Personality measurement and testing

An overview

Boyle, Gregory; Matthews, Gerald; Saklofske, Donald H.

Published in:
The Sage handbook of personality theory and assessment

1-1-2008

Personality measurement and testing: An overview

Gregory J. Boyle
Bond University, Gregory_Boyle@bond.edu.au

Gerald Matthews

Donald H. Saklofske

Personality Measurement and Assessment: An Overview

Gregory J. Boyle, Gerald Matthews, & Donald H. Saklofske

A colleague recently remarked: "Psychologists who specialize in the study of personality and individual differences spend a lot of time coming up with various descriptions of people, like Machiavellianism, external locus of control, openness to experience, and Neuroticism. Even more effort is spent trying to measure these ideas with tests like the MMPI-2, brief anxiety scales, and Rorschach Inkblots. But do they really tell us anything about human behaviour in general or about the individual? Does it make a difference in how we view people, select them for jobs, or guide therapy choices and assist in evaluating outcomes"?

This is a very loaded question and one that appears to challenge both the technical adequacy of our personality measures, but especially the construct and criterion validity or effectiveness of personality instruments in describing individual differences, clinical diagnosis and guiding and evaluating interventions. Technically; there are very few actual "tests" of personality—the *Objective-Analytic Battery* being an exception. Most so-called "tests" of personality are in fact self-report scales or reports of others' rating scales. Such scales quantify subjective introspections, or subjective impressions of others' personality make-up. At the same time, it is a relevant question and one that we will continue to face in the study of personality and the application of the findings, including assessment of personality, within psychological practice areas such as clinical, and school psychology, and within settings such as the military, business and sporting fields, among

others. Vol. 1 in this two volume series is devoted to a critical analysis of the theories, models and resulting research that drive the personality descriptions and assessment discussed in Volume 2. Demonstrating both the construct and practical validity of personality descriptions is essential to psychology as a scientific discipline and empirically grounded practice/profession.

The Status of Psychological Assessment

In a recently published paper focusing on psychological assessment, the following claim was made:

> Data from more than 125 meta-analyses on test validity and 800 samples examining multimethod assessment suggest four general conclusions: (a) Psychological test validity is strong and compelling, (b) psychological test validity is comparable to medical test validity, (c) distinct assessment methods provide unique sources of information, and (d) clinicians who rely exclusively on interviews are prone to incomplete understandings (Meyer et al., 2001, p. 128).

The authors also stated that multiple methods of assessment in the hands of "skilled clinicians" further enhanced the validity of the assessments so that the focus should now move on to how we use these scales in clinical practice to inform diagnosis and prescription.

This is a remarkable accomplishment, if accurate, and even a bold claim that has not gone unchallenged. Claims (a) and (b) have been attacked on various grounds (e.g., see critiques by Fernandez-Ballesteros, 2002; Garb et al., 2002; Hunsley, 2002; Smith, 2002). Furthermore the debate about the clinical or treatment validity of psychological assessment and the added or incremental value of multimethod assessment is argued by some to not rest on solid empirical ground (e.g., Hunsley, 2003), in spite of such carefully argued presentations on the utility of integrative assessment of personality with both adults (e.g., Beutler & Groth-Marnat, 2003) and children (e.g., Riccio & Rodriguez, 2007; Flanagan, 2007). In fact this is very much the argument put forward by supporters of RTI (respnse to intervention) in challenge to the view that diagnostic assessment, using multiple assessment methods, should point the way to both diagnosis and intervention planning (see special issue of *Psychology in the Schools*, 43,7, 2006).

While the Meyer et al. review focused on all areas of psychological assessment, it does suggest that the theories and models, as well as research findings describing various latent traits underlying individual differences have produced sufficient information to allow for reliable and valid measurement and in turn, application of these assessment findings to understanding, predicting and even changing human behaviour associated with intelligence, personality and conation (see Boyle & Saklofske, 2004). While there has been considerable progress, but certainly not a consensus in the models and measures used to describe intelligence and cognitive abilities, the other main individual differences' areas of personality and conation have traveled a somewhat different path to their current position in psychological assessment.

Calling this a remarkable accomplishment also has to be put in the context of time. Psychological science is only slightly more than 125 years old. As a profession that applies the research findings from both experimental and correlational studies in diagnosis, intervention and prevention in health care settings, schools, business, etc., psychology is even younger. Specializations that are heavily grounded in psychological assessment such as clinical, school, counseling, and industrial-organizational psychology only began to appear more or less in their present form in the mid-20th century. While it can be debated, the success of the Binet intelligence scales in both Europe and North America in the early 1900s, followed by the widespread use of ability and personality instruments for military selection during WW1 in the USA, and the growing interest in psychoanalysis complimented by development and use of projective measures to tap "hidden" personality structures, provided the strong foundation for the contemporary measurement and assessment of personality.

A Brief Historical Note on Psychological Assessment

However, history shows that the description and assessment of individual differences is not new to psychology. Sattler (2001) and Aiken (2000) have provided brief outlines of key events in cognitive and educational assessment during the several hundred years prior to the founding of psychology and one can clearly sense that the "tasks" of psychological measurement were being determined during this time. Prior to the creative scientific studies by Galton in the 19th century, the first psychological laboratory established in 1879 by Wundt, and psychology's earliest efforts at measuring the "faculties of the mind"

during the Brass Instruments era (e.g., James McK Cattell), there is a long history documenting efforts to describe the basis for human behaviour and what makes us alike all others and yet unique in other ways. As early as 4000 years ago in China, there is evidence of very basic testing programs for determining the "fit" for various civil servants followed by the use of written exams some 2000 years ago that continued in various forms through to the start of the 20th century. Efforts to understand and assess human personality also have a long history that predates the study of psychology. Centuries before the psychoanalytic descriptions of Freud, who argued for the importance of the unconscious and suggested that the putative tripartite personality structure of the id, ego, and superego were shaped by a developmental process reflected in psychosexual stages, the Roman physician Galen contended that human personality was a function of the body secretions (humors). Galen subsequently outlined the first personality typology characterized by the choleric, melancholic, sanguine, and phlegmatic types.

Interest in such processes as memory and reaction time, and efforts to assess and distinguish between mental retardation and mental illness were already underway before the establishment of Galton's psychometric laboratory in London and Wundt's and Cattell's psychophysical laboratories in Germany and the USA respectively. While much of this work was focused on the study of intelligence and cognitive abilities, it laid the the foundation for psychological testing and assessment that has shaped the face of psychology today. Probably the greatest impetus for test development came as a result of the success of the Binet intelligence tests, first in France and then in the USA. The use of tests to classify school children according to ability was followed by the development and

use of the Army Alpha and Beta tests to aid in the selection of recruits (in terms of their cognitive abilities) for military service in the US Army. However at that same time, it was also recognised that there was a need to identify military recruits who might be prone to, or manifest the symptoms of psychological disorders. Woodsworth (1919) created the *Personal Data Sheet* that presented examinees with a questionnaire not unlike those found on scales tapping psychiatric disorders to which a "yes-no" response could be made. While there was not a control or check for "faking good-faking bad" protocols, the measure was deemed to be a success. Thus, well before the *Minnesota Multiphasic Personality Inventory* (MMPI), constructed by Hathaway and McKinley (1940) and its revised version (MMPI-2), as well as the *California Psychological Inventory* (CPI), and other more recent personality measures, Woodworth's (1919) *Personal Data Sheet*, was followed shortly after by other personality scales such as the *Thurstone Personality Schedule* (Thurstone & Thurstone, 1930) and the *Bernreuter Personality Inventory* (Bernreuter, 1931), which may be considered the earliest personality measures, at least employing a contemporary questionnaire format. Of interest is that other measures being constructed around the same time highlighted the divergent views on personality assessment methods at the time including the *Rorschach Inkblot test* (Rorschach, 1921) and the *Human Figure Drawings* (Goodenough, 1926) and the *Sentence Completion Tests* (Payne, 1928).

Psychological Science vs. Pseudoscience

The basis by which current psychological assessment methods and practices can be separated from other attempts to describe the latent traits and processes underlying

differences in human behaviour is the very fact that they are grounded in scientific research, as outlined in the editors' introduction to Vol. 1. It is science that forces a method of study including, objectivity, experimentation and empirical support of hypotheses, and requires the creation and testing of theories. Psychology requires the operationalizing of variables and factors to be used in a description of human behaviour. In contrast to pseudosciences that operate outside of this framework and rest their case in beliefs, personal viewpoints, and idiosyncratic opinions, psychology also demands replication and, where possible, quantification of measures.

Measurement is the cornerstone of psychology and has spawned a number of methods for gathering the very data that may demonstrate the usefulness or lack of usefulness of a theory or provide the information needed to describe a particular human personality characteristic or even diagnose a personality disorder or clinical condition. Pseudosciences such as astrology, palmistry and phrenology, that compete with psychological views of personality, do not require such objective evidence to support their claims; rather vague "theories" are treated as fact and so-called evidence is often tautological. Thus a strength of psychology is that it has as its basis, measurement that includes varying methods of gathering data to test theoretical ideas and hypotheses, as well as strict adherence to psychometric measurement principles such as reliability, validity and standardization (cf. Boyle, 1985).

Foundations of Personality Measurement and Assessment

As mentioned above, it was concurrent with the advent of WW1 that a major effort to assess personality characteristics was first witnessed. Prior to that time, the closest measure of personality would likely be considered the word association techniques used by Jung. Today almost everyone is familiar with personality measures, self report questionnaires and rating scales that most often appear in the form of a statement or question (e.g., " I am a very nervous person"; "I enjoy activities where there are a lot of people and excitement") that the client answers with a "yes-no," "true-false" or an extended scale such as a 5 or 7 point or greater Likert-type scale with anchors such as "always true of me- never true of me" or "definitely like me- not at all like me." These highly structured measures contrast with the more ambiguous, subjective, and open ended techniques most often found in projective tests such as the *Rorschach Inkblot* or *Thematic Apperception Tests*.

Indeed, there is a longstanding tension between objective and subjective strategies for personality assessment (see Cattell & Johnson, 1986). Use of questionnaires based on subjective insights and self-reports has dominated the field, but one may wonder how much this dominance reflects the convenience and low cost of questionnaire assessment. Advocates of objective testing may legitimately question the validity of subjective experience and the apparent ease with which desirable responses may be faked. Table 1 sets out the key issues dividing the two camps; both have strengths and weaknesses. We do not take a position on which approach is ultimately to be preferred; the chapters in this volume illustrate the vitality of both subjective and objective measurement approaches. Ideally, multimethod measurement models in which subjective and objective indices

converged on common latent traits are to be desired, but current measurement technology remains some way from achieving this goal.

Table 1 about here

Given the more common use of standardized personality measures such as, for example, the MMPI-2, CPI, 16-PF, EPQ-R, and NEO-PI-R, a brief description of the strategies underlying their construction will be presented here. Kaplan and Saccuzzo (2005) provide a useful description of the various strategies employed in constructing personality measures. Deductive strategies employ both face validity (logical –content strategy) and theory driven views of personality. However the assumption that an item followed by a "yes" response, on the basis of content alone ("I am frequently on edge") taps anxiety or the broader neuroticism dimension found on scales assessing the Big Five (NEO-PI-R) or the three Eysenckian dimensions (EPQ-R) may or may not be accurate. And for instruments that employ a face-validity perspective, the rational approach to constructing items to measure particular characteristics may provide the client motivated by other alternative needs with the opportunity to provide inaccurate and biased responses (e.g., see Boyle et al., 1995; Boyle & Saklofske, 2004). For example, a scale purportedly tapping aggression with items such as " I often start fights" or "I have never backed down from a chance to fight" may be so transparent as to increase the likelihood that examinees will also be more able to create a "false" impression, depending on their motivation (e.g., early parole or lighter court sentence, malingering).

The foundational basis for many contemporary personality scales includes empirical strategies that employ the responses of various criterion groups (e.g., anxious vs. non-anxious adolescents) to determine how they differ on particular items and scales. For example, the very successful psychopathology scales, namely the MMPI (Hathaway & McKinley, 1940, 1943) and the revised MMPI-2 published in 1989, and the *Millon Clinical Multiaxial Inventories* I, II, and III (Millon, 1977, 1987, 2006) as well as the "normal" personality trait scale, the *California Personality Inventory* or CPI (Gough, 1987), are examples of instruments grounded in this approach to test construction and clinical use. Criterion-keyed inventories employ the approach that is less tied to what an item "says" or any *a priori* views of what it might be assessing, but rather whether the item discriminates or differentiates a known extreme group (e.g., clinical groups such as depressed, schizophrenic, etc.) from other clinical and normal respondants.

In other instances, statistical techniques, particularly factor analysis, are also used to infer or guide psychologists in determining the meaning of items and, thus, to define the major personality trait dimensions. Cattell's *Sixteen Personality Factor Questionnaire* or 16PF began as a large set of items based on a lengthy trait list that were then reduced to 36 "surface traits" and then further to 16 source traits, said to describe the basic dimensions of personality structure (see Boyle, 2006). In turn, structural equation modeling (see Cuttance & Ecob, 1987) then allows personality structure to be examined in the larger context of other psychological variables to portray a more comprehensive and integrated description of human behaviour.

Finally, theory driven measures draw from descriptions of "what should be" or "folk concepts" (e.g., CPI) and use this as the basis for constructing personality instruments, an example being the *Edwards Personal Preference Schedule* based on Murray's description of human needs. The major personality theories that have influenced the measurement of personality include psychoanalysis (e.g., *Rorschach Inkblot Test*; Vaillant's (1977) *Interview Schedule* for assessing defense mechanisms), Phenomenology (Rogers & Dymond's *Q-sort*), Behavioral and Social Learning (*Rotter's I-E Scale*) and trait conceptions (Cattell's 16 PF; Eysenck's EPQ-R; and Costa & McCrae's NEO-PI-R).

Certainly, the personality scales and assessment techniques most often employed today in both research and clinical practice include a combination of all the above approaches. The Eysenckian measures (e.g., MPI, EPI, EPQ-EPQ-R), the Cattellian measures (e.g., 16PF, HSPQ, CPQ; CAQ) as well as Big Five measures such as the NEO-PI-R have relied not only on theory, but also on empirical and factor analytic input into scale construction. Thus, the argument may be made that the NEO-PI-R, in spite of varying criticisms (see Boyle, Vol. 1), is a popular instrument for assessing putative trait dimensions labeled: Extraversion, Neuroticism, Conscientiousness, Agreeableness, and Openness to Experience. However, as Boyle et al. (1995, p. 432) reported, the NEO-PI-R accounts for less than 2/3 of the known trait variance within the normal personality sphere alone. Indeed, the proponents of any one of the major personality measures we have listed would claim that the measure concerned is based on theory, supported by research findings and of practical value in clinical psychology and other applied fields.

Types of personality assessment

When one thinks of personality assessment, what usually comes to mind is the self-report questionnaire. This almost exclusive reliance on questionnaires asking the respondent to answer a series of questions is showing signs of change and will continue to do so as genetic, biological, and neurological markers for particular personality traits come to the fore over time. At this time, the emphasis is on multimethod assessment approaches to ensure a convergence of results related to personality (and other) assessment as well as diagnosis of cognitive and affective disorders, includes case history and other extant data, interview, observation, behavioural and a *pot pourri* of informal assessment strategies, in combination with standardized tests and questionnaires. However, it would appear from the Meyer et al. (2001) review, that standardized, norm referenced measures (standard set of questions, method of administration, scoring) are the most valid and reliable of the currently available methods for assessing personality constructs.

The use of questionnaires and self-report inventories has dominated the field of personality measurement. In contrast to performance measures used in the assessment of cognitive ability (intelligence tests) and the assessment of skills through the use of, for example, driving tests, musical competitions, and electrical apprenticeship practica, personality assessment has largely employed somewhat subjective self-report techniques or reports of others using questionnaires, checklists and rating scales. While questionnaire methods predated projective scales, their development was spurred by the need for standardized scales that would minimize human error in administration, scoring and interpretation. Use of such measures also allowed quantification of the personality

dimensions being examined. Accordingly, psychologists could not only determine the direction (e.g., introversion vs. extraversion) but also the magnitude (e.g., very high score on extraversion at say, the 98th percentile) of a particular trait. This, in turn allowed for further refinements in assessment as well as replicability and cross-validation of the instruments themselves. Standardized personality instruments are most often associated with the assessment of personality traits (see Matthews et al., 2003) including those described by H.J. Eysenck, R.B. Cattell, P.T. Costa and R.R. McCrae, D.N. Jackson, and others.

Projective measures are grounded in the tenets of dynamic psychology, beginning with the early psychoanalytic work of Freud. These measures were developed as a way of probing into unconscious content and motivations and to give a "window" into the basic personality of the client. Here it is the subjectivity that is celebrated both in terms of the structure-free format that clients are given to respond, often to ambiguous stimuli, that will presumably allow for the expression of personality but also the openness of interpretation afforded the clinician who is well grounded in the views and "clinical experiences" of dynamic psychology. While there are a number of projective measures ranging from the *Szondi and Blacky Pictures* to the *Rosenzweig Picture Frustration Test*, *sentence completion techniques*, and *House-Tree-Person Test* that were created during the early and middle part of the last century, the *Thematic Apperception Test* (Morgan & Murray, 1935) and *Rorschach Inkblot test* (Rorschach, 1921*)* remain the most often used projective measures today. Even with some waning in the interest of subjective/projective measures, in recent years, the well known *Draw-A-Person* and

Bender Gestalt Tests, among others, have been further extended to include the assessment of psychopathology and affective indicators (e.g., *Draw A Person: Screening Procedure for Emotional Disturbance*; Naglieri et al., 1991). Langens and Schmalt (Vol. 1) discuss more recent work that builds on the TAT.

As a reaction to the psychodynamic influence in psychology and further drawing from the earlier success of Pavlov and Watson's work in describing and changing behaviour, Skinner's model of operant conditioning was extensively embraced following WWII and for the following 30 years. Here there is no interest in inferring latent traits underlying the expressions of human behaviour, or searching for unconscious mechanisms (the so-called "Black Box") that might help explain individual differences. Rather personality is viewed or operationalized as observable behaviour reflecting the interaction between the person and his/her environment. Thus a behaviour that has been identified as potentially relevant for intervention (e.g., hitting others; talking out of turn) is observed in terms of frequency, duration, etc. in the context of its antecedent and consequent conditions. Thus it can be determined if the behaviour requires change and if so, the prescriptive approach for doing so is to change those antecedent (environmental factors such as a noisy and distracting classroom) and/or consequent (e.g., reinforcement) conditions that would maintain the behaviour in question. Furthermore, this method has considerable predictive utility regarding the likelihood of the occurrence of particular behaviours. Based on systematic behavioural observation, there is no need to infer personality factors or an underlying personality structure. However, it is the use of observational data, most

salient in the behavioural approaches, that is also central to the clinical and research study of personality.

Interviews have been a mainstay of psychological information and continue to form the cornerstone of such specialized areas as counseling psychology. Clinical psychology, industrial/organization and many other branches of applied psychology employed both structured and more open-ended interviews to gather critical information about a client's personal history, worries and concerns, career aspirations, mental health problems, etc. While personality tests are essentially a form of structured interview, the use of interview techniques in general are considered to be less reliable and valid in diagnosis and treament planning. However, in the service of a multimethod approach to personality assessment, interview data can have both exploratory and confirmatory usefulness. To paraphrase Gordon Allport, if you want to know what people think or feel, ask them!

In more recent years, explorations of the biological and neurological bases of human behaviour, from fields such as behaviour genetics and neuropsychology have contributed significantly to the study of personality. These contributions are extensively described in Volume 1 in chapters by Stelmack and Rammsayer (psychophysiology) and also by Johnson et al. (behaviour genetics). While many personality theories are firmly grounded in brain-behaviour and genetic explanations (e.g., Eysenck's E and N factors), tests of these hypothetical links are now much more possible with the use of MRI and fMRI, as well as metabolic, neurotransmitter, and genetic measures.

In line with the dominant tradition of the field, many of the contributions to this volume are concerned with questionnaire assessments. The various uni- and multidimensional personality questionnaires may be evaluated against agreed standards for determining the efficacy of a given psychological measure (*AERA/APA/NCME Test Standards*, 1999). These standards lay out a framework for interpreting reliability and validity, so that the questionnaire developer has the following obligations:

(1) To provide evidence for the reliability of the measure in question and information on the standard error of measurement.

(2) To demonstrate that a meaningful relationship exists between the test's content and the construct that it is intended to measure (similar to 'content validity').

(3) To provide theoretical and empirical analyses supporting (or disconfirming) relationships between the construct and the responses provided by the test-taker (e.g., checking that responses are not driven by social desirability or other biasing factors).

(4) To demonstrate that the internal structure of the construct is as suggested by the underlying theoretical framework (e.g., whether it is uni – or multidimensional; whether it is hierachical in structure, etc).

(5) To localize the construct within a nomological net; i.e., other individual differences variables to which the assessment relates, as specified by theory. This criterion relates to 'construct validity', including establishing both convergent and discriminant evidence, test-criterion relationships, and investigating how validity generalizes across samples, situations and cultures.

Readers may determine for themselves how well the leading questionnaires match up to these test standards. We have indicated previously in this introduction the need for alternatives to questionnaires, including objective tests. This volume also addresses these alternatives, in reviewing psychophysiological techniques that may lend themselves to assessment, and also implicit, objective and projective tests. Historically, it has often proved difficult to obtain evidence for reliability and validity that matches corresponding evidence for questionnaire assessments, but the chapters here provide optimism that a new era of computer-interactive objective assessment may be at hand.

Introduction to Volume Two

Volume 2 contains a series of in-depth and critical chapters on the broad topics of personality measurement and assessment written by leading experts. The chapters are grouped into several themes including general methodological issues, multidimensional personality instruments, assessing biologically-base and self-regulative traits, followed then by projective and objective personality measures, and lastly by measures assessing abnormal personality.

Genereral Methodological Issues

It is often said in relation to psychological assessment that the key to moving forward with psychometrically sound measurement rests with the definitions that are determined to best represents a particular domain of behaviour, psychological disturbance (or well-being), or underlying trait such as extraversion or neuroticism. From the start, we realize that this is a daunting task for psychologists that will invariably require an

interdisciplinary perspective and effort. John et al. (1988) quite rightly asserted that "personality psychology has not yet established a generally accepted taxonomy of its subject matter which includes all variation in the overt social behaviour and the internal experiences of individuals" (p. 171). This is based on the view that personality attributes, like so many other psychological constructs, including intelligence, are abstract concepts that are not directly observable, but rather are inferred. The search for a generally supported taxonomy would provide the needed basis for personality research, in spite of differences in theoretical orientation by bringing "an order" to the huge collection of personality variables that have been created and studied over the years. In turn, this has direct relevance to what we measure in our personality instruments and how we can use this information in understanding individual differences.

Saucier contends that how we define, organize and measure personality can be guided from lexical studies of natural language. In turn, these studies have formed the basis for a personality structure that runs the gamut from a single factor solution somewhat akin to 'g' in intelligence theory and measurement, to seven lexical factors. While language may partly determine the number of factors that emerge in an examination of personality structure based on human lexicons, the issue becomes even more apparent when we attempt to develop measures to assess personality. The question can be asked: are personality characteristics universal? If so, then other than their expression or the actual behaviours observed to infer a personality characteristic, the universality of personality traits for example should allow for the translation and adaptation of an instrument from one language and culture to another. But as we search further into the cultural and

linguistic fabric of differing societies, we find unique examples of personality factors that do not seem to have an equivalent elsewhere. Comparing and contrasting cultures that are defined by an independent vs. interdependent view of the self, the Japanese concept of ”Omoiyari” would seem to exhibit some relationship with prosocial behaviour and empathy as defined in Western psychology. However it is also unique because of the intuitive aspect (”sasshi”) that is valued so highly in societies that are grounded in an interdependent view of the person. The chapter by van de Vijver and van Hemert describes the critical aspects of the methodology required in cross-cultural research and instrument construction and then follow this with some of the advances in the cross-cultural measurement of personality. It readily becomes apparent that the search for both universal personality factors and potentially unique clusters of personality variables will not be uncovered by simply comparing the responses to scales administered in two different countries or cultures, even if the measures are translated.

In response to the diversity of views on personality that have resulted from various theoretical, research and measurement perspectives in psychology and allied disciplines, Jackson proposes a “hybrid mode”’ that should serve both heuristic and practical functions. Integrating biological, experiential and social-cognitive theories, Jackson describes how this model departs from earlier views that appear to have fragmented rather than unified the study of personality (e.g., viewing approach-avoidance as orthogonal constructs; separating temperament and character). Of particular interest to practitioners is Jackson’s contention that the proposed hybrid model will guide the implementation of various psychological treatment interventions. This has been a major

concern of clinical, school/educational, counseling and I/O psychologists as well as those who practice psychology in health, military, sport, forensic and other venues. As one psychologist known to the authors quipped recently, "what good does it do for the psychologist and client to know the client's scores on the BIG 5 or to tell the client that they are a stable introvert"! Predicting successful and unsuccessful outcomes with and without interventions will provide personality psychology with the status accorded to intelligence and intelligence tests.

Multidimensional Personality Instruments

The second group of chapters is focused on an examination of some of best known and most often used measures of personality. In contrast to scales that are intended to assess psychopathology such as the MMPI-2, these measures reflect an eclectic underpinning of theory, trait descriptions, and factor analysis that rather describe the structure of personality. As the late Professor Hans Eysenck so often reminded us, the Psychoticism or the P factor in his theory of personality, and also assessed on the EPQ and EPQ-R, is not a measure of psychotic behaviour or psychopathology. Rather it reflects a tough minded-tenderminded dimension that may predispose a person to psychopathy or schizophrenia.

The well known *California Psychological Inventory* is now over 50 years old and is considered to be very much akin to a "folk description" of personality in contrast to instruments either driven by theory or derived empirically from factor analysis. There is

some disagreement about the actual factor structure and whether this measure is best described within the currently popular Five Factor Model (FFM). As well, some of the scales on the CPI 260 and 434 are less reliable than is minimally ideal. However exploration of the current 20 CPI scales has resulted in some new scales summarized by Boer et al. (Vol. 2). The CPI is one of the more often used measures in the business sector by I/O psychologists for personnel selection but has also been used extensively in counseling and forensic settings.

Factor analysis has been a driving influence on the development of both intelligence and personality instruments. For example, the widely recognized three strata structure of intelligence described by Carroll was based on an analysis of 456 factor-analytic studies of intelligence. On the other hand, models of intelligence proposed, for example by Spearman and Thurstone many years earlier, have been tested with factor analysis (both exploratory and confirmatory) to determine if the proposed structure can be replicated with large data sets. Certainly, many of the trait descriptions of personality are in part derived from factor analysis (e.g., Cattell's 16 PF, or the theoretical structure is supported with the aid of factor analysis (e.g., Eysenck's PEN model reflected in the EPQ/EPQ-R)

The *Comrey Personality Scales* (CPS) are a very good example of how factor analysis has been employed over time to create the eight factors found in this measure. As is now expected with all scales that employ a questionnaire format and self ratings, a validity and response bias scale are included to assist in determining various biases that would then challenge the accuracy of the report and its clinical usefulness. Comrey has

provided solid empirical evidence in support of the factor structure as well as the validity and clinical use of these scales. A question sometimes asked about personality scales is "what does it mean to be extraverted… what does this tell the psychologist or even the client?" Comrey has provided clear clinical descriptions of what it means to score high or low on measures such as Orderliness vs. Lack of Compulsion, or Trust vs. Defensiveness.

Probably one of the very best examples of the early use of factor analysis to define and measure personality characteristics as well as to expand and refine the scale is found in the 16PF developed by R.B. Cattell that was first published in 1949 and subsequently the 5th revision being published in 1993. Cattell (1973) described personality as comprised of three levels. Starting with the 16 primary personality traits, factor analysis produced the second-order global factors that very much interface with the current Big Five personality factors (Krug & Johns, 1986). In turn, two broad factors, akin to Active Outward Engagement and Self-Disciplined Practicality vs. Unrestrained Creativity, emerge from the five second order factors. Such a model does allow for a personality description at several levels but also contributes to an understanding of individual differences (described on p.XX), as can be seen, for example, with Global Extraversion. Such a scale permits both research on large scale population comparisons (e.g., cross-cultural comparisons) but also at the level of the individual who has requested counseling for interpersonal problems or work-related stress. Heather Cattell also portrays the significant part played by the 16PF in defining the Big Five but further delves into the debate surrounding the correlated vs. orthogonal relationship of these factors, drawing

our attention to the "power" of factor analysis (varimax vs. oblique solutions) in defining the relational structure of personality. Another interesting issue is the relationship between the psychometric cornerstones of reliability and validity and how they interact to an optimal level on measures such as the 16PF. Of particular interest to the readers of this chapter are the comprehensive references to the clinical use and applications of the past and current versions of the 16PF.

Boyle and Barton have extended the chapter by Heather Cattell to first remind us that Raymond B. Cattell (as indexed by journal citations) is one of the most influential psychologists of the 20th century (Haggbloom et al., 2002, p. 142). We are also reminded of the huge compendium of research and measurement instruments in the personality field alone that Cattell gave us including the *Sixteen Personality Questionnaire, High School Personality Questionnaire, Adolescent Personality Questionnaire, Children's Personality Questionnaire, Early School Personality Questionnaire, Preschool Personality Questionnaire, Central Trait-State Kit, Objective-Analytic Battery* and the *Clinical Analysis Questionnaire* along with its more recent version, the *PsychEval Personality Questionnaire*. This chapter then turns to an analysis of personality measures using Barton's nine parameter model that targets key "Who, What, How" questions and echos Cattell's call for the development of personality measures that go beyond the use of only L and Q- measurement but also draw from the observation of behaviour (T- data). Such an approach will provide psychologists with the multimethod assessment framework needed to converge on the most accurate and meaningful description of an individual's personality. On another note, as one reads through these two volumes and

possibly becomes concerned about the diversity of personality models and measurement approaches, we are reminded of a statement by Eysenck (1984), also named as one of the most influential contributors to 20th century psychology (see Haggbloom). In an analysis of Cattell's personality theory, Eysenck stated that, "the Cattell and Eysenck constructs and theories should be seen not as mutually contradictory, but as complementary and mutually supportive" (p. 336).

The Big Five personality factors have dominated the personality trait literature over recent years. More references are seen to Extraversion, Neuroticism, Openness to Experience, Agreeableness and Conscientiousness as measured by the NEO-PI-R. than to any other set of personality traits, in spite of the lack of agreement among psychologists (e.g., Cattell, 1995; Eysenck, 1992; McAdams, 1992; Schneider et al., 1995). However Costa and McCrae have provided a very detailed "inside" look at the construction of the NEO-PI-R in relation to the position of trait psychology, and criticisms of earlier personality measurement including other trait measures based on Eysenck's P, E, N model and Cattell's 16PF. Moving beyond N, E and O, and influenced by Norman and Goldberg's factors defining the structure of personality, Costa and McCrae engaged in an extensive research program that resulted in the NEO-FFI, NEO-PI and the more recent NEO-PI-R and NEO-PI 3 (see Boyle et al., 1995, pp. 431-433, for a critique of the factor analytic methodology employed in construction of the NEO-PI-R).

While considerable research pertaining to scale reliability and validity has been undertaken and some of the key findings are included in this chapter, the question of

accuracy in self-report measures has been addressed quite simply in the NEO-PI and NEO-PI-R questionnaires. A last item asks respondents to say if they have answered items honestly and accurately and if all items have been responded to and the answer sheet completed correctly. However, this response can easily be manipulated. Even if a person thinks at a conscious level that he/she has answered "honestly", it does not follow necessarily that the responses are accurate, particularly if the individual has poor self-insight. This simple approach is quite in contrast to many other scales such as the 16PF, MMPI, PAI and BASC that have included a number of "validity" checks; however the computer scored version of the NEO-PI-R does give further indications of such potentially relevant indicators of response accuracy and bias in reporing such as the number of missing items. Although in recent times the Big Five personality factors have tended to take "centre stage" in personality research, the use of the NEO-PI-R in applied settings is tempered by Costa and McCrae's view that "more research is still needed to optimize its application."

As stated above, the debate over the number of traits that would "best" define personality, and in turn, that will have the greatest application to "real world" settings, ranging from personnel selection to therapeutic intervention prescriptions, has been heard for many years. Based on both taxanomic descriptions and factor-analytic investigations, it would appear that the three most often cited positions are those reflected in Professor Hans Eysenck's (1991) paper "Dimensions of Personality: 16, 5 or 3? Criteria for a taxonomic paradigm." Eysenck's personality theory, while having undergone various revisions as outlined by O'Connor, has stayed true to the position that the three major personality

dimensions of Extraversion, Neuroticism and Psychoticism are sufficient to account for individual differences across a wide spectrum of human behaviours. Eysenck's model has resulted in the publication of a number of scales beginning with the MPI tapping Extraversion and Neuroticism followed by the EPI and eventually the EPQ and EPQ-R that included the P scale. Children's versions of the EPI and EPQ scales have been concurrently constructed by Sybil Eysenck and are referred to as the Junior versions (e.g., Jr. EPQ). While Eysenck's model clearly described a number of primary traits from which the second order factors of E, N, and P emerged, there was less effort invested in developing scales to measure each of these. However, over time the components of extraversion, initially focusing on impulsivity and sociability were split off with sociability remaining as part of E along with such other primary traits as sensation seeking and venturesomeness. While the Eysencks developed several scales to assess impulsivity, venturesomeness and empathy, there was also some effort to select those items from the EPQ defining the three-factor space to predict criminal propensity and antisocial behaviour. The EPI and EPQ are still used extensively in research studies and have been translated and adapted for use in many different countries.

Entering the taxonomic debate regarding the number of personality factors that are needed to account for individual differences in behaviour, the *Zuckerman-Kuhlman Personality Questionnaire* (ZKPQ) has redefined the factor space described by the Big Five. As Zuckerman explains in his chapter, the ZKPQ was developed to derive a personality structure appropriate for measuring basic personality traits with their roots in biological traits. Zuckerman provides a detailed discussion of the evolution of the ZKPQ

into its current five-factor structure and the labeling of the factors as ImpSS (Impulsive Sensation Seeking), N-Anx (Neuroticism-Anxiety), Agg-Hos (Aggression-Hostility) and Sy (Sociability). While the factor structure and psychometric integrity of the ZKPQ and also the short-form (Zuckerman, 2002) has been replicated in cross-cultural studies, some studies have also attested to the potential for use in a variety of settings ranging from risk taking in college students (Zuckerman & Kuhlman, 2000) to the reactions of migrant groups to moving into a new and different culture (Schmitz, 2004). However, because of its underpinnings in the psychobiology of personality, Zuckerman contends that possibly one of the greatest uses of the ZKPQ should be to explore the underlying basis of personality in the brain.

Ashton and Lee have offered yet another model comprised of six factors, assessed using the HEXACO-PI (Ashton et al., 2006). They have argued that studies of more than a dozen languages show that six personality factors appear common to all. The name of this model (HEXACO) serves as an acronym for the names of the factors including: Honesty-Humility (H), Emotionality (E), eXtraversion (X), Agreeableness (A), Conscientiousness (C), and Openness to Experience (O). In their chapter, Ashton and Lee write that, "despite its lexical origins, the HEXACO model uses the name Openness to Experience rather than Intellect/Imagination/Unconventionality" (p.XXX). The major addition to this model, in comparison with the Big Five factors, and resulting scale, again in comparison with the NEO-PI, is the H factor. However, Ashton and Lee further suggest that the six personality factors described in their model reflect two dimensions representing altruistic in contrast to antagonistic tendencies, and engagement within

different areas of endeavour. It is the contention of the authors that the Honesty-Humility and Emotionality factors are what give the HEXACO model an advantage over the currently popular Big Five. This chapter provides Croatian, Greek, and Filipino data supporting the cross-cultural factor structure.

Tellegen and Waller describe the process of constructing a measure of personality from both deductive and external approaches to scale construction. However they argue that an "exploratory" approach used during the 10 year construction period of the *Multidimensional Personality Questionnaire* (MPQ; Tellegen, 1982, 1995) has the advantage of permitting changes during the research and development phases of test construction. The authors state that, "the intent in constructing the MPQ was to clarify and demarcate major dimensions in the self-descriptive personality trait domain"(p.XXX) and offer strong support for their scale (e.g., heuristic virtues, substantial heritabilities, links with neurobehavioral personality models, shared factor loadings with other major personality scales such as the Big Five and Cattell's 16 PF, as well as respectable scale reliabilities. An interesting and controversial issue is raised in this chapter regarding the congruence between self and other ratings, as well as with external criteria. When comparing self-report personality ratings with ratings by knowledgeable others (e.g., spouse, friends/peers, employer/employees), it is not uncommon to see some and possibly considerable divergence. This could be construed as measurement error related to the varying reliabilities of the scales or that the descriptors (e.g., items and scales) are aimed more at assessing latent traits vs. overt behavioural manifestations that are less readily observed by significant others. Score discrepancies might also reflect biases in

responding (social desirability, malingering) by the "client" leading to a difference in self vs. other ratings. Tellegen and Waller in fact argue that the very nature of the MPQ will possibly result in "reliable discrepancies which are potentially informative and should not be dismissed as simply demonstrating the fallibility of self-report" (p.XXX). One has to appreciate the view taken in this chapter about the use of and feedback from personality measures such as the MPQ in everyday applied settings: "With use of appropriate norms, feedback contributes to self-clarification by translating discrete self-statements into coherent and telling characterizations" (p.xxx),

Assessment of Biologically-based Traits

The third section of this volume is titled "Assessment of biologically-based traits" and includes five chapters focusing on the biological underpinnings of personality structure and measurement. A major criticism of many personality measures is that they are often so phenotypical in content and purpose that they miss describing the causal and underlying correlates of key personality traits and factors. Saklofske and Eysenck (1994) stated that "trait models of personality are sometimes ciriticized for apparently pretending to explain differences in behaviour by simply postulating the existence of traits based on that behaviour….Factor analysis and other correlational methods are not meant to tell us anything about causality but to act as tools for the discovery of a proper personality taxonomy. Having solved the problem, we may then go on to carry out the more difficult task of finding out why some people are more sociable, others shy, why some people are extraverted, others introverted" (p.XXX).

Brain-behaviour studies began with the earlier EEG studies and are now driven by technologically sophisticated techniques such as fMRI used in neuropsychological research. The fields of psychobiology and psychophysiology using early measures such as EMG, GSR and HR, have more recently engaged in direct studies of neurotransmitters such as dopamine and serotonin. The specialty areas such as behaviour genetics, initially studying human behaviour between individuals of varying degrees of genetic relatedness (e.g., kinship studies) now have access to DNA data. All are essential for a full understanding of human personality that links the phenotypic expression of personality with the genotypic foundations.

Furedy's chapter provides an historical backdrop for the relevance of psychophysiological measurement in the study of personality. He then describes how the following psychophysiological measures may be considered by personality researchers: "peripheral vs. central measures, baseline vs. response-to-challenge measures, tonic vs. phasic measures; uniphasic vs. multiphasic measures; lo-tech vs. hi-tech measures; physiological 'respectability' vs. psychological validity; temporal vs. localization measures; specific vs. reactive sensitivity; psychophysiological vs. behavioural measures; reliability vs. validity" (p.XXX). But before psychologists who rely on self-report questionnaires, and observation and interview data begin to feel that their measures are less adequate and not a "direct" measure of personality, Furedy provides a very good example of the high reliability but low validity of the polygraph for classification purposes (e.g., truthful vs. deceptive individuals).

De Geus and Neumann provide two very good reasons for the significance of psychophysiological measurement in furthering our understanding of personality. They point to the limitations of an over-reliance on self- and other-report "paper and pencil" measures that may be prone to various biases, distortions and psychometric shortcomings. In contrast, psychophysiological indices have the advantage that "voluntary control over the recorded biological signals is limited if not absent" (p.XXX). At more of a construct validity level, these authors argue that personality may only be completely understood by also describing the biological processes underlying the major dimensions of personality such as the Eysenck's Three or the Big Five. Focusing on the two most agreed upon personality traits, Extraversion and Neuroticism, compelling evidence is provided to show their psychophsiological underpinnings, ranging from ERP to fMRI data. In a similar tone to Furedy however, De Geus and Neumann also agree that the "reliability of psychophysiological measures is currently less convincing than those for paper-and-pencil measures and validity has been far more rigorously tested for the latter. They suggest two reasons for the "shortfall" of psychophysiological data to provide a solid foundation on which to build comprehensive understandings of personality. First, many studies rely only on a single measure (e.g., cerebral blood flow, EEG asymmetry) rather than examining the complex and often interactive nature of multiple causal pathways. Second, "mainstream neuroscience is still very much focused on universal affective and cognitive brain processes at the expense of individual differences…by not taking individual differences into account, or considering them a mere nuisance variable, many neuroscience studies may have failed to detect a link

between a brain structure and the putative affective and cognitive processes in which it is involved" (p.XXX).

Congdon and Canli focus their analysis of the biological basis of personality on the "primary" personality factor of impulsivity. While impulsivity is considered a multidimensional construct found in many personality descriptions, included in many personality scales, and identified in various psychopathological (e.g., Bipolar Disorder, DSM-IV-TR Axis 11 disorders) and neurobehavioural (e.g., AD/HD) conditions, "the fact that patients are classified based on a taxonomy that is not biologically based poses a serious challenge to efforts to investigate the biological basis of impulsivity" (p.XXX). Noting the shortcomings of fitting impulsivity into the larger framework of personality, they describe research from noninvasive neuroimaging and molecular genetic studies that have separately provided support for a biological foundation. However, their reliance of these studies on heterogeneous diagnostic categories and self-report measures, "…obscure any effect that a genotype may have on the phenotype of interest, especially when the size of the effect is small". (p. XXX). Thus the authors argue for an "endophenotype" approach that would combine neuroimaging and molecular genetic approaches and the show its efficacy in an investigation of dopaminergic gene variation on impulsivity.

Strelau follows up on a lifetime of work on temperament which he defines as a stable set of personality traits, essentially present from birth or early infancy, Although temperament has a neurobiochemical basis, changes may occur due to external

conditions. This chapter briefly outlines the more significant theoretical views of temperament and describes some of the key measures that have evolved within the psychometric tradition. The very fact that there are some 30 temperament instruments and more than 80 temperament scales reported in the literature, suggests that either the construct is so broad as to not be particularly useful in theory, research and practice, or that there is considerable overlap. Strelau states that "the results of factor analysis confirm the expectation of a broad five-factor domain of personality with temperamental scales located mainly in two 'arousability' factors: 'emotionality/neuroticism' and 'extraversion/activity' "(p.XXX). However, Strelau argues that much work is still required to add specificity to the very broad concept of temperament and to develop reliable and valid measures.

Assessment of Self-Regulative Traits

"Styles of self-regulation are integral aspects of personality" (Matthews et al., 2000, p. 199). They further argue that the integration of personality traits and self regulation requires a resolution of two divergent viewpoints; "the trait approach views personality as stable across time and across different situations...much of the literature on self-regulation adopts a social-cognitive perspective that conceptualizes personality as the outcome of idiographic, contextually sensitive cognitive processes" (p. 171). While most primary (e.g., impulsivity) and higher-order personality traits (e.g., extraversion) relate to styles of self-regulation, the chapters in Section 3 highlight this critical feature of human behaviour.

Shoda's chapter complements the two chapters in Vol. 1 that examine social-cognitive views of personality. Shoda further provides a detailed analysis of the social cognitive perspective on key questions that have confronted the study of personality. Earlier views of behaviour in more simple terms were described as a function of persons interacting with their environment (PxE). However, the growing realization of the complexity of both these variables was most obvious when observing both individual and intraindivual differences. While there is a predictability to human behaviours that is surely grounded in personality, there is also the observation that a person's behaviours will vary across situations. It is not uncommon to hear expressions such as "he is a situational extravert" suggesting that under particular situations, and the demands arising from particular circumstances, a person may behave or act somewhat differently than they might under other conditions. Thus a purely trait perspective does not account for such variability across situations, but at the same time, human behaviour is not continuously random. Shoda argues that such variability can best be understood by knowing what features in a given situation are "psychologically active" for each or us. It is the psychologically important or "if features" of situations that activate both cognitive and affective processes which in turn result in thoughts, feeling and actions. This social cognitive perspective outlined by Shoda suggests "that that *if…then…* profiles provide clues for identifying individuality and personality coherence within individuals' cross-situational variability. This variability need not be considered a source of error to be eliminated (pXXX).

As highlighted in the chapter by Fernandez, anger, hostility and aggression have been studied in psychology since the early formulations of Freud that elevated aggression to one of the major human "instincts." The early work of Rosenzweig using the *Picture-Frustration Test*, the questionnaire analysis of aggression developed by Buss and Durkee (1957) , the theoretical analysis of aggression by Bandura (1983) and the more recent

cognitive-behavioural descriptions presented by Dodge (e.g.., Crick & Dodge, 1996) are but some examples of the interest in understanding and measuring aggression, anger, and hostility. Fernandez distinguishes the qualitative aspects of anger from other emotions such as sadness and also the quantitative aspects of low (annoyance) to high levels (rage) but further contends that "anger can assume the form of an emotion, a mood, or a temperament, depending on whether it is phasic, tonic, or cyclic" (p.XXX). While questionnaires have served as the major methods of assessing anger and hostility, the major contribution of this chapter by Fernandez is the description of six core dimensions in the expression of anger including direction, locus, reaction, modality, impulsivity, and objective of anger. Using this model, Fernandez shows how anger profiles are created using the anchor points of these six dimensions.

In contrast to the strict trait approach for defining and measuring personality, Horowitz et al. have drawn from the interpersonal model of personality and have identified four interpersonal measures (behaviours, traits, interpersonal goals, interpersonal problems) that are further organized around the two interpersonal dimensions of communion and agency. This allows for the creation of a profile (or "nomological net") describing the individual using eight interpersonal variables. The measures derived from this model and their application to personality assessment are illustrated in this chapter but more importantly serve to "show how the four interpersonal measures (the IMI, IAS, IIP, and CSIV) may be used together to clarify other concepts in clinical psychology" such as personality disorders.

Probably one of the major catalysts for stimulating an examining of the interface between personality and intelligence called for by Saklofske and Zeidner (1995) is the more recent examination of emotional intelligence (EI). The two somewhat divergent views of EI reflected in the trait formulation with its closer links with such personality traits as the Big Five and the ability model, proposed by Mayer and Salovey (1997) are described in the chapter by Austin et al. (see. Vol. 1). What sets the so-called ability approach apart from the trait EI view is the focus on the interaction between emotion and cognition. As Rivers et al. outline in their chapter, "Emotional intelligence (EI) refers to the capacity to both reason about emotion and use emotion to enhance thinking and problem solving" (p. XXX). It is the skills of perceiving, using, understanding, and managing emotions that are the foundation of EI. Furthermore, the method of measuring EI can be contrasted. While trait scales (e.g., Bar-On, 1997, Schutte et al. 1998) mimic traditional self-report personality questionnaires where a person's position on the scale(s) is usually determined using normative comparisons, the ability scales (MSCEIT and MSCEIT-YV) discussed in this chapter employ a problem solving approach applied to emotional situations, using both consensus and expert scoring (MSCEIT) and veridical scoring (MSCEIT-YV).

The MSCEIT is a departure from standard personality assessment using self-report measures but rather, like intelligence tests, one that employs problem situations to which the respondent's answers are compared to expert opinion. Thus the low correlations between the two forms of EI assessment may reflect differing conceptual underpinnings of EI or method variance or both. And the far from high correlations with intelligence tests for both trait and ability measures raises the interesting question of whether EI

should be considered “an intelligence” (Austin & Saklofske, 2005). These are key issues raised by Roberts et al. in their paper (see Vol. 1) which continues to look critically at EI models and measures following the first and more recent books by Matthews et al. (2002, 2007). Focusing on the “intelligence” aspect of EI, Roberts et al. argue that self report measures do not assess intelligence and thus should not be construed as measures of EI, in contrast to ability-based models that are “the only appropriate ones to delineate, and hence investigate, emotional *intelligence*” (p.XXX). Arguing for a constrained view of EI, the authors then suggest that the ability (or maximum performance) model reflected in the MSCEIT kind of measures hold the greatest promise for assessing EI. However, they are currently limited by their “mono-operation and mono-method biases” and will benefit as well from “using new paradigms from emotions research, and new test construction techniques from I/O psychology” (P. XXX).

Implicit, Projective and Objective Measures of Personality

The next section of Vol. 2 turns to an examination of implicit, objective and projective personality measurement. Probably no other topic in the personality assessment literature has generated the same level of debate as that seen between proponents of standardized vs. projective tests. On another level, cross-cultural issues have also risen to the fore in relation to both personality as well as intelligence measures. The “emic-etic” perspectives on cross-cultural comparisons have raised a number of questions about how well both the constructs used in one culture to operationalize and assess, say intelligence or personality, travel across national, cultural, and linguistic borders. The reader may

also refer to Saucier's chapter on the significance of the lexicon in determining how a culture describes and values various human characteristic.

Paunonen and Ryan describe two "non-verbal" personality measures, the NPQ and FF-NPQ. In contrast to the more psychodynamic measures such as the TAT, these two measures focus on explicit (rather than unconscious) personality characteristics, are samples of behaviour reflecting personality traits rather than "signs" of some underlying personality disposition, and use a structured response format to ensure objective scoring and increase scorer reliability. A key advantage argued for this format by Paunonen and Ryan is that these tests are likely more portable and flexible when assessing individuals from different cultural and language backgrounds because the problem of translation, but also reading skills level, is reduced. In particular, the FF-NPQ should provide an alternative measure for determining the robustness of the Big Five across cultures.

The basic difference in the assumptions posed by projective vs. standardized personality instruments relates to whether personality traits and factors are explicitly known to the person who is self-reporting or instead, that personality is more implicit and may be assessed with techniques referred to as *Implicit Association Tests* (IATs). Schnabel et al. argue that IATs have a number of advantages over traditional questionnaire methods for assessing personality. As described by Schnabel et al, IAT measures are designed to "assess automatic associations between a contrasted pair of target (such as 'me' versus 'others') and attribute (such as 'anxious' versus 'confident') concepts through a series of discrimination tasks that require fast responding" (p. XXX). A basic premise of these

measures is that such motivational distortion factors as faking good/bad etc. are less likely to confound or yield misleading results. However, while social cognitive research has provided a foundation on which to build IATs, explanations for IAT effects are still less than fully understood, and the psychometric properties, especially reliability, of standard IAT measures is somewhat lower than considered desirable. At the same time, there is growing evidence from validity studies that IATs may provide another "method" for assessing personality that would allow researchers and clinicians to potentially address the issue of contaminating method variance that likely occurs when exclusively relying on self-report questionnaires. However at this time, there is not sufficient evidence that IAPs should be used in clinical decision making related to diagnosis and treatment planning and/or selection.

Schuerger provides yet a further alternative to assessing personality based on the efforts of Cattell and Warburton (1967) to create actual performance tests (T-data) of personality, a careful selection of which have been included in the *Objective-Analytic Test Battery* or OAB. While Schuerger concedes that that original versions were very cumbersome and not widely adopted, the idea underlying the OAB is quite contemporary and one that clearly supports a multi-trait, multi-method, multi-modal approach to assessing human characteristics. A consensus is lacking regarding the factor structure of the OAB with Schuerger stating that only six of the factors originally reported by Cattell have been replicated in research conducted outside Cattell's laboratory. However Schuerger also contends that the OA tests still hold remarkable promise as demonstrated in both educational and clinical settings, and there may be even greater untapped potential in

individual OA variables. In contrast to the time when Cattell was developing the OA tests, the advent of modern computer technology and widespread computer use may yet be the format for reviving interest in such perfornce tests of personality structure. It would be surprising not to find more performance measures being presented by computer in the very near future. Our clinical laboratories already have this capability but microcomputers will also make this is a reality for the "travelling clincian" such as the school psychologist.

Standardized personality instruments such as the Eysenck and Cattell measures and projective techniques including the Rorschach inkblots and TAT have certainly dominated the field of personality assessment for much of the 20th century. While personality scales are still a mainstay in contemporary psychology, including those described in the chapters of this volume, both research and practicing psychologists are also interested in assessing the manifestations and related behaviours of underlying personality dimensions such as and anxiety, depression, and aggression. Thus, while the tendency towards aggressive behaviour can be plotted on a three-dimensional matrix defined by E, N, and P, of greater clinical utility to psychologists is to have more "direct" measures of the level and typeof aggressive behaviour. For example, the early Buss-Durkee scale was more focused on assessing the direct expression of aggression in its own right just as were the depression and anxiety scales developed by Beck. Not all scales are focused on the "negative" or pathological side and in particular we now see scales tapping happiness, life satisfaction and subjective well-being.

The past several decades have witnessed the development of multiscale measures that tap a wide range of psychologically important behaviours. Scales developed by Achenbach and Connors multidimensional measures paved the way for many of the new scales that tap a number of behavioural factors of relevance to clinical diagnosis and intervention planning. The first and now recently revised *Behavior Assessment System for Children* developed by Reynolds and Kamphaus (1992, 2004) is not a personality measure in the strict sense of tapping those traits thought to underlie behaviour but rather a more direct assessment of behaviour itself. The advantages they offer to the practicing psychologist are described by Rowe et al. who also remind us that this or any other scale should never "stand alone" as the sole basis for diagnosis or prescription. The current BASC-II provides statements that the respondent (child, parent, teacher) answers using a 4-point (or True-false) Likert-type format yielding composite, primary, and content scores. For example, the mixture of items on the primary self-report scales range from anxiety to attention problems, self-esteem to sense of inadequacy, and locus of control to sensation seeking, thus reflecting a very eclectic mix of scales all focusing on behaviour and the behaviours argued to describe, say, locus of control. Of interest is that the BASC-II has become the most often used behavioural measure by school and child-adolescent clinical psychologists, in part because of its solid psychometric properties and time-cost benefits, but more so because it provides a "direct" method of assessment of both "strengths and problem areas."

There is clearly consensus among psychologists for a multi-method approach to personality assesssment based on empirically supported models and methods. Blais and

Baity have critically assessed the position of the two most well known projective measures, the *Rorshach Ink Blot Test* and the TAT in the context of clinical diagnosis by examining the DSM-IV Axis-II personality disorders and diagnostic efficacy of current scoring methods for both measures. While there is certainly controversy and disagreement about the use of projective instruments in assessing not just personality, but also psychopathology (e.g., Gacono, 2002), psychologists must remember that these are empirical questions that remain to be decided by the evidence. More to the point, Blais and Baity also remind us that there is not a direct correspondence between the various Rorschach or TAT scoring systems and an acutal DSM-IV diagnosis. Rather the contribution of these "performance" measures can best be realized when they are integrated into systems and perspectives describing personality and psychopathology. In a recent paper that the reader may also wish to consult, Hughes et al. (2007) focused attention on the use of the Rorschach by school psyhologists and after an extensive review of the Rorschach and Exner's Comprehensive System for administration and interpretation, concluded that they, "meet current ethical and legal starndards for tests" (p. 288).

Abnormal Personality Trait Instruments

The last section of this volume examines several very specific measures for assessing abnormal personality traits either through an examation of those personality characteristics known or believed to underlie psychopathological behaviour, or by a more "behavioral" examination of particular clinical conditions and syndromes that have been described in DSM and ICD classifcations. Just as the *Wechsler Intelligence Scales* were

deemed the most often used measures for assessing cognitive ability in the 20th century, so too, the MMPI and more recently the MMPI-2 have been among the most popular and often used self-report measures for assessing psychopathology. In fact, the MMPI was so widely used that it found its way from primarily psychiatric and forensic settings, to personnel selection and university counselling settings.

Again, much like intelligence tests of the earlier part of the 20th century that attempted to assess the full complement of cognitive abilities, the MMPI was intended as a comprehensive measure of the gamut of psychopathological conditions. Thus, in some ways, as the chapter by Helmes points out, the MMPI does stand out as compared with the shorter and more specifically focused measures of abnormal personality of more recent years. Both as a screening instrument and for distinguishing broad types of psychopathology (e.g., Depression vs. Psychopathy), both versions of the MMPI have served us well. However the MMPI is more limited for differential diagnosis (e.g., anxiety vs. depression), but then that is an unrealistic expectation for any measure, even one as lengthy as the MMPI, since the diagnosis of psychological disorders requires the convergence of clincal data from a multimethod approach. A review of the literature (Helmes & Reddon, 1993) does not provide a great deal of evidence to support the use of the clinical scales for differential diagnosis. With any instrument that has survived as long as the MMPI, there is the tendency for some myth or beliefs to "trump" what the evidence actually tells us about the MMPI's clinical efficacy and empirically validated best practices use. However, Helmes states that "there is promise that the new RC scales

will be better able to make such distinctions than the traditional Clinical scales, but the relevant studies have yet to appear in print" (p.XXX).

The MMPI/MMPI-2 raises an interesting issue about the current role of such measures in clinical assessment and the issue of efficiency. The tendency now seems to be towards use of shorter, more time efficient and more focused measures, although the more recently published BASC/BASC-II, for example, are not so brief. The advantage of the MMPI/MMPI-2 is that it does serve as a broad screening measure for evaluating various broad types of psychopathology (or the lack thereof). In contrast, the use of, say, a brief depression inventory, would only be more useful if the psychologist was either attempting to rule out depressive symptomatology or had formed the hypothesis of depression, based on other indicators (e.g., interview and presenting symptoms, family and previous clinical history) and was adding confirmatory evidence.

What is in store for this "battleship" measure? Helmes summarizes, "The MMPI-2 does not represent a highly sophisticated approach to assessment that is based upon the state-of-the-art in diagnosis and conceptualizations of psychopathology. Successive introductions of new scales have modernized aspects of the interpretation of the test, at the cost of providing increased opportunities for conflicting scores that need to be reconciled during the overly complex interpretive process. The escalating collection of scales for the MMPI-2, with each successive set providing at best modest increases in incremental validity for some applications, simply multiply the number of potential

sources of interpretive conflict….the future of the MMPI/MMPI-2 thus remains difficult to predict" (p.XXX).

While new measures, including personality, behaviour, and psychopathology measures continue to abound in psychology and certainly present a challenge to formerly well-established measures, Krug reviews the *Cattellian Clinical Analysis Questionnaire* (CAQ) and its revised version, the PEPQ as measures of both normal personality and psychopathology. In spite of some support for the psychometric strengths of the separate and composite scales, they do not appear to be often used in either research or in clinical settings, although in military contexts the CAQ has received considerable use (e.g., the Australian Army Psychology Corps has used the CAQ extensively in its psychological research, assessment and selection procedures—see Boyle, 1989). This chapter raises the interesting issue of whether we have been too quick to abandon the theoretical, research, and measurement contributions of such key figures in psychology as Cattell and Eysenck. As noted above, Eysenck's E and N scales as found on the EPQ and EPQ-R are psychometrically sound and central to a trait description of personality. Similarly a thorough study of the Cattellian instruments is needed before we too quickly engage in an "out with the old and in with the new" attitude and later discover that we may have simply "reinvented the wheel" (e.g., see the number of different scales that assess risk taking, sensation seeking, thrill seeking, etc.)

In contrast to the MMPI and MMPI-2 that are not grounded in a contemporary model of either personality or psychopathology, the measures described in the remaining chapters

provide reassuring alternatives. The *Dimensional Assessment of Personality Pathology* (DAPP) measures employ a construct validation approach to arrive at a classification of personality disorder. Empirical evaluations leading to revisions in the initial theoretical description of personality disorder are supported by increases in the validity of the classification scheme. The DAPP has also evolved from a somewhat different approach than many of the current personality and clinical scales that start with a description of the personality trait (e.g., extraverion) or disorder (e.g., anxiety). Rather than beginning with this *a priori* view, Livesley and Larstone state that the DAPP, "incorporates a bottom-up approach in which diagnostic constructs evolve based on empirical evidence of the way the features of personality disorder are organized" (P. XXX). The advantage of this approach is that it provides for a dynamic rather than static view of personality criteria and categories that is forced to modify or change with new evidence from both research studies and clinical use. Even more compelling is the reconciliation between normal and abnormal personality that were treated quite separately even into the latter part of the 20th century. This may well be why measures such as the EPQ and NEO-FFI and NEO-PI appeared in stark contrast to the many separate pathology scales (e.g., MMPI), with the latter not grounded in underlying personality factors but rather collections of psychiatric symptoms. The DAPP and its counterparts have provided the foundation for the much needed reconciliation between basic descriptions of personality models and traits, on the one hand, and personality disorders and psychopathology on the other.

As noted above, while there has been a tendency towards constructing more specifically focused and brief measures of both normal and abnormal personality traits, in contrast,

the *Personality Assessment Inventory* (PAI) devised by Morey (1991, 2007), has been constructed on the basis of contemporary diagnostic classifications, and as an attempt to overcome the well-documented limitations of the MMPI/MMPI-2 instruments. The PAI comprises no fewer than 344 items with a mix of validity, clinical, treatment consideration and interpersonal scales. Morey and Ambwani summarize findings showing that this multidinesional inventory has received increasing attention in both research and clinical practice settings. Studies supporting the validity of the PAI subscales in the assessment of a wide range of psychological problem areas, ranging from eating disorders to emotional injury, will ensure its continued use in both applied and research settings.

This section ends with a summary of the Millon inventories and a view of personality assessment. All would agree that any single theory and measurement instrument, no matter how robust or narrow can ever give a complete description of an individual's personality structure. As Millon argues, our efforts to measure human personality with a predefinded set of traits that are reflected in our assessment tools is complicated by the very nature of examining a breakdown of these traits for each individual, and then reconstructing a description of personality; the "loop" from idiographic individuality to nomothetic commonality to nomothetic individuality is brought to closure" (p.XXX). Millon's chapter provides a detailed overview of the inventories that he has developed over the past several decades, highlighting the links with both the DSM and ICD taxonomies of personality disorders, as well as the theoretical basis for conceptualing both personality and abnormal behaviour. The critical question that has so often been

posed regarding the direct association between assessment, diagnosis and treatment planning or therapy is addressed by Millon who argues that his inventories provide a necessary basis for associating polarity schemes and clinical domains with corresponding theapies: “Any discussion of *personalized psychotherapy*…must take place at a level of abstraction or integration commensurate with that of personality itself. Personality disorders and clinical syndromes cannot be remedied if the person is thoroughly integrated while the therapy is not. Therapy must be as individualized as the person” (p.XXX).

Summary Comments

The chapters included in this volume are testimony to the incredible progress that has been made in the measurement and assessment of personality, particularly in more recent years. Guided by various theoretical models and research findings as well as extensive interdisciplinary collaboration, the sophistication of psychological measurement will continue to provide the necessary assessment tools to further our basic and applied analysis of human personality. Each chapter in this volume is a celebration of the the research contributions and clinical knowledge of leading experts in personality measurement and assessment. We thank all of the authors for sharing with us their critical analyses of the models and methods for measuring personality and especially their insights and creativity that will serve well the clinical assessment of personality.

References

Aiken, L. R. (2000). Rating scales and checklists: Evaluating behavior, personality, and attitude. New York: Wiley.

American Educational Research Association, American Psychological Association, National Council on Measurement in Education (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Ashton, M. C., Lee, K., de Vries, R. E., Perugini, M., Gnisci, A., & Sergi, I. (2006). The HEXACO model of personality structure and indigenous lexical personality dimensions in Italian, Dutch, and English. *Journal of Research in Personality, 40,* 851-875.

Austin, E. J. & Saklofske, D. H. (2005). Far too many intelligences? On the communalities and differences between social, practical, and emotional intelligence. In R. Schulze & R.D. Roberts (Eds.). *Emotional intelligence: An international handbook* (pp. 107-128). Cambridge, MA: Hogrefe & Huber.

Bandura, A. (1983). Psychological mechanisms of aggression. In R.G. Geen & E.I. Donnerstein (Eds.), *Aggression: Theoretical and empirical reviews* (Vol. 1, pp. 1-40). New York: Academic.

Bar-On, R. (1997). *BarOn Emotional Quotient Inventory (EQ-i): Technical manual.* Toronto: Canada: Multi-Health Systems.

Barrett, P. (2005). What if there were no psychometrics?: Constructs, complexity, and measurement. *Journal of Personality Assessment*, *85*, 134-140.

Bernreuter R. G. (1931). *The personality inventory.* Stanford, CA: Stanford University Press.

Beutler, L. E. & Groth-Marnat, G. 2003). *Integrative assessmetn of adult personality* (Eds., 2nd Ed.). New York, NY: Guilford.

Boyle, G. J. (1985). Self-report measures of depression: Some psychometric considerations. *British Journal of Clinical Psychology, 24*, 45-59.

Boyle, G. J. (1989). *A guide to use of the Sixteen Personality Factor Questionnaire and Clinical Analysis Questionnaire in military selection*. Canberra: Australian Army Directorate of Psychology, 1st Psychological Research Unit.

Boyle, G. J. (2006). *Scientific analysis of personality and individual differences*. Doctor of Science Thesis, University of Queensland.

Boyle, G. J., & Saklofske, D. H. (2004). (Eds.), *Sage benchmarks in psychology: The psychology of individual differences, Vols. 1-4*. London: Sage.

Boyle, G. J., Stankov, L., & Cattell, R. B. (1995). Measurement and statistical models in the study of personality and intelligence. In D. H. Saklofske & M. Zeidner (Eds.), *International handbook of personality and intelligence*. New York: Plenum.

Buss, A. & Durkee, A. (1957). An inventory for assessing different kinds of hostility. *Journal of Consulting Psychology, 21*, 343-349.

Cattell, R. B. (1973). *Personality and mood by questionnaire.* New York: Jossey Bass.

Cattell, R. B. (1995). The fallacy of five factors in the personality sphere. The Psychologist, May, 207-208.

Cattell, R. B., & Johnson, R. C. (1986). (Eds.), *Functional psychological testing: Principles and instruments*. New York: Brunner/Mazel.

Cattell, R. B., & Warburton, F. W. (1967). *Objective personality and motivation tests: A theoretical introduction and practical compendium*. Champaign, IL: University of Illinois Press.

Costa, P.T. & McCrae, R.R (1992). *The NEO Personality Inventory (Revised) manual.* Odessa, FL: Psychological Assessment Resources.

Crick, N. R. & Dodge, K. A. (1996). Social information-processing mechanisms in reactive and proactive aggression. *Child Development, 67,* 993-1002.

Eysenck, H. J. (1991). Dimensions of personality: 16, 5, or 3? Criteria for a taxonomic paradigm. *Personality and Individual Differences, 12*, 773-790.

Eysenck H. J. (1992). Four ways five factors are not basic. *Personality and Individual Differences, 13*, 667-673.

Fernandez-Ballesteros, R. (2002). How should psychological assessment be considered.*American Psychologist,57,* 138-139.

Flanagan, R. (2007). Comments on the miniseries: Personality assessment in school psychology. *Psychology in the Schools, 44*, 311- 318.

Gacono, C. B., Evans, F. B., & Viglione, D. J. (2002). The Rorschach in forensic practice. *Journal of Forensic Psychology Practice, 2*, 33-53.

Garb, H. N., Klein, D. F., & Grove, W. M. (2002). Comparison of medical and psychological tests. *American Psychologist, 57*, 137-138.

Goodenough, F. L. (1929). *Measurement of intelligence by drawings.* New York: Harcourt, Brace, & World.

Gough, H. G. (1987). *California Psychological Inventory manual.* Palo Alto, CA: Consulting Psychological Press.

Haggbloom, S. J., Warnick, R., Warnick, J. E., Jones, V. K., Yarbrough, G. L., Russell, T. M., Borecky, C. M., McGahhey, R., Powell III, J. L., Beavers, J., & Monte, E. (2002). The 100 most eminent psychologists of the 20th century. *Review of General Psychology, 6*, 139-152.

Hathaway, S. R. & McKinley, J. C. (1940). A multiphasic personality schedule (Minnesota): Construction of the schedule. *Journal of Psychology, 10*, 249-254.

Hathaway, S. R. & McKinley, J. C. (1943). *The Minnesota Multiphasic Personality Inventory (rev. ed.).* Minneapolis, MN: University of Minnesota Press.

Helmes, E. & Reddon, J. R. (1993). A perspective on developments in assessing psychopathology: A critical review of the MMPI and MMPI-2. *Psychological Bulletin, 113*, 453-471.

Hunsley, J. (2002). Psychological testing and psychological assessment: A closer examination. *American Psychologist, 57*, 139-140.

Hughes, T. L., Gacono, C. B., & Owen, P.F. (2007). Current status of Rorschach assessment implications for the school psychologist. *Psychology in the Schools, 44*, 281-291.

John, O. P., Angleitner, A., & Ostendorf, F. (1988). The lexical approach to personality: A historical review of trait taxonomic research. *European Journal of Personality, 2*, 171-203.

Kaplan, R. M. & Saccuzzo, D. P. (2005). *Psychological testing: Principles, applications, and issues (6th Ed.).* Belmont, CA: Thomson Wadsworth.

Krug, S. E., & Johns, E. F. (1986). A large scale cross-validation of second-order personality structure defined by the 16PF. *Psychological Reports, 59*, 683-693.

Matthews, G., & Gilliland, K. (1999). The personality theories of H.J. Eysenck and J.A. Gray: A comparative review. *Personality and Individual Differences, 26,* 583-626.

Matthews, G., Deary, I. J., & Whiteman, M. C. (2003). *Personality traits* (2nd. ed.). Cambridge, UK: Cambridge University Press.

Matthews, G., Schwean, V. L., Campbell, S. E., Saklofske, D. H., & Mohamed, A. A. R. (2000). Personality, self-regulation, and adaptation: A cognitive-social framework. In M. Borksrtyd, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp.171-207). San Diego, CA: Academic.

Matthews, G., Zeidner, M., & Roberts, R. D. (2002). *Emotional intelligence: Science and myth.* Cambridge, MA: MIT Press.

Mayer, J. D. & Salovey, P. (1997). What is emotional intelligence? In P. Salovey and D.J. Sluyter (Eds.), *Emtional development and emotional intelligence: Educational implications* (pp. 3-31). New York: Basic Books.

McAdams, D.P. (1992). The five-factor model in personality: A critical appraisal. *Journal of Personality, 60(2),* 329-361.

Meyer, G.J., Finn, S.E., Eyde, L.D., Kay, G.G., Moreland, K.L., Dies, R.R., Eisman, E.J., Kubiszyn, T.W., & Reed, G.M. (2001). Psychological testing and psychological assessment: A review of evidence and issues. *American Psychologist, 56,* 128-165.

Millon, T. (1977). *Manual for the Millon Clinical Multiaxial Inventory (MCMI).* Minneapolis, MN: National Computer Systems.

Millon, T. (1987). *Manual for the Millon Clinical Multiaxial Inventory-II (MCMI-II).* Minneapolis, MN: National Computer Systems

Millon, T., (2006). *Millon Clinical Multiaxial Inventory-III Manual (3rd Edition).* Minneapolis, MN; NCS Pearson, Inc.

Morey, L. C. (1991). The Personality Assessment Inventory professional manual. Odessa, FL: Psychological Assessment Resources.

Morey, L. C. (2007). The Personality Assessment Inventory professional manual, second edition. Odessa, FL: Psychological Assessment Resources.

Morgan, C. D. & Murray, H. A. (1935). A method for investigating phantasies: The Thematic Apperception Test. *Aarchives of Neurology and Psychiatry, 34,* 289-306.

Naglieri, J., McNeish, T. & Bardos, A. (1991). *Draw-A-Person: Screening Procedure ofr Emotional Disturbance.* Austin, TX: ProEd.

Paulhus, D. L. (2002). Socially desirable responding: The evolution of a construct. In H. I. Braun and D. N. Jackson (Eds.), *The role of constructs in psychological and educational measurement,* pp. 49-69. Mahwah, NJ: Erlbaum.

Payne, A. F (1928).*Sentence completions.* New York: New York Guidance Clinic.

Reynolds, C. R. & Kamphaus, R. (1992). *Behavior Assessment System for Children.* Circle Pines, MN: American Guidance Service.

Reynolds, C. R., & Kamphaus, R. W. (2004). *Behavior Assessment System for Children, Second Edition manual*. Circles Pines, MN: AGS.

Riccio, C. A. & Rodriguez, O. L. (2007). Integration of psychological assessment approaches in school psychology. *Psychology in the Schools, 44,* 243-255.

Rorschach, H. (1921). *Psychodiagnostik.* Berne: Birchen.

Saklofske, D. H. & Eysenck, H. J. (1994). Extraversion-introversion. In V. S. Ramachandran (Ed.), Encyclopedia of Human Behavior (pp. 321-332). New York: Academic.

Saklofske, D. H. & Zeidner, M. (1995). *International handbook of personality and intelligence*. New York: Plenum

Schuerger, J. M. (1986). Personality assessment by objective tests. In R. B. Cattell & R. C. Johnson (Eds.), *Functional psychological testing: Principles and instruments* (pp. 260-287). New York: Brunner/Mazel.

Sattler, J. M. (2001). *Assessment of children: Cogntive applications (4th ed.).* San Diego, CA: Author.

Schmitz, P. G. (2004). On the alternative five-factor model: Structure and correlates. In R. M. Stelmack (Ed.) On the psychobiology of personality: Essays in honor of Marvin Zuckerman (pp. 65-87). Amsterdam: Elsevier.

Schneider, R. J., Hough, L. M., & Dunnette, M. D. (1995). Broadsided by broad traits: How to sink science in five dimensions or less. *Journal of Organizational Behavior, 17,* 639-655.

Schutte, N. S., Malouff, J. M., Hall, L. E., Haggerty, D. J., Cooper. J. T., Golden, C. J. et al. (1998). Development and validation of of a measure of emotional intelligence. *Personality and Individual Differences, 25,* 167-177.

Tellegen, A. (1982). *Brief manual of the Multidimensional Personality Questionnaire.* Unpublished manuscript, Univeristy of Minnesota.

Tellegen, A. (1995). *Multidimensional Personality Questionnaire.* Unpublished document, Univeristy of Minnesota

Smith, D.A (2002). Validity and values: Monetary and otherwise. *American Psychologist, 57,* 136-137.

Tellegen, A. (1982). *Brief manual of the Multidimensional Personality Questionnaire.* Unpublished manuscript, Univeristy of Minnesota.

Tellegen, A. (1995). *Multidimensional Personality Questionnaire.* Unpublished document, Univeristy of Minnesota

Thurstone, L. L. & Thurstone, T. (1930). A neurotic inventory. *Journal of Social Psychology, 1,* 3-30.

Woodsworth, R. S. (1919). Examination of emotional fitness for warfare. *Psychological Bulletin, 16,* 59-60.

Zuckerman, M., & Kuhlman, D. M. (2000). Personality and risk-taking: Common biosocial factors. *Journal of Personality, 68,* 999-1029.

Zuckerman, M. (2005). *Psychobiology of personality* (2nd rev. ed.*).* New York: Cambridge University Press.

Table 1. Objective vs. subjective assessments of personality – some key issues

	Objective Testing Perspective	Subjective Testing Perspective
Meaningfulness of self-reports	People often lack insight into their true personalities. Personality may be shaped by unconscious forces (psychoanalysis) or by situationally-specific implicit learning processes.	Self-reports are a class of behaviours that may usefully index latent personality traits. As Cattell (1973) pointed out, self-reports may be treated as behaviours whose meaning can be established through research (Q' data) rather than as veridical insights (Q data).
Role of response bias	Self-reports often reflect no more than trivial response styles (e.g., acquiescence), or deliberate impression management (e.g., faking). Techniques for assessment of response bias may themselves be open to manipulation.	Response bias may be assessed independently from latent traits. Furthermore, some 'biases' may be integral to personality and worth investigating as substantive traits (Paulhus, 2002).
Biological basis of personality as the basis for measurement	If personality is biologically based, it is unlikely that self-reports map directly onto the brain systems controlling traits. Research should work towards direct assessment of	Traits may be higher-level emergent personal qualities that are not isomorphic with any single brain system (Zuckerman, 1991). Thus, it is difficult to capture traits

	individual differences in neural functioning and their molecular-genetic sources.	in their entirety using biological indices. Specific biological theories also have only mixed support from empirical tests (Matthews & Gilliland, 2005).
Status of objective, implicit and projective tests	It is questionable whether subjective experience possesses the scaling properties necessary for quantitative measurement models (cf., Barrett, 2005). Tests based on objective behaviours may be intrinsically superior to subjective reports in supporting adequate measurement.	Historically, the reliability and validity of leading projective tests has been controversial. The new generation of implicit measures do not yet have the extensive nomological net of traits assessed by questionnaire. Such traits currently possess superior criterion, construct and consequential validity.