. Oxford University Press, New York
- Pervin L A, John O P 1999 (eds.) *Handbook of Personality Theory and Research*, 2nd edn. Guilford Press, New York
- White R W 1952 *Lives in Progress: A Study of the Natural Growth of Personality*. Dryden Press, New York
- Wiggins J S 1996 (ed.) *The Five-factor Model in Personality: Theoretical Perspectives*. Guilford, New York

D. P. McAdams

## Personality Psychology: Methods

All empirical research methods in psychology are concerned with the measurement of variation and covariation. Three methods for studying three types of variation and covariation can be identified. Experimental methods discern how behavior and experience vary across different environments. Developmental methods describe how behavior and experience vary over time. Finally, differential methods measure relatively enduring differences in characteristics of persons and the covariation among these differences. Although personality psychology occasionally employs experimental and developmental methods, its primary use of differential methods defines its distinctive position in psychology (Cronbach 1957). The specific methods used by personality psychologists can be further defined by (a) the type of individual differences under study, (b) the source of information about these differences, and (c) the purpose of assessing individuals' distinguishing attributes.

### 1. Types of Differences Measured

Gordon Allport's (1937) definitive etymological analysis of the term 'personality' finds that *persona*, the Latin root of 'personality,' originally referred to a mask worn by actors, and thus connoted a false or superficial appearance. Before long, however, *persona*

also came to designate more substantial attributes that qualify a person to perform particular roles. Today, 'external' and 'internal' aspects of personality continue to be distinguished. Ordinary-language descriptions of personality often refer to external social impressions. Trait terms representing these impressions tend to be broad and evaluative (e.g., nice, charming, nasty, irritating), and sometimes metaphorical (cold, prickly, warm, slick). Social impressions are explained by observable behavioral traits. For example, to impress others as 'warm,' a person must consistently smile, make eye-contact, talk tenderly, express affection, and show kindness and sympathy. External (social and behavioral) aspects of personality are typically measured by observer judgment (see Sect. 2.2).

Behavioral consistencies, in turn, are explained by 'inner' traits that guide and motivate behavior. Inner traits that refer to persistent thoughts, expectations, attitudes, thinking styles, and mental abilities are called 'cognitive traits.' Inner traits involving recurring feelings and emotional dispositions are called 'emotional' or 'motivational traits.' Personality psychologists often rely on the self-report method (see Sect. 2.3) to assess inner traits. Cognitive styles and competencies (e.g., creativity), are typically assessed by performance tasks (see Sect. 2.4).

### 2. Types of Individual Differences Data

Differential methods can be organized according to the types of the data one can use to study individual differences. The acronym LOST conveniently summarizes the four major types of individual differences data: life events, observer judgments, self-reports, and tests, or L-data, O-data, S-data, and T-data (see Block 1977).

#### 2.1 L-data: Life Events

'Life events' refer to relatively objective facts about people that are often a matter of public record. Some examples are birth date, number of years of education and degrees obtained, marital status and number of children, occupational record, church attendance, hospitalizations, membership in societies, criminal convictions, leadership positions, and property ownership. Although life event information is often actually obtained by self-report methods (Sect. 2.3) its factuality can be independently verified. Research indicates a very high correlation between self-reported and independently verified life events, even under highly evaluative conditions such as applying for employment.

In both applied (see Sects. 3.1 and 3.2) and basic (see Sect. 3.3) personality research, life events are sometimes used as criteria to be predicted by other types of personality data and sometimes as predictors themselves. When gathered by self-report, L-data is interpreted according to self-report methods (Sect. 2.3).

## 2.2 O-data: Observer Judgments

Observer judgments are a fundamental source of information about external traits for two reasons (Block 1978, Hofstee 1994). First, an observer's judgment of external traits is direct and noninferential. Second, idiosyncratic biases and errors of judgment tend to cancel out when judgments are averaged across multiple observers. This means that even inner traits may be more validly assessed by observer judgments than self-reports, despite an individual's privileged access to his or her inner thoughts and feelings (Hofstee 1994).

*2.2.1 Retrospective, integrative judgments versus one-time, direct observation.* Observer judgments are often made by acquaintances who use their past knowledge to summarize their perceptions of someone's personality. Concern about potential distortions and memory limitations involved in making retrospective judgments lead some psychologists to favor direct behavioral observation. Despite potential distortions, retrospective judgments are more likely than single encounters to produce the representative sampling necessary for accurately assessing relatively enduring characteristics of individuals (Kenrick and Funder 1988).

*2.2.2 Normative and ipsative frames of reference for observer judgments.* For observer judgments to indicate what is distinctive about someone's personality, the judgments must be compared to some reference point. Take, for example, a study in which observers record from behind one-way glass instances of different types of behavior in a nursery school classroom over a period of one year. At the end of the year, a particular behavior count by itself—say 42 acts of aggression—is insufficient for describing a child as relatively aggressive or nonaggressive. A normative frame of reference would compare the number of aggressive acts for that child to the average number of aggressive acts across all children. An ipsative frame of reference compares the number of aggressive acts for the child to all other types of acts recorded for that child.

In retrospective observer judgments, judges are often instructed to compare a person to people in general with a rating scale. The middle point (e.g., '3' on a 1–5 rating scale) represents a theoretical normative reference point (i.e., people in general), while greater numbers represent higher levels of the trait and lesser numbers lower levels. Unipolar rating scales are anchored by a single trait word at the high end, while bipolar scales are also anchored with the psychological opposite at the low end (e.g., 'thoughtful' vs. 'inconsiderate'). Defining the anchors with concrete, descriptive phrases instead of abstract adjectives improves measurement validity. Often scores on several

related rating scales will be averaged to produce an overall score on a trait. For example, broad extraversion versus introversion might be assessed by the average of ratings on more specific scales such as talkative vs. silent, outgoing vs. reserved, and so forth.

When interpreting rating scores, psychologists need to consider whether to take the numerical score at face value. For example, a rating of 4 on a 1–5 scale of thoughtfulness is above the theoretical norm of 3, but if the computed average rating for a large group of people is 4.5, a rating of 4 could be interpreted as a relatively low level of thoughtfulness. This problem is further complicated when some judges restrict their ratings to one portion of the scale. A 5 might actually indicate a much higher value from a judge who assigns mostly 3s than from a judge who assigns mostly 4s and 5s.

Psychologists unwilling to accept scores at face value will recalibrate all of a judge's ratings with respect to the mean of all ratings made by that judge. This process is called 'ipsatizing' scores. Some rating procedures expressly call for ipsative assessment in the act of judgment itself. For example, judges using the California Q-set (Block 1978) are required to sort 100 personality descriptions into nine categories following a specified distribution (five descriptions in each of the extremely characteristic and uncharacteristic categories, eight descriptions in each of the quite characteristic and uncharacteristic categories, and so forth, to 18 characteristics in the relative neutral or unimportant category).

Formats other than rating scales for observer judgment of personality include questionnaires and adjective checklists. Items in some personality questionnaires are phrased in ways that allow either self-report (Sect. 2.3) or observer judgment. Adjective check lists contain a set of adjectives that judges check if they believe them to apply to the person being judged. Scores on adjective check lists are computed by counting the number of judges who check a particular adjective and/or by summing the checks for a group of adjectives considered to measure the same trait.

Factors that potentially limit the validity of observer judgments include misleading communications, stereotypes, and gossip about those being judged; unfairly positive or negative attitudes from judges who like or dislike the targets; and insufficient knowledge due to judges knowing the target only in certain roles and settings. These limitations and methods for overcoming them are discussed by Block (1978), Hofstee (1994) and Kenrick and Funder (1988).

## 2.3 S-data: Self-Reports

The two basic types of self-report instruments are projective tests and objective questionnaires. In projective testing, respondents are asked to finish incomplete phrases or construct stories about intentionally ambi-

guous images. Following guidelines developed by the test author and community of test users, psychologists score the respondent's protocol for psychological themes. Proponents of projective tests claim that these instruments are able to tap deep, unconscious needs and motives; critics insist that scoring projective protocols is too subjective and unreliable. Research indicates that carefully designed projective tests can be as reliable and valid as objective measures.

Personality questionnaires (Angleitner and Wiggins 1986) rarely consist of questions anymore. Instead, questionnaire items are statements about one's self and other people. Respondents express how much they agree with each statement or the degree to which they think the statement applies to them. The most comprehensive personality questionnaires contain several hundred items. Any subset of items within a personality questionnaire that is scored for a particular trait is called a 'personality scale.' The major personality questionnaires contain as many as several dozen different scales. Items are collected into scales and responses to items are scored according to one of four strategies outlined below: empirical, rational-intuitive, theoretical, or factor-analytic.

*2.3.1 Empirical scales.* Paul Meehl (1945) argued that psychologists would be naively optimistic to take questionnaire item responses at face value or to attempt to judge their hidden psychological meanings expertly. More prudent, he suggested, would be to treat each item response as a bit of behavior whose meaning must be determined by its empirical correlates. Empirical scales are constructed by locating all items on a questionnaire that tends to be answered differently by two groups known by other methods to differ on a particular personality trait. A person's score on an empirical scale is defined by the number of responses that match the responses given by one of the groups used in original scale construction. That is, if a person answers many items the same way as a group of people known to be aggressive, that person is considered likely to be aggressive also.

*2.3.2 Rational-intuitive scales.* The rational-intuitive approach to personality scales suggests personality traits can be assessed straightforwardly by items whose content, according to common sense, seems relevant to the trait (Wolfe 1993). Thus, a person who endorses items such as 'I am an aggressive person' and disagrees with items such as 'I never get in fights' would receive points on a rational-intuitive personality scale of aggressiveness. The obviousness of rational-intuitive scales perennially raises concerns about self-enhancement (exaggerating socially desirable traits and denying undesirable traits), but research indicates that respondents self-enhance on personality questionnaires no more than they do in everyday life. Furthermore, research indicates that

rational-intuitive scales validly predict relevant criteria as well as scales constructed by any other method.

*2.3.3 Theoretical scales.* Like rational-intuitive scales, theoretical scales are comprised of items whose content is judged to be relevant to the personality characteristic being assessed. The difference is that the relevance is not apparent to common sense and can only be seen by professionals versed in a particular theory. For example, the theoretical items 'I am fascinated by fire,' 'I would do anything on a dare,' 'My father and I always fought,' 'I own a gun,' and 'Women find me charming' seem unrelated to common sense, but, to a Freudian, these are items likely to be endorsed by a man with a phallic personality character resulting from an unresolved Oedipal conflict.

*2.3.4 Factor-analytic scales.* Factor-analysis is a statistical method for identifying clusters of items that tend to be answered the same way. This method, like the empirical method, begins with a large set of items that are administered to a group of respondents. If respondents who agree with item 'A' also tend to agree with items 'B,' 'C,' 'D,' and so forth, these items are deemed to measure the same psychological trait (Briggs and Cheek 1986). The nature of the trait is normally determined by rational-intuitive inspection of the content of the items. Factor analysis can be applied to scales as well as items, and factor analytic research has repeatedly indicated that much of the content of personality falls into five broad factor domains: extraversion, agreeableness, conscientiousness, emotional stability, and a fifth factor variously called intellect, imagination, or openness to experience. Many psychologists regard the 'Big Five' or 'Five-factor model' (Wiggins 1996) as a major integrating focus for future research.

#### 2.4 T-data: Laboratory Tests

Assessing personality by testing respondents in laboratories or other controlled conditions is motivated by two interests. The first is establishing objective, replicable procedures that cannot be affected by biases and errors potentially found in observer judgments or self-reports. In particular, the measurement of involuntary reactions such as changes in electrodermal conductivity or heart rate to assess anxiety, or pupil dilation to assess interest, is seen as a way of circumventing dissembling that can occur with self-reports. Unfortunately, as Block (1977, 1978) points out, T-data are always indirect measures of personality, and laboratory tests that might seem to be reasonable measures of personality have a record of failing unpredictably.

The second motivation for using laboratory tests is the particular suitability of such procedures for measuring certain traits. Laboratory tests are particularly apt for assessing cognitive performance variables. For example, a personality theory built around the mental accessibility of different concepts would naturally be tested by measuring reaction time to words representing different concepts. Cognitive styles are almost invariably measured by performance tasks.

Personality theories that attempt to explain mental or behavioral differences in terms of underlying biological differences also require laboratory facilities to assess those differences. In addition to traditional psychophysiological recordings, new laboratory tests for measuring the physical basis of personality include biochemical assaying, positron emission tomography and functional magnetic resonance imaging. Details of these methods can be found in sources such as Davidson (1999), and Pickering and Gray (1999).

### *3. Three Purposes of Measurement*

The three major purposes for measuring individual differences are: (a) making decisions about people, (b) helping people make decisions, and (c) conducting basic research. All three cases involve prediction (for further information see Wiggins 1973). In decision-making, personality assessments are used to predict how an individual person will think, feel, behave, or be perceived by others if various courses of action are pursued. In research, predictions are made about how different forms of individual differences are related to one another.

#### *3.1 Making Decisions About People*

In clinical and counseling psychology, personality scores are used to predict what course of therapy will best help persons with psychological problems. In personnel psychology, personality scores are used to predict which individuals will perform best if hired and placed in particular jobs. Some applied psychologists endorse what they call an 'assessment' approach to these decisions, in which the decision maker intuitively weighs all the information gathered about the person. In contrast, a 'statistical' approach inserts personality scores into a mathematical equation developed from past empirical research. Studies indicate that statistical predictions are almost invariably more accurate than assessment predictions.

#### *3.2 Helping People Make Decisions*

Methods for helping people make decisions with personality measures differ from making decisions

about people only in terms of who is making the decisions. Individuals seeking greater satisfaction in personal relationships, for example, may complete a personality questionnaire to increase self-insight, much as they might read a self-help psychology book. Likewise, individuals uncertain about career choice can complete questionnaires that predict which careers would be most satisfying. Traditionally, psychologists interpret and discuss personality scores with clients, but some self-help personality measures are completely self-administered.

#### *3.3 Basic Research: Uncovering Covariation Among Variables*

The usefulness of all applied personality methods depends upon the ability of researchers to construct valid personality measures and to ascertain reliable covariation among these measures. This process, called 'construct validation,' is identical with any other type of scientific hypothesis testing (Hogan and Nicholson 1988). In the typical case, construct validation takes the following form. A researcher's theory predicts that individual differences in a particular personality trait (say, conscientiousness) will covary with differences in some L-data (say, job performance). A method is devised to measure the personality trait with O-, S-, or T-data and the life event with L-data. Successful prediction supports both the validity of the measures and the theory that led to the prediction. Predictive failure means either the hypothesis was incorrect, a procedural error occurred (inappropriate research sampling, administration, or scoring), or one or both of the measures lack validity. Progress in personality research occurs when many successful predictions leads to the acceptance of a measure as 'well-validated.' Careful research with well-validated measures always advances knowledge because even predictive failures indicate the need to revise hypotheses.

Because valid measurement is crucial to the entire personality enterprise, a significant amount of personality research is directed at improving measurement methods. Some of this research aims to clarify the dynamics of the measurement process, that is, the psychological processes that occur during observer judgments and self-reports. A second line of research employs computers to administer, score, and interpret personality tests. When computer programs for analyzing personality data are combined with artificial intelligence, 'observers' with artificially constructed personalities will some day make 'observer judgments' of personality worldwide over the Internet.

*See also:* Personality Assessment; Personality Psychology; Personality Structure; Personality Theories; Projective Methods in Psychology

### Bibliography

- Allport G W 1937 *Personality: A Psychological Interpretation*. H Holt, New York
- Angleitner A, Wiggins J S 1986 *Personality Assessment via Questionnaires*. Springer-Verlag, Berlin
- Block J 1977 Advancing the psychology of personality: Paradigmatic shift or improving the quality of research? In: Magnusson D, Endler N S (eds.) *Personality at the Crossroads: Current Issues in Interactional Psychology*. Lawrence Erlbaum Associates, Hillsdale, NJ, pp. 37–64
- Block J 1978 *The Q-Sort Method in Personality Assessment and Psychiatric Research*. Consulting Psychologists Press, Palo Alto, CA
- Briggs S R, Cheek J M 1986 The role of factor analysis in the development and evaluation of personality scales. *Journal of Personality* **54**: 106–48
- Cronbach L J 1957 The two disciplines of scientific psychology. *American Psychologist* **12**: 671–84
- Davidson R J 1999 Biological bases of personality. In: Derlega V J, Winstead B A, Jones W H (eds.) *Personality: Contemporary Theory and Research*, 2nd edn. Nelson Hall, Chicago, pp. 101–25
- Hofstee W K B 1994 Who should own the definition of personality? *European Journal of Personality* **8**: 149–62
- Hogan R, Nicholson R A 1988 The meaning of personality test scores. *American Psychologist* **43**: 621–6
- Kenrick D T, Funder D C 1988 Profiting from controversy: Lessons from the person-situation debate. *American Psychologist* **43**: 23–34
- Meehl P E 1945 The ‘dynamics’ of structured personality tests. *Journal of Clinical Psychology* **1**: 296–303
- Pickering A D, Gray A D 1999 The neuroscience of personality. In: Pervin L A, John O P (eds.) *Handbook of Personality: Theory and Research*, 2nd edn. Guilford Press, New York, pp. 277–99
- Wiggins J S 1973 *Personality and Prediction: Principles of Personality Assessment*. Addison-Wesley Pub Co, Reading, MA
- Wiggins J S (ed.) 1996 *The Five-Factor Model of Personality*. Guilford Press, New York
- Wolfe R N 1993 A commonsense approach to personality measurement. In: Craik K H, Hogan R, Wolfe R N (eds.) *Fifty Years of Personality Psychology*. Plenum Press, New York, pp. 269–90

J. A. Johnson

Copyright © 2001 Elsevier Science Ltd.  
All rights reserved.

### Personality Structure

Personality structure is the branch of personality psychology that is concerned with quantitative representations of the ways in which persons consistently differ from one another in their characteristic adaptations to social environments (see *Personality and Adaptive Behaviors*). The basic unit of study is the ‘personality trait’ which is considered to be a moderately heritable disposition to behave consistently across different situations. Such traits are thought to

be manifest in the individual differences in thinking, feeling, and behaving reported in responses to personality questionnaires or to lists of trait-descriptive adjectives. Current research in the field of personality structure emphasizes two quite different structural representations of personality traits: the Five Factor Model (FFM) of personality traits and the Interpersonal Circumplex (IPC) model of interpersonal dispositions. These two models originated in different historical contexts, are based on different assumptions concerning the nature of traits, and employ different statistical models for representing dimensions of personality. Nevertheless, the FFM and the IPC models of personality structure may be viewed as complementary rather than competitive, and as being applicable to different but equally important realms of knowledge.

#### 1. Origins of the Five Factor Model of Personality Structure

The idea that variations in human characteristics have an evolutionary origin was suggested at the end of the eighteenth century by Erasmus Darwin (1731–1801) and it later bore fruit in the empirical researches of two of his grandsons, Charles Darwin (1809–82) and Francis Galton (1822–1911). That Galton was uniquely qualified to conduct research on ‘hereditary genius’ is suggested by the accomplishments and eminence of the Darwin family and by the fact that Galton was, by any criteria, a genius himself. Galton coined the term ‘eugenics’ (Greek *eugenes* ‘well-born’) for the field which studied the inheritance of desirable characteristics, especially intelligence. In 1904, Galton endowed a permanent chair in eugenics at University College, London and appointed the mathematician, Karl Pearson (1857–1936) as its first professor. Pearson developed the correlation coefficient to express the relation between the intelligence of parents and the intelligence of their offspring. His colleague, Charles Spearman (1863–1945) developed the method of factor analysis (see *Personality Psychology: Methods*) which, when applied to mental ability data, suggested a general factor of intelligence (‘g’) in addition to more specific factors. Raymond B. Cattell (1905–98), the founder of modern personality structure research, received his doctorate under Spearman and was the first to suggest extending the applications of factor analysis to the study of individual differences in personality characteristics.

Cattell (1967) had a master plan for the study of personality structure, which he and his associates implemented over a period of 65 years: (a) determine the totality of trait-descriptive terms in a given language that denote individual differences in thinking, feeling, and behaving (‘surface traits’); (b) reduce correlated clusters of these surface trait terms by