Personality Traits

SECOND EDITION

GERALD MATTHEWS
University of Cincinnati

IAN J. DEARY
University of Edinburgh

MARTHA C. WHITEMAN
University of Edinburgh

CAMBRIDGE
UNIVERSITY PRESS

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge, United Kingdom

CAMBRIDGE UNIVERSITY PRESS
The Edinburgh Building, Cambridge, CB2 2RU, UK
40 West 20th Street, New York, NY 10011–4211, USA
477 Williamstown Road, Port Melbourne, VIC 3207, Australia
Ruiz de Alarcón 13, 28014 Madrid, Spain
Dock House, The Waterfront, Cape Town 8001, South Africa

http://www.cambridge.org

First published 1998. Reprinted 1999, 2000, 2002
Second edition 2003

Printed in the United Kingdom at the University Press, Cambridge

Typefaces Times 10/13 pt. Formata *System* LaTeX 2$_\varepsilon$ [TB]

A catalogue record for this book is available from the British Library

Library of Congress Cataloguing in Publication data
Matthews, Gerald.
Personality traits / Gerald Matthews, Ian J. Deary, Martha C. Whiteman. – 2nd edn.
p. cm.
Includes bibliographical references and index.
ISBN 0 521 83107 5 – ISBN 0 521 53824 6 (pb)
1. Personality. I. Deary, Ian J. II. Whiteman, Martha C. III. Title.
BF698.M3434 2003
155.2′3 – dc21 2003046259

ISBN 0 521 83107 5 hardback
ISBN 0 521 53824 6 paperback

Contents

Figures

Tables

Boxes

1 The trait concept and personality theory

Introduction: conceptions of traits

Everyday conceptions of traits

The idea of personality traits may be as old as human language itself. Aristotle (384–322 BC), writing the Ethics in the fourth century BC, saw dispositions such as vanity, modesty and cowardice as key determinants of moral and immoral behaviour. He also described individual differences in these dispositions, often referring to excess, defect and intermediate levels of each. His student Theophrastus (371–287 BC) wrote a book describing thirty 'characters' or personality types, of which a translator remarked that Theophrastus's title might better be rendered 'traits' (Rusten, 1993). Basic to his whole enterprise was the notion that individual good or bad traits of character may be isolated and studied separately.

Contemporary English is replete with terms used to describe personal qualities. Table 1.1 shows some examples: the five words rated by American college students as the most and least favourable words in Anderson's (1968) survey of 555 personality terms, together with five words given a neutral rating. Allport and Odbert (1936) identified almost 18,000 English personality-relevant terms; more words than Shakespeare used! Nouns, sentences and even actions may also have personality connotations (Hofstee, 1990). The language of personality description permeates our everyday conversation and discourse.

Everyday conceptions of personality traits make two key assumptions. First, traits are stable over time. Most people would accept that an individual's behaviour naturally varies somewhat from occasion to occasion, but would maintain also that there is a core of consistency which defines the individual's 'true nature': the unchangeable spots of the leopard. In other words, there are differences between individuals that are apparent across a variety of situations. We might expect a student we have noted as a 'worrier' to be unusually disturbed and worried in several different contexts such as examinations, social occasions and group discussions. Stability distinguishes traits from more transient properties of the person, such as temporary mood states. Second, it is generally believed that traits directly influence behaviour. If a person spontaneously breaks into cheerful song, we might 'explain' the behaviour by saying that he or she has a happy disposition. Such lay explanations are, of course, on shaky ground because of their circularity. Aristotle

Table 1.1 *Ratings of likeableness of some favourable, neutral and unfavourable traits*

Favourable traits		Neutral traits		Unfavourable traits	
Trait	Rating	Trait	Rating	Trait	Rating
Sincere	5.73	Quiet	3.11	Dishonest	0.41
Honest	5.55	Impulsive	3.07	Cruel	0.40
Understanding	5.49	Changeable	2.97	Mean	0.37
Loyal	5.47	Conservative	2.95	Phony	0.27
Truthful	5.45	Hesitant	2.90	Liar	0.26

Note Each word was rated on a 0–6 scale by 100 US college students
Source Anderson, 1968

suggested a more subtle, reciprocal causal hypothesis: that it is through actions that dispositions develop, which in turn influence actions.

> It is by refraining from pleasures that we become temperate, and it is when we have become temperate that we are most able to abstain from pleasures. (Thomson's, 1976, translation of the *Ethics*, 1104a: 33–35)

One of the major tasks for a scientific psychology of traits is to distinguish internal properties of the person from overt behaviours, and to investigate the causal relationships between them. To avoid circularity, it is essential to seek to identify the underlying physiological, psychological and social bases of traits, which are the true causal influences on behaviour.

Scientific conceptions of traits

This book places the concept of the trait at centre stage in the scientific study of human personality because, 'if there is to be a speciality called personality, its unique and therefore defining characteristic is traits' (Buss, 1989). There is a large gap between the everyday concept of a trait, and a concept that is scientifically useful. Several distinct steps are necessary for developing a science of traits. The first step is the measurement and classification of traits. The simplest technique for personality measurement is just to ask the person to rate how well trait adjectives such as those shown in Table 1.1 apply to himself or herself. We can also ask questions about behaviours that are thought to relate to personality. Measures of the extraversion–introversion trait typically ask whether the person enjoys parties, meeting people and other social activities, for example. We can also have a person who knows the respondent well, such as a spouse or close friend, provide ratings of his or her personality. Traits need not be measured solely by verbal report: real-world actions and behaviour in the laboratory may be assessed too (Cattell, 1973). We would expect an extraverted person to belong to many clubs and societies, for example. Experimental tests of typically extraverted behaviours may also be devised, such as amount of laughter at jokes and willingness to respond rapidly but inaccurately. In practice, however, personality measures based on objective

Table 1.2 *Examples of experimental studies showing correspondences between traits and objective behavioural measures*

Study	Trait	Behavioural measure
Carment, Miles and Cervin (1965)	Extraversion	More time spent talking
Edman, Levander and Schalling (1983)	Impulsivity	Faster reaction time
De Julio and Duffy (1977)	Neuroticism	Greater distance from experimenter chosen
Ganzer (1968)	Test anxiety	More time spent looking away from the task during testing
Newman, Patterson and Kosson (1987)	Psychopathy	More persistence in gambling when consistently losing

behavioural tests have had only limited success, and few have been validated (see Kline, 1993). Verbal report has been the preferred method of trait assessment used by personality researchers.

As we have seen already there is a huge number of words which may be used to describe personality. Many of these words have rather similar meanings: precise, careful, meticulous and painstaking would all seem to relate to some common quality of conscientiousness. Such overlapping traits can be grouped together as a broad aspect or dimension of personality. The question then becomes: what is the number of broad dimensions needed to describe the main elements of any individual personality? Much research effort has been devoted to drawing up classificatory schemes of fundamental personality dimensions: estimates of the number required range from three to thirty or so.

There is no guarantee that people's self-descriptions are accurate. The second step in personality research is to test whether and how traits relate to behaviours. Table 1.2 gives some examples of correlations obtained empirically between personality traits and objectively assessed behavioural measures. In each case, the data imply that the person's self-ratings or questionnaire responses are at least partially accurate. Traits may also be useful in applied settings, in predicting a person's job performance, or the response of a patient to therapy, for example. A related research question is the consistency of behaviour in various situations. The implicit assumption of the trait approach is that people do in fact tend to behave consistently in different settings, an assumption which has been vigorously challenged, as we shall see in chapter 2.

A science of personality traits requires a final, but difficult step: development of a satisfactory theory of personality traits. We may be able to assess people's levels of extraversion and other traits, and show that our assessment predicts some aspects of their behaviour, but in themselves these observations tell us nothing about why the personality dimension predicts behaviour. One difficulty is that personality may be represented at a variety of levels of psychological description. For example, extraversion might be associated with simple properties of the central nervous system, such as the excitability of individual neurones, or with style of information processing, or with acquired social knowledge and beliefs. We can only

distinguish these broad possibilities by the normal, somewhat laborious scientific methods of formulating specific hypotheses and testing them rigorously against experimental and observational evidence.

There are also some more subtle conceptual problems to be overcome. There is some question over whether we can ever develop a general scientific theory of traits at all. The idiographic approach to personality (e.g., Lamiell, 1981) considers that all aspects of personality are fundamentally unique and idiosyncratic to each individual, so that no generalised theoretical statements are possible. In this book, we adopt the alternative nomothetic approach, which assumes that we can arrive at general hypotheses concerning stable individual differences through the normal scientific method. We cannot, of course, expect such hypotheses to predict all or even most of the person's behaviour; the uniqueness of individuals seems secure.

Causal primacy. There is uncertainty too over the causal status of traits. Suppose we have a person who obtains a high score on a measure of neuroticism, and also shows clinical symptoms of mild depression. Did neuroticism cause depression, did depression cause neuroticism, or are both qualities independently influenced by some additional causal factor such as a stressful life event? A traditional assumption of trait theorists has been the *causal primacy* of traits. Although, as suggested by Aristotle, there is probably some reciprocity of causal influence between traits and behaviours, it has often been supposed that the dominant direction of causality is from trait to behaviour. For example, Brody (1994) stated that 'I assume that personality traits are causal. They are genotypically influenced latent characteristics of persons that determine the way in which individuals respond to the social world they encounter.' That is, although measures of traits such as questionnaire scores are not causal agents themselves, they validly index underlying physiological or psychological structures which directly influence behaviour. One of the pioneering trait psychologists, Gordon Allport (1937), saw traits as organised mental structures, varying from person to person, which initiate and guide behaviour.

There are two important qualifications to this general principle. First, as Hettema and Deary (1993) pointed out, the explaining of behaviour requires different levels of analysis, including genetics, physiology, learning and social factors. Allport's notion that all the various manifestations of traits can be explained at a single level of 'mental structure' is simplistic. Hence, causal models of trait action will vary depending on the level investigated, although the ultimate research aim is to develop a trait theory that will interrelate the various levels. Second, the causal effects of traits on behaviour may be indirect. As discussed in chapter 2, traits interact with situational factors to produce transient internal conditions or states, which may sometimes be a more direct influence on behaviour than the trait. For example, trait anxiety may interact with an immediate situational threat to generate transient state anxiety, which in turn disrupts ongoing information processing and impairs performance (Spielberger, 1966).

Inner locus. A second traditional assumption is that of the *inner locus* of traits. The most important traits, such as extraversion and neuroticism (a broad

tendency to experience negative emotions), are assumed by some to relate to some fundamental, core quality of the person, which might be influenced substantially by genetic factors (Eysenck, 1967; McCrae et al., 2000). Again, even within theories that are sympathetic to the traditional view of traits, there has been some modification of the basic view. For example, Cattell and Kline (1977) distinguished 'surface' traits, which are simply clusters of overt responses which tend to be associated, from 'source' traits, which are deeper properties of the person with causal effects on behaviour. Modern developments of traditional theory seek to identify and explain underlying sources of consistency in behaviour, whether these are conceived of as genetic, physiological or cognitive in nature. The process of relating operationally defined measures such as questionnaire scores to theory is often referred to as construct validation, and is discussed further below.

Both assumptions of traditional trait theory – their causal primacy and inner locus – have been challenged more radically. The alternative to causal primacy is the view that traits are a construction with no independent causal status. For example, Buss and Craik (1983) argued that traits are simply descriptions of natural categories of acts. Wright and Mischel (1987) characterised traits as conditional statements of situation–behaviour contingencies. Furthermore, traits may be jointly constructed by two or more people in social interaction, according to the social dynamics of the situation (Hampson, 1988). Social psychological approaches to traits tend also to abandon the inner locus assumption. Even if traits represent genuine psychological structures, these structures may be no more than the superficial mask the person presents to the outside world, in order to present a socially acceptable self-image to other people. Such challenges to traditional views of traits are explored in more detail in chapters 5 and 8.

The upshot of these considerations is that there is no generally accepted scientific theory of traits. Some trait theorists have tended to take the relatively easy option of focusing on the dimensional structure and measurement of traits rather than investigating their underlying nature (Goldberg, 1993). However, it should be clear from the preceding discussion that we cannot accept trait descriptions at face value, and that there may be various qualitatively different types of explanation for consistencies in self-reports and behaviours. In recent years progress has been made in developing psychobiological information processing, and social psychological trait theories which are partly complementary and partly competing accounts. One of the major aims of this book is to show that trait psychology requires these theoretical endeavours as well as its traditional concern with psychometrics. Development of successful theories is necessary for the study of traits to take its rightful place as a fundamental area of psychological science.

A brief history of traits

The scientific study of traits develops two aspects of common-sense discourse on personality. First, it formalises the tendency in natural language to use

trait descriptors of individuals. Second, it formalises the popular awareness that there are generalities of personality, such that individuals of a similar disposition may be grouped together. This tendency is seen in folk psychology: astrology has twelve personality-based sun signs, and there is a Chinese custom of ascribing certain aspects of personality to the year in which a person was born; for instance, those born in the years of the cow are said to be conscientious and hardworking. Traits emerged from folk psychology and medicine, and from natural language. The history of traits is a story which may be told in various ways: through tracing the counterparts to extraversion and neuroticism identified in different epochs (Eysenck and Eysenck, 1969; Eysenck, 1981), or through emphasising the evolution of the currently dominant five factor model of personality (Goldberg, 1993). We confine ourselves to highlighting three aspects of the history of traits: the influence of classical thinking, the earliest scientific work on traits, and the emergence of current models of personality.

The four humours

Amongst the earliest progenitors of present-day trait theories, apart from Aristotle and Theophrastus, were Hippocrates (ca. 460–377 BC) and Galen of Pergamum (AD 130–200) (Stelmack and Stalikas, 1991). The Hippocratic conception of the aetiology of physical illnesses was based upon the theory of humours, or bodily fluids, notably blood, phlegm, black bile and yellow bile. It was in the writings of Galen, a Greek physician, that the humours became the bases of temperaments. Galen's temperamental terms, melancholic (tending towards low mood), choleric (tending toward anger), phlegmatic (tending towards stolid calmness) and sanguine (tending towards optimism and confidence), survive in today's English. When the humours were blended in a balanced fashion, an optimal temperament resulted:

> in his soul he is in the middle of boldness and timidity, of negligence and impertinence, of compassion and envy. He is cheerful, affectionate, charitable and prudent. (Stelmack and Stalikas, 1991, p. 259)

Imbalance led to physical illness, but also to mental disturbance. For example, the melancholic temperament, associated with feelings of depression and anxiety, resulted from an excess of black bile. In the seventeenth century, Burton's (1837; originally published 1621) description of the melancholic character has some resemblance to the high neuroticism scorer on a present-day personality questionnaire,

> that which is a flea-biting to one causeth unsufferable torment to another; and which one by his singular moderation and well-composed carriage can happily overcome, a second is no whit able to sustain; but, upon every small occasion of misconceived abuse, injury, grief, disgrace, loss, cross, rumour etc. (if solitary,

> or idle) yields so far to passion, that his complexion is altered, his digestion hindred, his sleep gone, his spirits obscured, and his heart heavy, his hypocondries misaffected; wind, crudity, on a sudden overtake him, and he himself overcome with melancholy. (vol. 1, p. 140)

The humoral terms exist today merely as descriptive metaphors. Their aetiological significance did not long outlast the Middle Ages. Immanuel Kant recast the four humoral temperaments along the dimensions of 'feeling' and 'activity' to yield a typology of four simple temperaments that emphasised their psychological nature. The humoral terms also appear in the writings of the father of modern psychology, Wilhelm Wundt. Wundt described the four temperamental types in terms of two dimensions: strong–weak emotions versus changeable–unchangeable activity. The relationships between the humoral terms and the schemes of temperament classification devised by Kant and Wundt are shown in figure 1.1. Stelmack and Stalikas (1991) described the relationship between these schemes and the present-day dimensions of neuroticism and extraversion as 'uncanny'. However,

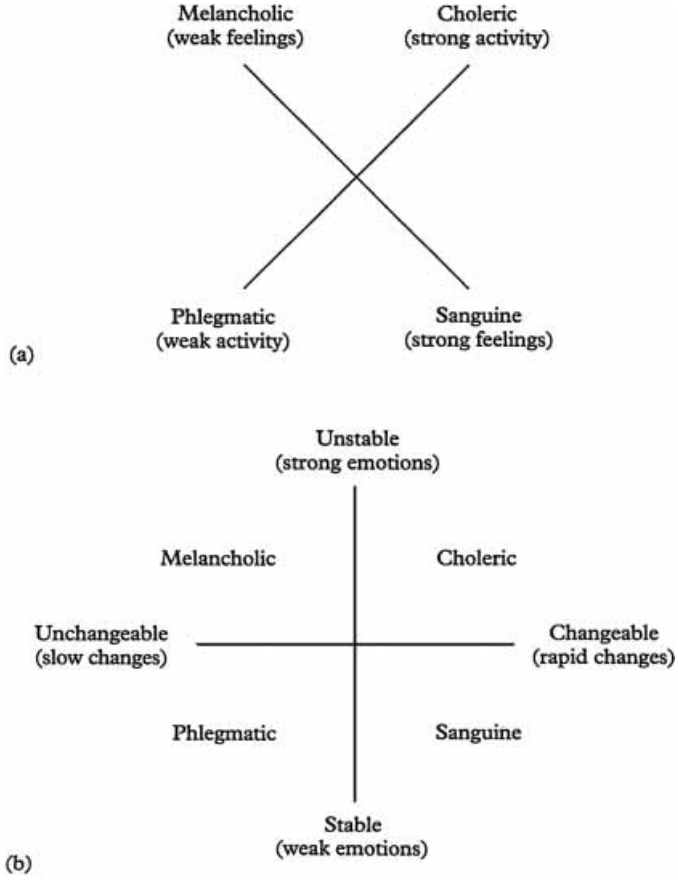

Figure 1.1 Humoral schemes of temperament proposed by (a) Kant and (b) Wundt

any veracity they have is owed to shrewd psychological observation and not the classical theory of the humours.

Beginnings of the science of traits

Three ingredients were required for the initiation of scientific research on traits: systematic data collection, statistical techniques for data analysis, and development of testable theories. These prerequisites became available around the beginning of the twentieth century. Of key importance were the new techniques of correlation and, somewhat later, factor analysis (Gorsuch, 1983). Before the introduction of factor analysis there was no objective method for reducing the huge numbers of trait terms to a manageable number of broad dimensions. Thurstone's (1947) introduction of multiple factor analysis proved particularly influential, and the systematic use of factor analysis began the modern research era in personality.

The first empirical studies

The raw materials, or stimuli, for some early researchers were gathered from the dictionary. Sir Francis Galton (1884) was prescient in hypothesising that individual differences in personality might be represented in natural language terms, and trawling Roget's Thesaurus for character-descriptive terms. This was later dubbed the 'lexical hypothesis', for which De Raad (2000) provides a history. Pioneers of empirical personality research included the Dutch psychologists Heymans and Wiersma who, in a series of papers between 1906 and 1909, obtained ratings of character for large numbers of subjects and attempted to reduce these to smaller numbers of factors or dimensions. They employed a statistical method that was conceptually related to factor analysis, though much more crude, and obtained three factors. Eysenck (1970) identified the first dimension with emotionality, and the other two with introversion–extraversion.

After Spearman's (1904) epoch-making study of mental ability, in which he discovered general intelligence and introduced an embryonic form of factor analysis, similar techniques were used under his supervision to analyse character. Webb (1915) collected detailed ratings of mental qualities on 194 students at a teacher training college and 140 younger schoolboys. The individual rating items were collected under the headings of intellect, emotions, sociality, activity, and self qualities. Webb used such statistical techniques as were available to deduce that, after general intelligence had been extracted, a second general factor of character could be identified. This second factor was called 'persistence of motives' or 'will'. There are many aspects of Webb's study which make it a good source of data: the subject sample was large, the ratings were performed consistently, by more than one rater, for each subject over an extended period of time, and the range of personality qualities assessed was broad. As a result, it has been re-analysed at intervals since its publication: these re-analyses are documented by Eysenck (1970). A comprehensive re-analysis showed that five or six factors existed in Webb's data,

and trait researchers consider them to be very similar to modern dimensions of personality (Deary, 1996).

The beginnings of trait theory

In addition to minimally adequate statistical procedures for dealing with traits, and some conception of where to begin to search for trait stimuli, there was a contemporaneous theoretical development of trait psychology. In part, this theoretical development was driven by an awareness of the fact that trait psychology was perforce beginning with commonsense terms in everyday use. Allport (1937) commented that:

> To use trait terms, but to use them cautiously, is, then, our lot. Nor need we fear them simply because they bear the age-long sanction of common sense.

Carr and Kingsbury's article from 1938 addressed many core issues of trait psychology at a conceptual level. They emphasised the predictive nature of traits, i.e., knowing the traits of an individual was predictive of that person's likely future behaviour. Moreover, they articulated the notion that traits were not directly observable – traits may only be inferred from behaviour. This continues to be the view of prominent trait theorists. For example, McCrae et al. (2000, p. 175) stated,

> Traits cannot be directly observed, but rather must be inferred from patterns of behaviour and experience that are known to be valid trait indicators.

Carr and Kingsbury emphasised the need for trait scales in order to compare individuals on a given characteristic. They lamented the blind progress of trait psychology and its lack of 'principles of orientation in reference to the concept'. This last continued to be one of the most contentious issues in the theory of traits (Pervin, 1994). One of their closing comments is ironic when one reflects on the pre-eminence of the dimensions of neuroticism (emotional stability) and introversion–extraversion today,

> We may note that abnormal and clinical psychology have evinced no interest in the popular traits, but have developed a new set of traits that are supposed to possess a distinctive value for their purposes. We refer to such traits as introversion and extraversion, submission and ascendancy, emotional stability, mal-adjustment, and integration. Perhaps a systematic psychology should likewise be concerned with the development and study of a set of new traits that are relevant to its purposes.

Perhaps the most comprehensive contribution to the conceptual development of trait psychology, and of personality psychology more generally, is Allport's (1937) book, *Personality: a Psychological Interpretation*. Much of present-day trait psychology may be considered as empirical footnotes to Allport's chapters 9–12, where he laid out the tasks for, and difficulties facing, the personality psychologist.

Allport's resounding 'Resume of the Doctrine of Traits' began with the famous sentence,

> In everyday life, no one, not even a psychologist, doubts that underlying the conduct of a mature person there are characteristic dispositions or traits.

In addition to the common traits that are emphasised in the present book (indicative of the nomothetic approach), Allport also emphasised those traits which are more specific to individuals and that are not prone to distribute normally in the population (indicating that an idiographic approach is necessary also). Allport's account of traits was able to embrace many disparate approaches. Thus, in addition to accommodating differential psychologists, his overall definition of traits moved Murray (1938) to indicate that his 'needs' – identified by a depth psychology approach using biographical interviews and projective tests – could also be conceptualised as traits, such as need for achievement (*n*Ach).

Psychometric approaches to identifying personality dimensions

Questionnaire construction and psychometrics

Contemporary views of traits are intimately related to the processes of measurement and assessment necessary to identify basic personality dimensions. Typically, the trait researcher has some hypothesis about the number and nature of the principal dimensions, and designs a questionnaire to measure them. Subsequent work investigates how useful a measuring device the questionnaire actually is, and modifies the questionnaire items in response to any shortcomings detected.

The initial development of a satisfactory questionnaire for measuring traits is not easy. Care must be taken in the composition of items: they must be easily understood and unambiguous, applicable to all respondents, and unlikely to cause offence (see Angleitner and Wiggins, 1986). There should also be some systematic sampling of the various expressions of the personality trait of interest. It is important also to check that items are not strongly contaminated by response sets or biases, such as social desirability, yea-saying or extreme responding (see also chapter 13). However carefully the questionnaire has been designed, it is still necessary to assess its adequacy formally, by application of psychometrics, the science of psychological measurement. Psychometrics provides statistical techniques which tell us how good a measuring tool a particular questionnaire is, just as we might assess the accuracy of a thermometer or balance in the physical sciences. The sophistication of modern techniques and the number-crunching power afforded by computers provide the contemporary researcher with powers of data analysis far beyond those envisaged by the pioneering trait researchers. Today's researcher is in some danger of becoming a sorcerer's apprentice though, as the increasing availability of powerful statistical packages raises the risk of misapplication and

abuse of statistics. Hence, understanding traits requires at least a rudimentary grasp of psychometrics. In this section, we provide a brief, non-technical overview of some of the key psychometric techniques applied to personality assessment. Of particular concern is factor analysis, because of its use in investigations of the fundamental structure of personality traits. For a more detailed review of psychometric statistics and personality measurement, Kline's (1993, 2000) accessible books are recommended. The reader should also note the importance of the Pearson correlation coefficient (r) in psychometrics. A thorough grasp of this statistic and its limitations is invaluable in understanding research on personality traits. Howell (2002) and Jensen (1980) offer good introductory accounts of Pearson's r.

Psychometrics of single scales

Any single trait scale must be satisfactory with respect to three essential criteria: *reliability*, *stability* and *validity* (for more detailed accounts, see Anastasi and Urbina, 1997; Cronbach, 1990; Jensen, 1980; and chapter 13).

Reliability. This refers to the accuracy with which the questionnaire measures a given quality. At this stage, we are not committing ourselves to specifying what that quality actually is. Reliability may be assessed by administering two alternative measures of the trait to a sample of subjects, and computing the correlation between them. If the correlation is high, the quality can be assessed consistently and the scale is reliable or *internally consistent.* If not, the two supposedly equivalent forms are not assessing the same quality, the scale is unreliable, and the items must be revised. The Cronbach alpha statistic is a widely used measure of reliability calculated from a single set of test items. It is, in effect, the correlation of the test with itself. In general, alpha tends to increase both as inter-item correlation increases, and as the number of items on the test increases.

Stability. Reliability should be distinguished from stability, which is the test–retest correlation of the scale over a given time interval. Personality is expected to change slowly as the person grows older, but it is expected that stabilities of trait measures will be fairly high over periods of a year or more. If we have a scale that is reliable, but has a low test–retest correlation, we may be assessing a mood or some other transient quality of the person, rather than a genuine trait.

Validity. The third essential quality for a personality questionnaire is validity: it must be shown that the measure actually does assess what it purports to assess. A scale may be reliable but not valid. For example, a fortune teller might use a highly consistent method for inferring a person's future from the lines on their palm, but the consistency of the technique would be no guarantee that the fortune teller's predictions were accurate. The most straightforward and convincing method for assessing validity is referred to as *criterion* or *predictive validity*. The trait measure is correlated with some independent index of a quality associated with the trait, as in the studies listed in table 1.2. Other external criteria frequently used in personality research include measures of job performance and behaviour, psychophysiological functioning and clinical abnormality. Establishing predictive validity is important

in the early part of questionnaire development and in applied settings. However, the ultimate goal of theory-driven trait research is to establish *construct validity.* The essence of construct validity is that correlations between the trait and external criteria are predicted in advance from an adequate scientific theory, rather than from common sense or a superficial analysis of trait characteristics. For example, we could use a psychobiological theory of personality to predict how a particular trait should correlate with measures of autonomic functioning, such as heart rate. Construct validity arises out of the total web of empirical data and theoretical analysis which builds up around a trait, sometimes referred to as its *nomological network* (Eysenck, 1981). The difficulties of construct validity are those of establishing scientific truth. Even 'good' theories are never fully satisfactory, and require periodic modification of hypotheses and concepts as new research findings are obtained (see Lakatos, 1976). Hence, construct validity is always somewhat provisional, and may be reduced or enhanced by fresh research. There are various other forms of validity, but they are of less importance than predictive and construct validity.

Psychometrics of multiple traits: factor analysis

The methods just described may be used to obtain a satisfactory scale for measuring a single trait, such as extraversion or agreeableness. However, we cannot arrive at a satisfactory model of personality simply by accumulating different traits. Inevitably, some of the traits will be positively correlated, and it will be uncertain whether the traits concerned are genuinely distinct, or simply different aspects of some unitary trait. The technique most widely used for the simultaneous identification of multiple traits is factor analysis, described in more detail by Gorsuch (1983) and, in a text for beginners, by Kline (1994). The input to a factor analysis is the matrix representing all possible correlations between the various items making up a questionnaire or questionnaires. The aim is to simplify the correlation matrix, by identifying one or more underlying dimensions or factors which account for most of the variation in individuals' item scores. Factors are defined by the individual items which correlate with or 'load' on them.

Let us look at an example of a simple factor-analysis, using trait data taken from a study by Matthews and Oddy (1993). One thousand and ten people working in British business occupations rated themselves on a set of personality-descriptive adjectives. Table 1.3 shows the correlation matrix for ratings on twelve of these adjectives, divided into three sets. Each set of four adjectives was thought to relate to a different broad personality trait: Conscientiousness, Agreeableness and Intellectance (self-rated intelligence and intellectual interest). The pattern of correlations seems to accord with this expectation. For example, correlations between the four conscientiousness items are moderately large, ranging from 0.35 to 0.54. Correlations between the conscientiousness items and the other adjectives are considerably smaller, ranging from 0.01 to 0.25. That is, if a person is hardworking, it is likely that they are also industrious, conscientious and meticulous, but we

Table 1.3 *Correlations between trait descriptive adjectives thought to relate to conscientiousness, agreeableness and intellectance (n = 1,010)*

Trait adjective	1	2	3	4	5	6	7	8	9	10	11	12
1 Hardworking	1.00											
2 Industrious	0.54	1.00										
3 Conscientious	0.47	0.47	1.00									
4 Meticulous	0.38	0.35	0.41	1.00								
5 Compassionate	0.24	0.12	0.21	0.16	1.00							
6 Tender-hearted	0.16	0.06	0.17	0.14	0.59	1.00						
7 Loving	0.21	0.12	0.19	0.15	0.42	0.51	1.00					
8 Mild	0.08	0.01	0.10	0.14	0.24	0.40	0.25	1.00				
9 Brainy	0.15	0.20	0.12	0.12	0.08	0.09	0.10	−0.06	1.00			
10 Knowledgeable	0.19	0.25	0.18	0.12	0.05	0.01	0.07	0.01	0.45	1.00		
11 Wise	0.22	0.21	0.25	0.22	0.14	0.13	0.15	0.07	0.38	0.38	1.00	
12 Intelligent	0.14	0.21	0.17	0.09	0.03	−0.00	0.10	−0.13	0.62	0.48	0.39	1.00

Source Matthews and Oddy, 1993

cannot predict whether they will also be agreeable or intellectual. Intuitively, we might say that there is an underlying dimension of conscientiousness, associated with all four related adjectives, together with distinct dimensions of agreeableness and intellectance. Factor analysis aims to show whether such intuitions are actually in agreement with the data, by re-describing the data in terms of hypothetical underlying constructs or factors. Its end-point is a listing of the correlations between each factor and each of the initial variables. Hence, if there is a 'conscientiousness' factor it should correlate with each of the four conscientiousness items, but it should be largely uncorrelated with the remaining items.

Table 1.4 shows the factor matrix obtained following extraction of three factors. The first factor is defined mainly by the intellectance items, the second by the conscientiousness items, and the third by the agreeableness items. We can now describe individuals' personalities in terms of three dimensions rather than twelve. (For the knowledgeable reader, we have run a principal components analysis, followed by varimax rotation. Note that there is a technical difference between 'factor analysis' and 'principal components analysis', which is not important in the present context.) Techniques exist for calculating factor scores that would describe any individual's intellectance, conscientiousness and agreeableness. Together, the three factors explain 59% of the variance in the original correlation matrix. This considerable gain in economy of description is bought at a moderate cost in loss of information about individual item responses. The assumption of factor analysis is that the information discarded is trivial, largely error and item-specific variance.

In a non-technical exposition of this kind, we cannot adequately explain the actual computation of the factor matrix (see Jensen, 1980; and Kline, 1993, 1994 for more detailed but accessible accounts). In brief, there are two stages to the

Table 1.4 *Factor solution obtained from correlational data of table 1.3*

	Factor 1	Factor 2	Factor 3
Hardworking	0.12	0.77	0.14
Industrious	0.19	0.78	−0.03
Conscientious	0.11	0.76	0.14
Meticulous	0.05	0.68	0.13
Compassionate	0.07	0.15	0.76
Tender-hearted	0.04	0.05	0.86
Loving	0.13	0.12	0.71
Mild	−0.12	0.05	0.60
Brainy	0.82	0.04	0.03
Knowledgeable	0.73	0.15	−0.02
Wise	0.62	0.21	0.15
Intelligent	0.84	0.07	−0.06

Note Factor solution obtained from principal components analysis, followed by varimax rotation

analysis, each of which produces a factor solution. The second-stage solution (shown in table 1.4) is usually preferred to the first-stage solution (not shown). At the first stage, the general principle is that the first factor extracted explains as much of the variation in data as possible. For the correlations shown in table 1.3, the first factor explains 28% of the variance. The next factor extracted then explains as much as possible of the remaining variance: 18% in the example. Subsequent factors are extracted on the same basis, with the third factor extracted from the table 1.3 data explaining 13% of the variance. In personality research, the principle of grabbing as much variance as possible for each successive factor does not usually give psychologically meaningful results. (The position is different in research on ability tests, where the first factor is typically an approximation to *g* or general intelligence.) The second stage of the analysis capitalises on the fact that there is an infinite number of mathematically equivalent factor matrices which may be extracted from a given correlation matrix. We can recompute the factor matrix to explain exactly the same amount of variance using different values for the factor loadings. This re-computation is referred to as rotation, because it can be illustrated geometrically (e.g., Kline, 1993, chapter 8). The principle used to guide rotation is that of simple structure, the assumption that the most meaningful factor solution is the one for which factor interpretation is most clearcut. The various methods of rotation aim to maximise the number of loadings which are either 1.0 or 0.0, so we can say unequivocally whether or not a given variable is associated with a given factor. The factor matrix shown in table 1.4 has been rotated, and approximates to simple structure: large loadings are all 0.60 or more, whereas small loadings do not exceed 0.21. Rotation re-assigns variance across factors more evenly: the three factors shown in table 1.4 explain 20%, 20% and 19% of the variance, respectively.

Limitations of factor analysis

No factor analysis should ever be accepted uncritically. Three questions should always be asked. The first is whether the data are actually suitable for factor analysis. Since the technique is based on Pearson correlation, its validity depends on whether the original correlations are satisfactory. For example, correlation does not represent non-linear relationships validly, and correlations will be reduced if measures are unreliable or if the range of variable scores is restricted (Jensen, 1980). It is important that there are sufficient items which relate to or 'mark' each hypothesised personality dimension. Factor analysis also requires large sample sizes, particularly when there are many items and when loadings of items on factors are expected to be small.

The second question is how much the results depend on the particular methods of analysis used. Factor analysis should really be seen as a family of related techniques, and the exact choice of method may profoundly influence the eventual solution. In the example of factor analysis described previously, the 'orthogonal' rotation that was used forced the factors to be independent, that is, uncorrelated. However, we could also have chosen an 'oblique' rotation that allowed the factors to be correlated if that gave better simple structure. Another key choice is the number of factors extracted (Zwick and Velicer, 1986). There is a number of rules for deciding how many factors should be extracted from a set of items, but none is definitive.

The third, and most difficult, question is what the results actually mean. Critics of factor analysis point out that the mathematical equivalence of alternative factor solutions make all of them suspect. This criticism is probably overstated. As we shall see, use of the simple structure criterion for rotation has led to real progress in identifying scientifically useful personality measures. The essential point is that factor analysis does no more than indicate structural relationships among sets of variables. Construct validity must be established for factorial dimensions just as it must for single scales, by relating factorial measures to external criteria, and developing a testable scientific theory.

Further techniques of factor analysis

The techniques discussed so far are *exploratory*: the researcher relies on simple structure or some other theory-neutral, empirical criterion to determine the eventual factor solution rather than any hypothesised target solution. Thus, exploratory factor analysis can only suggest hypotheses. A newer approach, *confirmatory factor analysis* (Jöreskog, 1973), allows hypothesis testing, because the pattern of factor loadings for a given set of items tested on a subject sample is specified in advance. The factor analysis calculates the factor solution which is closest to the hypothesised factor matrix, and computes the goodness of fit between actual and hypothesised matrices. The researcher can then gauge whether or not the data provide an acceptable fit to the initial hypothesis. Confirmatory factor analysis

is part of a larger group of techniques known as *structural modelling* (Bentler, 1995; Byrne, 2000). The researcher may specify any set of relationships between directly observed variables, and unmeasured or latent factors, and test whether the hypothetical model fits the data. Unlike conventional factor analysis, structural modelling may formally test for fit among competing models, so it is particularly useful for establishing construct validity.

If the investigator chooses an oblique rotation, which allows derived personality factors to be correlated, an intriguing possibility arises. If the factors are in fact correlated, we can run a further factor analysis of the correlations between the factors themselves. This second factor analysis will then identify *second-order* or secondary factors. For example, in cognitive ability research the initial factor analysis of test scores often gives us a set of 'primary' abilities, such as verbal, mathematical and spatial abilities, which are all positively intercorrelated. Factoring the correlations between these somewhat specific abilities then defines broader, higher-order ability factors, such as general intelligence or *g*. Similarly, in personality research, we may obtain secondary, or broader, personality factors by factoring correlated primary, or narrower, personality trait measures. In the next section of this chapter, we review attempts to establish a comprehensive set of primary trait dimensions, which could be used to provide a detailed description of an individual's personality. In the following section, we look at efforts to describe personality in terms of secondary traits such as extraversion and neuroticism.

Primary factors of personality: the 16PF and other questionnaires

The Sixteen Personality Factor Questionnaire (16PF)

Discussion of primary traits must begin with the work of Raymond B. Cattell. The Cattellian project is one of the most ambitious ever undertaken in psychology. It seeks to explain individual differences in every area of life from psychometrically sound measures of ability, motivation, personality and mood. Massive quantities of data have been generated by this enterprise (see, e.g., Cattell, 1971; Cattell and Kline, 1977), along with several widely used questionnaires and tests. Cattell (e.g., 1946) began his personality research with the lexicon of trait-descriptive words, but shifted the main focus of his work to questionnaire items early in his research career. He eventually identified twenty-three fundamental primary factors, one of which is an ability factor, general intelligence. The sixteen most robust of these dimensions are measured by the Sixteen Personality Factor Questionnaire (16PF: Cattell, Eber and Tatsuoka, 1970), which has been extensively used in research and applied settings over several decades. Cattell et al.'s (1970) version of the 16PF became a standard personality measure, but attracted a number of psychometric criticisms. Internal consistencies of some of the scales

were low, and several investigators (e.g., Barrett and Kline, 1982; Matthews, 1989) were unable to recover the Cattellian primary factors from factor analysis of the 16PF.

The latest version of the 16PF, the 16PF5 (Conn and Rieke, 1994), features improved internal consistency, with a mean Cronbach alpha for the sixteen scales of 0.74, although some alphas remain relatively modest (less than 0.70). However, internal consistency may have been increased at the cost of loss of comparability with previous 16PF versions. 51 per cent of the 16PF5 items are new or substantially revised, and correlations between equivalent scales on the 16PF5 and the previous version of the 16PF (Cattell et al., 1970) are small or modest in most cases (less than 0.6 for eleven scales, and less than 0.4 for four scales). The 16PF has a hierarchical factor structure, such that secondary factors may be derived from the intercorrelations of the sixteen primary factors (Chernyshenko, Stark and Chan, 2001). As we shall see subsequently, there is some correspondence between the 16PF secondaries and the personality factors of the five factor model, sometimes called the Big Five. Table 1.5 provides descriptions of the 16PF scales, together with examples of historical and literary figures who exemplify the qualities assessed. These should not be taken too seriously, in the absence of actual questionnaire data. The table also gives 16PF5 alpha coefficients. Note that in this and subsequent tables we adopt the common convention of omitting the decimal point from reliability and correlation coefficients.

Extensive evidence on the predictive validity of the various versions of the 16PF has been obtained. We provide two examples here. Barton, Dielman and Cattell (1971) found significant correlations between several 16PF primary scales and achievement in various school subjects. The high achiever at this level of education is outgoing (A+), conscientious (G+), venturesome (H+), self-assured (O–), and self-controlled (Q3+). None of the personality traits predicts achievement as much as intelligence (B) does, but other, similar research (Cattell and Butcher, 1968) shows that personality predicts achievement even when intelligence is statistically controlled. Figure 1.2 shows mean levels of the traits for three occupational groups, which differ as we might expect. Note the social reserve of physicists (low A and H), the high sensitivity (I) and imaginativeness (M) of artists, and the calmness of airline hostesses (high C, low Q4). A large study of the 16PF5 among Church of England clergy showed that, within this occupational group, many of the usual gender differences were reversed: female clergy were less outgoing (A), more emotionally stable (C), more dominant (E), less rule-conscious (G), less emotionally sensitive (I), less apprehensive (O), and more open to change (Q1) (Musson, 2001). The 16PF is also useful for discriminating various clinical groups from one another and from normal subjects.

Although the 16PF has good predictive validity, doubts remain about the construct validity of the 16PF scales. Cattell (1973) provides detailed descriptions of qualities associated with the scales, which include references to experimental and psychophysiological data. However, there has been little attempt to use this

Table 1.5 *The fifteen personality traits assessed by the 16PF, with examples of famous individuals exemplifying the traits, and 16PF5 alpha coefficients*

	Trait descriptions		Famous individuals		
Trait	High	Low	High	Low	Alpha
A	Outgoing	Reserved	Falstaff	Greta Garbo	69
	Warmhearted	Detached			
C	Unemotional	Emotional	Washington	Hamlet	78
	Calm	Changeable			
E	Assertive	Humble	Genghis Khan	Jesus	66
	Dominant	Cooperative			
F	Cheerful	Sober	Groucho Marx	Clint Eastwood	72
	Lively	Taciturn			
G	Conscientious	Expedient	Mother Teresa	Casanova	75
	Persistent	Undisciplined			
H	Venturesome	Shy	Columbus	Sylvia Plath	85
	Socially bold	Retiring			
I	Tough-minded	Tender-minded	James Bond	Robert Burns	77
	Self-reliant	Sensitive			
L	Suspicious	Trusting	De Gaulle	Pollyanna	74
	Sceptical	Accepting			
M	Imaginative	Practical	Van Gogh	Henry Ford	74
	Bohemian	Conventional			
N	Shrewd	Forthright	Machiavelli	Joan of Arc	75
	Discreet	Straightforward			
O	Guilt-prone	Resilient	Dostoevsky	Stalin	78
	Worrying	Self-assured			
Q1	Radical	Conservative	Karl Marx	Queen Victoria	64
	Experimental	Traditional			
Q2	Self-sufficient	Group-dependent	Copernicus	Marilyn Monroe	78
	Resourceful	Affiliative			
Q3	Controlled	Undisciplined	Margaret Thatcher	Mick Jagger	71
	Compulsive	Lax			
Q4	Tense	Relaxed	Macbeth	Buddha	76
	Driven	Tranquil			

Note Dimension B (Intelligence) is omitted. Examples of famous individuals are partly taken from Cattell (1973)

Sources Cattell, 1973; Conn and Rieke, 1994

descriptive information on scale correlates to derive detailed, testable hypotheses concerning the nature of the psychological constructs associated with the scales. Cattell's (1983) favoured theoretical approach is the construction of linear equations which predict behaviour from individual difference measures. However, most psychologists would see this approach as essentially descriptive; the nature of the constructs linked to behaviour remains obscure.

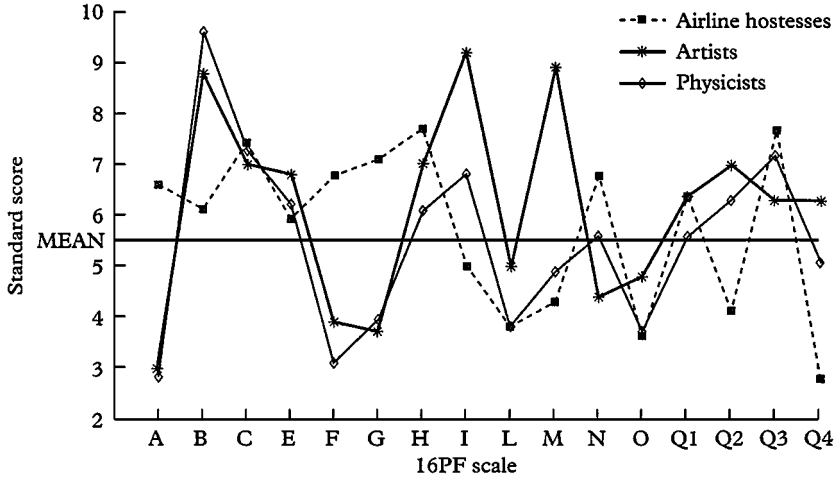

Figure 1.2 Mean scores obtained on the 16PF by three occupational groups
Source Cattell and Kline, 1977

Other systems of primary factors

Several other questionnaires attempt to assess primary traits comprehensively but most suffer from deficiencies more serious than those of the 16PF (see Kline, 1993, for a review). Perhaps the most popular is the California Psychological Inventory (CPI: Gough, 1987; Gough and Bradley, 1996) which assesses eighteen traits with moderately good reliability, and is widely used in industry. However, development of the CPI made no reference to factor analysis. Instead, the method of *criterion-keying* was used: items were chosen on the basis of their ability to discriminate criterion groups. This method has the serious disadvantage that scales may not correspond to those obtained by factor analysis, and, in the absence of systematic experimental studies, construct validity is lacking (see Kline, 1993). A more recent questionnaire is the Occupational Personality Questionnaire (Saville et al., 1984), which measures thirty-one traits relevant to personnel recruitment and selection, career development and training. Reliability of the scales is good, although, like the CPI, the thirty-one-trait model is not explicitly based on factor analysis. A recent re-analysis of the OPQ standardisation data (Matthews and Stanton, 1994) concluded that only about twenty dimensions could be identified through factor analysis of the items, although correspondences between these dimensions and the traits hypothesised by Saville et al. (1984) were good. There is also encouraging evidence for the validity of the OPQ traits (Saville et al., 1996).

Higher-order factors: the 'Big Five' or the 'Gigantic Three'?

In this section we describe two prominent personality schemes which advocate the usefulness of higher-order secondary factors, describing personality

Table 1.6 *Traits associated with the three dimensions of Eysenck's model of personality*

Neuroticism	Anxious, depressed, guilt feelings, low self-esteem, tense, irrational, shy, moody, emotional
Extraversion	Sociable, lively, active, assertive, sensation seeking, carefree, dominant, surgent, venturesome
Psychoticism	Aggressive, cold, egocentric, impersonal, impulsive, antisocial, unempathetic, creative, tough-minded

in broad, abstract terms. Within these schemes each dimension may be assumed to be significantly related to hundreds of basic trait terms. The proper identification of such higher-order factors, their validation, the discovery of their origins, and the demonstration of their value in predicting behaviour are the chief goals of trait researchers.

H. J. Eysenck's three factor model

According to the personality theory of Eysenck (1967, 1997), there are three broad personality factors, named neuroticism, extraversion–introversion, and psychoticism. These factors are assessed using a self-report questionnaire in which the testee is required to answer 'yes' or 'no' to a number of questions. The questionnaire has evolved through several different versions, culminating in the Eysenck Personality Questionnaire-Revised (EPQ-R: Eysenck and Eysenck, 1991). The EPQ-R, like some of its predecessors, also contains a 'Lie scale' intended to measure subjects' tendencies to 'fake good' when completing the questionnaire. Although Eysenck's higher-order dimensions are intended to be statistically uncorrelated, there are slight positive correlations, especially among male subjects, between psychoticism and the other two scales (Eysenck and Eysenck, 1991). The distribution of neuroticism and extraversion scores in the population approximates to a normal curve, whereas psychoticism scores are markedly skewed towards low scores.

Some of the lower-level traits captured by Eysenck's three dimensions are shown in table 1.6. Eysenck and Eysenck (1991) describe the typical extravert – a high scorer on the introversion-extraversion scale – as someone who is sociable, craves excitement, takes chances, is fond of practical jokes, is not always reliable, and can at times lose his temper. Their characterisation of the typical introvert is someone who is quiet and retiring, is fond of books rather than people, is serious, keeps feelings under close control, is reliable and has high ethical standards. The high neuroticism (N) scorer is someone who tends towards anxiety and depression, worries, has bad sleep and psychosomatic disorders, allows emotions to affect judgement, and is preoccupied with things that might go wrong. Unlike the high neuroticism scorer, the low N scorer recovers quickly after an emotionally upsetting experience and is generally calm and unworried.