

# 60 Psychological Assessment

---

by

Mitchell K Byrne Ph.D, M.App Psych, BA (Hons)

Kylie Patterson Ph.D, M Psych

Courtney Rudd M Psych, B Psych (Hons)

Rosa Nolan M Psych, BA (Hons)

## AUTHOR INFORMATION

### [60.05] Authors

#### **Mitchell Byrne**

Mitchell Byrne PhD is an Associate Professor at the University of Wollongong. He is an endorsed Clinical and Forensic Psychologist and in addition to teaching and research, has an active clinical/forensic practice. He has a particular interest in forensic assessment, trauma, and autism spectrum disorders.

##### *Contact details*

School of Psychology  
University of Wollongong  
Northfields Avenue  
Wollongong, NSW 2522  
Email: mbyrne@uow.edu.au

#### **Kylie Patterson**

Kylie Patterson PhD is a Forensic Psychologist employed by Juvenile Justice in NSW. She supervises Forensic Masters students and is a casual lecturer at the University of Wollongong.

##### *Contact details*

Wollongong Juvenile Justice Community Office  
Level 2, 80 Crown Street  
Wollongong, NSW 2500  
Email: Kylie.patterson@justice.nsw.gov.au

#### **Courtney Rudd**

Courtney Rudd M Psych is a recent Clinical Masters graduate of the University of Wollongong. She has a particular interest in working with vulnerable populations.

##### *Contact details*

School of Psychology  
University of Wollongong  
Northfields Avenue  
Wollongong, NSW 2522  
Email: clr171@uowmail.edu.au

#### **Rosa Nolan**

Rosa Nolan M Psych completed research at the University of Wollongong on neurofeedback and state regulation training in AD/HD. She studied developmental neuropsychology before completing her clinical psychology training. Her interests include trauma, psychosis and neurodevelopment.

*Contact details*

School of Psychology

University of Wollongong

Northfields Avenue

Wollongong, NSW 2522

Email: rosa.n@hotmail.co.uk

**[The next text page is 60-101]**



# CONTENTS

## AUTHOR INFORMATION

Authors .....	[60.05]
---------------	---------

## PART 1: INTRODUCTION

The Nature of Psychological Testing and Assessment .....	[60.10]
What are “Psychological Assessments”? .....	[60.20]
How Psychologists Use Tests When Undertaking a Psychological Assessment .....	[60.30]
Evaluating of Psychological Testing	
Introduction .....	[60.130]
Features of the Test .....	[60.140]
Test Selection and Administration .....	[60.160]
Summary .....	[60.200]

## PART 2: KEY PSYCHOLOGICAL CONCEPTS UNDERLYING PSYCHOLOGICAL ASSESSMENT AND TESTING IN FORENSIC ENVIRONMENTS

Introduction .....	[60.1250]
Mental Health .....	[60.1260]
Diagnosis .....	[60.1270]
Motivation	
Introduction .....	[60.1370]
Defining Motivation .....	[60.1380]
Intrinsic and Extrinsic Motivation .....	[60.1390]
Motivation and Forensic Psychology .....	[60.1400]
Motivation and Mental Health .....	[60.1500]
Intelligence	
Defining Intelligence .....	[60.1600]
Intelligence Testing .....	[60.1610]
Personality	
Personality Traits .....	[60.1710]
Personality Disorders .....	[60.1720]
Substances and Addiction	
Introduction .....	[60.1810]
Physiology .....	[60.1820]
Addiction and mental health .....	[60.1830]
Psychosis	
Psychosis and Schizophrenia .....	[60.1930]
Psychosis and the Law .....	[60.1940]
Stress-Vulnerability and Trauma .....	[60.1950]
Early Intervention .....	[60.1960]
Recovery Model .....	[60.1970]

## PART 3: HOW PSYCHOLOGICAL ISSUES AND ASSESSMENTS CONTRIBUTE TO JUDICIAL DECISION MAKING ACROSS JURISDICTIONS

Introduction .....	[60.3010]
Assessment across different jurisdictions	
Criminal Law .....	[60.3110]
Family Law .....	[60.3120]
Children’s Court .....	[60.3130]
Civil Litigation .....	[60.3140]
Coroner’s Court .....	[60.3150]
Tribunals .....	[60.3160]

## PART 4: DIFFERENT CATEGORIES AND METHODS OF TESTING

Introduction .....	[60.4210]
Categories of psychological tests .....	[60.4220]

Methods of assessment within psychological tests .....	[60.4230]
<b>PART 5: CATEGORIES OF PSYCHOLOGICAL ASSESSMENT AND COMMONLY USED TESTS WITHIN CATEGORIES</b>	
Introduction .....	[60.5210]
Tests of intellectual functioning	
Individual and group administered .....	[60.5220]
The Wechsler Scales .....	[60.5230]
Woodcock Johnson .....	[60.5240]
Stanford-Binet Intelligence Scales .....	[60.5250]
Differential Ability Scales (DAS) .....	[60.5260]
Kaufman Intelligence Scales .....	[60.5270]
Adaptive Behaviour .....	[60.5280]
Culturally Robust Intelligence Tests .....	[60.5290]
Tests of Neuropsychological Impairment	
Neuropsychological Impairment .....	[60.5410]
Tests of Personality	
Introduction .....	[60.6510]
Multivariate Trait Approach .....	[60.6520]
Empirical Approach .....	[60.6530]
Personality Disorder Diagnostic Assessment .....	[60.60540]
Tests of Neurodevelopmental Disorders	
Introduction .....	[60.7010]
Autism Spectrum Disorder	
Introduction .....	[60.7110]
Diagnosis .....	[60.7120]
Assessment .....	[60.7130]
Screening Measures .....	[60.7140]
Attention Deficit Hyperactivity Disorder (ADHD)	
Introduction .....	[60.7250]
Diagnosing AD/HD .....	[60.7260]
Assessment .....	[60.7270]
Specific Learning Disorder .....	[60.7370]
Language Disorder .....	[60.7470]
Tests of Mental Health	
Introduction .....	[60.8010]
Screening Measures .....	[60.8020]
Mental Health Disorders – Psychotic	
Introduction .....	[60.8120]
Assessment .....	[60.8130]
Bipolar Disorder I and II .....	[60.8140]
The Assessment of Mental Health Disorders – Non psychotic	
Introduction .....	[60.8240]
Mood Disorders	
Introduction .....	[60.8310]
Major Depressive Disorder .....	[60.8320]
Suicide & Self Harm .....	[60.8330]
Anxiety Disorders .....	[60.8340]
Obsessive Compulsive Disorder .....	[60.8350]
Tests of Trauma & Stressor Related Disorders	
Introduction .....	[60.9010]
Diagnosis .....	[60.9020]
Tests of Alcohol and Drug Use	
Introduction .....	[60.9510]
Assessment .....	[60.9520]
Tests of Eating Disorders	

Introduction .....	[60.10010]
Anorexia Nervosa .....	[60.10020]
Bulimia Nervosa .....	[60.10030]
Binge Eating Disorder .....	[60.10040]
Assessment of Eating Disorders .....	[60.10050]
Tests of Disruptive, Impulse Control and Conduct Disorders	
Introduction .....	[60.10510]
Oppositional Defiant Disorder (ODD) .....	[60.10520]
Conduct Disorder (CD) .....	[60.10530]
Pyromania .....	[60.10540]
Kleptomania .....	[60.10550]
Assessment of Disruptive, Impulse-Control and Conduct Disorder .....	[60.10560]
Tests of Somatic Disorders and Malingering Assessment	
Introduction .....	[60.11010]
Assessment .....	[60.11020]
Response Styles .....	[60.11030]
Malingering and Simulated Adjustment/ Defensiveness	
Introduction .....	[60.11210]
Psychological Tests to Identify Malingering and Deception .....	[60.11220]
Screening Measures .....	[60.11230]
Risk Assessment	
Introduction .....	[60.11710]
Tests of Risk of Violence, Sexual Offending and Recidivism	
Introduction .....	[60.11810]
Juvenile	
Violence Risk Assessments .....	[60.11860]
Juvenile Sex Offender Risk Assessments .....	[60.11870]
Risk of Recidivism .....	[60.11880]
Adult	
Introduction .....	[60.11930]
Risk of Recidivism .....	[60.11940]
Risk of Violence/Sexual Offending .....	[60.11950]
Tests of Attachment	
Introduction .....	[60.12510]
Assessment .....	[60.12520]
Tests of Anger and Emotional Intelligence	
Introduction .....	[60.13010]
Assessment .....	[60.13020]
Anger .....	[60.13030]
<b>BIBLIOGRAPHY</b>	
Bibliography .....	[60.15010]

## PART 1: INTRODUCTION

### [60.10] The Nature of Psychological Testing and Assessment

A psychological test is a sample of behaviour used to make inferences about an individual within a specific context, or to elicit signs of an underlying disposition. It is generally acknowledged that among the various mental health professionals, psychologists in particular are best equipped to administer and interpret psychological tests. Professional training programs are mandated to attend to core competencies in psychological testing and assessment in order to meet accreditation standards set by the Australian Psychology Accreditation Council, while new graduates pursuing registration through alternative training schemes must successfully sit the National Psychology Examination overseen by the Psychology Board of Australia on behalf of the Australian Health Practitioner Registration Agency. In general, assessment competencies are exclusive to psychologists (Ready & Veague, 2014).

The principles underlying the certification of psychologists, and the tests that they use, relate broadly to the concept of Psychology as a science. This is grounded in the philosophical approach of Positivism: a system recognizing only that which can be scientifically verified or which is capable of logical or mathematical proof (Beck et al, 2014). In basic terms, psychological assessment should be undertaken in the context of a coherent scientifically justifiable theory as to what question(s) the assessment will answer, how this relates to the issue or problem at hand, and using tests and measures to assist in that assessment which have a data pedigree lending confidence to the conclusions drawn from the assessment, referred to as the “psychometric properties” of the test.

Thus, when evaluating the probative value of psychological assessments, the legal practitioner should ascertain:

- whether the psychologist used a theoretically sound rationale to develop the questions to be addressed in the assessment;
- if s/he chose the right sort of test or tests to answer those questions; and,
- whether the specific test or tests chosen has been proven to provide reliable data relevant to the question at hand.

Each of these questions will be addressed in the remainder of this chapter, inclusive of the limits of confidence that may be applied to each in relation to general issues of psychological relevance as well as the specifics of individual assessment tools (tests). However, it is first necessary to discuss what is meant by “psychological assessment” and how psychologists conduct assessments to both determine the nature of the problem and derive strategies to address the problem.

### [60.20] What are “Psychological Assessments”?

Psychological testing and psychological assessment are not synonymous. A psychological test is a systematic procedure for observing an aspect of a person’s behaviour or performance, or the general characteristics of the person (Anastasi & Urbina, 1997). For example, a psychological test can describe a person’s aptitude for a job, or their general level of intelligence. Psychological tests are based on psychological theories that consider the importance of individual differences and seek to explain them. Psychological tests are different from other tests or quizzes in that they are standardised, meaning that the materials used and the scoring methods are always the same so that the same test can be given at different times and in different locations. The same (or nearly the same) result should be able to be produced irrespective of the psychologist who is administering the test, providing the psychologist

conforms to the testing procedures for the particular test and the person being tested is able to perform tasks with an equal level of competence on each occasion that testing occurs. This enables the development of norms, or what scores mean with reference to the performance of a large group of people on the same test. This is referred to as the “standardisation sample” or a “standardised score” (Groth-Marnet, 2016).

As noted above, a test may provide a sample of behaviour or a sign of an underlying disposition. Where the test is a sample of behaviour (eg working memory ability), interpretation of test performance is in terms of “criterion referencing” (eg average length of working memory), while tests providing signs of a disposition (eg anxiety) are interpreted using “norm referencing” strategies (eg proportion of the population endorsing anxiety items). Criterion referencing means that the standard of behaviour expected is predetermined and the test-taker’s performance is judged against that standard. However, in norm referenced interpretations, the results of a representative group of people similar to the test-taker are used to assess the extent to which the test-taker’s responses are similar or different (Shum et al, 2017).

Psychological assessment, on the other hand, involves a process by which the psychologist collects a variety of information from different sources in order to answer a question. Part of that information may include psychological tests; however other sources of information may be included, such as interview data, information relevant to the person’s age, gender, culture or ethnicity, medical information, personal history, and the observations of others. In general, the more complex the question, the less adequate is a single test in answering the question. For that reason, in addition to multiple sources of information, psychological assessments often include a multitude of psychological tests, which in combination may provide a clearer picture of what is going on.

### **[60.30] How Psychologists Use Tests When Undertaking a Psychological Assessment**

As noted above, the more complex the legal question or problem, the less adequate a single test to answer that question. To some extent the reverse may be true: a simple question may be answered by a single or few psychological tests. For instance, if the question is “*does this person have an intellectual disability?*” it may be sufficient to administer a valid and reliable standard measure of intellectual functioning. However, even when the question is constrained as such, there are a multitude of variables that the psychologist must take into consideration when interpreting the person’s test results. For example: was the test appropriate for the individual’s cultural group; did the person engage fully in the testing process and with effort; were there any physical or health factors affecting their performance; and is there reason to believe that the individual would benefit from under-performing or malingering illness. Additionally, the psychologist must consider whether the person’s current performance is consistent with, or incongruent with, past intellectual/academic performances, life achievements (such as occupations), general skills in living (referred to as “adaptive behaviour skills”), and the presence or absence of neurological injuries. In short, a test is rarely sufficient in and of itself to answer a question and forms but a part of a comprehensive psychological assessment.

That being said, psychological tests are increasingly (and appropriately) used in the provision of expert opinions, with the majority of psychologists deeming psychological testing as either necessary or recommended when undertaking a psychological assessment (Borum & Grisso, 1995). Psychological tests are designed to measure unobserved constructs, referred to as latent variables. Given that these constructs are “unobserved”, the test is used to confirm or challenge hypotheses drawn from other information sources, such as observed behaviour, interview data, or the opinion of others. The use of tests ensures a more holistic assessment by the psychologist, and the interpretation of the test result is undertaken in the context of those other information sources. The usefulness of the test to inform assessment and opinion is predicated

on the appropriateness of the test to the issue at hand, the psychometric properties of the test chosen, and the quality of the testing process undertaken by the psychologist.

## *Evaluating of Psychological Testing*

### **[60.130] Introduction**

Both the features of the test used and the process for test selection and administration are critical in determining the value of the data provided by any given test.

### **[60.140] Features of the Test**

The most critical aspect of a test is its validity - does it measure what it is intended to measure. A psychological test cannot be said to be valid in an abstract sense but must be valid in a particular context for a specific group of people (Messick, 1995). To be valid, a test must also be reliable (see below). Establishing validity is difficult because tests are designed to assess latent variables – unobservable abstract concepts that have no tangible reality, so that their existence must be inferred. Take for example depression: it can't be seen, smelt, touched or heard, yet we know of its existence through a multitude of tests which consistently infer its existence. A further difficulty in validity is that conceptualisation of constructs changes over time, requiring a continual process of validation and test refinement. This can be seen through version numbers of tests. Any estimate of test validity is concerned with the relationship between the test and some external, independently observed event. There are three main methods of establishing validity: content-related; criterion-related; and construct-related.

*Content validity* refers to the representativeness and relevance of the test to the construct being measured. This is usually ascertained by evaluating test items against theoretical conceptualisations of the variable or construct being measured. This is generally appraised by experts in the area and reported in technical manuals as the level of agreement between such experts on the test items. A related concept to content validity is *face* validity: the extent to which the test looks like it relates to the construct under investigation, as appraised by test users, policy makers and untrained personnel

*Criterion-validity* is also known as empirical or predictive validity and is determined by comparing test scores with some sort of performance on an outside measure. For example, does a test of intelligence relate to academic performance? Another way of looking at criterion-validity is termed *concurrent validity*. This is commonly used in cross-sectional psychological assessments as it is simpler and less time consuming. In concurrent validity, two tests may be administered which are purported to measure the same construct. Concurrent validity is more useful when assessing a person's current status, while predictive validity is more helpful when considering what may occur at some point in the future. Tests will often report validity coefficients, especially when predicting performance or behaviour. Because there are numerous factors that may influence the performance of a given behaviour, validity coefficients rarely exceed .40. As a guide, coefficients of .35 and above suggest that the test is very beneficial to answer the specific question, while coefficients between .21 and .35 indicate that the test is likely to be useful. Caution should be used with coefficients between .11 and .20, while coefficients below .11 indicate that the test is unlikely to be useful.

*Construct-validity* assesses the extent to which the test measures a theoretical construct or trait. In this case, the test results (the construct) should in theory relate in a predictable way to other variables. For example, a person scoring high on a depression measure should have a higher incidence of social withdrawal, or a lower incidence of social engagement. Construct validity is often reported as a consequence of intervention, such that an intervention based on the test score results in a change in scores post-intervention. Within construct-validity are the concepts

of sensitivity and specificity. Sensitivity refers to how well a test identifies a person with the particular variable under consideration (percent of true positives), whereas specificity refers to how well it excludes people without the variable under consideration (percent of true negatives).

The other critical aspect of test quality is *reliability*. Reliability refers to the degree of stability, consistency, predictability and accuracy of a test. In the simplest of terms, reliability refers to the extent to which the scores obtained by a person tested on two separate occasions, in the absence of intervention, remain the same. Underlying reliability is the range of error or error measurement. This accounts for random fluctuation in scores attributable to a myriad of causes, such as fatigue, illness, effort and so forth. The *standard error of measurement* is used to calculate the *confidence interval* of a given score. Essentially, the standard error of measurement and confidence interval indicate the range of possible true scores (above and below the obtained score) that may reflect the test-takers performance. All tests should report these statistics. In general, the smaller the standard error of measurement, the more accurate the test.

There are four areas of reliability within a test: the extent to which the test produces consistent results (test-retest reliability); the relative accuracy of a test at a given time (alternate forms of the same test); the internal consistency of items (split half reliability); and the degree of agreement between two different raters (inter-scoring reliability). Test manuals should be able to report these statistics to lend credence to the reliability of the test.

Finally, an evaluation of a test should consider the adequacy of the norms used by that test. The norms reflect the distribution of scores within a standardisation sample and the interpretation of an individual's scores depends upon how they relate to the standardisation sample. The more dissimilar the person is to the standardisation sample, the less useful the test for assessment (Groth-Marnat, 2016).

Interpretation of reported reliability can vary according to the nature of the test and the sample. However, it is generally agreed that a reliability (*r*) coefficient of .90 and up is considered excellent, .8 to .89 is good, and .7 to .79 is adequate. Reliability below .7 indicates that the test may have limited applicability. In terms of prediction, a statistic often cited is "Area Under Curve" or "ROC Curves". The accuracy of the test depends on how well the test separates those with and without the condition. Coefficients between .90 and 1.0 are considered excellent. Coefficients of .8 to .9 are good, and .7 to .8 fair. Between .6 and .7 is seen as poor while coefficients below .6 are seen as unreliable.

## [60.160] Test Selection and Administration

Clearly, the tests selected to inform a psychological assessment should be demonstrably reliable, valid and appropriately standardised for the cohort with which the person assessed belongs. However, there are additional issues external to the quality of the test that will influence its usefulness. One such factor, which relates to the standardisation sample, is cultural relevance. Generally speaking, most psychological tests have been constructed within a Western English speaking context and have embedded within them cultural understandings and language conventions appropriate to that context. When selecting a test, the psychologist should ensure that the test has included a standardisation sample relevant to the culture of the assessee. The academic journals are replete with articles whereby a modification or standardisation has been applied to an extant test to meet the needs of a particular nationality or cultural group. For example, the most reliable and valid assessment of intellectual functioning, the Wechsler Adult Intelligence Scale IV was developed in the United States but includes an Australian version. Where used with other nationalities, the psychologist should ensure that the adaptation or version for their particular cohort is used.

Related to the above, the test administrator should ensure that the language used in the test, especially in self-report assessments, is comprehensible to the assessee. English language

proficiency, reading skills and level of intellectual functioning will all influence the accuracy of results obtained. Further to this, many tests assume an absence of physical or sensory disability. Motor coordination problems, visual deficits or hearing loss can all influence test results and should be evaluated prior to test selection. A range of suitable tests for persons with impairments have been developed and should be selected where appropriate.

The assessor also needs to consider the context in which the assessment is taking place. This is particularly relevant in situations where the judiciary is involved and the results of the assessment may be perceived by the assessee as either delivering a benefit or a cost to that individual. In some circumstances the impact of context can be ameliorated by the way in which the test is introduced to the individual by the assessor. In other cases, the norms upon which interpretation takes place may include a standardisation sample in a similar situation as the individual being assessed. Finally, some tests contain scales that directly assess deliberate or accidental response distortion, enabling the assessor to either select another test or report the response distortion as part of the assessment process and conclusions. A good example of this is malingering illness, where if detected within a more broad assessment, such as the Personality Assessment Inventory-2, can be further investigated with a more specific assessment of malingering, such as the Miller Forensic Assessment of Symptoms Test.

Finally, the performance of an individual on any given test will be influenced by factors external to the assessee and the specific test used. The assessment environment can be crucial, especially for tests that take a substantial time to complete. Noise, comfort and privacy can all play a factor in influencing results. Some tests are suitable for the assessee to take home or complete independently in a location other than the assessor's office. However, other tests risk being distorted by the influence of third parties, or invalidated by confusion in the absence of an examiner to assist with test understanding. Finally, the assessing psychologist can also influence test outcomes, either through a poor rapport, internal bias (such as when the test is administered by a former provider of therapy), or as a result of non-standardised administration or errors in scoring and interpretation.

## [60.200] Summary

Psychological Assessment is a complex process of answering questions where psychological tests are often used as one method of collecting relevant data. When evaluating the value of a psychological assessment in contributing to judicial decision making, not only the quality of the tools (tests) used should be considered, but also the process of test selection and administration. While judicial officers are not expected to be able to assess and report on the psychometric properties of a test, whether it was appropriate to the case or the individual assessed, and whether or not process and procedure of testing was adequate, they should expect that the psychologist providing the assessment, either within the report or under cross examination, is able to answer these questions.

The following parts are designed to assist legal professionals in determining the quality of psychological assessments with afore mentioned issues in mind:

- Part 2 provides a broad summary of how psychological issues affect behaviour and how this may bring an individual into contact with the judicial system. A summary of the types of questions that psychological assessments can answer is provided.
- Part 3 summarises the applications of psychological knowledge across a range of legal jurisdictions and what may be contributed to the judicial decision making process.
- Part 4 outlines the process of psychological assessment and the types of information psychologists use in undertaking an assessment. We describe how we bring that information together and manage conflicting data.
- Part 5 describes categories of tests used in psychological assessments. This final

section of the chapter includes some of the most commonly used tests within 12 categories of mental health and psychological investigation. Basic information on reliability and validity are provided as a reference point for evaluating the strength of the test. The 12 categories of tests reviewed are as follows:

- Tests of Intellectual Functioning
- Tests of Neuropsychological Impairment
- Tests of Personality
- Tests of Neurodevelopmental Disorders
- Tests of Mental Health (Psychotic and Non-Psychotic)
- Tests of Trauma and Stressor Related Disorders
- Tests of Alcohol and Drug Use
- Tests of Eating Disorders
- Tests of Disruptive, Impulse Control, and Conduct Disorders
- Tests of Somatic Disorders and Malingering Assessment
- Risk Assessment
- Tests of Attachment.

## PART 2: KEY PSYCHOLOGICAL CONCEPTS UNDERLYING PSYCHOLOGICAL ASSESSMENT AND TESTING IN FORENSIC ENVIRONMENTS

### [60.1250] Introduction

An understanding of a number of core psychological concepts is integral to the interpretation of psychological assessments (and tests) as they relate to judicial processes. Neither psychology nor psychiatry is an exact science such as physics or chemistry, where elements and processes behave the same way over repeated observations. Psychological concepts are used to help explain behaviour and they develop a certain verisimilitude through their ability to repeatedly predict how an individual might act, or what might be done to change their behaviour, (where behaviour includes processes of thought and feeling). Therefore, what psychology means and the issues associated with the terms *Mental Health, Diagnosis, Motivation, Intelligence, Personality, Substance Abuse, and Psychosis* is important to an understanding of specific assessment processes employed by psychologists.

### [60.1260] Mental Health

*Mental health is defined as a state of well-being in which every individual realizes his or her own potential, can cope with the normal stresses of life, can work productively and fruitfully, and is able to make a contribution to her or his community.*

- World Health Organisation, August 2014

### Categorical and Dimensional

Psychological health can be thought of as being categorical or dimensional. Using a categorical medical model, psychological health is defined as “the absence of illness and ill health”. States of psychological illness are categorically distinct from psychological health and this is inferred from the cognitive, behavioural and emotional patterns. Explanation and treatment is the goal in a medical model (American Psychiatric Association, 2013).

A dimensional approach, on the other hand, fails to distinguish qualitatively between health and illness, considering various psychological phenomena (intelligence, personality, motivation) as being distributed in the population in the same way as other biological traits, such as height. Most people demonstrate moderate aspects of most traits, with fewer people showing extremely little or a great deal of any given trait. Those for whom a trait is problematic and interferes with their lives are presumed to be in these extreme ranges and are captured by the categories in diagnostic manuals. The dimensional approach is more apparent in psychological literature than diagnostic manuals.

### [60.1270] Diagnosis

Diagnostic classification in psychology is prototypical: a shorthand for symptoms that statistically cluster together. This allows people with very different patterns of experience or symptoms to be given the same diagnosis, and for that diagnosis to be relevant if the symptom expression changes over time. A concern with this system of classification is that psychological diagnoses are descriptions of symptoms rather than explanations of their cause, as is the case in medical diagnoses. Nonetheless, classification serves the purpose of providing a language for researchers and practitioners that allows the development of a body of literature including interventions and epidemiology.

In Australia, the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5: American Psychiatric Association, 2013) is the most commonly used system for diagnosis in

Australian Courts. It is noted that the International Statistical Classification of Diseases and Related Health Problems (ICD), currently in its tenth revision (with the 11<sup>th</sup> due for release in 2018), is also used in some jurisdictions and is commonly referred to in health services. The ICD is a medical classification list by the World Health Organization (WHO) generally used for classification outside of the United States, and particularly in Europe. However, as noted, the DSM-5 is more commonly referred to in Australia.

The DSM-5 (and its predecessors) has attracted a number of controversies and criticisms, including but not limited to diagnostic accuracy, cross cultural applicability, and pathologisation of otherwise normal expressions of human behaviour and emotion. Further, the DSM-5 does not provide guidance on treatment or intervention with any disorder. The text provides a “Cautionary Statement for Forensic Use of DSM-5” (page 25) which explicitly states that the DSM-5 “is primarily designed to assist clinicians in conducting clinical assessment, case formulation, and treatment planning” and is not designed to meet the “technical needs of courts and legal professionals”. However, the authors assert that the appropriate use of the DSM-5 can assist judicial determinations where a diagnosis is relevant to the understanding of past behaviour, or to the prediction of future behaviour. However, the DSM-5 notes that a DSM-5 diagnosis *does not* imply that an individual meets legal criteria for the presence of a mental disorder and for such a determination to be made, additional information is required beyond that of a DSM-5 diagnosis, such as information about functional capacity/impairment, and how these impairments affect the ability of the individual to behave in a specified situation. The authors state: “It is precisely because impairments, abilities, and disabilities vary widely within each diagnostic category that assignment of a particular diagnosis does not imply a specific level of impairment or disability” (APA, 2013, p 25). It is also noted that even when diminished control over behaviour is a feature of a specific disorder, this does not calibrate the extent to which a specific individual was able to exert control over their behaviour at a particular point in time.

## Motivation

### [60.1370] Introduction

As noted earlier, a number of variables may affect the outcome of testing and assessment, including but not limited to, cognitive ability of the assessee, cultural appropriateness of the test, assessee fatigue, and the competency of the assessor. However, the most common influence on test outcomes is assessee motivation. This encompasses both effort to perform to the best of the assessee’s abilities, as well as the potential to deliberately distort the test results in a given direction (see malingering below). For this reason, it is important for the assessor to evaluate assessee effort and where necessary, comment upon assessee motivation. It is therefore important that this chapter address the theoretical constructs associated with “motivation”.

### [60.1380] Defining Motivation

“To be motivated means *to be moved* to do something” (Ryan & Deci 2000). In psychology, motivation is understood as a fluctuating state of intent toward performing behaviours. Any behaviour can have multiple motivations, possibly conflicting, and each can vary according to thoughts, sensations, affect, and the external environment.

The structure of the causal relationship between a behaviour and the state of a person varies between theoretical and epistemological frameworks. For example, behavioural theories attribute causality to the external stimuli, holding that people perform behaviours in response to a stimulus and its associated value or meaning. In contrast, positive psychology gave rise to Maslow’s Hierarchy of Needs that attributes causality to a state internal to the person to satiate a need.

### **[60.1390] Intrinsic and Extrinsic Motivation**

One accepted distinction drawn from Self Determination Theory is that between intrinsic and extrinsic motivation (Deci & Ryan, 1985). Intrinsic motivation arises from within a person, without reference to the outcome of a behaviour or act. The act itself is rewarding or pleasurable; there is an autonomous valuing, interest or enjoyment to be had. Intrinsic motivation is associated with an internal sense of control or effectiveness in the environment, with the ability to affect the external world. Further, it is associated with mastery of skills and self-efficacy (one's belief in their own ability to complete a task).

Extrinsic motivation is motivation that takes as its referent the outcome of the behaviour rather than the behaviour alone. Working for income can be understood in these terms. The quality of extrinsic motivation can vary from resentment to willingness. At the point of willingness, the distinction between extrinsic and intrinsic motivation may be blurred. That is, when someone accepts and values the reason for the outcome of the behaviour, they are both extrinsically motivated by the outcome of the behaviour, while intrinsically motivated to engage in the behaviour because it is ego-syntonic (that is, behaviours, values, and feelings that are in harmony with or acceptable to one's ideal self-image).

Intrinsic motivation is understood to be more effective long-term than extrinsic, as well as more generalizable across different contexts (Cerasoli, Nicklin & Ford, 2014). It is also more Maslow closely associated with personal values, identity and empowerment (Benabou & Tirole, 2003) and with better performance on a task (Ryan & Deci, 2000). Extrinsic motivators (incentives), on the other hand, are limited to context and don't lead to behaviour change in multiple domains.

### **[60.1400] Motivation and Forensic Psychology**

Psychology considers motivation for a single act to be complexly determined and refers to multiple concurrent layers of causality. Generally speaking people seek to satisfy a series of hierarchically organised needs: physiological (hunger, thirst, sleep), safety (security, shelter, health), social (love, friendship), esteem (recognition, achievement) and self-actualisation (Maslow, 1943). The strategies that people develop to meet these needs may be complex and conflicting and develop from a lifetime of experiences of pleasure, displeasure, success and failure (Silvetti et al., 2013). Motivation in forensic psychology is most pertinent in relation to malingering and explaining the reasons for attempting to deceive relies on a thorough history and full assessment (see also Chapters 61, 63A and 68).

### **[60.1500] Motivation and Mental Health**

Motivation is different in multiple disorders including Attention Deficit/Hyperactivity Disorder (AD/HD), Oppositional Defiance Disorder (ODD) and Autism Spectrum Disorder (ASD). The overarching desires to live fulfilling lives, meet biological needs and enjoy pleasure are still present but the learning processes and neurobiology of pleasure that underpin what is subjectively valuable differ. For example, people with AD/HD have differential dopaminergic responses to pleasurable stimuli and difficulties with inhibition control that make sustaining attention on repetitive tasks difficult, including reading and exams. This in turn makes literacy a weakness for many people with AD/HD and contributes to a sense of failure or poor self-worth (Sonuga-Barke et al., 2016).

Motivation is also critically affected in Major Depressive Disorder; a key symptom is low motivation. Lack of energy, feelings of sadness and numbness, and cognitive difficulties with planning coalesce into a functional inability to be motivated and to execute goal-directed behaviours. In contrast, in the Anxiety Disorders people are highly motivated by fear to avoid

certain stimuli, such as objects (phobias), memories (Post Traumatic Stress Disorder: PTSD), ambiguous prospects (Generalised Anxiety Disorder) or intrusive negative thoughts (Obsessive Compulsive Disorder: OCD).

## *Intelligence*

### **[60.1600] Defining Intelligence**

Most helpfully referred to as a “complex of diverse and numerous components”, intelligence is a “multifaceted, polymorphous” phenomenon defined by the adaptive or functional value of a person’s behaviour (Wechsler, 1975). A behaviour is intelligent if it is subjectively successful according to the intended goal of the person enacting it. Intelligence is not a unitary, single cognitive factor but a result of the operation of many abilities. Success or value will be culturally defined and intelligence will therefore be judged by a person’s capacity to meet their needs and those of the people and society with which they live.

### **[60.1610] Intelligence Testing**

Despite the multifaceted nature of intelligence psychology commonly refers to a single intelligence factor, “g”, that reflects the aggregate of their adaptive, functional and cognitive skills. This general intelligence factor is distributed normally through the population and strongly related to academic and occupational success. Please see Chapters 68, Neuropsychology by Crowe, and 63A Intellectual Disability by Hayes for more on Intelligence Testing, including common tests of general intelligence and specific cognitive skills.

## *Personality*

### **[60.1710] Personality Traits**

Personality in psychology refers to enduring dispositions or traits in people that organise their behaviour and motivations. Such dispositions inform all other psychological phenomena. They are themselves dimensional and distributed throughout a population; each person will be more or less extreme in the degree to which they experience each trait. A trait based model of personality includes both healthy, adaptive, and resilient personality traits, with a dimensional perspective leading to the polar opposite, where dysfunction and distress becomes evident.

Contemporary research on personality has focused on the Five-Factor Model of Personality. Statistical analysis has extracted the five dimensions on which people vary: neuroticism (tendency to experience unstable emotionality); extraversion (sociability and emotional expressiveness); openness to experience; agreeableness (co-cooperativeness); and conscientiousness, thoughtfulness and self-discipline; (McCrae & Costa, 2008). Each trait is understood to develop in dynamic relationship to the external environment and to genetics. Genetics are thought to explain 50% of the variance in each trait (Carr & McNulty, 2016, p 52).

Personality traits may change across the lifespan, however they tend to show a relative consistency when compared with specific behaviours. When behaviours aggregate across time and circumstance, this is reflective of a “trait”. A trait may make a person more susceptible to *symptoms*, such as depressivity or the likelihood of responding to a challenging event with depression, without connoting that the person has a mood disorder or that all adverse events will promulgate a depressive episode. Importantly, both symptoms and traits are amenable to intervention.

## **[60.1720] Personality Disorders**

Personality traits or dispositions are fundamental characteristics of all of us which give rise to the expectation of how a person will perceive and act upon a given event or circumstance. A personality disorder is qualitatively and quantitatively different from a disposition. The DSM-5 defines a personality disorder as “an enduring pattern of inner experience and behaviour that deviates markedly from the expectations of the individual’s culture, is pervasive and inflexible, has an onset in adolescence or early adulthood, is stable over time, and leads to distress or impairment” (DSM-5, p. 645). The key to determining that a person’s personality is “disordered” is the impact that their perceptions and behaviours have upon them and others, and the pervasiveness of their behaviour. Personality Disorders are often severe, impacting all areas of a person’s life including relationships, employment, cognition, impulse control and parenting. Relationships are often intense, confusing and volatile, and difficulty maintaining employment is common. While recovery from personality disorders may be possible, with several evidence-based treatments reporting clinically significant improvements in affect and behaviour, these treatments usually involve intensive therapy in multiple settings for a period of at least two years, necessitate a commitment by the disordered individual to treatment, and generally have limited availability (due to their expense and the level of training that the clinician requires to deliver the treatment).

### **Integrating the Five-Factor Model of Personality with Personality Disorders**

In DSM-5, an Alternative Model for Personality Disorders is included to aid research while the diagnostic criteria from the previous edition remained largely unchanged. The Alternative Model represents an attempt to integrate the Five-Factor Model of Personality with the diagnostic criteria of personality disorders. The ten disorders were organised into three clusters based on shared Five Factor Model factors: Cluster A (schizoid, schizotypal, paranoid) share extreme introversion; Cluster B (borderline, antisocial, narcissistic, histrionic) share extreme extraversion and low agreeableness; and Cluster C (avoidant, dependent, obsessive-compulsive) share high neuroticism (Carr & McNulty, 2016). At the time of writing, the “Alternative DSM-5 Model for Personality Disorders” is still under investigation and a significant body of research evidence is yet to be established.

## *Substances and Addiction*

### **[60.1810] Introduction**

Addiction is a “primary and chronic disease of brain reward, motivation, memory and related circuitry.” Dysfunction in these circuits leads to the behaviours, cognitions and physiological symptoms characteristic of addiction including: inability to abstain, loss of behavioural control, craving and inability to recognise deleterious effects (American Society of Addiction Medicine, 2011). The DSM defines numerous Substance Use Disorders as well as disorders related to addiction to repetitive behaviours which take on rewarding properties (APA, 2013). All substances taken in excess produce neurochemical reward. Addiction ensues when engaging in one substance or behaviour is to the neglect of other potentially or previously rewarding activities, and despite significant substance related problems.

### **[60.1820] Physiology**

The physiology of addiction is a pattern of tolerance and withdrawal. Tolerance is the need for gradually greater amounts of the drug to achieve intoxication or desired effect and a parallel diminished effect if using the same amount of the substance. Withdrawal is a set of physiological responses (withdrawal syndrome) when the substance use is ceased sometimes indicated by the continued use of the substance to relieve these symptoms. The neurological

and physiological mechanisms underlying these remain changed even after cessation of substance use, making repeated relapses more likely and long-term treatment and management necessary.

### **[60.1830] Addiction and mental health**

A person who is addicted to a substance is likely to focus their energy on obtaining and using it, to experience instability and conflict in relationships, to neglect other areas of their life, to increase high risk behaviours either while under the influence, or to obtain the substance and to feel shame about their patterns of use.

## *Psychosis*

### **[60.1930] Psychosis and Schizophrenia**

The Schizophrenia Spectrum Disorders are a family of illnesses (including Schizophrenia, Delusional Disorder and Brief Psychotic Disorder) that involve states of psychosis. Psychosis refers to any state involving delusions, hallucinations or disorganised thinking and behaviour. Delusions are stubbornly maintained beliefs for which there is no evidence, while hallucinations are vivid perceptions without an external stimulus. Such states vary in duration, frequency, intensity, distress, cause and relapse or re-occurrence. In most psychotic disorders active symptoms are accompanied or preceded by instability in mood such as anxiety or depression. The shared aetiology and symptom profiles of the psychotic disorders suggest dimensional distribution of psychotic symptoms in the general population (DSM-5; Carr & McNulty, 2016).

Schizophrenia is a chronic condition usually involving numerous psychotic episodes over the course of a lifetime. Remission between episodes is often incomplete and experiences of comorbid anxiety and depression are common, especially in prodromal phase before an episode. The disease burden of schizophrenia is significant with higher likelihood of long-term unemployment, lower academic and employment success, homelessness and imprisonment. Treatment differs according to needs and phase but is as relevant during remission as during episodes. Commonly, treatment involves a multi-modal combination of anti-psychotic medications, psychological treatments (CBT, Social Skills Training, Interpersonal Therapy) and functional supports (employment rehabilitation).

### **[60.1940] Psychosis and the Law**

Psychosis is more common in people interacting with or encountering the justice system than in the general population and the prevalence of psychosis in inmates is 30 times that of non-incarcerated Australians. Eight percent of incarcerated men and fifteen percent of incarcerated women have a psychotic illness, and people with Schizophrenia are more likely to commit seriously violent crimes and homicide (Senate Select Committee on Mental Health, 2006).

### **[60.1950] Stress-Vulnerability and Trauma**

The Stress-Vulnerability (diathesis-stress) Model of aetiology applies to all mental illnesses including the psychotic disorders. The basic tenant of the Stress-Vulnerability model is that all individuals lie on a continuum of vulnerability for any given disorder, where some will be highly vulnerable to a disorder, while others may be highly resilient to the same disorder. The disorder is expressed (experienced) when a stressor or multiple stressors impact the individual. For those with a high vulnerability to the disorder, fewer or less intense stressors are required to activate symptoms, while those resilient to the disorder can endure multiple and intense

stressors before symptoms of illness are activated. A psychotic illness is expected to emerge when people who have a neurobiological vulnerability to psychosis experience adverse life events causing psychological stress (Rudnick & Lundberg, 2012). The onset, course and severity of schizophrenia has been statistically associated with psychosocial stress and severity of psychotic symptoms has been correlated with severity of childhood trauma (neglect, physical and sexual abuse, family violence: Carr & McNulty, 2016).

### **[60.1960] Early Intervention**

Early treatment during the prodromal phase or first episode of schizophrenia has developed as best practice following evidence suggesting poorer outcomes are associated with later intervention (Jackson & McGorry, 2009). Early intervention targets people who are assessed as ultra-high risk for developing schizophrenia, promotes early recovery, reduces the duration and number of episodes, and prevents relapse (Carr & McNulty, 2016).

### **[60.1970] Recovery Model**

The recovery model is commonly used as a framework for service provision in mental health following advocacy from service users and consumer groups. Regarding the psychotic illnesses, groups such as Hearing Voices have argued that the stress of stigma and disempowerment through treatment are contradictory to recovery and therefore problematic. The Recovery Model prioritises autonomy, trauma informed care, optimism, strengths-based care and civil rights.

## PART 3: HOW PSYCHOLOGICAL ISSUES AND ASSESSMENTS CONTRIBUTE TO JUDICIAL DECISION MAKING ACROSS JURISDICTIONS

### [60.3010] Introduction

Psychologists seek to explain (and predict) behaviour. As such, any aspect of psychological functioning related to an outcome of relevance within a given jurisdiction may be relevant to judicial decision making. Examples include competency to stand trial, level of intellectual functioning, neurological impairment, parental capacity, malingering, personality disorder, mental health status (during assessment and/or at time of a past event), risk of a future behaviour, substance abuse, and psychological consequences of exposure to injury or violence. The psychologist will assess the client with a clinical lens, and formulate their findings in a report with language that is appropriate for the court room. By evaluating behaviour using interview, assessments and collateral information, and with reference to professional literature, the psychologist is able to provide the court with a scientifically informed opinion on a matter before the Court. The assessment strategy and tests used by a psychologist is informed not only by the specific question asked, but also by the jurisdiction and the relevant legislation appropriate to that jurisdiction.

### *Assessment across different jurisdictions*

### [60.3110] Criminal Law

Psychological assessment may be used for the prosecution, defence or sentencing processes. It may also serve to inform the court of areas that assist in determining whether criminal responsibility and competence to stand trial are worthy of consideration. The psychologist conducting assessment remains objective and impartial, regardless of whether defence or prosecution are requesting assessment. Their job is to inform the court of the relevance of assessments to help in the decision making process. Mental health, personality, cognitive capacity and adaptive functioning are often examined in the context of criminal cases.

### [60.3120] Family Law

Psychological assessment can assist in the identification of issues that relate to child wellbeing and the best interests of the child. The focus of family assessments is often to establish the extent of any emotional harm suffered by the child and explore the functioning of the family when together or apart. Parenting capacity may be assessed with a variety of instruments and psychologists may inform the court of their professional opinion in relation to this issue. The use of psychological tests provides valid and reliable evidence that support the psychologists professional opinion in a sensitive area of law.

Assessments will be accompanied by parent or carer interview, and observations of interactions between each parent and the child/ren. Where the child's wellbeing is enhanced or threatened by other family members or a parents extended network, further interviews and observations may be made. Interviewing teachers and significant others may be useful to the assessment process, and they may be asked to complete assessments that seek to obtain data regarding the behaviour of child/ren. Assessing motivation to change (parenting course, counselling, drug and alcohol treatment) may also feature in treatment recommendations. Psychologists will combine all data and with their knowledge of theories of development, attachment, mental health, personality, drug and alcohol abuse and the impact of violence or exposure to risks/neglect, in order to provide an expert opinion on the best circumstances for the child/ren.

### **[60.3130] Children's Court**

Psychological assessment in the Children's Court may deal with criminal cases and cases involving the care and protection of children. Similar to criminal law, psychological assessment can aid in the detection of mitigating factors such as cognitive functioning, mental health issues or social factors that have influenced behaviour. Chapter 62 looks further into the intricacies of assessment within the Children's Court.

### **[60.3140] Civil Litigation**

Psychologists may inform the court on issues such as impairment, personality profile and malingering. The use of psychological tests in this area of law complements clinical interview by identifying possible conscious or unconscious efforts to fabricate results for secondary gain. At times the client may exaggerate their symptoms and this may be identified through use of psychometric testing. Robust psychological assessments such as the MMPI-2 (Ben-Porath & Tellegen, 2008) or the PAI (Morey, 1996; 2007) have built-in measures that help identify patterns of inconsistency and efforts to fabricate results. Chapter 61 provides comprehensive detail of the role of the psychologist in civil matters. Chapter 68 is also relevant in civil litigation matters if the employment of neuropsychologist is warranted (brain injury).

### **[60.3150] Coroner's Court**

Psychological autopsy may be considered for this Court. Psychologists may be asked to review all available information pertaining to the deceased and formulate an assessment based on interviews of relatives, review of medical records or any information that might assist the psychologist in formulating a report about the behaviour of the deceased prior to death. In the case of suicide, the Court may ask a psychologist to review all available information to assess the likelihood that the deceased took their own life, as opposed to the actions of others. Medical records and evidence of seeking help from professionals or others will add weight to the assessment, alongside reference to contemporary literature on predicting suicidality.

### **[60.3160] Tribunals**

Numerous tribunals and matters before the tribunal benefit in their decision making by consideration of psychological assessments. For example, a specialist Mental Health Review Tribunal may utilise psychological assessment and test data to determine whether a patient should remain in a secure facility or be integrated back to the community with necessary support. Similarly, a Guardianship Board relies on formal intelligence testing and the results of functional assessments to determine whether an individual is able to manage their financial affairs. Other tribunals, such as Professional Boards (eg, a Medical Board), may benefit from psychological appraisals of professional competence or the characteristics associated with unethical behaviour.

## PART 4: DIFFERENT CATEGORIES AND METHODS OF TESTING

### [60.4210] Introduction

There are several different categories of structured psychological tests. Each category has a specific purpose (ie a certain question it seeks to answer) and thus also a context in which tests within that category should and should not be used. Due to these differences, each category of test has different strengths and weaknesses. Additionally, within these different categories of psychological tests, there is also a variety of different methods and formats of testing. Again, each method also has its strengths and weaknesses, which contribute to the overall reliability and validity of the test (see earlier parts of this chapter for a more detailed discussion regarding reliability and validity). Therefore, in order to understand and thus critically evaluate the use of a test in a legal setting, it is important to understand the broader category of the test, the methods it uses, and consequently the strengths and weaknesses of the test relative to its contextual parameters.

### [60.4220] Categories of psychological tests

The categories of structured psychological tests include:

- **Broad measures of general symptoms/general dispositions**

These tests are used when there is a need to screen for a range of psychological issues. Often these tests are used early in the assessment process when the referral question is either unclear or ambiguous, or as a general measure of the client's overall wellbeing. Whilst broad measures differ significantly in length (ie some are brief checklists, whilst others are lengthy personality inventories), they are generally considered as screening tools only, providing a cross-sectional snapshot of the person's current experience. Consequently, broad measures are often used in conjunction with more specific measures, as broad measures on their own do not provide enough information from which to draw absolute conclusions ie they are not designed for diagnostic purposes. An example of a broad measure which is commonly used is the Symptom Checklist-90-Revised (SCL-90-R; Derogatis, 1994).

- **Brief measures of symptoms**

These tests provide more in depth assessment of syndrome specific symptomology. Often these tests are used when there is a clearer indication of a problem in a certain area, or particular symptomology of an identifiable mental health issue. These tests provide clinicians with valuable information regarding the specifics of the client's presentation, and also function as excellent measures to track progress/change across time. An example of a brief measure of symptoms is the Beck Depression Inventory (BDI-II; Beck, Steer & Brown, 1996).

- **Structured diagnostic interviews**

As the name would suggest, these tests are designed for diagnostic purposes and thus are structured in relation to the diagnostic criterion of either the DSM-5 (APA, 2013) or ICD-10 (WHO, 2010). These interviews can take several different forms including, a) omnibus interview schedules which assess for the presence of symptoms across a variety of disorders; b) specific schedules which differentiate symptomology within a group of disorders and c) schedules which identify the presence of a single disorder.

Although such interviews were designed to increase diagnostic reliability, they are scarcely used in clinical practice due to large administrative burden. Some examples of structured diagnostic interviews include the Structured Clinical Interview for DSM-5 Disorders (SCID-5-CV; First et al, 2016), the Anxiety Disorder Interview Schedule for DSM IV (ADIS; DiNardo, Brown & Barlow, 1994) and the Clinician Administered PTSD Scale (CAPS; Weathers et al, 2013).

- **Neuropsychiatric/cognitive functioning tests**

These tests are used to assess intelligence and cognitive abilities such as memory, executive functioning, learning, attention, concentration and problem solving. For more information on these tests please see chapter 68 in current volume, which provides extensive discussion of neuropsychological testing in the context of the legal system.

### [60.4230] Methods of assessment within psychological tests

As mentioned earlier, there are also many different *methods* of assessment used within psychological tests. The most commonly used method of psychological testing is self-report, which requires clients to complete the psychological measure themselves without interference from a clinician. This format is favoured due to its relative ease of administration (in terms of time and cost) and because it directly captures the clients experience of constructs that can not necessarily be observed ie the client rates how much they endorse an internal symptom, such as depression. However, the validity of self-report measures is often criticised due to issues with bias, as these measures rely on the client's introspective ability and honesty. Clients can easily distort their responding by exaggerating symptoms or under-reporting the severity/frequency of symptoms, particularly if there is an identifiable secondary gain based on the outcome of the measure (see the assessment of Malingering below). Similarly, self-report measures also rely on client understanding, memory and insight, which in some cases may be lacking. However, in order to address some of these issues, strategies are employed when designing self-report questions to limit the extent to which clients can engage in distortion. These include techniques such as reverse questions, malingering scales, and rating scales with forced choice questions. Additionally, by their nature, self-report questionnaires are generally repeatable measures, meaning they can be administered on multiple successive occasions and thus are excellent measures to assess change across time.

Similar to self-report measures, psychological testing can also involve observer-administered measures, in which an individual completes a measure regarding the behaviour of another individual. This is most often used in regards to the assessment of children i.e., parents complete measures of a child's behaviours. Such measures also include those where clinicians record client presentations. These tests have similar strengths and limitations to those discussed above ie they are open to response bias and distortion. However, such measures often include a variety of different versions that are completed by different individuals to allow for comparison across parties. For example, the Conner's Rating Scale (Connors, 2008), which assesses for ADHD in children, has a parent, teacher and child version and includes positive/negative impression indices to delineate distorted responses between parents and teachers.

Another method of psychological testing is psychologist/clinician-administered measures. These tests involve the psychologist asking the client to respond to specified questions or tasks and their response or performance is recorded by the clinician. This method is most notably used within neuropsychological testing (see chapter 68 for more details), but is also employed in structured diagnostic interviews, assessments of malingering, and some personality testing. Tests which use this method often involve very strict administration instructions, including verbatim questions and/or directions in order to ensure consistency and thus increase

reliability. However, given these strict parameters, this method of testing requires large amounts of time for administration (including both testing and scoring), and often requires specialised training on behalf of the clinician. These tests are generally hand scored (which takes considerable time), although they can also be computer scored with appropriate software. As a result, given the administrative burden, clinicians will often administer shortened versions of a test, or will only administer subtests in a test battery that relate directly to the referral question (for further discussion on this practice refer to Chapter 68). Finally, unlike self-report and observer measures, clinician administered measures are generally not repeatable measures, and require time to elapse before they can be re-administered.

## PART 5: CATEGORIES OF PSYCHOLOGICAL ASSESSMENT AND COMMONLY USED TESTS WITHIN CATEGORIES

### [60.5210] Introduction

A vast array of tests exist traversing a broad spectrum of areas of psychological inquiry. While this chapter cannot discuss the entirety of either categories of psychological assessment or indeed all the individual tests within each or any category, the following section contains the most commonly used forms of assessment within the most frequent areas of psychological inquiry that may be of interest to the judicial system.

### *Tests of intellectual functioning*

#### [60.5220] Individual and group administered

These tests may be divided into two categories: individual; and group administered. For most legal assessments, an individual test is preferable because of its higher level of accuracy in measuring intelligence. Also, by its nature, it provides more information through the psychologist being able to observe the subject's performance. The reader should note that a low score on an Intelligence Test does not in and of itself determine that the assessee is Intellectually Disabled: this must be combined with an assessment of adaptive functioning (daily living skills – see below) and supported by corroborative information.

#### [60.5230] The Wechsler Scales

The most recognised intelligence tests that are utilised world wide are the Wechsler scales. The Wechsler scales have been translated, standardized and adapted for many countries. They vary according to the age range being assessed but in the forensic context, the most widely used are the Wechsler Intelligence Scale for Children – the fifth edition (WISC V) and the Wechsler Adult Intelligence Scale – fourth edition (WAIS-IV) (Weiss et al, 2015)<sup>1</sup>. As with most tests, they are constantly updated and revised according to the current literature on measuring intelligence.

The WAIS-IV is used for adults (aged 17 +) and WISC-V is used with children (aged 5-16). At the time of writing, these tests are current versions of Wechsler scales that are used by psychologists to assess intelligence. Both tests provide a range of scores that indicate one's level of functioning across a variety of ability areas that include verbal comprehension, working memory, processing speed, and perceptual reasoning. The WISC V also includes the addition of fluid reasoning as one of the more recent advances in the expansion of the view of intelligence as a 5 factor model. Each test provides an overall estimate of IQ as well as an indication of strengths or weaknesses across each of the 4 or 5 scales. Psychologists will often express results in a general bandwidth (see Table 1 below) and compare results with similar age ranges to provide a percentile score (Weiss et al. 2015). When a psychologist is asked to give expert evidence regarding one's level of intellectual functioning, they will usually refer to the bandwidth and percentiles, so others can gain an understanding of how that person might function compared to others the same age. Table 1 shows the band of scores and their corresponding labels:

**TABLE 1 Descriptors for IQ Ranges within the Wechsler Scales**

<b>IQ Range</b>	<b>Descriptor (WISC-V)</b>	<b>Descriptor (WAIS-IV)</b>	<b>Percentile Range</b>
130 and above	Extremely High	Very Superior	98 <sup>th</sup> and above
120 to 129	Very High	Superior	91-97
110 to 119	High Average	High Average	75-90
90 to 109	Average	Average	25-74
80 to 89	Low Average	Low Average	9-24
70 to 79	Very Low	Borderline	3-8
69 and below	Extremely Low	Extremely Low	2 <sup>nd</sup> and below

Various circumstances may dictate that a full (and lengthy) assessment of intelligence is either unnecessary or impracticable. In such cases, where an estimate of intellectual functioning will still assist the court, shortened or abbreviated measures may be appropriate. In this case, the “gold standard” Wechsler scales are still applicable. An example of such is the shortened version of the Wechsler scales called the Wechsler Abbreviated Scale of Intelligence (WASI-III), which is a screening tool that indicates intellectual capacity but only takes about 15-20 minutes to complete. It does not provide a comprehensive profile of intelligence but can indicate lower levels of functioning that might require further investigation.

*Statistical Properties* : The statistical properties of all Wechsler scales are impressive and lend support for their use. The overall reliability coefficients for WISC V composite scores range from 0.88 (processing speed) to 0.96 (Full Scale IQ). The WAIS IV is a well-established scale and has high consistency. The test manual states test-retest reliabilities ranged from 0.70 (7 subscales) to 0.90 (2 subscales) over a 2-12 week period. Inter-rater reliability coefficients are also very high, all being above 0.90. The WAIS IV correlates highly (concurrent validity) with the Stanford-Binet IV test (0.88) and has high concordance with various measures of memory, motor speed, attention, and cognitive ability.

<sup>1</sup> The Wechsler Pre School Intelligence (WPPSI) is the scale used for children below the age of five.

## **[60.5240] Woodcock Johnson**

This Woodcock Johnson intelligence tests are commonly used by psychologists who prefer the Cattell-Horn-Carroll (CHC) model of intelligence. The CHC model integrates 2 established models (the Gf-Gc theory of fluid and crystallised intelligence Cattell, 1941; Horn 1965) with Carroll’s three-stratum theory of intelligence (1993). There were many similarities between the theories thus they were amalgamated to produce the CHC model. At present, the fourth edition is the latest in the series of tests that were first developed in 1977. The Woodcock Johnson Fourth Edition – Australasian Adaption Test of cognitive ability (WJ-IV; Schrank, McGrew & Mather, 2015) measures intelligence similar to the Wechsler scales with a variety of subtests that measure 10 ability areas such as oral vocabulary, number series, verbal attention, letter-pattern matching, phonological processing, story recall, visualisation, general information, concept formation and numbers reversed. Like the Wechsler scales, skill areas are grouped into 3 ability cluster scores: Comprehension-Knowledge (Gc); Fluid Reasoning (Gf); and Short Term Working Memory (Gwm). All combine to produce a General Intellectual Ability score

(GIA). The GIA provides a comprehensive assessment of functioning that may be representative of the general intellectual ability of the test-taker (like the FSIQ in Weschler tests). For the screening version of the Woodcock Johnson, a Brief Intellectual Ability (BIA) is produced. The BIA is used to screen for possible intellectual difficulties (as per the WASI in Weschler scales) or re-evaluations when a comprehensive assessment is not required. The WJ-IV is an individually administered test.

*Statistical Properties:* The WJ-IV has reported reliability coefficients that range above 0.80 for subtests and above 0.90 for each of the cluster scores. For the GIA, a median reliability coefficient of 0.85 is reported. Correlations with other measures of intelligence are 0.86 for the WISC-IV (GIA versus FSIQ), 0.84 for the WAIS-IV, 0.72 for the Kaufman Assessment Battery for Children second edition (KABC-II) and 0.80 with the Stanford Binet Intelligence Scales.

### **[60.5250] Stanford-Binet Intelligence Scales**

The Stanford-Binet Intelligence Scale is an individually administered standardized test that measures intelligence in children and adults, from ages 2 plus. The current version of the Stanford Binet Scales is the fifth edition (SB5; Roid, 2003). Unlike the Weschler scales, there are no variants for children or adolescents; all age groups use the same set of tasks. This is sometimes a criticism of the scale because different age ranges complete different tests. For pre-school age children, it is not uncommon for them to score 0 for an item due to difficulty or unwillingness to co-operate. This makes it difficult to compare younger age ranges. The SB5 provides a score for each of the 10 subtests, divided into verbal IQ (VIQ) and Non-verbal IQ (NVIQ) tasks. All combine to provide an overall estimate of IQ name Full Scale IQ (FSIQ), as per the Weschler scales.

The SB5 is marketed as assisting early predication of learning disabilities in children.

*Statistical properties:* The SB5 boasts reliability coefficients that are very high. For FSIQ, NVIQ and VIQ reliabilities range between 0.95 and 0.98 (average across all age groups). For each of the 10 subtests, reliability ranges between 0.84 and 0.89. Concurrent and criterion validity data were obtained via the Woodcock Johnson, Bender gestalt and Weschler scales.

### **[60.5260] Differential Ability Scales (DAS)**

The Differential Ability Scales are standardised batteries of cognitive and achievement tests for children only. The tests are individually administered depending on the age range of the child. For those aged 2 years 6 months (2:6) - 6 years 11 months (6:11), they will complete the Early years Cognitive battery. For 7 -17:11, completion of the School-Age Cognitive Battery is appropriate. The current version of the test is the DAS Second Edition (DAS-II; Elliot, 2007). The diagnostic subtests measure a variety of cognitive abilities including verbal and visual working memory, immediate and delayed recall, visual recognition and matching, processing/naming speed, phonological processing, and understanding of basic number concepts. Some of these subtests can be used with children aged 2:6–17:11, while others have specific age ranges. Each of the subtests are combined to produce a General Conceptual Ability Score (GAC).

*Statistical Properties:* Average internal consistency reliability coefficients for the GCA and non verbal composites are greater than 0.90 for both age range batteries. The range for subtests is 0.77 – 0.95. In relation to validity, the GAC score on the DAS II scale correlates well with the full Scale IQ (FSIQ) scores on the WAIS-IV (0.84) and WPPSI II (0.87).

### **[60.5270] Kaufman Intelligence Scales**

The Kaufman intelligence scales are standardised measures assessing intelligence across age ranges. Like the Weschler tests, there is a screening version named the Kaufman - Brief Intelligence test, with the latest version being the K-BIT 2 (Kaufman & Kaufman, 2004).

The Kaufman Adolescent and Adult Intelligence test (KAIT) (Kaufman & Kaufman, 1993) is designed for use with clients aged 11-85 years. The Kaufman Assessment Battery for Children (KABC-II: Kaufman & Kaufman, 2004) is for use in the 3-18 age range. All tests are individually administered. The KABC-II is considered culturally fair due to administration flexibility according to linguistic and cultural backgrounds. It gives the psychologist the option to incorporate or exclude verbal ability scores, which would likely provide a score more representative of the assessee's ability, if English is considered their second language and less developed. A normative update was issued in 2018 and reports should reference the new norms, although the test battery remains unchanged.

*Statistical Properties:* The KBIT-2's IQ Composite internal consistency coefficient of 0.93 across ages (0.89 to 0.96), with reliabilities increasing with age. The Verbal (0.91) and Nonverbal (0.88) coefficients are somewhat lower but within acceptable ranges, although the Nonverbal scale coefficients are only 0.78 at ages 4 and 5. IQ Composite test-retest stability was .90 over mean intervals of 22.5 to 30.8 days, with a mean performance increase of 4 points.

Concurrent validity evidence included data on the relationship with several other measures of intelligence. Comparisons showed Full Scale and Performance IQs about 4.5 points and 7 points higher on the Wechsler Abbreviated Scale of Intelligence (WASI) than on the corresponding KBIT-2 scales, even though the correlations were strong. Correlational studies were reported with several comprehensive Wechsler scales. The Full Scale-IQ Composite correlations with the WISC-III and WISC-IV were 0.76 and 0.77, respectively. The WAIS-III correlation was 0.89.

The KAIT and KABC-II has similar statistical properties as the K-BIT2. Each had very good levels of reliability coefficients demonstrated across each of the subtests and Full scale IQ and validity studies showed correlation with Wechsler tests of intelligence.

## **[60.5280] Adaptive Behaviour**

When a psychologist assesses intelligence and the levels are considered significantly below average for the assessee's age, a functional assessment may be completed to provide more information about the impact of cognitive functioning upon day to day living skills. These assessments are required before a diagnosis of Intellectual Disability can be made. The most commonly used tests of adaptive functioning include the *Adaptive Behaviour Assessment System – Third Edition* (ABAS-3: Harrison & Oakland, 2015) and the *Vineland Adaptive Behaviour Scale – Second Edition* (VABS-II: Sparrow, Cichetti & Balla, 2006).

Functional skills that are generally assessed include communication skills, living skills, academic skills, self-care, home living skills, community use, and motor skills. Often a parent, teacher or adult who knows the client well enough to answer the required items will complete the form. For determination of qualifying for an intellectual disability the diagnostic criteria includes a score of <70 IQ on a standardised intelligence scale (such as the WAIS-IV), combined with at least 2 areas of significant limitation in adaptive functioning. Meeting both criteria enables qualification for assistance from government agencies, or more recently the National Disability Insurance Scheme (NDIS). The NDIS recognises intellectual disability as one of the criteria that leads to assistance from relevant agencies for a number of supports dependant on need.

Both scales produce an overall estimate of functional abilities. For the ABAS 3 it is named the General Adaptive Composite (GAC). For the VABS II, it is demonstrated via the Adaptive Behaviour Composite (ABC).

*Statistical properties:*

ABAS 3 - Internal consistency reliability scores range from 0.97 - 0.99 for the General Adaptive Composite (GAC) and 0.91- 0.98 for each of the adaptive domains (Conceptual,

Social etc). Test – retest reliability is placed in the excellent range (above 0.90) for the GAC, while 0.80 - 0.90 for adaptive domain scores and 0.70 - 0.90 for the skill areas. Inter-rater reliability ranged from 0.82 - 0.91 for the GAC, 0.78 - 0.84 for adaptive domain scores and 0.70 – 0.82 for each skill area. Concurrent validity with the VABS II is 0.75 for the teacher/day care provider form and 0.84 for the teacher form (Harrison & Oakland, 2015).

*VABS II* - Internal consistency reliability scores range from 0.84 – 0.93 for Communication domain, 0.86 – 0.91 for Daily Living Skills domain, 0.84 – 0.93 for Socialization domain and 0.77 – 0.90 for Motor Skills domain. The Maladaptive Behaviour Index range is 0.85 – 0.91 and Adaptive Behaviour Composite (ABC) 0.93 – 0.97. Test – retest reliability indicated average correlations between 0.76 and 0.92 across domains. Interrater reliability ranged between 0.71 – 0.81 for each of the domains. The maladaptive behaviour subscales demonstrate correlations ranging from 0.59 – 0.83 for the survey interview form and between 0.39 – 0.87 for the parent/caregiver rating form across age groups.

### [60.5290] Culturally Robust Intelligence Tests

Intelligence tests are sometimes criticised because they are primarily sampled and normed on populations that do not include indigenous samples. Thus caution in interpreting tests is important for psychologists testing those with a language barrier or coming from a remote community that is very different to the general population (APS Guidelines).

Intelligence tests that are considered more appropriate for cultural minorities or give an indication of intellectual capacity for those with language barriers (or hearing impairment) include the Raven’s Standard Progressive Matrices (Raven 1998). Although the latest versions of the Wechsler tests for children are more inclusive of indigenous samples in their norming process (WISC-V), cultural variance needs to be considered by the assessing psychologist (Weiss et al, 2015).

When the client’s second language is English and they have difficulty conversing, the psychologist may employ the Raven’s or the Wechsler Non-Verbal Scale of Ability (WNV: Wechsler & Naglieri, 2006) for an indication of cognitive capacity. The Raven’s is a non-verbal assessment tool that measures novel reasoning and visual abstract concept formation, thus providing an estimate of IQ. As mentioned previously, the KABC-II is considered culturally fair due to administration flexibility according to linguistic and cultural backgrounds.

Ethnicity and cultural factors need to be considered by the psychologist during assessment. If the assessment does not consider culture, it could affect the admissibility of the evidence in court because the validity of the collected information may be questionable (Chiu, 2014).

Chapter 68 discusses finer details of intelligence testing and the subtests within the Wechsler scales (Crowe, 2010).

## *Tests of Neuropsychological Impairment*

### [60.5410] Neuropsychological Impairment

This area of psychology generally considers cognitive impairment, organic brain damage (resulting from injury or illness, alcohol abuse) and the wider impact of brain dysfunction. Within the DSM-5, the section “Neurocognitive Disorders” subsumes a range of disorders for which the underlying pathology and often aetiology can potentially be determined, such as is the case with dementias. Neurocognitive disorders are evidenced by deficits in one or more of the following cognitive domains: complex attention; executive function; learning and memory; language; perceptual-motor; and social cognition. Disorders associated with aging, disease, trauma, and the effects of substances are all included.

There are a number of tests that are commonly used for the forensic context. Close analysis of subtests within the Wechsler intelligence tests can indicate certain aspects of functioning. This gives the psychologist an understanding of a client's skills or weaknesses in a specific area of cognitive (neurological) functioning. If a psychologist views a profile indicative of possible brain damage, they may refer to a neuropsychologist for further investigation, as this will have major implications for competency to stand trial or criminal responsibility.

Referral to, or assessment by, a clinical neuropsychologist entail areas such as attention, memory, language, executive functioning (planning), visuospatial skills, dementia screens and any evidence of acquired brain injury. Note however that the use of neuropsychological tests is not limited to specifically trained neuropsychologists. Psychologists are trained to assess a variety of areas and their skill sets will vary dramatically depending on training and experience.

Chapter 68 provides a comprehensive review of the application of neuropsychology to the criminal justice system.

## *Tests of Personality*

### **[60.6510] Introduction**

A personality test is a standardised questionnaire or procedure designed to elicit information about a person's character or "psychological make-up". There are many such tests and self-assessment tests have proliferated on the Internet. Early personality tests were largely projective (eg the Thematic Apperception Test and the Rorschach Ink Blot test), however the lack of standardisation and psychometric weaknesses make such tests unreliable. Personality tests vary greatly according to the underlying theory of the test developers, with major theoretical perspectives including the Psychoanalytic approach, the Interpersonal approach, the Personological approach, the Social-Cognitive approach, the Multivariate (trait) approach, and the Empirical approach. Among these approaches, the most contemporary and predictive tests are yielded by the latter two approaches: Multivariate and Empirical approaches.

### **[60.6520] Multivariate Trait Approach**

The multivariate (trait) approach proposes that there are a number of dimensions of individual difference that people have in common, which may be referenced to identify an individual's personality (Shum et al, 2017). Trait theories propose that a number of commonly identifiable descriptors (traits) exist within individuals that serve to predict how they will respond in any given situation. The combination of various traits results in a personality type. By examining the behaviour of people with the same combinations of traits, an understanding of past behaviour and predictions about future behaviour can be made. The most commonly used "trait" based personality tests are Cattell's 16PF Questionnaire and the more recent NEO Five Factor Personality Inventory-3. Both assessments are conducted using a self-report questionnaire format and thus are susceptible to response distortion and social desirability bias.

*Sixteen Personality Factor – Fifth Edition* (Cattell, Cattell & Cattell, 1993)

The 16 PF measures sixteen primary factors (including intelligence) and five secondary factors, known as global scales. The Global Scales correspond closely to the "Big Five Personality Traits" (Openness, Conscientiousness, Extraversion, Agreeableness and Neuroticism) identified in the NEO (see below), although there are differences in the meanings attributed to the traits by virtue of differing factor analytic strategies used. The 16 PF can be used to assist in both psychiatric diagnosis and treatment planning, and is thought to provide some prognostic value. The 16 PF contains 185 items, is suitable for ages 16 years and above, and is available in 35 different languages. The 16 PF has an extensive research literature, with internal consistency

varying from 0.66 to 0.86 across the sixteen factors, and an average of 0.87 across the global scales. A large-scale (n= over 10,000) factor-analytic study confirmed validity and there is sound evidence for the predictive validity of the 16 PF in organisational and educational settings (Schum et al, 2017; Cattell & Mead, 2008).

*NEO Personality Inventory – Third Edition* (McCrae & Costa, 2010)

The NEO-PI-3 consists of 240 items and is suitable for use from ages 12 years and above. It measures thirty “facet” scales, six for each of the five factors: Neuroticism (worried versus calm, insecure versus secure, self-pitying versus self-satisfied); Extraversion (sociable versus retiring, fun-loving versus sober, affectionate versus reserved); Openness (imaginative versus down-to-earth, preference for variety versus preference for routine, independent versus conforming); Agreeableness (soft-hearted versus ruthless, trusting versus suspicious, helpful versus uncooperative); and Conscientiousness (well organised versus disorganised, careful versus careless, self-disciplined versus weak-willed). These factors are also referred to by some authors as “domains”. A brief version, the NEO Five Factor Inventory (NEO-FFI-3) contains 60 items and provides a quick measure of the five factors, but not the facet scales. The internal consistency of the NEO PI-3 ranges from 0.89 to 0.93 and is thus considered very good. While shorter (and thus less reliable) internal consistencies of the five factors in the NEO-FFI are good, ranging from 0.74 to 0.89, and test-retest reliability is also sound, ranging from 0.86 to 0.90 (Costa & McCrae, 2008). Test-retest reliability of the NEO-PI-3 is satisfactory, ranging from 0.63 to 0.83, and both convergent and discriminant validity has been demonstrated against measures of employee career suitability. The NEO has been translated into multiple languages, with little loss to psychometric properties. The NEO PI-R was developed for use with persons without overt psychopathology

### [60.6530] Empirical Approach

The empirical approach is concerned with describing personality for the purposes of predicting socially relevant criteria, such as behaviours relevant to social outcomes (mental illness, criminality, academic achievement and work performance: Schum et al, 2017). Like the trait approach, the empirical approach uses questionnaires containing items (individual questions), which discriminate between respondents thought to demonstrate, or not demonstrate, a particular attribute. The empirical approach investigates the relationship between measures of individual difference and measures of socially relevant criteria, without necessarily inferring the cause of that relationship (Schum et al, 2017). Consequently, item development in the empirical approach commonly relate to known or observed criteria associated with specific diagnoses or behaviours. The three most commonly used empirical based measures of personality are the Minnesota Multiphasic Personality Inventory-2 Restructured Form (MMPI-2 RF), the Millon Clinical Multiaxial Inventory – IV (MCMI-IV) and the Personality Assessment Inventory-2 (PAI-2).

*Minnesota Multiphasic Personality Inventory* (MMPI)

The MMPI is a standardised psychometric test of adult personality and psychopathology. It is commonly used in forensic psychology. The original MMPI was developed in the 1940’s to assist in the differential diagnosis of patients presenting with a range of symptoms at a large psychiatric hospital (Hathaway & McKinley, 1942). Over 500 questions were developed from previous questionnaires and from questions routinely asked by psychiatrists. The MMPI was then administered to patients with known diagnoses and item endorsements (against diagnostic profiles) were computed. The questionnaire was also delivered to “normal” participants and item endorsement was similarly tallied. By comparing items routinely endorsed by patients with a psychiatric illness compared with item endorsement for persons without a diagnosis, the authors identified magnitude of difference and possible predictors of specific diagnostic categories. The specific content of items was not important; rather the frequency of

endorsement was seen as significant. Quantification of differences in endorsement between “normal” and psychiatric patients was expressed in terms of “T” scores, where T=50 was considered average and two standard deviations (a measure of distance from the average and computed as 1 standard deviation being 10 T) considered significant or “pathological”. Overtime, the measure was administered to multiple different identified groups, resulting in a range of additional scales being developed on the basis of item endorsement configurations.

The original MMPI had many problems. As a result, new versions of the MMPI were developed. Most recently, the MMPI-2-RF improved the psychometric sophistication of the restructured clinical scales of the MMPI-2. Given that the MMPI-2-RF is briefer than the MMPI-2 (338 versus 567 items) and uses a different theoretical approach to personality assessment development, many assessors continue to use the 2003 (Restructured Clinical Scales) version of the MMPI-2. However, multiple studies have shown that the MMPI-2-RF performs as well as, and in some cases better than, the MMPI-2. A feature of the MMPI-2-RF is improved discriminant validity, providing more distinct symptom constellations and thus greater diagnostic accuracy. An adolescent version of the MMPI, the MMPI-A was developed by Butcher and his colleagues (Butcher et al, 1992).

The ten clinical scales in the MMPI comprise: Hypochondriasis (concern with bodily symptoms); Depression (depressive symptoms); Hysteria (awareness of problems and vulnerabilities); Psychopathic Deviate (conflict, struggle, anger, respect for society’s rules); Masculinity/Femininity (stereotypical masculine or feminine interests/behaviours); Paranoia (level of trust, suspiciousness, sensitivity); Psychasthenia (worry, anxiety, tension, doubts, obsessiveness); Schizophrenia (odd thinking and social alienation); Hypomania (level of excitability); and Social Introversion (people orientation). Against this, the MMPI-2-RF contains 9 Restructured Clinical Scales, which, as stated above, purport to be more diagnostically distinct. These scales are: Demoralisation (a general measure of distress); Somatic Complaints (tendency to evidence medically unexplained symptoms); Low Positive Emotions (measures anhedonia); Cynicism (negative world view associated with interpersonal conflict); Antisocial Behaviour (social deviance); Ideas of Persecution (paranoid delusions, suspiciousness, alienations); Dysfunctional Negative Emotions (fear, anxiety and resentment); Aberrant Experiences (risk for psychosis, thought disorder); and Hypomanic Activation (features of mania).

The main strength of the MMPI from the beginning of its development and through to current versions is the inclusion of the validity scales. These scales address many of the problems associated with self-report measures. Three types of validity measures are included: those that detect non-responding or inconsistent responding (three scales); those that detect over-reporting of symptoms (four scales); and those that detect under-reporting of symptoms (three scales). These scales are often used in the detection of malingering (see Malingering section below).

The MMPI-2 has additional scales which provide incremental validity to the clinical scales, known as “Content Scales”, as well as numerous “Supplemental Scales”, useful in the detection of specific behaviours/risks (such as substance misuse). In addition, the MMPI-2 contains the Personality Psychopathology Five Scales (PSY-5) which measures dimensional trait of personality disorders, which demonstrate moderate internal consistency and comparability with the NEO-PI-R (see above).

Interpretation of the MMPI-2 conventionally involves what are known as “codetypes”. Codetype refers to the highest elevations on two (and sometimes 3) clinical scales, where both (or all three) are elevated by at least 1½ standard deviations (T greater than or equal to 65). Through the collection of vast numbers of MMPI-2 profiles, similarities in presentation for people sharing codetype configurations enable the clinician to describe what an individual is like and how they might be expected to behave. In a forensic context, this information may be used to confirm data from other sources (such as interview data), predict risk, or identify treatment needs.

The MMPI-2 has been criticised as being vulnerable to response variation based on ethnicity and/or socioeconomic status.

*Millon Clinical Multiaxial Inventory (MCMI-IV: Millon, Grossman, & Millon, 2015)*

The MCMI-IV provides information on personality traits and psychopathology, including specific psychiatric disorders identified within the DSM-V. Intended for use with adults aged 18 years and over, the MCMI-IV was developed and standardised on clinical populations and should not be used with the general population or adolescents. It is not suitable for use with persons of low intelligence or with very poor reading ability. It comprises 195 true/false items yielding 15 personality patterns, 10 clinical syndrome scales, 5 validity scales and 45 personality facet scales. The clinical symptom scales are subdivided into 3 severe clinical syndromes and 7 clinical syndrome scales. The severe clinical syndrome scales are: Thought Disorder; Major Depression; and Delusional Disorder. The clinical syndrome scales are: Generalised Anxiety; Somatic Symptom; Bipolar Disorder; Persistent Depression; Alcohol Use; Drug Use; and Post-Traumatic Stress. The 15 personality patterns are further subdivided into 12 Clinical Personality Patterns and 3 Severe Personality Pathology scales. The personality patterns are: Schizoid; Avoidant; Melancholic; Dependent; Histrionic; Turbulent; Narcissistic; Antisocial; Sadistic; Compulsive; Negativistic; and Masochistic. The severe personality pathology scales are: Schizotypal; Borderline; and Paranoid.

The MCMI-IV is based on Millon's Evolutionary Theory of Personality, which is but one of many theories of personality, and as such is limited to this specific conceptualisation. Psychometrically, the MCMI-IV demonstrates good internal consistency with Cronbach alpha's of 0.84 for personality pattern, 0.83 for clinical syndrome, and 0.80 for the personality facet scales (Millon, Grossman & Millon, 2015). Test-retest reliability has been reported as ranging from 0.73 to 0.93, however this data is over a relatively short time period and no longer term data are available.

*Personality Assessment Inventory-2 (PAI-2: Morey, 2007)*

The PAI-2 is a self-administered objective test of personality and psychopathology for adults aged 18 years and over. It includes measures that facilitate treatment planning, implementation, and evaluation. Scales comprise 22 non-overlapping scales: 4 validity scales; 11 clinical scales; 5 treatment scales; and 2 interpersonal scales. Ten of the scales contain conceptually derived subscales designed to facilitate interpretation and coverage of the full breadth of complex clinical constructs. Scales were developed from both a "rational" as well as "quantitative" framework using theory-informed strategies for item selection and attending to item stability and correlates. The clinical syndromes assessed by the PAI-2 were guided by contemporary mental health classifications and diagnostic practice, such that items reflected components central to the disorder in question, as identified in the clinical literature. The PAI-2 items survey the full range of severity of each construct, from mild presentations through to severe, ensuring that scales are useful across a number of different applications/cohorts (Morey, 1996).

The PAI-2 contains 344 items answered on a four point scale: totally false; slightly true; mainly true; and very true. Each item response is weighted according to the intensity of the feature of the particular disorder, such that the item may attract a "score" ranging from 0 through to 3. Item scores are summated to provide a construct or disorder score, which is then compared with the norm population. The PAI was normed against three participant samples: a U.S. census-matched standardisation sample of adults living in the community (1,462); a sample of adult patients from various clinical settings (1,265); and a sample of college students from various universities (1,051). Direct stratification or representative sampling was used to ensure that each sample population was representative, including strategies to minimise the risk of bias arising from demographic variable (eg age, gender, and ethnicity: Morey, 2007).

Individual scale scores are usually reported as *T* scores. A *T* score refers to a fixed, standardized distribution where the mean or average score is 50 and the standard deviation (a measure of the variability away from the mean) is 10. Conventionally, *T* scores less than 1.5

standard deviations away from average (15 *T* points or *T* = 65 and above) are not considered significant. The higher the *T* score, the less likely that the variation away from average is an error in measurement and the greater the confidence in interpretation. On the PAI a *T* of 70 or above is considered significant, although the authors provide minor interpretative adjustments with specific population groups or on specific scales.

The authors report mean alpha's of 0.81 to 0.86 for scales from normative data. Studies using non-US samples, low socioeconomic groups, and patients suffering physical (non-mental health) disorders returned internal consistencies ranging from 0.70 to 0.80 (Morey, 2007). The author also reports on numerous studies examining test-retest reliability, with coefficients ranging from 0.71 to 0.82. However, the test-retest reliability dropped to 0.65 in a study of the PAI translated into Spanish (Rogers et al., 1995). Morey (2007) also provides extensive analysis of convergent and discriminant validity within the manual, citing medium to strong correlations with measures of similar scale constructs, and strong effect sizes in discriminating between clinical and control groups. Overall, the PAI-2 is considered a psychometrically sound assessment.

### **[60.60540] Personality Disorder Diagnostic Assessment**

While the assessment of personality and the prediction of behaviours are often most relevant to psychological assessment, the diagnosis of a specific personality disorder cannot be made through the use of personality questionnaires in isolation from additional collateral data (such as the reports of significant others, observations of social interactions, work or academic performance, etc). However, disorder specific assessments are available which provide data consistent with diagnosis, key amongst which are the Structured Clinical Interview for DSM-V Personality Disorders (SCID-PD; First et al, 2016).

*Structured Clinical Interview for DSM-5 Personality Disorders* (SCID-5-PD: First et al, 2016)

The SCID-5-PD is a semi structured interview that assists the clinician to make diagnoses for the ten personality disorders identified in the DSM-V. The SCID-5-PD is intended for use by mental health professionals who are familiar with DSM-5 Personality Disorder classifications and diagnostic criteria. While predominantly worded for an adult population, the SCID-5 can be used with adolescents, with some modifications of the language used in interview. In general, the SCID-5-PD can be seen as a diagnostic checklist and decision tree, but is best used with support from collateral sources. The SCID-5-PD has been translated into multiple languages and requires between 30 and 120 minutes to administer. At the time of writing, no research was available detailing the inter-rater reliability of the SCID-5-PD. However, data on previous versions of the SCID kappa values (a statistic referring to agreement between two raters, while accounting for chance), as ranging from 0.63 (fair) to 1.0 (excellent) for the clinical disorders, and test-retest reliability ranging from 0.35 (poor) to 0.78 (good) (Zanarini et al, 2000).

## *Tests of Neurodevelopmental Disorders*

### **[60.7010] Introduction**

Neurodevelopmental disorders constitute a group of conditions which have their onset during the developmental period (0 to 18 years). They are characterised by developmental deficits that affect personal, social, academic or occupational functioning. Neurodevelopmental disorders often co-occur, such that an individual might meet diagnosis for more than one neurodevelopmental disorder. Intellectual Disability, Communication Disorders, Autism Spectrum Disorders, Attention Deficit/Hyperactivity Disorder, Motor Disorders (such as Tic disorders), and Specific Learning Disorders are all examples of neurodevelopmental disorders.

Specific tests for these disorders are contained elsewhere in this chapter, or in the case of Intellectual Disability, in chapter 63A, Intellectual Disability in this volume. However, Autism Spectrum Disorder, which may affect as many as one in sixty eight persons (*Autism Speaks*), influences behaviour in a manner of relevance to the Court and is covered here.

## Autism Spectrum Disorder

### [60.7110] Introduction

Autism Spectrum Disorder (ASD) is associated with a persistent impairment in reciprocal social communication and social interaction and with restricted, repetitive patterns of behaviour, interests, or activities (DSM-5). As with other neurodevelopmental disorders, the symptoms of ASD are present from early childhood and impact upon everyday functioning. In previous versions of the DSM, Autism was included in a group of disorders called “Pervasive Developmental Disorders”, indicating that the symptoms of the disorder “pervade” every aspect of life and are present across time and circumstance. This remains true. ASD includes disorders previously referred to as Kanner’s syndrome, high functioning autism, atypical autism and Asperger’s disorder. As a spectrum disorder, the degree of impairment within and across symptom areas will vary, such that no two persons with ASD are the same. As a result, circumstances, compensatory strategies and social supports may mask the difficulties that the person with ASD has with activities of day-to-day living.

### [60.7120] Diagnosis

Diagnosis of ASD is most valid and reliable when based on multiple sources of information, including developmental data from previous assessments (including school reports), care giver data, formal assessments and, where possible, self-report. The quintessential deficit likely to manifest in a legal setting will be difficulties with, and abnormalities of, verbal and non-verbal communication. While language is often impaired, even those with well-developed formal language skills (vocabulary, grammar) are likely to experience impairment in the use of language for reciprocal communication. This may include behaviours that others, such as police, will misinterpret, such as poor eye contact, speech intonation, and unusual gestures and facial expressions. Deficits in social-emotional reciprocity and understanding are also likely to manifest in legal settings, particularly as they relate to navigating interpersonal relationships and conflict. Unusual or atypical interests are common, with many persons with ASD appearing obsessed with objects or activities unusual within the individual’s socio-cultural context. Furthermore, persons with ASD commonly experience sensory sensitivities or insensitivities to varying stimuli, including noise, temperature, texture or light. This can promote distress and agitation if the individual cannot control their exposure to these stimuli.

### [60.7130] Assessment

Assessment of ASD has been predominantly developed around the DSM-IV criteria, although research indicates that well validated DSM-IV assessments identify those also meeting DSM-V criteria.

*Autism Diagnostic Interview-Revised* (ADI-R: Lord, Rutter & Le Couteur, 1994; Rutter, Le Couteur & Lord, 2003)

The ADI-R, alongside the Autism Diagnostic Observation Schedule (ADOS: Lord et al., 2012; Lord et al., 2001), is increasingly becoming the “gold standard” in the assessment of ASD (Gray, Tonge & Sweeny, 2008). The ADI-R is a standardized, semi-structured, investigator based interview for parents and caregivers of people referred for possible ASD. The ADI-R comprises 93 items in three domains of functioning: language/communication; reciprocal

social interactions; and restricted, repetitive, and stereotyped behaviours and interests (Kim & Lord, 2014). In general items are coded 0 (no evidence of abnormality), 1 (some evidence of abnormality), and 2 (evidence of marked abnormality). Higher scores are indicative of greater abnormality. The ADI-R produces an algorithm which is linked to ICD-10 (World Health Organisation, 1992) and DSM-IV (American Psychiatric Association, 2000) diagnostic criteria (Gray et al., 2008). Algorithms have been developed with good sensitivity (0.85) and specificity (0.85) for children as young as 12 months. The ADI-R requires considerable training and takes about 2 to 2½ hours to administer.

*Autism Diagnostic Observation Schedule (ADOS: Lord et al., 2012; Lord et al, 2001)*

The ADOS is a semi-structured, standardized observational assessment of the child's communication, social interaction and play. The ADOS has four modules reflecting the different levels of an individual's expressive language level. It provides a series of standardized contexts in which the child's social, communication and repetitive, stereotyped behaviours can be observed. Standardized toys and activities are used to present opportunities for social and communicative interaction with the examiner. During these activities, absence or presence of the child's behaviour of interest is recorded. Higher scores are indicative of greater abnormality (Gray et al, 2008). An ADOS Toddler Module has been developed (ADOS – T: Lord et al, 2012) which is relevant for use in minimally verbal children aged 12 to 30 months. The original validation study of the ADOS demonstrated internal consistency across modules ranging from alpha .47 to alpha .91. Inter-rater reliability exceeded 80% (Hus & Lord, 2014).

*Diagnostic Interview for Social and Communication Disorders (DISCO: Wing et al, 2002)*

The DISCO is a semi-structured, standardised developmental history interview that guides the investigator in collecting a detailed profile of an individual's strengths and difficulties. Generally, the parent or caregiver is interviewed and the questions focus on seven broad areas: Family and medical background; Infancy; Developmental skills; Repetitive and stereotyped activities; Emotions; Maladaptive behaviour; and Interviewer's judgement of quality (Carrington et al, 2015). The DISCO has been validated against both the ADI-R (Nygren et al, 2009) and the ADOS (Maljaars et al, 2012). The DISCO takes a dimensional approach, obtaining a profile of development and behaviour, and it takes about two to three hours to complete. The DISCO was developed independently of both the DSM and ICD and thus is robust to changes in diagnostic systems (Leekam et al, 2002). The DISCO has high reported inter-rater reliability (kappa .75 or higher) for over 80% of items (Wing et al, 2002) and an abridged form of the DISCO has recently been validated for the purposes of screening (Carrington et al, 2015).

## **[60.7140] Screening Measures**

Screening measures can assist in identifying children and adults who require a more comprehensive diagnostic assessment for ASD. Many of the screening measures rely on parent/caregiver and/or teacher report, however they should be administered by the clinician to ensure correct interpretation of items, and clarification where needed. While there are many ASD screeners, the following have the greatest empirical support:

*The Autism Detection in Early Childhood (ADEC: Young, 2007)*

The ADEC is an effective, validated screening tool for identifying autism and autistic tendencies in children as young as 12 months through to 3 years of age. The ADEC takes 10 to 15 minutes to administer and can be used with little training by general practitioners, maternal healthcare nurses, allied healthcare professionals, psychologists, therapists and special needs educators. The ADEC also assists in identifying specific behaviours that require intervention and is able to discriminate autistic behaviours from learning disabilities and other developmental disorders with a high level of sensitivity and specificity. The ADEC has sound

psychometric properties (including inter-rater reliability: Nah, Young, Brewer & Berlinger, 2014) and is able to predict symptom severity in later childhood (Nah, Young and Brewer, 2014).

The *Modified Checklist for Autism in Toddlers – Revised, with Follow up* (M-CHAT-R/F: Robins et al, 2014; Robins, Fein & Barton, 1999)

The M-CHAT-R/F is a 2-stage parent-report screening tool to assess risk for Autism Spectrum Disorder (ASD). The M-CHAT-R/F is available for free download for clinical, research, and educational purposes. Download of the M-CHAT-R/F and related material is authorized from [www.mchatscreen.com](http://www.mchatscreen.com). The M-CHAT-R/F has been translated into many languages. The M-CHAT-R can be administered and scored as part of a child care visit, and also can be used by specialists or other professionals to assess risk for ASD. The primary goal of the M-CHAT-R is to maximize sensitivity, meaning to detect as many cases of ASD as possible. Therefore, there is a high false positive rate, meaning that not all children who score at risk will be diagnosed with ASD. To address this, the authors developed the Follow-Up questions (M-CHAT-R/F). However, even with the Follow-Up, a significant number of the children who screen positive on the M-CHAT-R will not be diagnosed with ASD; however, these children are at high risk for other developmental disorders or delays, and therefore, evaluation is warranted for any child who screens positive. The M-CHAT-R can be scored in less than two minutes. Robins (2008) reports that the Positive Predictive Value (PPV) of an earlier version of the M-CHAT-R/F (M-CHAT: Robins et al., 2001) varies between samples classified as Level 1 (general paediatric populations, generally employing caregiver reports) and Level 2 (those at risk of ASD, usually involving direct observation by specialists), ranging from .11 for Level 1 and .60 for Level 2. However, Robins (2008) notes that these PPV's increase substantially to .65 and .76 for Level 1 and Level 2 samples respectively when a follow up interview was included. More recently Chlebowski et al (2013) reported that, using the M-CHAT/F (Kleinman et al., 2008) with a Level 1 sample of 18,989, 54% of children screened positive presented with ASD, while 98% screening positive presented with a clinically significant developmental concern warranting intervention.

*Social Responsiveness Scale – Second Edition* (SRS-2: Constantino & Gruber, 2012)

The Social Responsiveness Scale - Second Edition (SRS-2) is a highly regarded autism assessment that offers the convenience of a screener and the power of a diagnostic tool. The SRS-2 identifies social impairment associated with autism spectrum disorders (ASDs) and quantifies the severity. It's sensitive enough to detect subtle symptoms, yet specific enough to differentiate clinical groups, both within the autism spectrum and between ASD and other disorders. Administered in 15 to 20 minutes, the SRS-2 is applicable from age 2 1/2 through to adulthood. The pre-school form is generally completed by the parent or teacher while for school aged children, the teacher's observations are sufficient. Correlations between SRS scores and the Autism Diagnostic Interview – Revised (ADI-R: Lord et al., 1994; Rutter, Le Couteur & Lord, 2003) algorithm scores for DSM-IV criterion sets were of the order of 0.7. SRS scores were unrelated to I.Q. and exhibited inter-rater reliability of approximately 0.8. The SRS-2 is among the most favoured screening tools used by Autism Spectrum Australia (ASPECT).

*Social Communication Questionnaire* (SCQ: Rutter, Bailey & Lord, 2003)

The SCQ is a screening tool based on the ADI-R. Previously known as the Autism Screening Questionnaire (ASQ), this brief instrument helps evaluate communication skills and social functioning in children who may have autism or autism spectrum disorders. Completed by a parent or other primary caregiver in less than 10 minutes, the SCQ is a cost-effective way to determine whether an individual should be referred for a complete diagnostic evaluation. The questionnaire can be used to evaluate anyone over age 4.0, as long as his or her mental age exceeds 2.0 years. It is available in two forms, Lifetime and Current, each composed of just 40

yes-or-no questions. Both forms can be given directly to the parent, who can answer the questions without supervision. Validation studies indicate that the SCQ discriminates well between ASD and non-ASD cases with a sensitivity of 0.85 and a specificity of 0.75 (Berument et al, 1999). Using the SCQ with children with and without ASD developmental disorders, Chandler et al, (2007) reported strong discrimination between ASD and non-ASD cases (sensitivity 0.90 and specificity 0.86). The SCQ is also one of the most favoured screening tools used by Autism Spectrum Australia (ASPECT) and can be used with adults.

*The Autism Spectrum Quotient (AQ: Baron-Cohen et al, 2001)*

The AQ is a widely used self-report measure of autistic traits in adults with a normal IQ. The AQ comprises 50 questions assessing 5 areas: social skill; attention switching; attention to detail; communication; and imagination. The scale has been shown to have good specificity, test-retest and inter-rater reliability, and moderate construct validity.

*The Ritvo Autism and Asperger Diagnostic Scale- Revised (RAADS-R: Ritvo et al, 2011; Ritvo et al., 2008)*

The RAADS-R is a self-report measure designed specifically to screen for ASD in adults with average intelligence. The RAADS-R comprises 80 items informing 4 scales/symptom areas: language; social relationships; sensory-motor; and circumscribed interests. The authors report excellent concurrent validity with the SRS-2 (see above), as well as high sensitivity (97%) and specificity (100%). Test-retest reliability was also high ( $r = .99$ ) and internal consistency was good (0.79 to 0.91 across symptom areas).

*Adult Autism Subthreshold Spectrum (AdAS Spectrum: Dell'Osso et al, 2017)*

The AdAS Spectrum is a recently developed measure of 160 items exploring the wide spectrum of ASD across seven domains: childhood/adolescence; verbal communication; non-verbal communication; empathy; inflexibility and adherence to routine; restricted interests and rumination, and hyper- and hypo- reactivity to sensory input. The AdAS Spectrum is not a diagnostic instrument but rather seeks to identify the broad variety of manifestations associated with ASD across the lifespan which may be present in people who do not meet diagnostic threshold. The authors report excellent internal consistency for the total score (Kuder-Richardson coefficient of 0.96) and across the seven domains (0.76 to 0.80). Test-retest reliability was reported as above ICC = 0.98, and convergent validity was reported for the AQ (see above:  $r = 0.76$ ) and the brief version of the RAADS-R (see above: 0.83).

## Attention Deficit Hyperactivity Disorder (ADHD)

### [60.7250] Introduction

Attention Deficit and/or Hyperactivity Disorder (AD/HD) is a relatively common, lifelong condition consisting of developmentally inappropriate patterns of impulsivity and/or inattention (APA, 2013). AD/HD is highly heterogeneous in both the combination of symptoms a person demonstrates, and in the profile of relevant neuropsychological deficits. People with the inattention profile demonstrate a marked difficulty sustaining attention on repetitive, long tasks that are not intrinsically rewarding. They experience difficulty completing tasks, planning, staying focused, organising tasks and are often forgetful or distracted. People with the hyperactivity symptoms tend to move or fidget, leave their seat, and when children, run or climb inappropriately. Adults with AD/HD might talk excessively and interrupt. A diagnosis at any age requires reports of significant symptoms before the age of 12 in multiple environments.

AD/HD interferes with occupational and educational outcomes and is associated with poorer levels of attainment in both as well as with low frustration tolerance, irritability and a higher

risk for anxiety, depression and substance use disorders (DSM). Just 14% of adults with AD/HD have it as a single diagnosis (Bramham, 2016). Difficulties in reading comprehension are common and reading acquisition is particularly difficult. The prevalence of AD/HD in prisoner populations is greater than that in non-custodial populations and is associated with greater levels of anger and violence (Byrne et al, 2015).

### **[60.7260] Diagnosing AD/HD**

A diagnosis of AD/HD can be made based on an interview reporting sufficient symptoms. Corroborating evidence should be sought from educators and cognitive and neuropsychological testing but is not required. First line treatment is pharmacological in combination with psychological interventions to manage problem behaviours.

### **[60.7270] Assessment**

Neurocognitive assessment of AD/HD targets Working Memory, Inhibition Control and Processing Speed (for more please see Chapter 68, Neuropsychology, by Crowe). However, several specific assessments have been developed:

#### *Brown Attention-Deficit Disorder Scale (BADDs)*

The BADDs is a short self-report form with parent, teacher forms for ages 3-12 and additional self-report scales for 8-18 years. The BADDs takes 15 minutes to complete, has variable convergent validity (0.48-0.94) and can discriminate people with AD/HD from those without. The strength of the BADDs is the availability of normative data and the fact that it measures the underlying executive function deficits of AD/HD rather than symptom frequency and severity (Brown, 2001). The BADDs provides five clinical subscales and a Total score. Internal consistency of the scales is high (between 0.70 and 0.98) and inter-rater reliability of the parent and teacher forms is moderate to high (0.61-0.93) while the inter-rater reliability of the self-report forms for ages 8-12 is moderate (0.40-0.60; Collett, Ohan & Myers, 2003). The BADDs also has an adult self-report form which consists of 40 items. The adult version is reported as producing 4% false negatives and 6% false positives (Brown et al., 2011), and internal consistencies are reported as being as high as  $\alpha = .97$  (Peter et al, 2016).

#### *Conners Rating Scale – Revised (CRS-S)*

The CRS-S is a self-report measure of AD/HD symptoms and related problems with parent, teacher and adolescent versions for ages 3-17 (Conners, 2008; 1997). The CRS-S provides a clinical cut-off to aid diagnosis, referring to the subscales of the DSM-IV, and a Global Index which is sensitive to the effects of treatment. The CRS-S has a psychometrically supported factorial structure including the following factors: Cognitive Problems/Inattention, Hyperactivity, Opposition, Anxious-Shy, Perfectionism, Social Problems and Psychosomatic. There is a large body of normative data available and the psychometric properties are generally strong. Internal consistency is high across all forms (0.73-94), but inter-rater reliability is poor (0.08-0.53 across all forms). Test-retest reliability ranges from poor to high (0.13-0.78) for parent and teacher forms (0.47-0.88) and high for adolescent forms (0.73-0.89). The CRS-S can discriminate AD/HD from nonclinical samples with high sensitivity, specificity and positive predictive power across all forms but is not sufficient to assess comorbidity (Collet, Ohan & Myers, 2003). An adult version, the Conners Adult ADHD Rating Scale has also been developed; however, item-level concordance rates are reported as slight to fair and the test has demonstrated a poor balance of sensitivity and specificity in predicting ADHD diagnosis (Voorhees, Hardy & Collins, 2011).

### **[60.7370] Specific Learning Disorder**

Specific Learning Disorder (SLD) is a distinct deficit of fundamental academic skills that interferes with learning in the developmental period and includes disorders previously referred

to as dyslexia (inaccurate reading, poor comprehension, poor spelling, poor writing) and dyscalculia (poor number sense, poor maths reasoning). These deficits are quantifiably below age expectations and significantly impede educational and employment achievement. SLD varies from difficulty with accurate reading that can be overcome with specific strategies and effort, through to a severe inability to comprehend writing (DSM-5, 2013). Most SLD diagnoses are for reading difficulties (dyslexia), with prevalence estimated at 10% of the population (Australian Dyslexia Association). SLD is a lifelong deficit for which compensatory skills can mitigate impairment, but will not resolve with age. A diagnosis of SLD requires standardised clinical assessment of specific cognitive skills (phonological), ability (IQ) and academic achievement, and possibly of functional abilities. For more information on relevant tests see Chapter 68, Forensic Neuropsychology.

### **[60.7470] Language Disorder**

A Language Disorder (LD) is an apparent difficulty in the acquisition and use of language in all modalities (spoken and written) regardless of language spoken (DSM-5, 2013). Deficits can be in comprehension or production of speech with specified subtypes including vocabulary, grammar or discourse. A LD will become apparent as developmental milestones are failed and language use is below age expectations. Assessment and potential treatment is conducted by a Speech Pathologist. LD is associated with significant functional difficulties and with psychiatric comorbidity especially in late adolescence and young adulthood, which is remediated, however, by educational support (Bao, Brownlie & Beichman, 2015). Data on prevalence is limited in Australia but an estimated 12-13% of school-aged children have a Communication Disorder (Senate Standing Committee on Community Affairs, 2014; McLeod & McKinnon, 2007).

## *Tests of Mental Health*

### **[60.8010] Introduction**

Mental health disorders are commonplace in the criminal justice system. Although highly varied in their presentation, all such disorders involve cognitive, affective, and somatic changes, which influence an individual's behaviour. Hence, understanding the nuances of such disorders and how they influence behaviour is paramount to forensic assessment, as the presence of a mental health disorder can help to explain why an individual may have acted/or failed to act in a certain situation. Therefore, the key features of the most common mental health disorders and some of the psychological tools used to assess for their presence are discussed below. Please note this is not a comprehensive discussion of every disorder listed in the DSM-5, nor is it a compendium of all the assessments tools related to each diagnosis.

### **[60.8020] Screening Measures**

*Kessler Psychological Distress Scale (KPDS: Kessler et al, 2002)*

Designed to assess non-specific psychological distress in the general population the Kessler Psychological Distress Scale is commonly used by government in censuses and surveys as well as in health and mental health settings (Kessler et al, 2002). Both the K10 (10 items) and the K6 (6 items) assess mood, arousal, fatigue and common beliefs indicating mental distress (Kessler et al, 2003). Both have high internal reliability (Cronbach's alpha=0.92 for the K10 and 0.89 for the K6; Kessler et al, 2002); the K10 is the gold standard screening assessment for psychological distress (Kessler et al, 2003). High correlations indicate the validity of the measure across age, sex and educational backgrounds (Kessler et al, 2002). High scores on the K10 and K6 are significantly associated with an anxiety or mood, and to a lesser degree with

other mental disorders (Andrews & Slade, 2001). Despite this and a second association between high scores on the K10 and symptom severity (Kessler et al, 2003) there are, as yet, no agreed cut-offs indicating a clinical syndrome (Andrews & Slade, 2001). The K10 can distinguish the degree of psychological distress of a mentally unwell person from that of a well person but is insufficient for diagnosis (Kessler et al, 2002).

*Symptom Checklist 90 Revised (SCL-90-R: Derogatis & Unger, 2010; Derogatis, 1994)*

The SCL-90-R is a 90-item self-report inventory of symptoms and psychological distress. Normative data is available for people aged 13 and above concerning the 9 symptom dimensions assessed: Somatization; Obsessive-Compulsive; Interpersonal Sensitivity; Depression; Anxiety; Hostility; Phobic Anxiety; Paranoid Ideation; and Psychoticism. The SCL-90-R also provides a Global Severity Index, a Positive Symptom Distress Index, and the Positive Symptom Total (PST; Derogatis & Unger, 2010). High internal consistency for all subscales was found in both primary care and outpatient clinical samples indicating that the SCL-90-R is a useful screen for various psychological symptoms (Cronbach's  $\alpha=0.78-0.97$ ; Schmitz et al, 2000).

*Depression Anxiety and Stress Scale (DASS; Lovibond & Lovibond, 1995 a&b).*

Perhaps one of the most widely used screening measures; the DASS is a set of three self-report scales that assess for the presence of symptoms associated with depression, anxiety and stress. The basic version contains 42 items. Each item is scored on 4-point severity/frequency scale and individuals are asked to rate the extent to which they have experienced each symptom over the past week. The totals from each scale are calculated producing three final scores. These scores are then compared to clinical cut-offs indicating the severity of each set of symptoms (Lovibond & Lovibond, 1995a). The DASS is normed for use with individuals 17-69 years old; however, research suggests that it could also be used in adolescents as young as 12 years old (Carr, 2015). A shorter version, with 7 items per scale is also available (DASS 21); however, the DASS-42 is recommended as the more reliable measure (Henry & Crawford, 2005).

The psychometric properties of the DASS are strong. Reliability coefficients for each scale are high, (depression  $\alpha = .91$  to  $.97$ ; anxiety  $\alpha = .81$  to  $.92$ ; and Stress  $\alpha = .88$  to  $.95$ ) indicating excellent internal consistency (Gloster et al, 2008). The DASS has also been shown to have excellent construct validity, discriminating well between symptoms of depression, anxiety and stress in clinical and non-clinical samples and each scale has shown to correlate well with other measures of similar symptomology (Crawford & Henry, 2003). Note these reported statistics refer to adult populations only, and do not include adolescent samples.

## **Mental Health Disorders – Psychotic**

### **[60.8120] Introduction**

Psychotic disorders refer to the category of mental health disorders that are characterised by disturbances of perception. This includes symptoms such as delusional beliefs, hallucinations, disorganised thought and behaviour, which cause substantial distress and impairment in functioning. Mental health diagnoses included in this category include Schizophrenia, Schizoaffective Disorder, Schizophreniform Disorders, Substance-Induced Psychosis and Delusional Disorder. Each of these diagnoses varies significantly in their actual clinical presentation, and within each diagnosis significant variability exists in terms of symptom severity and functional disturbance. For example, symptom presentation is different in schizophrenia depending on the stage of illness. For further information on the aetiology and presentation of psychosis and schizophrenia spectrum disorders refer to Part 2 of this chapter.

Discussed below are some commonly used assessment tools to ascertain the presence/absence of psychotic disturbance, as well as the severity of symptoms. Please note that a majority of these assessment tools are cross diagnostic that is they are designed to assess for the presence/absence of psychotic symptomology which is common across many Schizophrenia spectrum disorders. Importantly, the assessment of psychotic illness requires collateral information – a single assessment measure is not enough to warrant a diagnosis. Similarly, to ascertain whether symptomology is the result of drug use, assessment must be conducted following a period of sobriety.

### [60.8130] Assessment

*Assessment Schedule for Affective Disorders and Schizophrenia* (SADS; Spitzer & Endicott, 1978).

The SADS is a semi structured diagnostic interview which assesses for the presence or absence of mood and/or psychotic disorders. The SADS is based on the Research Diagnostic Criteria (RDC), which is a different diagnostic schedule to the DSM, however, it is still considered valid. In the SADS, screening questions based on the RDC diagnostic criterion are asked for all disorders, and positive responses indicate further questioning is warranted. With positive responses, respondents are also asked to rate the severity and duration of their symptoms, as well as whether such symptoms represent a marked change from normal functioning. The SADS is divided into two parts – part one assesses current psychopathology and part two assesses historical psychiatric functioning. Overall the SADS produces eight summary scores which correspond to different symptom clusters (*ie, depressive mood and ideation, endogenous features, depression –associated features, suicidal ideation and behaviour, anxiety, manic syndrome, delusions-hallucinations, formal thought disorder*). In addition the SADS also provides detailed description of the individual's current episode of illness and severity of disturbance, a description of the severity of symptoms in the week preceding the interview, progression of symptoms over time, and details of past psychopathology and functioning.

There are three versions of the SADS available, each with a specific focus; the regular version which produces a cross sectional diagnosis, the lifetime version which looks for the presence of disorders across the lifetime (SADS-L) and the change version which is designed to assess for changes in symptoms (SADS-C).

The psychometric properties of the SADS are sound, with numerous studies reporting the strong reliability and validity of this measure (for a comprehensive review see Rogers, Jackson & Cashel, 2001). Across all summary scales the SADS demonstrates good internal reliability ( $\alpha = 0.58- 0.97$ ) excellent inter-rater reliability ( $r = 0.80 - 0.99$ ) and good test re-test reliability at between 48 hrs and one week intervals ( $r = 0.49 - 0.93$ ; Endicott & Spitzer, 1978). The SADS has also demonstrated sufficient convergent validity with other broad measures of symptom psychopathology (eg, SCL-90). Although, notably some of the SADS summary scores show better convergence than others (e.g., correlations ranges from 0.15 – 0.68; Endicott & Spitzer, 1978).

A child version is also available, the *Kiddie-SADS* (K-SADS), which assess for similar symptomology in individuals aged 6-18 years old. The KSADS combines interviews with both the parents and the child as well as reports from clinical observations during the interviews. The KSADS displays good psychometric properties, with research reporting high inter-rater agreement and high test-retest reliability across a majority of diagnoses ( $k = .77- 1.00$ ). Of note, the test-retest reliability for the KSADS is not as strong for PTSD and ADHD, although still good ( $k = .63-.67$ ; Kaufman et al, 1997). The KSADS has also demonstrated good concurrent validity and moderate-good predictive validity across the range of diagnoses (Jarbin et al, 2017).

*Structured Clinical Interview for DSM-5 Disorders – Clinical Version* (SCID-5-CV; First et al.

Please see discussion of SCID in later part of this section.

*Brief Psychiatric Rating Scale* (BPRS; Overall, & Gorham, 1988; Luckoff et al, 1986; Ventura et al, 1993a&b).

The BPRS is a clinician administered rating scale which assesses the severity of a variety of psychiatric symptoms, including those which characterised psychotic disorders (e.g., bizarre behaviour, unusual thought content). The BPRS is based on an interview with the client, as well as observation/reports of the client's behaviour from the previous 2-3 days. The standard BPRS contains 18 items and each item is rated on a 7-point scale (1 – not present, 7 extremely severe). Items are then summed to produce an overall score, with higher scores indicating more pathology. Although exact clinical cut-offs are debated, a score of 50 or more is generally considered indicative of psychopathology.

More recently an updated version, the *Brief Psychiatric Rating Scale – Expanded* (BPRS-E; Dingemans et al, 1995) has been published, which contains 24 items and covers a more comprehensive range of psychotic symptoms (eg, includes more questions related to affect; Dazzi, Shafer, & Lauriola, 2016). This version also contains more detailed administration instructions as well as additional probe questions for each symptom. The psychometric properties of the BPRS-E are robust, demonstrating excellent reliability, validity and sensitivity to change (see work by Burlingame et al. 2006 for review).

*Positive and Negative Syndrome Scale* (PANSS; Kay, Opler, & Fiszbein, 2000).

The PANSS is a clinician-administered measure of symptom severity in Schizophrenia. Overall, the PANSS contains 30 items, which are rated on a 7-point scale (1 = absent, 7 =extreme). The PANSS contains three subscales; *positive (P)* and *negative (N)* scales which assess the core symptomology of schizophrenia, as well as a *general psychopathology scale (G)*, which assesses more general symptoms such as anxiety, disorientation and insight. Scores are summed for each scale, as well as totalled overall. On the positive and negative scales, overall scores can range from a minimum 7 to maximum 49, and on the general psychopathy scale, overall score can range from minimum 16 to maximum 112. The PANSS can also produce five psychopathology factor scores from summing constituent items. These include *anergia, thought disturbance, activation, paranoid/belligerence and depression*. Finally, a composite score can also be produced, which shows the difference between negative and positive scores and thus reflects which symptom cluster is predominant.

For scoring, total raw scores for each scale, plus the five psychopathology clusters are converted to percentile ranks based on age based norms. Percentile ranks and the associated ranges are presented below.

Percentile Rank	Clinical Range
95 or above	Very High
75-94	High
26-74	Average
6-25	Low
5 or below	Very Low

The PANSS has demonstrated sufficient psychometric properties. Early research shows that across the three subscales, the PANSS demonstrates good internal consistency ( $\alpha = 0.73, 0.83,$  and  $0.79$ ), good test re-test reliability across a 3-month period ( $r = 0.80, 0.68,$  and  $0.60$ ) and

sufficient inter-rater reliability ( $r = 0.83, 0.85, \text{ and } 0.87$ ; Kay, Opler, & Lindenmayer, 1987). Similarly, the PANSS shows excellent convergent, discriminant and predictive validity (Kay et al, 2000).

*Birchwood Insight Scale (BIS; Birchwood et al., 1994)*

The BIS is a self-report measure designed to assess presence of insight in individuals with psychotic illness. The BIS contains 8 items, which cover three factors of insight including *awareness of illness* (ie, awareness of having a mental health illness), *need for treatment* and *relabelling of symptoms* (ie, attribution of symptoms as part of one's disorder). Individuals are required to respond to each question choosing between "agree", "unsure" and "disagree," which are scored on a 3-point scale range from 0 to 2. Subscales items are summed together to produce three subscale scores. These scores can also be added together to produce a BIS total score (range 0 -12). Higher scores indicate greater insight.

In terms of psychometric properties, the BIS shows good overall internal consistency ( $\alpha = 0.75$ ) and good test re-test reliability at 1 week interval for each subscale (awareness of illness  $r = 0.80$ , need for treatment  $r = 0.96$ , relabelling of symptoms  $r = 0.65$ ; Birchwood et al., 1994). Similarly, the BIS has also demonstrated good concurrent validity and good sensitivity to individual differences and change across time (Birchwood et al, 1994).

*Browns Assessment of Beliefs Scale (BABS; Eisen et al 1998).*

The BABS is a clinician administered semi structured scale, designed to assess insight and delusional content in psychiatric disorders. The scale identifies the dominant belief the client has had over the past week, and evaluates such belief across seven items. Each item is rated on a 5-point scale, (0 non-delusional, least pathological, to 4 delusional or most pathological) and cover aspects such as conviction, perception of others' view of the belief, rigidity, explanation of difference, attempts to disprove belief, insight, and ideas of reference.

The BABS shows good psychometric properties in a variety of different samples (eg, body dysmorphic disorder, OCD, mood disorders and schizophrenia spectrum disorders). Early research shows good internal consistency (cronbach  $\alpha = 0.87$ ), and good inter-rater and test re-test reliability ( $r = 0.95$ ; Eisen et al, 1998). In terms of validity, the BABS displays strong correlation with other measures of delusional content indicating good convergent validity. However, discriminant validity is poor, indicating the BABS is not an appropriate measure of severity in a specific diagnosis, but rather is appropriate for evaluating delusional content in general (Eisen et al, 1998).

## **[60.8140] Bipolar Disorder I and II**

Bipolar disorder is characterised by episodes of mania or hypomania (ie, elevated mood) in combination with periods of depression (ie, lowered mood; APA, 2013). Mania refers to a distinct period of abnormally and persistently elevated or irritable mood, whereby the individual displays symptoms such as inflated self-esteem/grandiosity, decreased need for sleep, pressured speech/racing thoughts, distractibility, increased goal directed activity, and excessive involvement in pleasurable yet risky activities (APA, 2013). To meet diagnostic criteria for Bipolar I the individual must display three or more symptoms of mania nearly every day for a period of at least 1 week and these symptoms must cause impairment in functioning (manic episode; APA, 2013). For Bipolar II, the individual must display the same symptoms, however the length and severity of symptoms is much less and the symptoms are not severe enough to cause significant impairment in functioning (hypomanic episode; APA, 2013). In both cases, mania and hypomania must be accompanied by a subsequent depressive episode. In some severe episodes of mania, psychotic symptoms can develop.

Clinically, manic episodes are distinguished by periods of elevated mood and behaviour, which appears incongruent with the individual's current circumstances. Similarly, individuals with

bipolar disorder display an increase in goal directed behaviour, whereby they freely express overly optimistic ideas, and embark on extravagant projects which are impractical and reckless based on their general lifestyle. Generally, this presentation also includes increased energy, increased sociability, talkativeness and over familiarity, whereby the individual is highly confident in themselves and in their ability to succeed in their goal. Consequently, these symptoms can cause significant functional impairment in terms of occupation, financial status, relationships and education, and can have significant impact on individual reputation as often behaviours are uncharacteristic to the individual (ie, overt sexual promiscuity). Episodes of mania are normally triggered by stressful life events, however, the length, frequency and pattern of episodes varies highly between individuals. On average periods of depression last longer than that of mania (Carr, 2015).

Discussed below are several measures designed to assess for the presence and severity of bipolar disorders. Note measures of depression (as discussed below) are also routinely used in bipolar assessment, to discern the nature of the individual's depressive presentation.

*Schedule for Affective Disorders and Schizophrenia – Lifetime Version* (SADS; Spitzer & Endicott, 1978)

Please see discussion of the SADS in the section on assessment of psychotic mental illness.

*Mood Disorder Questionnaire* (MDQ; Hirschfeld et al., 2000)

The MDQ is a self-report questionnaire designed to screen for the presence of bipolar disorder. It focuses specifically on the presence of mania or hypomania, in an attempt to distinguish symptoms from general depression. The MDQ contains 5 central questions, however, question 1 contains 13 parts, totalling 17 items overall. Questions are structured in a Yes/No format. In order to score a positive result, the individual has to report “yes” to seven or more of the 13 items in question 1, “yes” to question 2, and rate question 3 at least a “moderate or serious problem.” If the individual screens positive on the MDQ further testing is required before a diagnosis can be considered. Of note, as the MDQ is a screening questionnaire, it assesses lifetime history of symptoms, and thus should not be used as a measure of symptom severity.

The psychometric properties of the MDQ are sound. Demonstrated internal consistency is good ( $\alpha = 0.89$ ), and test re-test reliability at a one-month interval is fair ( $k = 0.79$ ; Weber-Rouget et al, 2005; Miller, Johnson, & Eisner, 2009). However, validity estimates are more variable. In clinical samples, the MDQ performs well, with specificity rated 0.90 and sensitivity reported at 0.73 (Hirschfeld et al, 2000). However, in community samples this sensitivity score decreases to 0.28, suggesting the MDQ is much better at identifying bipolar in clinical samples (Hirschfeld et al, 2003). Similarly, the MDQ has been shown to be much more sensitive to bipolar I than bipolar II (Weber-Rouget et al, 2005).

*Young Mania Rating Scale* (YMRS; Young, Biggs, Ziegler, & Meyer, 1978)

The YMRS is a clinician-administered interview designed to assess the severity of manic symptoms. It combines the individual reports of their symptoms over the previous two days, as well as observations from the clinician during the interview. Overall it includes 11 items which cover the central manic symptoms, as well as one item which assess presence of insight. Items are rated on either a 4 or 8-point severity scale and items which pertain to the four core symptoms of mania (irritability, rapid speech, bizarre content and disruptive/aggressive behaviour) are double weighted. Of note, the YMRS does not cover all the DSM criteria for mania. However, the YMRS can also be used in conjunction with self-reported measures of symptom severity, such as the Altman Self-Rating Mania (ASRM) Scale (Altman, Hedeker, Peterson, & Davis, 1997) or the Self-Report Manic Inventory (SRMI; Braunig, Shugar, & Kruger, 1996) to increase coverage of symptoms.

The YMRS has excellent psychometric properties, with high internal consistency ( $\alpha = 0.88$ ), test re-test reliability ( $r = 0.76$ ; Colom et al., 2002) and inter-rater reliability ( $r = 0.93$ ; Young et al., 1978). The YMRS also has good construct validity, correlating highly with other mania rating scales (Young et al, 1978)

## The Assessment of Mental Health Disorders – Non psychotic

### [60.8240] Introduction

Non - psychotic disorders refer to high prevalence mental health issues which do not have a psychotic component, such as mood, anxiety and stress disorders. These disorders are of particular relevance for forensic assessment, as research indicates high prevalence rates of such disorders in criminal populations (Lynch et al, 2014). Further, although mental health disorders vary considerably in their presentation, all such disorders involve cognitive, affective, and somatic changes, which influence an individual's behaviour. Hence, understanding how the various aspects of such disorders influence behaviour is paramount to understanding forensic assessments. The key features of each of these disorders and some of the psychological tools used to assess for their presence are discussed below.

### Mood Disorders

#### [60.8310] Introduction

Mood disorders refers broadly to a category of mental health disorders that are classified by primary disturbances in mood, including prolonged and persistently low mood (e.g., depression or dysthymia), and/or elevated and rapidly cyclic mood (eg, bipolar disorder type I and II).

#### [60.8320] Major Depressive Disorder

Major depressive disorder (MDD; American Psychiatric Association [APA], 2013), is marked by episodes of low mood, negative thinking patterns, and disturbances in normal sleep and appetite. To receive a diagnosis, the symptoms must be present nearly every day for a two-week period, demonstrate a marked difference to the individual's normal functioning, and be causing the individual significant distress, and impairment in social, occupational or educational functioning (APA, 2013). Depending on the severity of symptoms and impairment, depression can be classed as mild, moderate or severe. Similarly, if symptoms persist beyond two years, depression is then classified as dysthymia (APA, 2013).

In terms of presentation, depression is characterised by negative thinking styles in which the individual comes to view themselves, the world around them and the future in a hopeless, worthless and pessimistic fashion. Associated with this worldview the individual becomes stuck focusing only on the negative aspects of situations around them, which seemingly reinforces their perceptions. As a result, the individual often experiences intense sadness, loneliness, emptiness and despair. Consequently, their behaviour often becomes increasingly withdrawn or slowed, they tend to lack motivation or interest in daily activities, and find it difficult to concentrate or think clearly. Additionally, self-harm and suicidal ideation/behaviours are common, due to the claustrophobic sadness/hopelessness associated with this disorder (Carr, 2015).

Discussed below are several measures designed to assess for the presence and severity of major depressive disorder.

*Structured Clinical Interview for DSM-5 Disorders – Clinical Version (SCID-5-CV; First et al., 2016).*

The SCID-5 is an omnibus interview schedule designed to produce DSM-5 diagnoses for a wide variety of mental health disorders, including mood, anxiety and stress disorders. Organised into diagnostic modules, the SCID can be administered according to specific sections to confirm a suspected diagnosis, or overall as a broad measure. Each module contains interview probes which cover the core symptomology of each disorder. If there is an indication of disturbance from these questions, further supplemental questions are asked to assess symptom threshold and exclusion criteria, and thus clarify presence/absence of a disorder.

Given the nature of the SCID, it has demonstrated sufficient reliability across most of the modules. However, due to the amount of clinical judgement required in this interview test re-test and inter-rater reliability varies considerably (Williams et al., 1992). Validity estimates are also varied, with some modules demonstrating stronger concurrent and construct validity than others (Kranzler et al., 1996). Of note, however, this research is based on the previous versions of the SCID, with less evidence available regarding the psychometric properties of the current DSM-5 version.

*Depression Anxiety and Stress Scale (DASS; Lovibond & Lovibond, 1995).*

See discussion of DASS above

*Beck Depression Inventory (BDI-II; Beck, Steer & Brown, 1996)*

The BDI-II is a widely used 21 item self-report measure which assesses for the presence and severity of depressive symptoms in individuals 13 years and older. The BDI-II produces a single total score which is then compared to clinical ranges. Score cut-offs include 0 - 13 (normal range), 14 - 19 (mild depression), 20 - 28 (moderate depression) and > 29 (severe depression). The BDI-II is designed for use with individuals 13 to 80 years old and has been translated into multiple languages. A shorter 7-item version is also available (BDI-SF; Beck & Steer, 1993).

The statistical properties of the BDI-II (full version) are very strong, owing to its widespread clinical use. Reviews report excellent reliability, with internal consistency as high as  $\alpha = 0.90 - 0.92$  and test re-test estimates ranging from  $r = 0.73 - 0.96$  (Wang & Gorenstein, 2013). Similarly, the BDI-II has been shown to accurately discriminate depressed and non-depressed populations (construct validity) and correlates highly with other measures of depression (concurrent validity; Beck et al, 1996; Wang & Gorenstein, 2013).

*Hamilton Depression Rating Scale (HDRS or HAM-D; Hamilton, 1960).*

The HAM-D is a clinician rated questionnaire that assess for severity of depression in adults. Although several versions exist, the original HAM-D comprises 21 items, which are scored on a 3 or 5-point scale. Scores for the first 17 items are totalled to produce an overall score, which is then compared to clinical ranges. The remaining 4 items are used as clinical descriptors only. Score cut-offs are <10 (no depression), 10-13 (mild depression), 14-17 (moderate depression), and > 17 (severe depression). Administration of the HAM-D can be complemented with a structured interview guide (SIGH-D; Williams, 1988), which is recommended to increase reliability.

As one of the earliest scales to be developed for the assessment of depression, the psychometric properties of the HAM-D have been thoroughly investigated. Research indicates that reliability coefficients vary widely between different versions of the HAM-D (0.48 – 0.92), and reviews suggest the HAM-D has only adequate internal reliability (Bagby, Ryder, Schuller, & Marshall, 2004). Similarly, estimates of test re-test and inter-rated reliability are highly variable; however, the use of the structured interview guide has been shown to achieve

sufficient coefficients (Iannuzzo et al, 2006). Finally, the scales construct validity is also reported to be questionable. Hence, the HAM-D should not be used in isolation.

## [60.8330] Suicide & Self Harm

Although not categorised as mental health disorders in and of themselves, suicidality and self-harm behaviours present as key symptoms and comorbid factors in a variety of mental health disorders. Suicidality is broadly understood as any thoughts and behaviours related to the deliberate act of ending one's own life. Suicidality is conceptualised on an ideation-intention continuum, whereby there is a distinction made between thoughts about death versus deliberate intent to end one's life. Thoughts about death (i.e., suicide ideation), can be passive, active or persistent, and can fluctuate significantly across time and situations. Suicidal thoughts by themselves are not considered indicators of immediate high risk, *unless they are accompanied by intent and preparatory behaviours*. Suicide intention, whereby the individual has indicated a deliberate desire and motivation to end their life, is of more concern. Similarly, either suicidal ideation or intent which is accompanied by preparatory behaviours such as organising means, and/or personal belongings indicates highest risk. Importantly, suicidality is considered categorically different to deliberate self-harm (e.g., cutting or burning wrists), as this type of behaviour generally serves other functions (e.g., relieve tension), and occurs *without thoughts of death/intent to die*.

The assessment of suicidality requires a comprehensive risk assessment which broadly aims to ascertain the client's current level of suicidality, as well as their current needs for intervention (eg safety). This ranges from low risk, whereby the client can be appropriately managed in their daily life with support from family and mental health services, through to high risk, whereby the client is deemed unsafe and requires immediate inpatient support. Suicide risk assessments are part of routine clinical practice, and at the very least require a thorough assessment of: the client's current suicidal thoughts (including any evidence of planning), access to any identified means (eg do they have a weapon/medications), the client's level of intent to act upon thoughts, and the client's immediate support networks. Many clinical measures exist that can be used to augment this process (some which will be discussed below). However, such assessment is NOT an exact science and should be supplemented by a comprehensive clinical history, including previous suicide attempts, mental health diagnoses and current/past contextual factors. Collateral information from family/partners is important. Where there is justifiable concern regarding risk of harm to self, psychologists are required by law to ensure client safety and thus can breach client confidentiality to inform family of current risk.

*Personality Assessment Inventory (PAI; Morey, 2007) & Millon Clinical Multi-axial Inventory – IV (MCMI-IV; Millon, Grossman, & Millon, 2015).*

Both the PAI and MCMI contain subscales related to the assessment of suicidality and self-harm behaviours. For more detailed discussion of these measures please see section on personality assessment in this chapter.

*Beck Hopelessness Scale (BHS; Beck & Steer, 1988)*

The BHS is a 20-item self-report measure designed to assess level of hopelessness, or negative attitudes about the future, in adults aged 17 – 80 years old. The scale covers three aspects of hopelessness, including feelings about the future, loss of motivation and future expectations. Respondents are asked to rate each item either true or false, and a total score is produced by adding the number of items endorsed which indicate the presence of hopelessness. Scoring cut-offs include, 0-3 normal range/asymptomatic, 4-8 mild hopelessness, 9-14 moderate hopelessness and 14 + severe hopelessness. In general, scores above 8 are considered indicative of a higher risk of suicide.

In terms of statistical properties, the BHS has been normed in a variety of both clinical and non-clinical populations. According to the manual, the BHS demonstrates good internal

consistency in a variety of different clinical populations ( $a = 0.87 - 0.93$ ) and modest test re-test reliability at one week and six week intervals ( $r = 0.69$  and  $0.66$ ; Beck & Steer, 1988). Similarly, the BHS shows good construct validity with a variety of research indicating the BHS is a better predictor of suicidal intent ( $r=0.47$ ), in comparison depression and suicidal intent ( $r = 0.26$ ). The BHS also shows good convergent validity, correlating highly with overall BDI scores in a range of different clinical populations ( $r= 0.46 - 0.76$ ). Finally, the BHS has excellent predictive validity, with studies showing that in inpatient and outpatient samples, higher scores on the BHS reliability predicted eventual suicide at 5 – 10 year follow up (Beck & Steer, 1988).

*Beck Scale for Suicide Ideation (BSS; Beck & Steer, 1991)*

The BSS is a 21-item self-report questionnaire that identifies the presence and severity of suicide ideation in individuals 17 years and older. Broadly the BSS covers suicidal thoughts, plans and intent, as well as deterrents to suicide, the individual's openness to revealing suicidal thoughts, and presence and severity of previous suicide attempts. Five screening items are also included at the beginning of the scale to determine whether completion of the scale is necessary. All items are rated on a 3-point scale (0 = no desire and 2 = strong desire). All items are then summed to determine an overall score (ranging from 0 to 38), with higher scores indicating greater suicide risk. Generally, scores above 24 are considered clinically significant and indicate the client is at risk for suicide. However, positive endorsement on any items is considered worthy of further clinical investigation.

The BSS has demonstrated good psychometric properties. In clinical samples the BSS shows high internal consistency ( $a= 0.93 - 0.96$ ) and sufficient test retest reliability ( $r = 0.54$ ; Beck & Steer, 1991). Similarly, the BSS also shows strong concurrent and construct validity, with strong correlations shown between the BSS and clinician rated scales of suicidal ideation ( $r = .90 - 0.94$ ) as well as the BSS and other measures of similar symptomology (e.g., BDI and BHS; Beck, Steer, & Ranieri, 1988)

*Columbia-Suicide Severity Rating Scale (C-SSRS)*

The C-SSRS is a risk assessment questionnaire designed to assess lifetime suicide ideation and behaviour in individuals 12 years and over. The measure contains six yes or no questions, which cover desire to die, thoughts around suicide, suicidal intent, and suicidal behaviours. The number and sequence of questions depends on the individual answers, with more questions being asked following positive endorsements (e.g., questions examine how recently the thoughts/behaviour occurred and their severity). Overall the C-SSRS rates an individual's degree of suicidal ideation on a scale ranging from *wish to be dead* to *active suicidal ideation with a plan and intent*.

The C-SSRS is a highly adaptable tool, which can be used as a clinician-administered measure, or as a self-report measure. There is also a variety of different versions of the C-SSRS each designed for use in specific populations/contexts. This includes a version for use with individuals with cognitive impairment, a paediatric version, and brief screening versions designed for triage purposes (ie that assess suicide ideation/behaviour in previous three months or since last contact).

The C-SSRS has been used widely in clinical research and is considered by some as the "gold standard" for suicide assessment. Due to this widespread use, the psychometric properties of the C-SSRS have been explored in a variety of different populations. In general, the C-SSRS has demonstrated strong internal consistency ( $a=0.95$ ) and inter-rater reliability in terms of both suicidal ideation and behaviours (Posner et al., 2011; Kerr et al, 2014; Madan et al, 2016) Similarly, the C-SSRS is reported to show good predictive validity, good sensitivity and specificity for suicidal behaviour, and demonstrates good convergent validity with other similar scales (Kerr et al., 2013). However, more recently the C-SSRS has been criticised for both

under-identifying many cases of suicidal ideation and mis-classifying or over identifying others (e.g., low sensitivity and specificity; Giddens et al, 2014).

## [60.8340] Anxiety Disorders

Anxiety disorders refer to a category of mental health disorders which are distinguished by a) the existence of fear(s) or worry which are *irrational*, or *out of proportion* to an actual threat and b) in response to this fear the individual engages in avoidance behaviours to prevent facing their fear (APA, 2013). This is notably different from stress. Stress is the body's physiological reaction to a threat, which occurs naturally when the individual is under pressure. Anxiety, however, is a more complex cognitive, behavioural and affective reaction to the perception of stress. Additionally, whilst many people experience anxiety, this is distinct from anxiety disorders, whereby the combination of fear and avoidance must also be causing the individual a significant level of distress and impairment in social, occupational or education functioning (APA, 2013). Anxiety disorders are also distinct from disorders related to trauma, which are instead characterised by a fear in relation to an *actual* threat.

As there is a wide variety of anxiety disorders, this section is unable to detail the specific diagnostic and behavioural presentation of each disorder. However, while the stimulus which causes fear differs across anxiety disorders, the presentation of cognition, affect, arousal and behaviour is similar. In anxiety disorders, an individual's patterns of thinking become fixated around detecting and/or avoiding their feared outcome. As a result, they are generally hyper-vigilant to their environment and constantly preoccupied with the possibility of facing their fear. This constant state of hyper focus is associated with heightened affect, such as restlessness, tension, and uneasiness. Consequently, in order to avoid such uncomfortable affect, individuals often engage in avoidance behaviours to prevent the possibility of encountering their fear object/situation. Over time, this avoidance tends to increase, as it is immediately reinforcing. For example, individuals who are afraid of social judgment (social phobia) often withdraw from social contact in order to prevent having to engage with others.

Discussed below are several measures designed to assess for the presence of anxiety and anxiety disorders.

*Depression Anxiety and Stress Scale* (DASS; Lovibond & Lovibond, 1995).

Please see discussion regarding the DASS above.

*Anxiety and Related Disorders Interview Schedule for DSM-5 – Lifetime Version* (ADIS-5L; Brown & Barlow, 2014).

The ADIS-5 is a semi-structured clinical interview designed to diagnose past and present anxiety, mood, obsessive-compulsive, trauma and related disorders in accordance with the DSM-5. This interview provides in-depth assessment of symptoms as well as details on the functional impact of the problem eg, what situations are avoided. The ADIS-5L also briefly screens for a variety of other mental health disorders, and provides a diagnostic timeline in order to establish onset, remission and sequence of current and past problems.

The ADIS-5 is a relatively new version of this measure and as such the psychometric properties of this update have not yet been as extensively studied as its predecessor the ADIS-IV (DiNardo et al, 1994). However, the ADIS-IV has demonstrated fairly strong inter-rater reliability across the range of diagnoses (see Brown et al, 2001 for a comprehensive review).

A child version is also available, the *Anxiety Disorder Interview Schedule for Children- IV* (ADIS-IV-C/P; Silverman, & Albano, 1996), which assesses for similar symptomology in individuals aged 7 – 17yrs. This version includes both a child and parent form to increase reliability. The ADIS-IV-C/P is well validated; demonstrating strong internal consistency, test

re-test reliability and construct validity (Silverman et al, 2001; Wood et al, 2002). Similarly, the ADIS-IV-C/P has also been shown to correlate highly with other well-validated measures of child anxiety (concurrent validity; Simon & Bogels, 2009).

*Beck Anxiety Inventory (BAI; Beck & Steer, 1990)*

The BAI is a 21-item self-report measure that assesses the severity of cognitive, emotional and physiological symptoms of anxiety. Individuals are asked to rate the extent to which they have been bothered by each symptom over the previous week, on a four-point scale. All items are summed to produce an overall score. Score cut offs include 0-7 (minimal anxiety), 8-15 (mild anxiety), 16-25 (moderate anxiety) and 26-63 (severe). Generally, this item is used as a screener, from which the clinician can then administer more specific measures to understand the specific nature of the individual's anxiety. Although designed for use in adults aged 17 – 80 years old, the BAI also has acceptable reliability and validity in populations as young as 14 years old (Osman et al., 2002) and has been used across a variety of cultures and languages.

The psychometric properties of the BAI are excellent. Research indicates high internal consistency ( $\alpha = .90 - .94$ ) across a variety of different samples, and test re-test reliability ranging from .62 (7 week interval) to 0.93 (1 week interval) (Julian, 2011; Bardhoshi, Duncan, Erford, 2015). Similarly, the BAI displays good construct validity, producing scores consistent with other well-validated measures of anxiety (Julian, 2011).

*State- Trait Anxiety Inventory (STAI; Spielberger et al., 1983)*

The STAI is a 40 item self-report measure of anxiety for adults. The STAI contains two subscales: temporary state anxiety (S-Anxiety) which assesses current apprehension, arousal, tension, and worry; and trait anxiety (T-anxiety) which examines more long standing stable aspects of anxiety proneness, such as calmness and security. Both scales are scored on a 4 point system, however, state related questions are based on intensity of symptoms, whereas, trait related questions are based on frequency in general. Items are summed to produce two final subscale scores, with higher scores indicating greater anxiety. Scores can range from 20 – 80. Final raw scores are then converted to T scores, which are compared to age based norms (19-39, 40-49, 50-69). Notably the original manual contains norms for male prison inmates.

The psychometric properties of the STAI are sound. The original STAI manual indicates good internal reliability, with alpha coefficients ranging from .86 – .95 and test re-test reliability between  $r = .65 - .75$  across a 2 month period (Spielberger et al, 1983). Notably, given the S-anxiety is a transitory scale, the re-test reliability for this subscale is lower compared to T-anxiety, which is designed to be more resistant to change (Julian, 2011). However, despite good reliability, the overall validity of the STAI has been questioned, as research shows that the STAI has poor construct validity, in that it has trouble discriminating between anxiety and depression (Kennedy, Schwab, Morris, & Beldia, 2001), particularly in older populations (Kabacoff et al, 1997).

## **[60.8350] Obsessive Compulsive Disorder**

Obsessive compulsive disorder (OCD) is categorised in the DSM-V under the category of Obsessive-compulsive and related disorders (APA, 2012). The characteristic feature of OCD is the presence of recurrent, persistent intrusive thoughts/urges/images which the individual finds highly distressing and unable to control (obsessions). Similar to anxiety disorders, these obsessions centre on a feared outcome such as potential harm to self or others. In response to such obsessions the individual feels compelled to perform a response behaviour or ritual (compulsion) which aims to either alleviate immediate distress or neutralise the individual's feared outcome. These compulsions are often excessive and preservative and/or not connected in any realistic way to the feared event (Carr & McNulty, 2016). Contrary to public perception, OCD related obsessions and compulsions can relate to a wide variety of different

objects/behaviours and are not exclusive to cleanliness, including order/symmetry, counting/hoarding, safety/checking, and sexual/religious or moral issues. As with all mental health disorders, obsessive-compulsive thoughts/behaviours must cause the individual considerable distress and functional impairment in order to classify for a diagnosis (eg, repeated compulsive behaviours prevents the individual from leaving the house/attending work etc.).

Due to the nature of OCD, such psychopathology is associated with high levels of shame and often individuals perform covert rituals to avoid questions. Similarly, levels of insight into behaviours vary (some compulsions are not recognised by the individual) and sometimes individuals will have beliefs that talking about obsessions will cause them to come true (Carr, 2016). OCD is also highly comorbid with other mental disorders, mostly notably mood disorders, specifically major depression and other anxiety disorders (Adam et al, 2012). Therefore, comprehensive assessment of OCD requires a thorough investigation of an individual's overall mental health as well as their specific OCD related behaviours.

#### Assessment

*Yale-Brown Obsessive Compulsive Scale* (Y-BOCS: Goodman et al., 1989a,b)

The Y-BOCS is a clinician administered semi-structured interview designed to measure OCD presence and severity. It consists of the Yale-Brown Obsessive-Compulsive Scale Symptom Checklist (Y-BOCS-SC) and the Yale-Brown Obsessive Compulsive Scale-Severity Scale, which results in a detailed symptom checklist and a 10-item scale assessing the severity of obsessions and compulsions. The Y-BOCS-SC assesses 54 OCD symptoms across 17 categories of obsessions and compulsions using a dichotomous rating format to indicate the presence or absence of a given symptom. There are five factors with scales that have been developed from the Y-BOCS-SC: contamination/washing; aggressive/checking; religious/sexual; symmetry/ordering/repeating; and hoarding (Cullen et al., 2006). Variable reliability has been reported, with three of the five scales demonstrating good alphas (0.80) but two reporting inadequate alphas (Sulkowski et al, 2008), Validity has been variable, with small to moderate correlations with other measures of OCD (Mataix-Cols et al, 2004).

*Obsessive-Compulsive Inventory-Revised* (OCI-R: Foa et al., 2002)

The OCI-R is an 18-item self-report scale. It has six factors: washing; checking; ordering; obsessing; hoarding; and mental neutralising. The OCI-R assesses severity of each symptom on a 5-point Likert scale. Research has indicated sound internal consistency (alphas from 0.81 to 0.88: Abramowitz & Deacon, 2006) and test-retest stability (ranging from  $r = 0.70$  to  $0.91$ : Foa et al, 2002). Moderate to strong correlations with other measures of OCD have been reported, supporting construct validity (Sulkowski et al., 2008).

## *Tests of Trauma & Stressor Related Disorders*

### **[60.9010] Introduction**

Trauma and stressor related disorders refer to broad a category of mental health disorders in which an exposure to a traumatic event(s) or life stressor(s) serves as the basis for the individual's symptoms. Although previously characterised as anxiety disorders due to similarities regarding fear, arousal and avoidance, trauma/stress related disorder are now considered separate, as they present with distinct mood related disturbance i.e., anger, aggression, dysphoria and dissociation. For brevity this section will discuss the presentation and assessment of post-traumatic stress disorder (PTSD) and acute stress disorder only.

### **[60.9020] Diagnosis**

The diagnostic criterion of PTSD in DSM-5 is extensive and very specific. To receive a diagnosis, the individual must have been exposed to a traumatic event whereby there was

*actual or threatened death/injury/ sexual violence*; and as a result of this exposure the individual developed a variety of cognitive, affective and behavioural symptoms, which have persisted for longer than 1 month and are causing significant distress and impairment in functioning (APA, 2013). These symptoms are characterised by four core components: *intrusion* (eg re-experiencing symptoms such as nightmares/flashbacks), *active avoidance* (e.g., internal or external avoidance of trauma related triggers), *passive avoidance/emotional numbing* (eg detachment from others, persistent negative cognitions/mood regarding the trauma) and *hyperarousal* (e.g., persistent irritability, hyper-vigilance, sleep disturbance; Carr, 2015). If an individual displays these symptoms following an exposure to trauma but the duration of disturbance does not reach 1 month, the individual is given a diagnosis of Acute Stress Disorder (APA, 2013).

Clinically, the presentation of PTSD and acute stress disorder is highly varied among individuals and differs immensely depending on the nature and severity of the trauma. However, in general, trauma related disorders present as an intense reaction to fear, where the individual is unable to switch off after the threat has subsided. This presentation involves strong affective content, eg helplessness, guilt, shame, and anger, and extreme safety seeking behaviours, such as chronic hypervigilance to possible threat and/or emotional detachment. Rates of comorbidity are high, with estimates suggesting 80 – 85 % of individuals with PTSD meet the criteria for at least one other psychiatric disorder, most commonly depression, anxiety or substance use disorder (Creamer et al, 2001; Carr, 2015).

*The Clinician Administered PTSD Scale (CAPS; Weathers et al., 2001; 2013).*

Considered the gold standard in PTSD assessment, the CAPS is a structured clinical interview designed to produce a current or lifetime diagnosis of PTSD, and assess PTSD symptom severity, as well as the presence of PTSD symptoms over the past week. The current version, the CAPS-5 (Weathers et al., 2013), has 30 items. These items cover the twenty PTSD symptoms listed in the DSM-5, as well as questions pertaining to onset and duration of symptoms, level of distress, impact on functioning, response validity, overall PTSD severity and presence/absence of dissociative subtype. Each of the 20 symptoms are assessed for intensity and frequency and where possible are behaviourally defined.

In scoring the CAPS, the frequency and intensity of each symptom is combined into a single severity rating (0-4). Each of these ratings is then summed to provide an overall severity score. Cluster scores are also created by summing the severity ratings of the items which correspond to the different diagnostic criterion e.g., criterion B (items 1-5), criterion C (items 6-7), criterion D (items 8 -14), and criterion E (items 15-20) (Weathers et al., 2013). Severity scores are rated as *absent* (denied problem/report does not fit DSM-5 criteria), *mild/sub-threshold* (consistent with criterion, but problem not severe enough to warrant diagnosis), *moderate/threshold* (consistent with criterion and problem is clinically significant and warrants diagnosis), *severe/ markedly elevated* (problem is above threshold, difficult to manage and requires prominent intervention) and *extreme/ incapacitating* (symptoms far beyond threshold, problem is pervasive, unmanageable and high priority for intervention; Weathers et al., 2013).

Historically, the CAPS is a psychometrically robust measure and research demonstrates that the most recent version, CAPS-5, follows this trend. For diagnosis, CAPS-5 shows strong inter-rater ( $k = 0.78 - 1.00$ ) and test-re-test ( $k = 0.83$ ) reliability, as well as, strong correspondence ( $k = 0.84$ ) with diagnosis based on the previous versions of the CAPS. Similar statistics were also observed for severity scores. Finally, the CAPS-5 also demonstrates good convergent and discriminant validity when compared to other well-validated measures of PTSD and psychopathology (Carr, 2015; Weathers et al., 2017).

#### *Note*

The Life Events Checklist for DSM-5 (LEC-5; Weathers et al., 2013) is recommended to complement the criterion A questions in CAPS-5. The LEC-5 is a self-report measures that assess for potentially traumatic events across the lifetime.

*Trauma Symptoms Inventory (TSI; Briere, 1995)*

The TSI is a 100 item self-report measure designed to assess psychological sequelae associated with traumatic events in individuals 18 to 88 years old. The TSI is global measure of trauma and does not focus on any given event. The TSI has 10 clinical scales which correspond with symptoms related to trauma (anxious arousal, depression, anger/irritability, intrusive experiences, defensive avoidance, dissociation, sexual concerns, dysfunctional sexual behaviour, impaired self-reference, and tension reduction behaviour) and 3 validity scales, which assess for responses which invalidate results (atypical responses, response level, and inconsistent responses). Individuals are asked to rate the frequency with which they have experienced each item within the last 6 months on a 4-point scale. Scores are summed to produce an overall raw score. Raw scores are converted to T scores which are then compared to norms which are based on gender and age.

The TSI is a well validated psychological measure which has been normed in samples of both the general population and military personnel. The clinical scales of the TSI have demonstrated good internal consistency (mean  $\alpha = .87$ , range  $\alpha = .74- .90$ ) across a range of different samples (Norris & Raid, 1997). Similarly, validity reports are good, with the TSI demonstrating sufficient construct validity, as well as, high sensitivity (.92) and specificity (.91; Briere, 1995,1996). Finally, the TSI also demonstrates good convergent validity, as both the clinical and validity subscales on the TSI correlate with similar scales on other measures (e.g., the ATR correlated at  $r = .52$  with the Negative Impression Management scale on the PAI) (Briere, 1995, 1996; Norris & Raid, 1997).

A new version, the TSI-2, is available (Briere, 2011). The TSI-2 uses a similar format to its predecessor, although it has 12 symptoms scales (nine of the above plus 3 additional symptom scales; insecure attachment, somatic preoccupations, suicidality) and 2 new subscales (hyperarousal and other directedness). Similarly, the validity scales have been re-designed to now also assess for potential misrepresentation of PTSD. The TSI-2 has been standardised to the normal population as well as in a variety of clinical population (victims of domestic violence, sexual abuse, incarcerated women and combat veterans). Given the TSI-2 is a relatively new measure the psychometric properties of this version are less established than the original and do require further study. However, the TSI-2 manual suggests strong internal consistency across all clinical and validity scales ( $\alpha = .80 - .94$ ) and good test re-test reliability at 1 – 13 day intervals ( $r = .66 - .94$ ) on all subscales, with the exception of the suicidality-behaviour scale ( $r = .15$ ; Briere, 2011). Similarly, the TSI-2 demonstrates good predictive validity, demonstrating high sensitivity (1.00) and high specificity (.88; Briere, 2011), including for samples of immigrants evaluated for the courts, where excellent test-retest reliability, validity and internal consistency has been reported (Filone & DeMatteo, 2017).

## *Tests of Alcohol and Drug Use*

### **[60.9510] Introduction**

Although the diagnostic conceptualisation of alcohol and drug use has been debated in recent years, the simplest way to understand substance use is by using a dimensional framework. That is, substance use occurs across a spectrum, ranging from abstinence, through social use, hazardous or harmful use, to substance “addiction.” The difference between “normal” use and that which classifies as a “substance use disorder” can be blurry, although is generally measured according to the presence of impaired control of use, social impairment, risky behaviours related to use, and tolerance and/or withdrawals.

In the DSM-V (APA, 2013) Substance Use Disorders is an umbrella term which encompasses a range of specific disorders related to certain substances (i.e., Alcohol Use Disorder, Cannabis

Use Disorder, Hallucinogen Use Disorder, Inhalant Use Disorder, Opioid Use Disorder, Sedative-, Hypnotic- or Anxiolytic Use Disorder, Stimulant Use Disorder, Tobacco Use Disorder). Broadly, to meet the criteria of any of the above, the individual's substance use must be causing distress and impairment in functioning, and be characterised by at least two of the following; using more than they intended or for longer than intended, repeated unsuccessful attempts to cut-down or quit, cravings for the substance, devoting a lot of time to use of and obtaining the substance, inability to control frequency or extent of use, failure to fulfil obligations (eg working, parenting, study), continued use despite known consequences (e.g interpersonal issues, occupational problems), recurrent use in hazardous situations, and/or the presence of tolerance or withdrawals.

Substance use issues co-occur often with mental health disorders, specifically depression and anxiety, and can also have a significant impact on an individual's physical health. Alongside this, substance use is also associated with a range of psychosocial vulnerabilities including homelessness, family violence, relationship breakdown, financial issues, unemployment, and criminal offending. Hence, a comprehensive assessment of substance use involves not only assessing substance related behaviours, but also the impact such behaviours are having on the individual's social and psychological functioning. This also includes conducting a thorough risk assessment, which covers both harm to self and others, including the impact on proximal children.

### [60.9520] Assessment

Listed below are three commonly used assessment tools related to alcohol and drug use. Please note there are many more available, which cover different substances of choice.

*Addiction Severity Index* (ASI-6; McLellan et al., 2006; 1992).

The ASI is a clinician administered semi-structured interview designed to assess impairments in functioning associated with substance use in adults. The ASI examines the individual's current functioning in seven broad areas including alcohol and drug use, medical status, employment/support, legal status, family and social relationships, and psychiatric/psychological status. In each area, lifetime history and current functioning (past 30 days) is assessed, both in terms of objective items (e.g., type, duration and frequency of problem) as well as subjective items (eg individual's level of distress, and motivation towards change in relation to the problem).

The ASI produces severity ratings (ISRs) on a 10-point scale (0= no real problem, treatment not indicated to 9 = extreme problem, treatment absolutely necessary) for each of the seven areas. Severity ratings are estimates of the client's current status, but cannot measure change over time. However, the ASI can also produce composite scores (CSs), which are calculated from certain items, which are capable of showing change, in each problem area. Composite scores, therefore, measure change in symptoms relative to the last 30 days, but are not indicative of current addiction severity.

The ASI is one of the most commonly used measures in the drug and alcohol field and as such has been normed to a variety of different populations. Despite this, the psychometric properties of the ASI are somewhat controversial as reviews suggest the inter-rater and test-retest reliabilities of both the severity ratings and composite scores range from poor to excellent, and high internal consistency is only shown for three of the seven composite scores (medical status, alcohol use, and psychiatric status; Makela, 2004). Alternatively, more recent research using the current ASI-6 suggests good internal consistency across sub scores ( $\alpha = .64 - .95$ ) as well as sufficient concurrent validity between the ASI-6 alcohol and recent research using the current ASI-6 suggests good reliability and validity (Kessler, et al, 2012).

*Note*

in the current manual of the ASI-6 the authors state that the ASI is NOT appropriate (due to the nature of questions) for use in adolescent samples, despite the fact that several versions have been produced by other authors.

*Alcohol Use Disorder Assessment Test (AUDIT-2; Babor et al. 2001).*

The AUDIT is a brief 10-item screening tool used to assess for alcohol-related problems, including alcohol consumption and drinking behaviours. The AUDIT is based on the ICD-10 conceptualisation of alcohol dependence, which differs to the DSM-V, however, is still valid and widely accepted. The AUDIT has both a clinician-administered version, which takes the form of a brief structured interview, as well as a self-report version. On the self report version, questions are scored on either a 3 or 5-point scale. A score of 8 or more indicates hazardous or harmful levels of alcohol use.

The AUDIT has been standardised across six different countries and has been studied in a variety of clinical settings and diverse populations. In general research shows that the AUDIT has strong reliability, showing high internal consistency ( $\alpha = .80$ ; de Meneses-Gaya et al, 2009) and high test re-test reliability at one month intervals ( $r = .84$ ; Selin et al., 2003). Similarly, the AUDIT displays acceptable sensitivity (.90) and specificity (.80) for current ICD-10 alcohol use disorders and correlates well with other measures of alcohol use suggesting good construct validity (Babor et al, 2001). For a comprehensive review of the psychometric properties of the AUDIT see de Meneses-Gaya et al. (2009).

*Substance Abuse Subtle Screening Inventory (SASSI; Miller, 1985; 1999).*

The SASSI-4 is a self-report screening tool designed to identify adults and adolescents who are at a high probability of having a substance use disorder. The SASSI aims to accurately assess the individual's current substance use status, by using both direct and indirect questions to circumnavigate lack of insight, denial or dishonesty. Direct questions ask the individual to rate how frequently on a 4-point scale (never to repeatedly) they have had certain experiences related to the use of alcohol and other drugs. Indirect questions, appear unrelated to substance use, and are scored as either true/false. Overall, the SASSI produces ten subscale scores (three face valid scales, four subtle scales, one reliability scale and two scales related to interpersonal functioning and risk for contact with the legal system) which are then converted to T-scores and compared to a normative sample.

Older versions of the SASSI have demonstrated sufficient psychometric properties, including high internal consistency ( $\alpha = .74$ ) and good convergent and discriminant validity (Sadeghi et al., 2010). Additionally, the current version, the SASSI-4 has demonstrated high sensitivity (93%) and specificity (90%) in large samples (Lazowski, & Geary, 2016). An adolescent version of the SASSI is also available (SASSI-A2). It uses the same format, however, is designed for clients 12- 18 years old.

## *Tests of Eating Disorders*

### **[60.10010] Introduction**

Eating disorders are characterized by a persistent disturbance of eating or eating-related behaviours that result in disrupted absorption or consumption of food which can lead to impairment of physical health or psychological functioning. Body image is often distorted by the condition and this contributes to a range of difficulties within eating disorders. Over one million Australians suffer from an eating disorder and it is estimated that approximately 15% of women will experience an eating disorder at some stage of their life. Females are generally over represented in this category of disorders (approximately 64% according to Butterfly foundation, 2012) but males are increasingly at risk of suffering the condition, especially over

the age of 15. Those who suffer an eating disorder experience higher rates of comorbid conditions, the most common being depression and anxiety, followed by substance abuse and personality disorder.

Eating disorders are not attention seeking efforts but a serious mental health condition that has the highest mortality rate of any psychiatric illness.

### **[60.10020] Anorexia Nervosa**

The diagnostic criteria (DSM V) for Anorexia Nervosa includes: reduced intake of food (energy) leading to significantly low body weight in relation to height, age and sex; an intense fear of gaining weight or becoming fat, or behaviour that interferes with gaining weight, even though weight is significantly low; disturbance in body image and a lack of acknowledgement of the serious nature of their current low body weight.

Due to the distortion in body image, patients are unable to recognize their condition worsening or developing because they do not consider themselves thin. Often family members will seek professional help when significant weight loss has occurred.

Anorexia nervosa often begins during adolescence or emerging adulthood when the body is experiencing hormonal and physical change. The onset of the disorder is commonly associated with a stressful life event. The course of the disorder is highly variable. Many have periods of eating, combined with relapses, some have a single episode and it does not reoccur. Many have a life long battle with varying degrees of eating problems that require professional help. Suicide risk is elevated in Anorexia Nervosa with rates as high as 12 per 100 000 per year. Assessment of risk of suicide is thus a crucial element to psychological assessment.

### **[60.10030] Bulimia Nervosa**

Diagnostic criteria (DSM V) for Bulimia Nervosa (BN) includes: recurrent episodes of binge eating; recurrent behaviours that prevent weight gain (laxative use, diuretics, medication, fasting, excessive exercise); binge eating and recurrent behaviours occur on average at least once a week for three months; self-evaluation is dictated by body shape and weight; and the disturbance does not occur during episodes of Anorexia Nervosa.

Individuals with Bulimia are often within the normal weight range or overweight according to Body Mass index (BMI). Females are far more likely than males to suffer from the disorder with an approximate 10:1 ratio. Like Anorexia, it usually begins in adolescence or emerging adulthood. The binge eating usually occurs after an episode of dieting or attempts to lose weight. Stressful life events can also trigger binge episodes. The course may be chronic or intermittent, with often periods of remission when weight remains stable or controlled. Like Anorexia, there is an increased risk of suicide so this must be examined during assessment.

### **[60.10040] Binge Eating Disorder**

The most common category of eating disorders is Binge Eating Disorder (BED). Unlike Bulimia Nervosa (BN), those suffering BED will not employ compensatory behaviours such as over-exercising or vomiting after they binge eat. Often those suffering BED are overweight or obese.

The diagnostic criteria for BED includes: Recurrent episodes of binge eating (amount of food larger than usual consumption for a meal and lack of control over eating during that episode ie. Unable to control what or how much they are eating); the binge eating episodes are associated with:

1. Eating more rapidly than normal;
2. Eating until feeling bloated and uncomfortable;

3. Eating large amounts of food when not feeling hungry;
4. Eating in isolation because embarrassed regarding the volume of food;
5. Feeling repulsed by their behaviour, depressed and guilty after the binge eating episode.

Other criteria include distress regarding their binge eating; the binge eating occurs on average, at least once a week for three months. The binge eating is not associated with the use of compensatory behaviours as per bulimia and does not occur exclusively during the course of BN or AN (DSM V).

Severity levels vary from Mild (1-3 binge episodes per week), Moderate (4-7 episodes per week), Severe (8-13 binge eating episodes per week) and Extreme (14 + episodes per week).

Remission rates for natural course and treatment in BED are higher than BN and AN. BED appears to be a persistent pattern of problematic eating that is comparable to BN in terms of severity and duration. Compared to other eating disorders, males have comparable rates to females for binge eating disorder.

### **[60.10050] Assessment of Eating Disorders**

Clinical assessment often involves a comprehensive interview that examines weight, eating patterns and a description of possible factors contributing to the development of the disorder (biological, social, psychological). Medical examination is also crucial due to the physical impact of eating disorders. The following psychological tests may be used to supplement clinical interview and further examine intricacies of the eating disorder.

*Eating Disorders Inventory 3rd edition* (EDI-3: Garner, 2004).

The EDI-3 is a 91 item self-report questionnaire that is divided in to 12 subscales examining items specific to eating disorder research. The current version of the EDI is a standardised test based on the literature examining eating disorders such as Anorexia nervosa and Bulimia. Research conducted by Clausen and colleagues (2011) demonstrated the clinical utility of the EDI-3 in measuring eating problems. Apart from 1 scale, all differences between patient and non-clinical control group yielded high effect sizes. Discriminative validity is considered good, as is internal consistency.

*Eating Disorders Examination* (EDE: Fairburn & Cooper, 1993).

The EDE is a semi-structured interview that prompts the clinician with a line of enquiry that seeks to examine symptoms indicative of an eating disorder. The EDE is considered the gold standard for assessing those at risk of an eating disorder in clinical practice. The EDE has utility for assessing a number of binge episodes (Bulimia). It has demonstrated very good inter-rater and test-retest reliability for assessing eating disorders.

*Eating Disorders Examination – Questionnaire* (EDE-Q: Fairburn & Beglin, 1994). The EDE-Q was adapted from the EDE. It is a 41 item self-report measure that examines concerning eating habits over a 4-week period. It provides a global score based on four subscales (Restraint, Eating concern, Shape concern, Weight concern). It is used to assess the presence of an eating disorder when the interview (EDE) is not practical. Studies examining test-retest reliability of the EDE-Q show very good reliability for Bulimic patients but less reliable results for binge eating disorders.

## ***Tests of Disruptive, Impulse Control and Conduct Disorders***

### **[60.10510] Introduction**

This category of disorders includes conditions that involve difficulties with self-control of emotions and behaviour, often violating the rights of others and conflicting with societal norms

and expectations. Many of the features of these disorders occur in developing individuals, thus it is critical that diagnosis is considered relative to what is normal for one's age, gender and culture before labelling with a condition that tends to be documented on school, medical and other files. Often violating social norms leads to contact with the criminal justice system and as a result, many of these disorders feature heavily in psychological reports for the courts.

### **[60.10520] Oppositional Defiant Disorder (ODD)**

The DSM V diagnostic criteria for ODD includes: a pattern of angry/ irritable mood, argumentative/ defiant behaviour/ vindictiveness; disturbance in behaviour is associated with distress in the individual or others in their immediate social context (family, work, friends) or impacts on important areas of functioning (employment, education, social); the behaviours do not occur due to a psychotic episode, substance use, depressive or bipolar disorder.

### **[60.10530] Conduct Disorder (CD)**

Diagnosis requires the following pattern of symptomology: repetitive and persistent pattern of behaviour that violates the rights of others/ or the age-appropriate norms. Three of the following criterion must be present in the past 6 months:

- Aggression to people or animals;
- Destruction of property;
- Deceitfulness or theft;
- Serious violation of rules.

*Childhood-onset type* includes symptoms of CD before age 10, *Adolescent-onset type* demonstrates symptoms after age 10, *Unspecified-onset* includes cases when there is minimal evidence of the age of onset of symptoms.

The above social violations often attract criminal charges (especially if repeated) and as a result CD features heavily for those involved in the Juvenile Justice System.

Often ADHD and Oppositional Defiance Disorder (ODD) are common for those diagnosed with CD. The comorbid presentation predicts worse outcomes for the individual.

The criteria for Anti-Social Personality Disorder (ASPD) requires evidence of CD before the age of 15 and disregard for social norms after the age of 15.

### **[60.10540] Pyromania**

Pyromania refers to a pattern of deliberate and purposeful fire setting. Individuals with this disorder experience arousal or tension leading up to, and after the fire setting episode. There is a fascination with fire, usually coupled with curiosity and attraction to the fire and its situational contexts. Often those with the disorder will watch fires in their neighbourhood, may set off false alarms and find pleasure associated with organisations that manage fires. Often they will spend time at fire departments and may become volunteer or paid fire-fighters themselves. Diagnoses are made if not better explained by conduct disorder, manic episode or antisocial personality disorder (DSM V).

### **[60.10550] Kleptomania**

The main feature of kleptomania is the recurrent failure to resist the impulse to steal items even though they are not needed for personal use or monetary value. The individual will experience rising tension before the theft and pleasure, relief or gratification after the theft. The theft is not better explained by conduct disorder, manic episode or antisocial personality

disorder. The objects are taken although they have little value to the kleptomaniac who often could afford to purchase the item. The theft is completed without assistance from others.

## **[60.10560] Assessment of Disruptive, Impulse-Control and Conduct Disorder**

*Barrett Impulsiveness Scale* (BIS-11: Patton, Stanford & Barratt, 1995)

The Barrett Impulsiveness scale is a 30 item self-report measure of the personality construct, Impulsiveness. The BIS-11 is the 11<sup>th</sup> revision of the scale that measures three sub traits: attentional impulsivity; motor impulsivity; and non-planning impulsiveness. The Barratt scale is considered the gold standard in the domain of impulsivity. Fifty years after the original scale was released a comprehensive review of the 11<sup>th</sup> edition (BIS-11) revealed new data supporting use of the scale (Stanford et al, 2009). Stanford et al. (2009) report internal consistency rates for the total score on the BIS-11 all within an acceptable range (0.71-0.83) suggesting the scale is reliable across diverse cultures.

*Strengths Difficulties Questionnaire* (SDQ: Goodman, 1997).

The SDQ is a brief behavioural screening questionnaire for children and adolescents (aged 3-16). There is a self-report scale and a teacher/parent form of the SDQ. For 3-11 years, the scale is to be completed by a parent or teacher, for 11-16 there is a self-report version. Each of the four scales examines emotional symptoms, conduct difficulties, hyperactivity, peer problems and pro-social issues. The four scales combine to produce a total score, which is compared against mean scores within groups. Analysing information obtained from teachers, parents and other health professionals can assist in the identification of issues to be referred for further assessment and/or intervention for a range of issues such as conduct disorder, hyperactivity or social difficulties.

There are Australian norms for the SDQ (Mellor, 2005) and the psychometric properties of the test are good. The test boasts good internal reliability, test re-test reliability and sound informant reliability across all three forms (teacher/parent/self). In NSW, Child and Adolescent Mental Health services are using the SDQ as a screen for behavioural difficulties and to suggest whether there is a need for further investigation of identified needs.

*Child Behaviour Checklist* (CBCL: Achenbach, 1991)

The Child Behaviour Checklist (CBCL) is a rating form provided to caregivers to assist in identifying problematic behaviour in children. The form may be repeated over time (at least six monthly intervals) and thus provides an indication of change in problematic behaviours.

The CBCL consists of three forms. One for parents of children aged 6-18 (CBCL 6-18), one for teachers to complete for children 6-18 (TRF 6-18) and a youth self-report form for 11-18 year olds (YSR 11-18). There is a preschool version of the CBCL (CBCL 1 1/2 -5) and this may be used by parents or caregivers who regularly interact with the child and thus able to rate behaviour.

Each of the checklists examines eight syndrome scales that examine aggressive behaviour, anxious/depressed behaviour, attentional problems, conduct problems, somatic complaints, social problems, thought problems and withdrawn/depressed behaviours. Reliability coefficients across the scales vary between 0.71(somatic) -0.89 (conduct).

## *Tests of Somatic Disorders and Malingering Assessment*

### **[60.11010] Introduction**

“Somatic Symptom & Related Disorders” (a new section of the DSM V), which are identified by prominent somatic symptoms associated with significant distress and impairment. They

include distressing somatic symptoms associated with abnormal thoughts, feelings and behaviours in response to these symptoms and can accompany medically diagnosed disorders. Examples of these disorders include Somatic Symptom Disorder, Illness Anxiety Disorder, Conversion Disorder and Factitious Disorder.<sup>1</sup>

---

<sup>1</sup> The reader is referred to Chapter 61 (61.670, 61.680, & 61.690) and Chapter 68 (68.2140) for additional information on Malingering

### [60.11020] Assessment

Considered by many forensic psychologists as the “bible” of dissimulation assessment, Rogers’ *Clinical Assessment of Malingering and Deception* (2008, Third Edition<sup>1</sup>) observes in the very first sentence: “*Complete and accurate self-disclosure is a rarity even in the uniquely supportive context of a psychotherapeutic relationship ... Even in intimate relationships, willingness for self-disclosure is variable and multi-determined*” (p.3). In forensic environments, the assessor should not only expect response distortions, but must also assess for such distortions, especially where there is a secondary gain. Decisions about disclosure/honesty/deception/malingering are often rational and multi-determined. They are responses to interpersonal or contextual variables and not always trait based. This means that the detection of distortion can usually be assumed to be a deliberate act by the assessee, reflective of the situation in which the person finds themselves and their personal goals. For example, a claimant in a personal injury case may employ two response styles in the same evaluation, minimising problems (defensiveness) prior to the accident, and exaggerating symptoms (malingering) after the accident. Further, in evaluative contexts, both internal (mental health diagnosis) and external (social desirability) influences will affect the way the person presents themselves and responds to the assessment process. Response styles are influenced by mental health variables, such as Personality Disorders or severe psychopathology, and by identity issues, such as avoidance of stigmatisation.

---

<sup>1</sup> The Fourth Edition of Rogers’ *Clinical Assessment of Malingering and Deception* is due for release in March 2018.

### [60.11030] Response Styles

When undertaking and reporting psychological assessments, psychologists will often comment on the individual’s response style when undertaking a test or engaging in interview. Given the universality of response distortion, not all response styles should be assumed to be indicative of intentional deception. Response styles that may be observed include:

- Unreliability - Describes the accuracy of information provided but *does not* infer intent or reason for inaccuracy. May be due to disinterest, psychotic disturbance, lack of comprehension or attempts to feign confusion;
- Nondisclosure - Withholding information (omission), again with no assumptions about intentionality;
- Self-disclosure - Refers to how much information the person wishes to reveal about themselves. A lack of self-disclosure does not imply dishonesty but rather an unwillingness to share personal information;
- Deception - An all-encompassing term used to describe any consequential attempts by an individual to distort or misrepresent their self-reporting. Includes acts of deceit, often accompanied by nondisclosure, with this response style not necessarily related to the person’s psychological functioning;

- Dissimulation - Describes an individual who is deliberately distorting or misrepresenting *psychological symptoms*. Distortions do not tend to fall into a coherent pattern, unlike defensiveness or malingering.

While each of these response styles may impact upon the accuracy of the information derived during assessment, they do not in and of themselves indicate either malingering or defensiveness. Furthermore, the Diagnostic and Statistical Manual of Mental Disorders – Fifth Edition (DSM V: 2013) identifies several diagnosable mental health conditions which are distinct from purposeful malingering and defensiveness. Encompassed within these are “Somatic Symptom & Related Disorders” (see above). When diagnosing malingering, the psychologist should rule out any of these disorders or the response styles listed above.

One disorder that can appear to be malingering is Factitious Disorder, formerly known as Munchausen’s Disorder. Factitious Disorder is included within “Somatic Symptom & Related Disorders” (above) as similar to Somatic Symptom Disorder, Illness Anxiety Disorder, and Conversion Disorder. It involves persistent problems related to illness perception and identity. Most sufferers present with somatic symptoms and medical disease conviction. There are two types: Factitious Disorder Imposed on Self; and Factitious Disorder Imposed on Another (formerly Munchausen’s by Proxy). Each involves exaggeration, fabrication, simulation or induction of symptoms but is differentiated from malingering in that the deceptive behaviour is evident even in the absence of obvious external rewards.

Two further considerations should be accounted for before making a Malingering or Defensiveness determination. First, many mental health conditions and brain injuries result in confabulation. Confabulation is evidenced when people “fill in” missing pieces of a story. Their beliefs are usually held with high conviction, however their memories tend to be more detailed than normal memory would allow. Confabulation can be produced by the influence of others (including therapists) and accounts for the phenomenon of False Memory Syndrome. The other condition to rule out is Delusional Disorder, which lies within Schizophrenia Spectrum and Other Psychotic Disorders section of DSM-V (see above).

## **Malingering and Simulated Adjustment/ Defensiveness**

### **[60.11210] Introduction**

Malingering is the intentional production of false or grossly exaggerated physical or psychological symptoms, motivated by external incentives such as avoiding work, obtaining financial compensation, evading criminal prosecution, or obtaining drugs. Malingering should always be considered in the following circumstances:

- Medico-legal context of presentation;
- Marked discrepancy between the person’s claimed stress or disability and the objective findings;
- Lack of cooperation during the diagnostic evaluation and in complying with the prescribed treatment regime; and,
- Presence of a Personality Disorder, especially “cluster B” Personality Disorders (see above).

Simulated adjustment is best described as a person seeking to look better than they are, or concealment of undesirable characteristics. Defensiveness is the most common term and can be viewed as the polar opposite of malingering. Defensiveness involves the deliberate denial or gross minimisation of physical and/or psychological symptoms. In assessing defensiveness, the

psychologist should account for the influences of “Social Desirability” - presenting the self in the most favourable manner relative to social norms. Social Desirability involves *both* the denial of negative characteristics *and* the attribution of positive qualities.

Specific strategies have been developed for the detection of malingering and tests used to identify malingering use a combination of strategies. The first class of strategies are based primarily upon *unlikely* presentations:

1. Rare Symptoms: looks for symptoms infrequently reported (<5%) by bona fide clinical populations. Malingers over-report such symptoms. Assessments that identify rare symptoms are considered a robust technique in the assessment of malingering.
2. Quasi-Rare Symptoms: similar to Rare symptoms but some conditions may result in legitimate reporting, such as PTSD and schizophrenia. These approaches are confounded by the possibility of true reporting.
3. Improbable Symptoms: an extreme variant of Rare Symptoms, this strategy looks for features that are fantastical or preposterous. These strategies have limited usefulness with sophisticated malingerers.
4. Symptom Combinations: looks for features that are common in clinical populations but which rarely occur together (e.g. grandiosity & increased sleep). This strategy is considered resistant to coaching.
5. Spurious Pattern of Psychopathology: an elaboration of symptom combinations, it relies on scale configurations that are characteristic of malingering but very uncommon in clinical populations. Requires extensive cross validation to ensure results are not reflecting chance variance.

A second class of strategies are based on *amplified* presentations, with an emphasis on frequency and intensity, rather than presence or absence:

1. Indiscriminate Symptom Endorsement: malingerers tend to endorse a large proportion of symptoms.
2. Symptom Severity: capitalises on the fact that even severely impaired patients experience only a discrete number of symptoms as “unbearable” or “extreme” in intensity.
3. Obvious Symptoms: malingerers endorse prominent symptoms indicative of serious mental disorders and include everyday problems that are not necessarily indicative of a major mental disorder.
4. Reported vs. Observed Symptoms: discrepancies between the patient’s account of their symptoms and those observed by the assessor.
5. Erroneous Subtypes: capitalises on the fact that people have common misconceptions about which clinical characteristics are associated with specific mental disorders.

## **[60.11220] Psychological Tests to Identify Malingering and Deception**

### **Minnesota Multiphasic Personality Inventory-2 (MMPI-2)**

The MMPI-2 (Butcher et al, 2001) is a widely used psychological assessment instrument in forensic psychology (Helmes, 2008). While originally developed to assess psychopathology, it now has many varied uses. It contains 557 items. A new version (MMPI-2-RF: Ben-Porath &

Tellegan, 2008/2011) has fewer items (338) but is still under examination (although the research to date has provided strong support to its utility). A strength of the MMPI-2 is that it has a number of scales that can be used to detect a variety of types of response distortion. Specifically, the MMPI-2 measures attempts to appear more disturbed than is the case (Fake bad: eg the *F* and the *Fp* scales) and attempts to appear *less* disturbed than is the case (Fake good: eg the *L* and *K* scales). Elevations on scales are assessed with reference to “*T* scores”. With respect to the MMPI-2, a *T* score of 65 indicates that 92% of the general population would score less than this. *T* scores of 70, 75 and 80 indicate the following levels of “confidence” related to the general population: 96%, 98% and >99%. The MMPI-2 is predominantly used to detect Malingering through *Fp* (F-psychiatric) scale (Rare Symptoms) and/or *F* scale (Quasi-rare symptoms). Additional scales to assess malingering using scale combinations, such as the *F – K Index* (which couples unusual items with a low willingness to admit weakness), and the Fake Bad Scale (*Fbs*: Lees-Hayley, English & Glenn, 1991), have also been developed and attracted research, especially with specific malingered conditions, such as PTSD (Lees-Hayley, 1992).

The MMPI-2 also contains several scales for the detection of defensiveness. The Lie Scale (*L*) purports to identify persons deliberately trying to be defensive, however it has been observed that respondents with any degree of psychological sophistication easily identify the 15 items in this scale and as such the utility of the *L* scale is limited (Greene, 2008). The MMPI-2 also contains the Correction (*K*) scale, purported to identify an unwillingness to acknowledge any type of psychological distress. The *K* scale is more often used to “correct” scores on clinical scales so as to provide a more accurate picture of the respondent’s psychopathology than as a direct measure of defensiveness. Another measure, the Other Deception scale (*ODEcep*) was developed by Nichols and Greene (1991: cited in Greene, 2008) as an update of the Positive Malingering scale (*Mp*), originally developed for earlier versions of the MMPI. The *Mp* scale has been found to evidence large effect sizes in discriminating between honest and defensive responding (Baer et al, 1992). The Superlative scale (*S*), which seeks to identify people who present as outstanding individuals, and the Wiggins Social Desirability scale (*Sd*) have also been developed, however the distribution of scores tend to be similar between “normal” individuals and those with mental disorders, especially where marked elevations are observed (Greene, 2008).

There are four commonly accepted stages of interpretation of any MMPI-2 Profile evaluation: (1) assessment of the completeness of the profile (item omissions); (2) assessment of the consistency of item endorsement, which is independent of item content, determining whether the individual has provided a reliable pattern of responding to items throughout the inventory; (3) assessment of the accuracy of item endorsement (the stage where malingering, deception or response distortion is observed); (4) assessment of personality profile. In reporting the results of an MMPI-2 assessment, the assessor should detail the respondent’s performance, and hence the validity of the results, across each of these stages.

The MMPI-2 has adequate reliability and validity as reported by Butcher et al, (2001), although there is considerable inter-scale correlation. The strength of the MMPI-2 is the extensive research associated with the test, including its translation into over 50 different languages and normative and validity studies across several different cultural groups. However, it has also been observed that demographic variables, such as age, gender, intelligence, socio-demographic status, etc. influence responses such that scale elevations can have different meanings according to these variables (Groth-Marnat & Wright, 2016). This means that the interpretation of the MMPI-2 requires considerable expertise of the assessor and that expertise should be scrutinised in order to determine the probative value of any evidence that relies upon the MMPI-2 data.

### **Personality Assessment Inventory - Revised**

The PAI-2 (Morey, 2007) is a self-administered, objective test of personality and psychopathology. The PAI provides information central to treatment planning, implementation and evaluation. The PAI has sound psychometric properties (reliability and validity: Siefert et al, 2009). The PAI has two scales directly relevant to defensiveness and malingering, as well as four index scales (determined by combinations of scale configurations); two assessing defensiveness and two used to detect malingering.

The Negative Impression Management (NIM) Scale detects a negative response set due to either a pessimistic worldview and/or intentional dissimulation. The NIM is an example of the "Rare Symptoms" detection strategy, similar to the Fp scale on the MMPI-2. While not specifically a malingering scale, the NIM contains items that generally have a low endorsement rate in both normal and clinical samples. The NIM is most effective at detecting malingering of severe mental disorders and is robust to patient sophistication. Significant elevations on the NIM may be indicative of malingering, but also careless responding, or an elevation consistent with the clinical presentation of the respondent. As such, the use of the NIM by itself as a measure of malingering is unwise.

The Malingering Index (MAL) is an example of the "Spurious patterns of psychopathology" strategy to detect malingering. The MAL was designed to detect those likely to be simulating psychopathology rather than legitimate clinical disorders. The MAL consists of eight criteria with scores greater than four suggesting a strong likelihood of malingering. However, the sensitivity of this index is generally limited to the malingering of severe psychopathology, such as psychosis, while the malingering of milder disorders, such as anxiety, may not be detected by the MAL.

The Rogers Discriminant Function Index (RDF) provides another score designed to differentiate those with genuine psychopathology from those feigning a disorder, and is an example of the "Spurious patterns of psychopathology" strategy to detect malingering. The RDF consists of the weighted *T* scores of 20 items, with the final sum score representing the likelihood of simulated pathology. The RDF is reported as the most effective tool for the identification of malingering within the PAI. The RDF has been found effective and to be robust when trialed with sophisticated fakers such that those knowledgeable of psychopathology were unable to simulate mental ill-health on the RDF. The RDF does not correlate with either the NIM or psychopathology, suggesting that it is free from the negative response set of clinical populations and thus a good example of malingering. It is noted that when the RDF is used in combination with the F and the Fp of the MMPI-2, the detection of malingered psychopathology is robust.

The PAI-2 also contains scales to measure the under-reporting of symptoms, often referred to as defensiveness or social desirability. The primary scale is the Positive Impression or PIM scale. The PIM reflects the respondent's predilection to present a very favourable impression of themselves or deny relatively minor faults. However, the tendency to seek to make a favourable impression is fairly common in the normal population (Morey, 1996) and is even more likely in circumstances where the judgement of an individual may have significant consequences, such as in forensic environments. As a result, two configural index scales have been developed: The Cashel Discriminant Function (CDF) and the Defensiveness Index (DEF). The CDF has a reported sensitivity ranging from 79% to 87%, with a specificity of 88% and has been reported as superior to the DEF (Cashel et al, 1995). However, Moray (1996) advises that the detection of defensiveness should entail consideration of all three scales (PIM, CDF and DEF) when seeking to identify attempts at positive impression management.

### **Structured Interview of Reported Symptoms (SIRS)**

The SIRS (Rogers, 1992; Rogers, Bagby, & Dickens, 1992) is the only well validated structured interview of response styles (Rogers, 2008). As a fully structured (as opposed to semi-structured) interview, the SIRS does not allow clinicians to make their own clinical inquiries to clarify a respondents answer. The SIRS utilizes eight strategies to discriminate between feigned and genuine responses: Rare Symptoms; Symptom Combinations; Improbable and Absurd Symptoms; Blatant Symptoms; Subtle Symptoms; Selectivity of Symptoms; Severity of Symptoms; and Reported vs Observed Symptoms (Rogers, 2008). The SIRS has been reported to have good to excellent internal reliability (Vitacco et al, 2007), strong convergent reliability (Rogers, 2008), and robust discriminant validity (Rogers, 1997). The SIRS has been found to be generalizable across diagnostic groups and is less vulnerable to the effects of limited respondent literacy (Rogers, 1998). However, the SIRS is not intended to evaluate cognitive impairment, has limited utility with persons experiencing mental retardation and is vulnerable to the effects of coaching. Notwithstanding these limitations, the SIRS is considered the strongest measure of feigned mental disorders (Rogers, 2008).

### **[60.11230] Screening Measures**

A number of brief screening measures have been developed for the assessment of feigned psychopathology, key amongst which are the Miller Forensic Assessment of Symptoms (M-FAST: Miller, 2001) and the Structured Inventory of Malingered Symptomology (SIMS: Widows & Smith, 2005). The M-FAST is a brief 25-item structured interview designed specifically for forensic settings (Guriel et al, 2004). The M-FAST includes seven scales traversing detection strategies: Reported vs Observed; Extreme Symptomology; Rare Combinations; Unusual Hallucinations; Unusual Symptom Course; Negative Image; and Suggestibility. The M-Fast items vary in structure and the suggested cut score for suspected malingering is scores of six and over. The usefulness of the M-FAST is enhanced when combined with additional measures of malingering (Smith, 2008). The M-FAST has acceptable convergent and discriminant validity (Miller, 2004) and is not overly affected by respondent race (Guy & Miller, 2004), gender (Jackson et al, 2005), or coaching (Guriel et al., 2004). Given that the M-FAST was developed for feigned mental disorders, it is not suggested that it be used as a general assessment of malingering (Smith, 2008).

The SIMS is a 75-item true/false self-administered screen for the detection of malingering across a variety of clinical and forensic settings, as well as the detection of the feigning of both psychiatric and neuropsychological symptoms. The SIMS is intended for use with adults only (over age 18 years) and requires a reading age at least at fifth-grade level. The SIMS *is not* intended to be used beyond screening and a definitive statement of malingering will require additional sources of information. The SIMS screens for commonly malingered conditions: low intelligence; affective disorders; neurologic impairment; psychosis; and amnesia. Items are scored either "0" or "1" and are summed to provide sub-scale and total scores. A score >14 indicates a need for a more comprehensive assessment of malingering. The SIMS demonstrates good internal reliability across scales, ranging from .80 to .88, and has demonstrated a high level of sensitivity (95.6%) and good specificity (87.9%).

## *Risk Assessment*

### **[60.11710] Introduction**

Risk assessment involves determining quantitative or qualitative estimates of risk to a defined situation and threat. Actuarial risk assessment tools, being those based on statistical analyses of probability, are often employed by psychologists to predict the likelihood of future violence,

offending behaviour (recidivism), sexual offending, suicide and various mental health conditions. In the forensic context there is much focus on predicting likelihood of recidivism.

There is a general consensus about the classes of variables that are valid predictors of recidivism. Most of the research has examined the predictors in general, rather than exclusively violent recidivism (Taylor, 2015).

Most actuarial instruments use static, or unchanging, variables rather than dynamic variables. Predictors that change over time (dynamic) are potentially of value in the assessment of likelihood of recidivism. The issues involved in dynamic prediction are much more complex than those involved in predictions made from static variables. There are considerable difficulties in identifying dynamic predictors and assessing their accuracy. Over the past few years a new generation of instruments have been developed which include both static and dynamic factors and these enable changes in risk of recidivism due to time, or other factors such as treatment, to be assessed (Taylor, 2015).

## **Tests of Risk of Violence, Sexual Offending and Recidivism**

### **[60.11810] Introduction**

Risk assessments have been developed for different types of risks and for different populations of offenders. In most cases, a number of risk assessments exist for the same type of offence within the same offender populations. The following examples survey the most common risk assessment tools used by psychologists.

### **Juvenile**

#### **[60.11860] Violence Risk Assessments**

The following risk assessments are likely to be used by psychologists assessing juvenile offenders in Australia:

*The Structured Assessment of Violence in Youth* (SAVRY: Borum, Bartel & Forth, 2006) was developed to assist clinicians in assessing violence risk in adolescents, primarily 12-18 years of age. The SAVRY is an example of a “structured professional judgement” assessment, where such assessments are consistent, systematic and grounded in research (Borum et al, 2006). The SAVRY was designed for use in institutional and community settings to aid in the development of case plans and treatment plans. It is a structured rating scale that relies on the test administrator to gather the relevant information about the youth to rate items as per the instruction manual (Borum et al, 2006). By examining the research regarding violence and its predictors, the test developers generated a structured approach that assists psychologists in assessing an adolescent’s degree of dangerousness.

Structured professional judgements are consistent, systematic and grounded in research (Borum et al, 2006). The SAVRY is anchored in relevant research, considers developmental factors, and emphasises the dynamic (ever changing) and contextual nature of risk. As such, the test designers advocate for the test to be repeated at regular intervals because of the dynamic nature of adolescents and their rapid rate of change. For example, if a client was assessed during a custodial sentence, it may be advisable for another assessment just prior to their release to a community setting where significantly less supervision will be available for the young person (Borum et al, 2006).

The majority of research has been conducted in the United States. Research conducted in detention centres in Australia lends support to the use of SAVRY with moderate to strong support for its predictive validity of general offending and violent recidivism (Shepherd et al., 2014). Studies in Canada (Viljoen et al., 2017) and Singapore (Chu, Goh & Chong, 2016) also demonstrate empirical support for the use of the SAVRY and the YLSI/ CMI (see below) in assessing risk of violent and general recidivism. Comparison of ethnic subgroups did see a high degree of variability in predictability scores and thus more research across culturally and linguistically diverse populations needs to occur as the measure does not seem to cater for such populations.

Statistical properties: Reliability coefficients for SAVRY Risk Total were found to be 0.82 for offenders and 0.84 for the community sample. An independent study found a reliability coefficient of 0.83 for the SAVRY Risk Total (Borum, Bartel, & Forth, 2006). Interrater reliability coefficients are reported as 0.81 for the SAVRY risk total. Concurrent validity with the YLS/CMI and the PCL-YV found a significant correlation with both instruments' total scores among community and offender populations. (0.89 with YLSI/CMI, 0.78 with PCL: YV young offenders). As mentioned previously, research conducted by Shepherd et al, (2014) lends support for its use with juvenile offender populations in Australia due to moderate to strong support for its predictive validity of general offending and violent recidivism (Shepherd et al, 2014).

## **[60.11870] Juvenile Sex Offender Risk Assessments**

*The Juvenile Sex Offender Assessment Protocol* (JSOAP-II: Prentky & Righthand, 2003) and the Estimate of Risk of Adolescent Sexual Offence Recidivism (ERASOR: Worling, 2004) are the most researched tools used by psychologists working in this field. As with other assessments, a clinical interview and review of collateral information will take place and often add information to make sense of the JSOAP II or ERASOR.

### **JSOAP-II**

The JSOAP-II: Best practice involves a comprehensive assessment that examines the risk factors associated with sexual and criminal offending. Each scale of the assessment tool examines static and dynamic risk factors that are derived from the research. Close examination of protective factors (those that will assist the offender to avoid further offences) such as community support, participation in activities that will expose them to prosocial peers, stable homelife and an educational/employment pathway will assist the psychologist in their assessment of the risks the young person presents to the community in relation to reoffending. Often assessments for sexually based offenders will be comprehensive due to the seriousness of the crime and the risks future offending places on the community. The psychologist will use much collateral information so as to suggest the most suitable treatment options for the offender. Juvenile sex offenders often do not go on to offend as adults, despite popular belief that they do, hence ensuring they receive appropriate services that meet their individual needs is essential to the assessing psychologist (Nisbett, Smallbone & Wortley, 2010). As with other assessments, a clinical interview and review of collateral information will take place and often add information to make sense of the JSOAP II. Statistical Properties: The published research literature on J-SOAP-II documents five studies which have assessed the predictive accuracy of the instrument with non-institutionalized Juvenile Sex offenders. Four of these studies demonstrated moderate predictive validity for the J-SOAP-II in outpatient settings, with area under the curve (AUC) estimates ranging from .7 to .8 (fair) for the prediction of sexual re-offense. Additionally, one study of juveniles convicted of sexual offenses but whose ultimate placement was unknown demonstrated some support for the predictive accuracy of the

J-SOAP-II's Scale IV (Community Stability/Adjustment) in predicting sexual re-offense when included as part of a model along with a measure of sex offense severity, AUC = .82 (good) (Aebi et al, 2011).

*Estimate of Risk of Adolescent Sexual Offense Recidivism (ERASOR: Worling, 2004)*

This is an empirically-guided checklist to assist evaluators to estimate the short-term risk of sexual reoffending for youth aged 12-18. The ERASOR was designed as a single-scale instrument, consisting of 25 risk factors under 5 headings: Sexual Interests, Attitudes, and Behaviours, Historical Sexual Assaults, Psychosocial Functioning, Family/Environmental Functioning, and Treatment. All risk factors are coded as Present, Possibly/Partially Present, Not Present, or Unknown. As with other assessments, a clinical interview and review of collateral information will take place and often add information to make sense of the ERASOR.

Statistical Properties: The ERASOR has adequate to excellent interrater reliabilities (0.86-0.88) and internal consistency for the total score (Viljoen et al, 2009). Studies that found the ERASOR to be a moderate to strong predictor of sexual, nonsexual, and general recidivism include Rajlic & Gretton (2010) and Worling (2004).

## [60.11880] Risk of Recidivism

*Youth Level of Service/Case Management Inventory (YLS/CMI; Hoge and Andrews, 2011)*

This is a widely used, structured risk assessment for assessing the risk of recidivism and criminogenic needs for young offenders. Similar to the SAVRY, the YLS/CMI is composed of static and dynamic risk factors that are associated with reoffending. In Australia, the Australian Adaptation is utilized (YLS/CMI-AA: Hodge & Andrews, 2011). Like the Level of Service Inventory used in adult corrections, the youth version relies upon the risk, needs and responsivity principles when trying to comprehend and address adolescent criminal behaviour. The YLS/CMI-AA enables services to be matched to client's level of risk, ensuring that service relates to their need areas, and that the methods used are consistent with the adolescent's learning style, ability and individual characteristics for optimum rehabilitative results. Psychologists will report a Low, Medium, Medium-High or High risk rating of the YLSI/CMI-AA. Low risk offenders will receive less attention as they pose less of a risk of reoffending due to protective factors such as access to education, low level of drug and alcohol use, and possible community support. High-risk offenders require intensive supervision and higher levels of service delivery as they are more likely to have more contact with the criminal justice system (Hoge & Andrews, 2011).

Statistical properties: McGrath and Thompson (2012) report that the static and four dynamic domain scores independently predicted recidivism and that the combination of those domain scores yielded a small improvement in prediction. Meta analytic studies have found the predictive validity for general recidivism is good. In 11 studies of predictive validity the Area Under the Curve (AUC) obtained was .641, which is considered as poor.

*Psychopathy Checklist – Youth Version (PCL-YV: Forth, Kosson & Hare, 2003).*

Identifying youth with psychopathic traits is critical to understanding the factors that contribute to the development of adult psychopathy. This test is a 20-item rating scale for the assessment of psychopathic traits in young offenders aged 12 to 18. The PCL-YV assesses youth on several behavioural and personality characteristics associated with the construct of psychopathy. Based on the Hare Psychopathy Checklist (PCL) for adults, the PCL-YV uses a semi-structured interview, clinical observations and collateral information. Often the PCL-YV is used in conjunction with measures that look at predicting future recidivism. Some argue that criminal behaviour is not a core feature of psychopathy and is simply a correlate or an outcome

of those who present with strong psychopathic traits (Skeem & Cooke, 2010). An Australian study conducted by Shepherd and Strand (2016) emphasises the importance of the PCL-YV to identify psychopathic traits but not necessarily for predicting recidivism. While their study demonstrated that the ability of this test to predict youth recidivism has generally been satisfactory in the literature, the strength of association with recidivism is likely due to the antisocial domains which encompass criminal behaviours.

**Statistical Properties:** In the standardization sample Forth et al (2013) showed good internal consistency for the four facets of psychopathy (Facet 1 = 0.75; Facet 2 = 0.71; Facet 3 = 0.70 and Facet 4 = 0.78).

Because of the complexity of administering the PCL:YV, all those who administer the checklist should complete relevant training. Inter rating reliability studies have shown intra-class correlation coefficients (ICCs) with excellent rates of agreement for total scores (ICC 0.91).

Important to note: Adolescence is a time of constant change thus care should be taken in labelling youth with value laden term such a “psychopathy” during a period of rapid change, when often impulsive behaviours and risks are common within this developmental timeframe.

## Adult

### [60.11930] Introduction

Risk assessment tools utilised by psychologists for adult offenders (>18 years) vary widely.

### [60.11940] Risk of Recidivism

*Level of Service Inventory – Revised (LSI-R: Andrews & Bonta, 1995)*

The LSI-R is a 54-item actuarial risk assessment tool used to measure the likelihood of general reoffending. It assesses criminogenic needs that contribute to reoffending by examining dynamic and static risk factors. The LSI-R is the standard tool used by correctional staff to measure risk and needs for offenders based on Risk Needs Responsivity (RNR) principles mentioned with the juvenile measure YLS/CMI-AA. Like the juvenile measure, the LSI-R gives an indication of risk (Low, Medium, Medium-High, High) and this determines the level of service required for the offender. After assessing level of risk, need areas will be examined (eg, drug and alcohol problems, accommodation, criminal attitudes) and methods employed are adjusted according to the clients learning style, ability level and individual characteristics for optimum rehabilitative results. As per juvenile offenders, low risk offenders will receive less attention and therapeutic interventions as they pose less of a risk of reoffending. Greater resources will be employed for high-risk offenders.

**Statistical Properties:** Reliability analysis revealed that the internal consistency estimates for LSI-R subscales range from adequate (Cronbach’s alpha for the accommodation subscale = 0.509) to good (Cronbach’s alpha for the Education/Employment subscale = 0.784). Specific normative data has been generated across nationalities (e.g. Australia: Hsu, Caputi & Byrne, 2009), although criticisms of the scales sensitivity to the unique needs of Aboriginal offenders has been raised (Hsu, Caputi & Byrne, 2010). Hsu, Caputi and Byrne (2011) argue that the LSI-R be recalibrated for specific offender cohorts and that the original normative data may lack sensitivity and specificity in populations other than the normative sample.

### [60.11950] Risk of Violence/Sexual Offending

*Historical Clinical Risk – 20 Version 3 (HCR-20 V3: Douglas et al, 2013)*

The HCR-20, version 3, is a 20-item checklist based on established risk factors of violent behaviour. It is applicable to adults aged 18 and above and is often used in correctional facilities, forensic populations and general or in-patient psychiatric settings. Whether in a secure setting or not, it is applicable to adults who pose future risk of violence. It examines historical factors, clinical variables and future risk management issues.

The HCR -20 is considered a Structured Professional Judgement (SJP) model of violence and risk assessment. Extensive research in the violence field has enabled the checklist to be adopted or evaluated by over 35 countries in the world. It has been translated into 20 different languages.

Statistical Properties: In addition to good inter-rater reliability, the HCR-20 V3 discriminates well between violent and non-violent participants, is strongly associated with frequency of violence, and is moderately predictive of post-discharge violence.

*Violence Risk Scale – Second Edition (VRS-2: Wong & Gordon, 2006).*

The VRS-2 is a 26-item semi-structured interview that assesses violence risk, identifying treatment targets linked to violence, readiness for change, post-treatment improvement and post-treatment violence risk. The VRS scale relies heavily on risk, needs and responsivity principles previously discussed with YLSI and LSI.

Statistical Properties: During the validation study, good internal consistencies were demonstrated (Cronbach alpha .93) and interrater reliability of .92- .97 (intraclass correlations) were obtained. Correlations with the Psychopathy Checklist- Revised (.83) and the Level of Supervision Inventory (.82) were all highly significant ( $p < .001$ ) and in the expected directions.

A sex offender version of the VRS also has been developed. *The Violence Risk Scale-Sex Offender Version (VRS-SO: Wong et al, 2003)* is organized in identical format as the VRS. The rationale for the development, organization and rating instructions, etc. for the VRS also apply to the VRS-SO. The VRS:SO incorporates static and dynamic variables specifically relevant to sexual offending such as sexual deviancy, sexual compulsivity, offence planning etc.

*STATIC-99 (Hanson & Thornton, 1999)*

The STATIC-99 is a 10 item actuarial instrument used to assess adult male sex offenders who are at least 18 years of age when they are released in to the community. This instrument is used worldwide in correctional facilities and community based corrections to assist in predicting long-term likelihood for sexual and violent recidivism. The final score obtained relies on objective, obtainable information such as their criminal history, victim characteristics, age of victim, number of victims, gender of victim and other measurable details.

Statistical Properties: Previous research has demonstrated high rater reliability (Langton et al., 2007) and moderate to high predictive accuracy (Hanson & Morton-Bourgon, 2007).

*Psychopathy Checklist Revised – 2<sup>nd</sup> Edition (Hare, 2003)*

The PCL-R is a 20-item interview that assesses individuals for evidence of psychopathic personality traits in adult forensic populations. Like the PCL-YV, the PCL-R uses semi-structured interview, clinical observations and collateral information. Scores obtained on the PCL-R have been found to be strong predictors of recidivism thus the test is often used to supplement inventories that assess risk, despite is being designed to assess psychopathy.

Statistical properties: The PCL-R has been reported in various studies of prison and forensic psychiatric populations across many countries. There is much evidence to suggest the PCL-R is a very reliable test when it is administered and scored by a trained and experienced rater who has available data to complete the items. Internal consistency is high (. 0.80). Intra-class correlation (ICC generally exceeds 0.80 for a single rater and 0.90 for the average of 2 raters).

The PCL also has a screening version of the test – PCL-SV which takes far less time to administer and can give an indication of the presence of psychopathic traits.

*Violence Extremism Risk Assessment 2* (Vera 2: Pressman & Flockton, 2010).

The assessment of terrorism based violence is complicated by the diverse range of causes and goals of such violence. Despite the increasing threat imposed by violent extremism, few measures have been developed which specifically address this type of offence and these offenders. Some have argued that “actuarial” approaches to the assessment of risk of extremist violence are fatuous due to the complex interplay of political, ideological and individual variables (Sarma, 2017). Not-with-standing this concern, Pressman (2009) developed the VERA which was later revised by Pressman and Flockton (2010) and referred to as the VERA 2. The VERA 2 consists of 31 indicators designed to be assessed by psychologists and other trained professionals who have knowledge of the field of terrorism. It is a form of Structured Professional Judgement (SPJ) where “empirically based indicators related to terrorism (are applied) in a systematic protocol that rates a set of indicators according to criterion-defined levels” (Pressman & Flockton, 2012, p 244).

The VERA 2, while intended for use with terrorists, can be used across the spectrum of violent extremists where there is a conviction that violent action is “morally justified and/or sanctioned to bring about the desired ideological or political goals” (ibid). The indicators for risk on the VERA 2 are categorised into four areas: beliefs and attitudes; context and intent; history and capability; and commitment and motivation. Additionally, a fifth area, *protective* factors, is included to account for any attenuation of risk. The final risk decision is discretionary such that the assessor can “over-ride” ratings to either escalate or reduce assessed risk level. The authors do not provide statistical data on the VERA 2.

Recently (Sadowski et al, 2017) the Violent Extremism Risk Assessment version 2-Revised (VERA-2R) has been reported. According to the authors, the VERA-2R uses more specified *dynamic* (changeable) indicators known to be consistent with the radicalisation process to violent extremism. Monitoring of these indicators, which can change over time, can establish “risk trajectories” for an individual. The VERA-2R contains 34 indicators divided into five domains: beliefs, attitudes and ideology; social context and intention; history, action and capacity; commitment and motivation; and protective/risk-mitigating indicators. There are 31 additional indicators based on professional literature, also divided into five domains: criminal history; personal history; radicalisation; personality traits; and psychiatric characteristics. Each VERA-2R indicator has criteria for rating low, medium or high, consistent with the general SPJ approach. The final risk decision is not made by “score”, but rather by professional judgement after weighting all available information. The authors assert that users of the VERA-2R must be trained in the methodology and interpretation of the VERA-2R before using it. At the time of writing this chapter, there was no published reliability or validity data associated with the VERA-2. Not-with-standing the apparent absence of alternative extremism risk assessments, the VERA-2 should be considered a “test in development” at this time.

## *Tests of Attachment*

### **[60.12510] Introduction**

Attachment is understood as an enduring emotional bond between a child and a primary caregiver, which begins at birth and continues across the lifespan (Carr, 2015). Attachment is considered a biological necessity, in that in order to ensure survival, the child is driven to form a connection with a primary caregiver for safety and security (Bowlby, 1969; 1973). In good attachment (ie named secure attachment), the caregiver responds to this drive by providing a secure base in which they are attuned and responsive to the biological and emotional needs of the child (Ainsworth et al, 1978). For example, when the child cries, the caregiver responds by

attempting to soothe. Through this process, the child learns how to regulate distress, how to express needs and emotion appropriately, and how others reciprocate emotion in secure relationships (Carr & McNulty, 2016).

Unfortunately, caregivers are not always able to provide a secure base for the child, either due to emotional or physical absence (eg abuse, neglect) or due to their own poor attachment experience. In this scenario, an *insecure* attachment forms, in which the child's experience of safety is absent, inconsistent and unreliable. Consequently, the child will often do whatever it takes to get their needs of care and protection met (eg act out behaviourally, or in the case of violence become invisible to divert unwanted attention), and thus does not learn how to appropriately demonstrate or respond to emotional experience.

Thus, the early attachment bond between a child and the primary caregiver provides the child with an internalised "blueprint" of intimate relationships (Bowlby, 1969; 1973). It sets the scene for their expectations of how relationships work, both in terms of their role and the role of others. Therefore, for psychological assessment, understanding early attachment is vital in conceptualising the individual, in terms of how they see themselves, their relationships and the others around them (Zeanah, Berlin, & Boris, 2011). Similarly, attachment security has been shown to be a key predisposing and/or maintaining factor in the development of variety of clinical and subclinical psychopathology (Groh et al, 2017).

### [60.12520] Assessment

The assessment of attachment differs according to the age of the participant as well as the nature of relationship under question. Additionally, due to the complex nature of attachment, it cannot be comprehensively understood with a single measure. Rather, attachment measures must be used in a combination with standardised measures of related psychopathology (ie personality assessment, mental health status, relationship functioning), as well as alongside collateral reports from others (eg teachers, extended family members) in order to form a comprehensive picture of individual attachment (Carr & McNulty, 2016). It is also important to note that the role of attachment assessment in forensic settings has been discussed extensively due to ongoing issues with the operationalization of attachment measurement (See Garber, 2009; Ludolph, 2009 for examples). Discussed below are the two most commonly used assessments of attachment.

*Ainsworth Strange Situation Paradigm* (SSP; Ainsworth et al.,1978; Ainsworth & Witting, 1969)

In children, attachment security can be assessed using a framework called Ainsworth's Strange Situation Paradigm. Essentially, this involves direct observation of the child and caregiver(s) together, and focuses on the dynamics between the two in response to a variety of events (for more in-depth information on the SSP see Chapter 62, Assessment in the Childrens Court). Although the "gold standard" for assessing childhood attachment, the SSP has been criticised due to practicality, as it requires intensive and specialised training to administer reliably and is based around a laboratory, non-normative situation. However, the psychometric properties of the SSP remain sound, with Ainsworth et al, (1978) original work showing sufficient reliability and validity, and more contemporary research corroborating such results (Thompson & Raikes, 2003).

*Adult Attachment Interview* (AAI; George, Kaplan, & Main, 1984; Main, Goldwyn, & Hesse, 2002)

The AAI is a semi-structured interview designed to assess adult representation of attachment, by examining early attachment with caregivers. It contains approximately 20 questions which focus on relationships within the individual's family of origin. Interviews are transcribed verbatim and coded. Scoring focuses both on *what* is said as well as *how* things are said. In

general, the AAI's main determinant of attachment security is coherence in discourse. The AAI produces four major classifications: *autonomous* (which corresponds to secure attachment); *preoccupied*; *dismissing*; and *unresolved* (each of these correspond to different types of insecure attachment). According to the AAI, these classifications reliably predict a) the quality of the current infant-parent attachment relationship and b) how responsive parents will be towards their own children's attachment signals. The predictive validity has been corroborated in contemporary research, with a large meta-analysis demonstrating strong effect sizes (van Ijzendoorn, 1995). For a more comprehensive discussion of the psychometric properties of the AAI see Hesse (2008).

## *Tests of Anger and Emotional Intelligence*

### **[60.13010] Introduction**

Emotional Intelligence (EI) is the ability to accurately recognise emotions, manage them, and use them to guide behaviour. There are numerous models of EI and little agreement on the construct and its components which limits assessment (Salovey & Grewal, 2005). Broadly speaking the Ability Model of EI concerns emotional reasoning and problem solving (Salovey et al., 2004) and the Trait Model of EI focuses on self-perceived emotional skilfulness as a personality trait (Petrides & Furnham, 2001).

### **[60.13020] Assessment**

The primary assessment of the Ability Model is the *Mayer-Salovey-Caruso Emotional Intelligence Test* (MSCEIT V2.0; Mayer et al, 2003) which consists of tests comparable to intelligence tests in four domains: Perceiving, Facilitating, Understanding and Managing emotions. These domains have shown high internal reliability (Cronbach's  $\alpha=0.90, 0.76, 0.77, 0.81$ ) in an Expert population with similar results in the General population (Mayer et al., 2003). Mayer and colleagues also found evidence supporting the four-factor structure of their model (2003). Use of the MSCEIT is limited, however, by ceiling effects in its norms, by the use of consensus to determine correct responses (Roberts, Zeidner & Matthews, 2001), and by evidence that performance is predicted by general intelligence and personality (Schulte, Ree, Carretta, 2004).

The psychometric properties of the predominant Trait EI assessment, the Trait Emotional Intelligence Questionnaire (TEIQue) indicate that while highly correlated with some personality traits (extraversion, agreeableness, openness and conscientiousness) the Trait EI model explains unique variance (Andrei et al., 2016). The TEIQue consists of four factors (well-being, self-control, emotionality, sociability) for which there is good psychometric support, which are normally distributed, and which are moderately to highly reliable (Cronbach's  $\alpha=0.59-0.91$  across 20 subscales; Mikolajczak et al, 2007).

### **[60.13030] Anger**

Anger and aggression are not considered pathological; however, they are behavioural components of numerous disorders including depression, anxiety, PTSD, ODD and substance abuse disorder. Literature on anger and aggression informs research and interventions targeting interpersonal violence and domestic and family violence, and anger management is incorporated into numerous CBT treatments (Reilly & Shropshire, 2002).

*State Trait Anger Expression Index II* (STAXI-2; Spielberger, 1999)

The most common assessment of Anger is the State Trait Anger Expression Index II (STAXI-2; Spielberger, 1999). It is a 57 item self-report measure that takes between 10 and 25 minutes to

complete and assesses State Anger (intensity of feeling in a particular moment), Trait Anger (characteristic tendencies of a person in relation to anger), Anger Expression and Anger Control. The STAXI has strong face validity and there is evidence for good construct validity: a small positive correlation with the Overt Hostility scale of the MMPI and a moderate positive correlation with the Hostility Index of the same (Spielberger, 1999). While considered valid and reliable for use with forensic populations (Etzler, Rohrman, & Brandt, 2014) judicious interpretation of test results is required given evidence for a social desirability effect (Schamborg, Tully & Browne, 2016).

*The Novaco Anger Scale and Provocation Inventory* (NAS-PI: Novaco, 1994; 2003)

The NAS-PI assesses anger with respect to psychological functioning and physical health. It describes both how a person experiences anger and what situations provoke anger. It is a self-report questionnaire with the revised version (Novaco, 2003) comprising 48 anger items and 25 provocation items. The scale has been used and normed with both “normal” and “clinical” samples and has undergone extensive cross cultural validation. It has also been validated for use in male offenders with a developmental disability (Novaco & Taylor, 2004). The anger scale (NAS) three sub-scales: cognition; arousal; and behaviour. The provocation index (PI) examines the *intensity* of anger across a range of situations that a person might find provocative.

Statistical properties: Novaco (1994) reported internal reliability for the NAS of alpha .95 and test-retest reliability of .84. With respect to the PI, Novaco (1994) reports alpha .95 and test-retest reliability of .86. Mills et al., (1998) reported NAS total alpha as .95 and test-retest reliability of .89, while the PI yielded an alpha of .96 and a test-retest reliability of .85.

## BIBLIOGRAPHY

### [60.15010] Bibliography

Abramowitz JS, and Deacon BJ, "Psychometric properties and construct validity of the Obsessive-Compulsive Inventory-Revised: Replication and extension with a clinical sample". (2006) *Journal of Anxiety Disorders*, 20(8), pp 1016-1035.

Achenbach TM, *Manual for the Child Behavior Checklist/4-18 and 1991 Profile* (1991) Burlington: University of Vermont, Department of Psychiatry.

Adam Y, Meinschmidt G, Gloster AT and Lieb R, "Obsessive-compulsive disorder in the community: 12 moth prevalence, comorbidity and impairment" (2012) *Social Psychiatry and Psychiatric Epidemiology*, 47, pp 339-349.

Aebi M, Plattner B, Steinhausen HC and Bessler C, "Predicting sexual and non-sexual recidivism in a consecutive sample of juveniles convicted of sexual offences" (2011) *Sexual Abuse: A Journal of Research and Treatment*, 23, pp 456-473.

Ainsworth M and Wittig B, "Attachment and exploratory behavior of one year-olds in a strange situation", in B Foss (ed) *Determinants of infant behavior IV* (pp 111-136) (1969) London: Methuen.

Ainsworth MDS, Blehar MC, Waters E and Wall S, *Patterns of attachment: A psychological study of the strange situation* (1978) Hillsdale, NJ: Lawrence Erlbaum Associates.

Altman EG, Hedeker D, Peterson JL and Davis JM, "The Altman Self-Rating Mania Scale" (1997) *Biological Psychiatry*, 42, pp 948-955.

American Psychiatric Association, *Diagnostic and statistical manual of mental disorders* (2013) American Psychiatric Pub.

American Society of Addiction Medicine (ASAM) 2011, *Public Policy Statement: Definition of Addiction*. <http://www.asam.org/research-treatment/definition-of-addiction> (<http://www.asam.org/research-treatment/definition-of-addiction>)

Anastasi A and Urbina S, *Psychological Testing* (7<sup>th</sup> ed) (1997) Upper Saddle River, NJ: Prentice Hall.

Andrei F, Siegling AB, Aloe AM, Baldaro B and Petrides KV, "The incremental validity of the Trait Emotional Intelligence Questionnaire (TEIQue): A systematic review and meta-analysis" (2016) *Journal of personality assessment*, 98(3), pp 261-276.

Andrews DA and Bonta J, *The Level of Service Inventory-Revised* (1995) Toronto, Canada: Multi-Health Systems.

Andrews G and Slade T "Interpreting scores on the Kessler psychological distress scale (K10)" (2001) *Australian and New Zealand journal of public health*, 25(6) pp 494-497.

Autism Speaks: <https://www.autismspeaks.org/what-autism/prevalence> (accessed 31 January 2018).

Australian Dyslexia Association (2016), *Dyslexia in Australia*, retrieved at <http://www.dyslexiaassociation.org.au/what-is-dyslexia> (accessed 31 January 2018).

Babor TF, Higgins-Biddle, Saunders JB and Monterio MG, *The Alcohol Use Disorders Identification Test (AUDIT)* (2001) 2<sup>nd</sup> ed, Department of Mental Health and Substance Dependence; World Health Organization.

Baer RA, Wetter MW, and Berry DTR, "Detection of underreporting of psychopathology on the MMPI: A meta-analysis" (1992), *Clinical Psychology Review*, 12, pp 509-525.

Bagby RM, Ryder AG, Schuller DR and Marshall MB, "The Hamilton Depression Rating Scale: Has the gold standard become a lead weight?" (2004), *The American Journal of Psychiatry*, 161(12), pp 2163-2177.

Bao L, Brownlie EB and Beitchman JH, "Mental health trajectories from adolescence to adulthood: Language disorder and other childhood and adolescent risk factors" (2016) *Development and psychopathology*, 28(02), pp 489-504.

Bardhoshi G, Duncan K and Erford BT, "Psychometric meta-analysis of the English version of the Beck Anxiety Inventory" (2015) *Journal of Counselling and Development*, 94, pp 356-373.

Baron-Cohen S, Wheelwright S, Skinner R, Martin J and Clubley E, "The Autism-Spectrum Quotient: Evidence from Asperger syndrome/high functioning autism, males and females, scientists and mathematicians" (2001) *Journal of Autism and Developmental Disorders*, 31(1), pp 5-17.

Beck A, *Beck Anxiety Inventory* (1990) San Antonio, TX: Psychological Corporation.

Beck AT and Steer RA, *Beck Depression Inventory Manual* (1993). San Antonio, TX: Psychological Corporation.

Beck A and Steer RA, *Beck Hopelessness Scale* (1988) San Antonio, TX: Harcourt Brace.

Beck A and Steer RA, *Beck Scale for Suicide Ideation* (1991) San Antonio, TX: Harcourt Brace.

Beck A and Steer RA and Ranieri WF, "Scale for Suicide Ideation: Psychometric properties of a self report version" (1988) *Journal of Clinical Psychology*, 44(4), pp 499-505.

Beck A and Steer R and Brown G, *Beck Depression Inventory – Second Edition (BDI-II)*. (1996) San Antonio, TX: Psychological Corporation.

Beck JG, Castonguay LG, Chronis-Tuscano A, Klonsky ED, McGinn LK and Youngstrom EA "Principles for training in evidence-based psychology: Recommendations for the graduate curricula in clinical psychology" (2014) *Clinical Psychology: Science and Practice*, 21, pp 410-424.

Benabou R and Tirole J, "Intrinsic and extrinsic motivation" (2003) *The review of economic studies*, 70(3), pp 489-520.

Ben-Porath YS and Tellegen A, *MMPI-2RF: Manual for administration, scoring, and interpretation* (2008/2011) Minneapolis, MN: University of Minnesota Press.

Berument SK, Rutter M, Lord C, Pickles A and Bailey A, "Autism Screening Questionnaire: Diagnostic validity" (1999) *The British Journal of Psychiatry*, 175, pp 444-451.

Birchwood M, Smith J, Drury V, Healy J, Macmillan F and Slade M, "A self-report insight scale for psychosis: Reliability, validity and sensitivity to change" (1994) *Acta Psychiatrica Scandinavica*, 89(1), pp 62-67.

Borum R, Bartel P and Forth A, *Manual for the Structured Assessment of Violence Risk in Youth (SAVRY), version 1.1* (2006) Tampa, FL: University of South Florida.

Borum R and Grisso T, "Psychological test use in criminal forensic evaluations" (1995) *Professional Psychology: Research & Practice*, 26(5), pp 465.

- Bowlby J, *Attachment and loss: Attachment (Vol 1)* (1969) New York: Basic Books.
- Bowlby J, *Attachment and loss: Separation (Vo. 2)* (1973) New York: Basic Books.
- Bramham J, "Attention Deficit-Hyperactivity Disorder in adults", in A Carr and M McNulty (eds) (2016) *The Handbook of Adult Clinical Psychology: An Evidence Based Practice Approach*, London, UK: Routledge.
- Braüinig P, Shugar G and Krüger S, "An investigation of the Self-Report Mania Inventory as a diagnostic and severity scale for mania" (1996) *Comprehensive Psychiatry*, 37, pp 52–55.
- Briere J, *Trauma Symptom Inventory Professional Manual* (1995) Odessa, FL: Psychological Assessment Resources.
- Briere J, "Psychometric review of Trauma Symptom Inventory (TSI)" in BH Stamm (ed), *Measurement of stress, trauma, and adaptation* (pp 381-383) (1996) Lutherville, MD: Sidran Press.
- Briere J, *Trauma Symptom Inventory-2 Professional Manual* (2011) Odessa, FL: Psychological Assessment Resources.
- Brown TA, DiNardo PA, Lehman CL, and Campbell LA, "Reliability of DSM-IV Anxiety and Mood Disorders: Implications for the classification of emotional disorders". (2001) *Journal of Abnormal Psychology*, 110(1), pp 49-58.
- Brown TA and Barlow DH, *Anxiety and Related Disorders Interview Schedule for DSM-V – Adult Version* (2014).
- Brown TE, *Brown Attention-Deficit Disorder Scales (Brown ADD Scales): For Children and Adolescents* (2001) Psychological Corporation.
- Burlingame GM, Seaman S, Johnson JE, Whipple J, Richardson E et al, "Sensitivity to change of the Brief Psychiatric Rating Scale-Extended (BPRS-E): An item and subscale analysis" (2006) *Psychological Services*, 3(2), pp 77-87.
- Butcher JN, Dahlstrom WG, Graham JR, Tellegen A, and Kaemmer B, *The Minnesota Multiphasic Personality Inventory-2 (MMPI-2): Manual for administration and scoring*. (1989) Minneapolis, MN: University of Minnesota Press.
- Butcher JN, Graham JR, Ben-Porath YS, Tellegen A, Dahlstrom WG and Kraemer B, *Minnesota Multiphasic Personality Inventory-2: Manual for administration, scoring and interpretation* (Rev ed) (2001) Minneapolis: University of Minnesota Press.
- Butcher JN, Williams CL, Graham JR, Archer R, Tellegen A, Ben-Porath YS and Kaemmer B, *Minnesota Multiphasic Personality Inventory-Adolescent Version(MMPI-A): Manual for administration, scoring and interpretation* (1992) Minneapolis, MN: University of Minnesota Press.
- Byrne M.K, Parletta N, Webster D, Batterham M and Meyer BJ, "Adult Attention Deficit Disorder and aggressive behaviour: An exploration of the relationship between Brown Attention-Deficit Disorder Scales and the Aggression Questionnaire" (2015) *Psychiatry, Psychology and Law*, 22(3), pp 407-416.
- Carr F, "Psychiatric disorders: Modelling lithium responsiveness in a dish". *Nature Reviews Neuroscience*, 16(12), (2015) pp 702-703.
- Carr A and McNulty M, *The Handbook of Adult Clinical Psychology: An evidence based practice approach* (2<sup>nd</sup> ed) (2016) London: Routledge.

Carrington S, Leekam S, Kent R et al, "Signposting for diagnosis of Autism Spectrum Disorder using the Diagnostic Interview for Social and Communication Disorders (DISCO)" (2015). *Research in Autism Spectrum Disorders*, 9, pp 45-52.

Cashel ML, Rogers R, Sewell K and Martin-Cannici C, "The Personality Assessment Inventory and the detection of defensiveness" (1995) *Assessment*, 1, pp 333-342.

Cattell RB, "Some theoretical issues in adult intelligence testing" (1941) *Psychological Bulletin*, 38, pp 592.

Cattell HEP and Mead AD, "The Sixteen Personality Factor Questionnaire", in GJ Boyle, G Matherws and DH Saklofske (eds), *The Sage Handbook of Personality Theory and Testing: Vol 2, Personality Measurement and Testing* (2008) Los Angeles, CA: Sage Publications.

Cattell RB, Cattell AK and Cattell HEP, *16PF Fifth Edition Questionnaire* (1993) Champaign, IL: Institute for Personality and Ability Testing.

Cerasoli C, Nicklin JM and Ford MT, "Intrinsic motivation and extrinsic incentives jointly predict performance: A 40-year meta-analysis" (2014) *Psychological bulletin*, 140(4), p 980.

Chandler S, Charman T, Baird G, Simonoff E, Loucas T et al, "Validation of the Social Communication Questionnaire in a population cohort of children with Autism Spectrum Disorders" (2007) *Journal of the American Academy of Child & Adolescent Psychiatry*, 46(10), pp 1324-1332.

Charman T, Baird G, Simonoff E, Loucas T, Chandler S, et al, "Efficacy of three screening instruments in the identification of autistic-spectrum disorders" (2007) *The British Journal of Psychiatry*, 191(6), pp 554-559.

Chiu E, "Psychological testing in child custody evaluations with ethnically diverse families: Ethical concerns and practice recommendations". (2014) *Journal of Child Custody*, 11, pp 107-127.

Chlebowski C, Robins DL, Barton M.L and Fein D, "Large-scale use of the modified checklist for autism in low-risk toddlers" (2013) *Pediatrics*, 131(4), e1121-e1127.

Chu CM, Goh ML and Chong D, "The predictive validity of SAVRY ratings for assessing youth offenders in Singapore: A comparison with YLS/CMI ratings" (2016) *Criminal Justice and Behavior*, 43(6), pp 793-810.

Clausen L, Rosenvinge J, Friborg O and Rokedal K, "Validating the Eating Disorder Inventory – 3(EDI-3): A comparison between 561 female eating disorder patients and 878 females from general population" (2011) *Journal of Psychopathology Behaviour Assessment*, 33(1), pp101-110.

Collett BR, Ohan JL and Myers KM, "Ten-year review of rating scales. V: scales assessing attention-deficit/hyperactivity disorder" (2003) *Journal of the American Academy of Child & Adolescent Psychiatry*, 42(9), pp 1015-1037.

Colom F, Vieta E, Martinez-Aran A, Garcia-Garcia M, Reinares M, Salamera M "Spanish version of a scale for the assessment of mania: validity and reliability of the Young Mania Rating Scale" (2002) *Medicina Clinica*, 119(10), pp 366-371.

Conners C, *Conners' Rating Scales-Revised Technical Manual*. (1997) North Tonawanda, NY: Multi-Health Systems.

Conners C, *Conners Third Edition (Conners-3)* (2008) Toronto: Multihealth Systems.

- Constantino JN and Gruber CP, *Social Responsiveness Scale – 2<sup>nd</sup> Edition (SRS-2)* (2012) Torrance, CA: Western Psychological Services.
- Costa PT and McCrae RR, “The revised neo personality inventory (neo-pi-r)”. *The SAGE handbook of personality theory and assessment*, 2, (2008) pp 179-198.
- Crawford JR and Henry JD, “The Depression Anxiety Stress Scales (DASS): Normative data and latent structure in a large non-clinical sample” (2003) *British Journal of Clinical Psychology*, 42: pp111–131.
- Creamer M, Burgess P and McFarlane AC, “Post-traumatic stress disorder: findings from the Australian National Survey of Mental Health and Wellbeing”. (2001) *Psychological Medicine*, 31(7), pp 1237-1247.
- Crowe S, “Forensic Neuropsychology”, in Ian Freckleton and Hugh Selby (eds), *Expert Evidence* (2010) (pp 63A.50) Australia: Thomson Reuters.
- Crowe S, “Neuropsychological Assessment”, in Ian Freckleton and Hugh Selby (eds.) *Expert evidence* (pp 63A.50). (2015) Australia: Thomson Reuters.
- Dazzi F, Shafer A and Lauriola M, “Meta-analysis of the Brief Psychiatric Rating Scale – Expanded (BPRS-E) structure and arguments for a new version” (2016) *Journal of Psychiatric Research*, 81, pp 140-151.
- De Meneses-Gaya C, Zuardi AW, Loureiro SR and Alexandre J (2009), “Alcohol use disorders identification test (AUDIT): An updated systematic review of psychometric properties”. *Psychology & Neuroscience*, 2(1), pp 83-97.
- Deci EL and Ryan RM, “The general causality orientations scale: Self-determination in personality” (1985) *Journal of research in personality*, 19(2), pp 109-134.
- Deci EL and Ryan RM, *Self-determination* (2010) John Wiley & Sons, Inc.
- Dell’Osso L, Gesi C, Massimetti E, Cremone I.M, Barbuti M et al, “Adult Autism Subthreshold Spectrum (AdAS Spectrum): Validation of a questionnaire investigating subthreshold autism spectrum”. (2017) *Comprehensive Psychiatry*, 73, pp 61-83.
- (Derogatis L, *Symptom Checklist-90 Revised (SCL-90-R) and Brief Symptom Inventory Manual* (1994) Minneapolis, MN: National Computer.
- Derogatis LR and Unger R, *Symptom checklist-90-revised* (2010) John Wiley & Sons, Inc.
- DiNardo P, Brown T and Barlow, *Anxiety Disorder Interview Schedule for DSM-IV* (1994) San Antonio, TX: Psychological Corporation.
- Dingemans A, Linszen D, Lenior M and Smeets R, “Component structure of the Brief Psychiatric Rating Scale (BPRS-E)” (1995) *Psychopharmacology*, 122, pp 263-267.
- Douglas KS, Hart SD, Webster CD and Belfrage H., *HCR-20V3: Assessing risk of violence – User guide* (2013) Burnaby, Canada: Mental Health, Law, and Policy Institute, Simon Fraser University.
- Eisen MB, Phillips KA, Baer L, Beer DA, Atala KD and Rasmussen SA, “The Brown Assessment of Beliefs Scale: Reliability and validity”, (1998), *American Journal of Psychiatry*, 155, pp 102-108.
- Elliott CD, *Differential Abilities Scales – II (DAS – II)* (2007) PsychCorp: Pearson.

Etzler SL, Rohrman S, Brandt H, "Validation of the STAXI-2: A study with prison inmates". (2014) *Psychological Test and Assessment Modeling*, 56, pp 178-194.

*Expert witness code of conduct 2005*, updated 2016.

Fairburn CG and Beglin SJ, "Assessment of eating disorders: Interview or self-report questionnaire?" (1994) *International Journal of Eating Disorders*, 16(4), pp 363-370.

Fairburn SG and Cooper Z, "The Eating Disorder Examination (12<sup>th</sup> edition) (1993)", in CG Fairburn & GT Wilson (eds), *Binge Eating: Nature, Assessment and Treatment*. Guilford Press, pp 317-360.

Feldstein SW and Miller WR, "Does subtle screening for substance abuse work? A review of the Substance Abuse Subtle Screening Inventory (SASSI)" (2006) *Addiction*, 102(1), pp41-50.

Filone S and DeMatteo D, "Assessing "credible fear": A psychometric examination of the Trauma Symptom Inventory-2 in the context of immigration court evaluations" (2017) *Psychological Assessment*, 29(6), pp 701-709.

First MB, Williams JBW, Benjamin LS and Spitzer MD, *Structured Clinical Interview for DSM-5 Personality Disorder: User's Guide* (2016) Arlington, VA: American Psychiatric Association.

First MB, Williams JBW, Karg RS and Spitzer RL, *Structured Clinical Interview for DSM-5: Clinician Version User's Guide* (2016) Arlington, VA: American Psychiatric Association.

Foa EB, Huppert JD, Leiberg S, Langer R, Kichic R et al, "The Obsessive-Compulsive Inventory: Development and validation of a short version" (2002) *Psychological Assessment*, 14(4), pp 485-496.

Forth AE, Kosson DS, and Hare RD, *Hare Psychopathy Checklist: Youth Version (PCL:YV). Professional Manual* (2003) Toronto, Canada: MHS Assessments.

Forth AE, Bo S and Kongerslev M, "Assessment of psychopathy: The Hare Psychopathy Checklist measures", in KA Kent and WP Sinnott-Armstrong (eds) (2013) *Handbook on Psychopathy and Law* Oxford University Press.

Garber BD, "Attachment methodology in custody evaluation: four hurdles standing between developmental theory and forensic application" (2009) *Journal of Child Custody*, 6 (1-2), pp 38-61.

Garner DM, *Eating Disorder Inventory-3. Professional Manual* (2004) Lutz, FL: Psychological Assessment Resources, Inc.

George C, Kaplan N and Main M, *Adult Attachment Interview. Unpublished interview* (1984) Berkeley, CA: University of California.

Giddens JM, Harnett-Sheehan K and Sheehan DV, "The Columbia –Suicide Severity Rating Scale (C-SSRS): Has the "gold standard" become a liability" (2014) *Innovations in Clinical Neuroscience*, 11(9-10), pp 66-80.

Gloster AT, Rhoades HM, Novy D, Klotsche J, Senior A, Kunik M, Stanley MA, "Psychometric Properties of the Depression Anxiety and Stress Scale-21 in Older Primary Care Patients" (2008) *Journal of Affective Disorders*, 110(3), pp 248–259.

Goodman R (1997). "The Strengths and Difficulties Questionnaire: A research note". *Psychology and Psychiatry and Allied Disciplines*, 38(5), 581-586.

- Goodman WK, Price LH, Rasmussen SA, Mazure C, Fleischmann RL et al, "The Yale-Brown Obsessive Compulsive Scale: I. Development, use, and reliability" (1989a) *Archives of General Psychiatry*, 46(11), pp 1012-1016.
- Goodman WK, Price LH, Rasmussen SA, Mazure C, Delgado P et al, "The Yale-Brown Obsessive Compulsive Scale: II. Validity" (1989b) *Archives of General Psychiatry*, 46(11), pp 1006-1011
- Gray KM, Tonge BJ and Sweeney DJ, "Using the Autism Diagnostic Interview – Revised and the Autism Diagnostic Observation Schedule with young children with developmental delay: Evaluating diagnostic validity" (2008) *Journal of Autism and Developmental Disorders*, 38(4), pp 657-667.
- Greene RI, "Malingering and Defensiveness on the MMPI-2", in R Rogers (ed) *Clinical Assessment of Malingering and Deception – Third Edition* (2008) Guilford Press: New York, NY.
- Groh AM, Pasco Fearon RM, van Ijzendoorn MH, Bakermans-Kranenburg MJ and Roisman GI, "Attachment in the early life course: meta-analytic evidence for its role in socio-emotional development" (2017) *Child Development Perspectives*, 11(1), pp70-76.
- Groth-Marnet G and Wright AJ *Handbook of Psychological Assessment – Sixth Edition* (2016) Hoboken, NJ: John Wiley & Sons.
- Guriel J, Yanez T, Fremouw W, Shreve-Neiger A, Ware L et al, "Impact of coaching on malingered posttraumatic stress symptoms on the M-FAST and the TSI" (2004). *Journal of Forensic Psychology Practice*, 4, pp 37-56.
- Guy LS and Miller HA, "Screening for malingered psychopathology in a correctional setting: Utility of the Miller-Forensic Assessment of Symptoms Test (M-FAST)" (2004) *Criminal Justice and Behavior*, 31(6), pp 695-716.
- Hanson RK and Morton-Bourgon K, *The accuracy of recidivism risk assessments for sexual offenders: A meta-analysis* (User Report 2007-01) (2007) Ottawa, ON: Public Safety and Emergency Preparedness Canada.
- Hanson RK and Thornton D, *Static 99: Improving actuarial risk assessment for sex offenders* (Use Rep No 1999-02) (1999) Ottawa, Ontario, Canada: Department of the Solicitor General of Canada.
- Hare RD, *The Hare Psychopathy Checklist-Revised* (2003) Toronto, Ontario, Canada: Multi-Systems Health.
- Harrison P and Oakland T, *Adaptive Behavior Assessment System – Third Edition* (ABAS-3) (2015) Pearson: Australia.
- Hathaway SR and McKinley JC "A multiphasic personality schedule (Minnesota): III The measurement of symptomatic depression" (1942) *Journal of Psychology*, 14, pp 73-84.
- Hayes S, "63A Intellectual Disability", in Ian Freckleton and Hugh Selby (eds), *Expert Evidence*, (pp 63A.50) (2015) Australia: Thomson Reuters.
- Helmes E, "Response distortion of the Minnesota Multiphasic Personality Inventory-2 (MMPI-2) in offender rehabilitation" (2008). *Journal of Offender Rehabilitation*, 47(1/2), pp 101-120.

Henry JD and Crawford JR, "The 21-item version of the Depression Anxiety Stress Scales (DASS-21): Normative data and psychometric evaluation in a large non-clinical sample". (2005) *British Journal of Clinical Psychology*, 44, pp 227-239.

Hesse E, "The Adult Attachment Interview: protocol, method of analysis and empirical studies", in *Handbook of attachment: theory, research and clinical applications* (2<sup>nd</sup> ed) (2008) New York, NY: Guildford Press.

Hirschfeld RM, Holzer C, Calabrese JR, Weissman M, Reed M et al, "Validity of the mood disorder questionnaire: a general population study" (2003) *American Journal of Psychiatry*, 160(1), pp 178-180.

Hirschfeld RM, Williams JB, Spitzer RL, Calabrese JR, Flynn L et al, "Development and validation of a screening instrument for bipolar spectrum disorder: the mood disorder questionnaire" (2000) *The American Journal of Psychiatry*, 157 (11), pp 1873-1875.

Hoge RD and Andrews DA, *Youth Level of Service/Case Management Inventory (YLS/CMI): Professional Manual* (2011). Toronto, Ontario: MHS Assessments.

Hogan T, Tsushima W, Psychometrics and testing. *APA handbook of clinical psychology: Applications and methods, Vol 3* [e-book] (2016) Washington, DC, US: American Psychological Association; 2016:31-54. Available from: PsycINFO, Ipswich, MA. Accessed March 22, 2017.

Hus V and Lord C, "The Autism Diagnostic Observation Schedule, Module 4: Revised algorithm and standardized severity scores" (2014) *Journal of Autism and Developmental Disorders*, 44(8), pp 1996-2012.

Hsu C-I, Caputi P and Byrne MK, "The Level of Service Inventory Revised (LSI-R): A useful assessment measure for Australian offenders?" (2009) *Criminal Justice and Behavior*, 36(7), pp 728-740.

Hsu C-I, Caputi P and Byrne MK, "The Level of Service Inventory - Revised: Assessing the risk and need characteristics of Australian Indigenous offenders". (2010). *Psychiatry, Psychology and Law*, 17(3), pp 355-367.

Hsu C-I, Caputi P and Byrne MK, "The Level of Service Inventory Revised (Lsi-R) and Australian offenders: Factor structure, sensitivity, and specificity". (2011) *Criminal Justice and Behavior*, 38(6), pp 600-618.

Iannuzzo RW, Jaeger J, Goldberg JF, Kafantaris V and Sublette ME, "Development and reliability of the HAM-D/MADRS Interview: An integrated depression symptom rating scale". (2006) *Psychiatry Research*, 145(1), pp 21-37.

Jackson RL, Rogers R and Sewell KW, "Forensic applications of the Miller-Forensic Assessment of Symptoms Test (M-FAST): Screening for feigned disorders in competency to stand trial evaluations" (2005) *Law and Human Behavior*, 29, pp 199-210.

Jackson HJ and McGorry PD, *The recognition and management of early psychosis: a preventive approach* (2009) Cambridge University Press.

Jarbin H, Andersson M, Rastam M and Ivarsson T "Predictive validity of the K-SADS-PL 2009 version in school-aged and adolescent outpatients" (2017) *Nordic Journal of Psychiatry*, 71(4), pp 270-276.

- Julian LJ “Measures of anxiety: State-Trait Anxiety Inventory (STAI), Beck Anxiety Inventory (BAI), and Hospital Anxiety and Depression Scale-Anxiety (HADS-A)” (2011), *Arthritis Care Res*, 63, pp 467–472.
- Kabacoff RI, Segal DL, Hersen M and Van Hasselt VB, “Psychometric properties and diagnostic utility of the Beck Anxiety Inventory and the State-Trait Anxiety Inventory with older adult psychiatric outpatients”. (1997). *Journal of Anxiety Disorders*, 11, pp 33–47.
- Kaufman AS and Kaufman NL, *Kaufman Adolescent and Adult Intelligence Test*. (1993) American Guidance Service Inc.
- Kaufman AS and Kaufman NL, *Kaufman Assessment Battery for Children – Second Edition (KABC-II)* (2004) PsychCorp: Pearson.
- Kaufman AS and Kaufman NL, *Kaufman Assessment Battery for Children – Second Edition Normative Update (KABC-II NU)* (2018) PsychCorp: Pearson.
- Kaufman AS and Kaufman NL, *Kaufman Brief Intelligence Test – Second Edition (KBIT-2)* (2004) PsychCorp: Pearson.
- Kaufman J, Birmaher B, Brent D, Rao U, Flynn C et al “Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Versions (K-SAD-PL): Initial reliability and validity data” (1997) *Journal of the American Academy of Child and Adolescent Psychiatry*, 36(7), pp 980-988.
- Kay SR, Opler LA and Fiszbein A, *The Positive and Negative Syndrome Scale (PANSS) Manual* (2000). Multi-health Systems Inc: Toronto.
- Kay SR, Opler LA and Lindenmayer J, “Reliability and validity of the positive and negative syndrome scale for Schizophrenics” (1987) *Psychiatry Research*, 23, pp 99-110.
- Kennedy BL, Schwab JJ, Morris RL and Beldia G, “Assessment of state and trait anxiety in subjects with anxiety and depressive disorders” (2001) *Psychiatric Quarterly*, 72(3), pp 263-276.
- Kerr DCR, Gibson B, Leve LD and DeGarmo DS, “Young adult follow-up adolescent girls in juvenile justice using the Columbia suicide severity rating scale” (2014) *Suicide and Life Threatening Behaviour*, 44(2), pp 113-129.
- Kessler RC, Andrews G, Colpe LJ, Mroczek E, Normand S-LT et al, “Short screening scales to monitor population prevalences and trends in non-specific psychological distress” (2002) *Psychological Medicine*, 32(6), pp 959-976.
- Kessler K, Cacciola J, Alterman A, Faller S, Souza-Formigoni ML, Cruz MS, Brasiliano S and Pechansky F, “Psychometric properties of the sixth version of the Addiction Severity Index (ASI-6) in Brazil” (2012) *Revista Brasileira de Psiquiatria*, 42(1), pp 24-33.
- Kessler RC, Barker PR, Colpe LJ, Epstein J F, Gfroerer JC, Hiripi E and Zaslavsky AM, “Screening for serious mental illness in the general population” (2003) *Archives of general psychiatry*, 60(2), pp 184-189.
- Kleinman JM, Ventola PE, Pandey J, Verbalis AD, Barton M et al, “Diagnostic stability in very young children with autism spectrum disorders” (2008) *Journal of Autism and Developmental Disorders*, 38(4), p 606-615.
- Kranzler HR, Kadden RM, Babor TF and Tennen H, “Validity of the SCID in substance abuse patients” (1996) *Addiction*, 91(6), pp 859-868.

Langton CM, Barbaree HE, Hansen KT, Harkins L and Peacock E J, "Reliability and validity of the Static-2002 among adult sexual offenders with reference to treatment status" (2007) *Criminal Justice and Behavior*, 34, pp 616-640.

Lazowski LE and Geary BB, "Validation of the Adult Substance Abuse Subtle Screening Inventory – 4 (SASSI-4)", (2016) *European Journal of Psychological Assessment. Advance Online Publication*.

Leekam SR, Libby SJ, Wing L, Gould J and Taylor C, "The Diagnostic Interview for Social and Communication Disorders: Algorithms for ICD-10 child autism and Wing and Gould autistic spectrum disorder" (2002) *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 43, pp 327-342.

Lees-Hayley P, "Efficacy of MMPI-2 validity scales and MCMI-II modifier scales for detecting spurious PTSD claims: F, F-k, fake bad scale, ego strength, subtle-obvious scales, DIS, and DEB" (1992) *Journal of Clinical Psychology*, 48(5), pp 681-689.

Lees-Hayley P, English L and Glenn W, "A fake bad scale on the MMPI-2 for personal injury claimants" (1991) *Psychological Reports*, 68, pp 203-210.

Lord C, DiLavore PC, Risi S et al, *Aurism Diagnostic Observation Schedule, Second Edition (ADOS-2)* (2012) Pearson: Australia.

Lord C, Rutter M and Le Couteur A, "Autism Diagnostic Interview-Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders" (1994) *Journal of Autism and Developmental Disorders*, 24(5), pp 659-685.

Lovibond PF and Lovibond SH, *Depression Anxiety Stress Scales*. (2<sup>nd</sup> ed) (1995a) Sydney: Psychology Foundation.

Lovibond PF and Lovibond SH, "The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories" (1995b) *Behaviour Research and Therapy*, 33, pp 335-343.

Luckoff KH, Nuechterlein J and Ventura J, "Symptom monitoring in the rehabilitation of schizophrenic patients". (1986). *Schizophrenia Bulletin*, 12, pp 578-602.

Ludolph PS, "Answered and unanswered questions in attachment theory with implications for children of divorce" (2009) *Journal of Child Custody*, 6(1-2), pp 8-24.

Lynch SM, DeHart DD, Belknap JE, Green BL, Dass-Brailsford P et al, "A multisite study of the prevalence of serious mental illness, PTSD and substance use disorders in women in jail". (2014) *Psychiatric Services*, 65(5), 670-674.

Madan A, Frueh BC, Allen JG, Ellis TE, Rufino KA, Oldham JM and Fowler JC, "Psychometric reevaluation of the Columbia-Suicide Severity Rating Scale: findings from a prospective, inpatient cohort of severely mentally ill adults". (2016) *Journal of clinical psychiatry*, 77(7), pp 867-873.

Main M, Goldwyn R and Hesse E, *Adult attachment scoring and classification systems. Unpublished manuscript* (2002) Berkeley, CA: University of California.

Makela K, "Studies of the reliability and validity of the Addiction Severity Index", (2004), *Addiction*, 99(4), pp 398-410.

- Maljaars J, Noens I, Scholte E and van Berckelaer-Onnes I, "Evaluation of the criterion and convergent validity of the Diagnostic Interview for Social and Communication Disorders in young and low functioning children" (2012) *Autism: The International Journal of Research and Practice*, 16, pp 487-497.
- Maslow AH, "A theory of human motivation" (1943) *Psychological Review*, 50, pp 370-396.
- Mataix-Cols D, Fullana MA, Alonso P, Menchon J and Vallejo J, "Convergent and discriminant validity of the Yale-Brown Obsessive-Compulsive Scale symptom checklist" (2004) *Psychotherapy and Psychosomatics*, 73(3), pp 190-196.
- Mayer JD, Salovey P, Caruso DR and Sitarenios G "Measuring emotional intelligence with the MSCEIT V2.0" (2003) *Emotion*, 3: 97-105. doi:10.1037/1528-3542.3.1.97
- McCrae RR and Costa PT, "Five factor personality model" in P. Oliver, R.W. Robins, & L.A. Lawrence (eds) *Handbook of Personality: Theory and Research – 3<sup>rd</sup> Edition*. (2008) Guilford Press, pp 159-181.
- McKinley JC and Hathaway SR, "A multiphasic personality schedule (Minnesota): IV. Psychasthenia" (1942) *Journal of Applied Psychology*, 26, pp 614-624.
- McLellan AT, Cacciola J, Alterman AI, Rikoon SH and Carise D, "The Addiction Severity Index at 25: Origins, contributions and transitions" (2006) *American Journal of Addictions*, 15, pp 113-124.
- McLellan AT, Kushner H, Metzger D, Peters R, Smith I et al, "The fifth edition of the Addiction Severity Index" (1992) *Journal of Substance Abuse Treatment*, 9(3), pp 199-213.
- McLeod S and McKinnon DH, "Prevalence of communication disorders compared with other learning needs in 14 500 primary and secondary school students" (2007) *International journal of language & communication disorders*, 42(sup1), pp 37-59.
- Mellor D, "Normative data for the Strengths and Difficulties Questionnaire in Australia" (2005) *Australian Psychologist*, 40, pp 215-222.
- Messick S, "Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning" (1995) *Psychological Assessment*, 7, pp 741-749.
- Mikolajczak M, Luminet O, Leroy C and Roy E, "Psychometric properties of the Trait Emotional Intelligence Questionnaire: Factor structure, reliability, construct, and incremental validity in a French-speaking population" (2007) *Journal of personality assessment*, 88(3), pp 338-353.
- Miller CJ, Johnson SL and Eisner L, "Assessment Tools for Adult Bipolar Disorder" (2009) *Clinical Psychology: A Publication of the Division of Clinical Psychology of the American Psychological Association*, 16(2), pp 188-201.
- Miller GA, *The Substance Abuse Subtle Screening Inventory (SASSI) Manual, Second Edition* (1985, 1999) Springville, IN: The SASSI Institute.
- Miller HA, *M-FAST: Miller Forensic Assessment of Symptoms Test: Professional Manual*. (2001) Odessa, FL: Psychological Assessment Resources.
- Miller HA, "Examining the use of the M-FAST with criminal defendants incompetent to stand trial" (2004) *International Journal of Offender Therapy and Comparative Criminology*, 48(3), pp 268-280.

Millon T, Grossman S and Millon C *MCMI-IV: Millon Clinical Multiaxial Inventory Manual* (1st ed) (2015) Bloomington, MN: NCS Pearson, Inc.

Mills JF, Kroner DG and Forth AE “Novaco Anger Scale: Reliability and validity within an adult clinical sample” (1998) *Assessment*, 5, pp 237-248.

Moray L, *An Interpretive Guide to the Personality Assessment Inventory* (1996) PAR, Florida: USA.

Moray LC, *Personality Assessment Inventory (PAI): Professional Manual* (2007) 2<sup>nd</sup> edition, PAR, Florida: USA.

Nah Y-H, Young RL and Brewer N, “Using the Autism Detection in Early Childhood (ADEC) and Childhood Autism Rating Scales (CARS) to predict long term outcomes in children with autism spectrum disorders” (2014) *Journal of Autism and Developmental Disorders*, 44(9), pp 2301-2310.

Nah Y-H, Young RL, Brewer N and Berlingeri G, “Autism detection in early childhood (ADEC): Reliability and validity data for a level 2 screening tool for autistic disorder” (2014). *Psychological Assessment*, 26(1), pp 215-226.

Nisbet IA, Smallbone SW and Wortley R, “Developmental, Individual and Family Characteristics of Specialized, Versatile and Short-Duration Adolescent Sex Offenders” (2010) *Sexual Abuse in Australia and New Zealand*, 2(2), pp 86-97.

Norris FH and Raid JK, “Standardised self-report measures of civilian trauma and posttraumatic stress disorder”, in JP Wilson and TM Keane (eds) *Assessing psychological trauma and PTSD* (1997) The Guilford Press: New York.

Novaco RW, “Assessing anger among mentally disordered persons as a risk for violence”, in J Monahan and H Steadman (eds) (1994) *Violence and Mental Disorder: Developments in Risk Assessment* Chicago: University of Chicago Press, pp 21-59.

Novaco RW, *The Novaco Anger Scale and Provocation Inventory: Professional Manual*. (2012) Western Psychological Services.

Novaco RW and Taylor JL, “Assessment of anger and aggression in male offenders with developmental disabilities” (2004) *Psychological Assessment*, 11(4), pp 233-246.

Nygren G, Hagberg B, Billstedt E, Skoglund A, Gillberg C and Johansson M, “The Swedish version of the Diagnostic Interview for Social and Communication Disorders (DISCO-10. Psychometric properties” (2009) *Journal of Autism and Developmental Disorders*, 39, pp 730-741.

Osman A, Hoffman J, Barrios FX, Kopper BA, Breitenstein JL and Hahn SK, “Factor structure, reliability, and validity of the Beck Anxiety Inventory in adolescent psychiatric inpatients” (2002) *Journal of Clinical Psychology*, 58(4), pp 443-456.

Overall JE and Gorham DR, “The Brief Psychiatric Rating Scale” (1962) *Psychological Reports*, 10, pp 790-812.

Patton JH, Dtanford MS and Barratt ES, “Factor structure of the Barratt Impulsiveness Scale”, (1995) *Journal of Clinical Psychology*, 6, pp 768-774.

Peter Samuel C, Whelan James P, Ginley Meredith K, Pfund Rory A, Wilson Katherine K, and Meyers Andrew W, “Disordered gamblers with and without ADHD: the role of coping in elevated psychological distress” (2016) *International Gambling Studies*, 16(3), pp 455-469.

- Petrides K and Furnham A, "Trait Emotional Intelligence: Psychometric Investigation with Reference to Established Trait Taxonomies" (2001) *European Journal of Personality*, pp 425-448.
- Posner K, Brown GK, Stanley B, Brent DA, Yershova KV, Mann JJ, "The Columbia-Suicide Severity Rating Scale: Initial Validity and Internal Consistency Findings from Three Multisite Studies with Adolescents and Adults" (2011) *American Journal of Psychiatry* 168(12), pp 1266-77.
- Prentky RA and Righthand S, *Juvenile Sex Offender Assessment Protocol-II: J-SOAP-II manual* (2003) Bridgewater, MA; Author.
- Pressman DE, "Risk assessment decisions for violent political extremism", *Risk Assessment decisions for Violent Extremism 2009-02*, *Public Safety Canada* (2009) Government of Canada, Ottawa, Cat No PS3-1/2009-02-1E-PDF.
- Pressman DE and Flockton JS, "VERA-2 Violent extremism risk assessment, version 2: The manual" (2010) *Unpublished manuscript*.
- Pressman DE and Flockton J, "Calibrating risk for violent political extremists and terrorists: The VERA 2 structured assessment" (2012) *The British Journal of Forensic Practice*, 14(4), pp 237-251.
- Rajlic G and Gretton HM, "An examination of two sexual recidivism measures in adolescent offenders: Moderating effect of offender type". (2010) *Criminal Justice and Behaviour*, 37(10), pp 1066-1085.
- Raven JC, *Raven's Standard Progressive Matrices (SPM)* (1998) Pearson: Australia.
- Ready RE and Veague HB, "Training in psychological assessment: Current practices of clinical psychology programs" (2014) *Professional Psychology: Research & Practice*, 45(4), pp 278-282.
- Reilly PM and Shropshire MS, *Anger management for substance abuse and mental health clients: cognitive behavioral therapy manual* (2002) US Department of Health & Human Services.
- Ritvo RA, Ritvo ER, Guthrie D, Ritvo MJ, Hufnagel DH et al, "The Ritvo Autism Diagnostic Scale-revised (RAADS-R): A scale to assist the diagnosis of autism spectrum disorder in adults: An international validation study" (2011) *Journal of Autism and Developmental Disorders*, 41(8), pp 1076-1089.
- Ritvo RA, Ritvo ER, Guthrie D, Yuwiler A, Ritvo M.J and Weisbender L, "A scale to assist the diagnosis of autism and Asperger's disorder in adults (RAADS): A pilot study". (2008). *Journal of Autism and Developmental Disorders*, 38(2), pp 213-223.
- Roberts RD, Zeidner M, Matthews G, "Does emotional intelligence meet traditional standards for an intelligence? Some new data and conclusions" (2001) *Emotion* 1, pp 196-231. doi:10.1037/1528-3542.1.3.196.
- Robins DL, Fein D and Barton, *Modified Checklist for Autism in Toddlers Follow Up (M-CHAT/F)* (1999) Self Published.
- Robins D, Casagrande K, Barton M, Chen CMA, Dumont-Mathieu Y, and Fein D, "Validation of the Modified Checklist for Autism in Toddlers, Revised with Follow-up (M-CHAT-R/F)". *Pediatrics*, 133(1), (2014) pp 37-45.

Rogers R, *Structured Interview of Reported Symptoms* (1992) Odessa, FL: Psychological Assessment Resources.

Rogers R, "Structured interviews and dissimulation", in R Rogers (ed) *Clinical Assessment of Malingering and Deception* (1997) (2<sup>nd</sup> ed) New York: Guilford Press, pp 301-327).

Rogers R, *Clinical Assessment of Malingering and Deception – Third Edition* (2008) Guilford Press: New York, NY.

Rogers R, "An introduction to response styles", in R Rogers (ed) *Clinical Assessment of Malingering and Deception – Third Edition* (2008) Guilford Press: New York, NY pp 3-13

Rogers R, Bagby RM, and Dickens SE, *Structured Interview of Reported Symptoms (SIRS) and Professional Manual* (1992). Odessa, FL: Psychological Assessment Resources.

Rogers R, Flores J, Ustad K and Sewell KW, "(Initial validation of the Personality Assessment Inventory – Spanish version with clients from Mexican American communities)" 1995 *Journal of Personality Assessment*, 64, pp 340-348.

Rogers R, Jackson RL and Cashel M, "The Schedule for Affective Disorders and Schizophrenia (SADS)", in Rogers R, (ed) *Handbook of diagnostic and structural interviewing* (2001) New York: Guilford Press.

Roid GH, *Stanford-Binet Intelligence Scales (SB5) – Fifth Edition* (2003) Pro-ED: USA.

Rutter M, Bailey A And Lord C, *Social Communication Questionnaire (SCQ)* (2003) Los Angeles: Western Psychological Services.

Rutter M, Le Couteur A and Lord C, *Autism Diagnostic Interview-Revised (ADI-R)* (2003) Pearson: Australia.

Ryan RM and Deci EL, "Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being" (2000) *American Psychologist*, 55(1), pp 68-78.

Rudnick A and Lundberg E, "The stress-vulnerability model of schizophrenia: A conceptual analysis and selective review" (2012). *Current Psychiatry Reviews*, 8(4), pp 337-341.

Sadeghi V, Najafi S, Rostami R and Ghorbani N, "The evaluation of validity and reliability of substance abuse subtle screening inventory (SASSI-3)" (2010) *Procedia – Social and Behavioral Sciences*, 5, pp 1129-1134.

Sadowski VF, Rossegger A, Pressman DE, Rinne T, Duits N AND Endrass J, "Das Violent Extremism Risk Assessment Version 2 Revised (VERA-2R): Eine Skala zur Beurteilung des Risikos extremistischer Gewalt/Deutsche Übersetzung (The Violent Extremism Risk Assessment Version 2 Revised (VERA-2R): A scale for assessing the risk of extremist violence / German translation)" (2017). *Kriminalistik*, 71(5), pp 335-342.

Salovey P and Grewal D, "The Science of Emotional Intelligence" (2005) *Current Directions in Psychological Science*. 14: 6. doi:10.1111/j.0963-7214.2005.00381.x.

Salovey Peter, Mayer John, Caruso David, "Emotional Intelligence: Theory, Findings, and Implications", (2004) *Psychological Inquiry*, pp 197–215.

Sarma KM, "Risk assessment and the prevention of radicalization from nonviolence into terrorism" (2017). *American Psychologist*, 72(3), pp 278-288.

- Schamborg S, Tully RJ and Browne KD, "The use of the State-Trait Anger Expression Inventory-II with forensic populations: a psychometric critique" (2016) *International journal of offender therapy and comparative criminology*, 60(11), pp 1239-1256.
- Schmitz N, Hartkamp N, Kiuse J, Franke GH, Reister G and Tress W, "The symptom check-list-90-R (SCL-90-R): a German validation study" (2000). *Quality of Life Research*, 9(2), pp 185-193.
- Schrank FA, McGrew KS and Mather N, *Woodcock Johnson Fourth Edition (WJ IV) – Australian Adaptation* (2015) Psychological Assessments Australia.
- Schulte M J, Ree M J and Carretta TR "Emotional intelligence: Not much more than g and personality" *Personality and Individual Differences*, (2004) 37, pp 1059–1068.
- Selin KH, "Test-retest reliability of the Alcohol Use Disorder Identification Test in a general population sample" (2003). *Alcoholism: Clinical and Experimental Research*, 27, pp 1428-1435.
- Senate Select Committee on Mental Health, 2006, *A National Approach to Mental Health – from crisis to community: First Report*. Retrieved from: [http://www.aph.gov.au/Parliamentary\\_Business/Committees/Senate/Former\\_Committees/mentalhealth/report/c13](http://www.aph.gov.au/Parliamentary_Business/Committees/Senate/Former_Committees/mentalhealth/report/c13) (accessed 6 February 2018).
- Senate Standing Committee on Community Affairs, 2014, *Prevalence of different types of speech, language and communication disorders and speech pathology services in Australia*. Retrieved from: [http://www.aph.gov.au/Parliamentary\\_Business/Committees/Senate/Community\\_Affairs/Speech\\_Pathology/Report](http://www.aph.gov.au/Parliamentary_Business/Committees/Senate/Community_Affairs/Speech_Pathology/Report) (accessed 6 February 2018).
- Shepherd SM, Adams Y, McEntyre E and Walker R, "Violence risk assessment in Australian Aboriginal offender populations: A review of the literature", (2014) *Psychology, Public Policy, and Law*, 20, pp 281-293.
- Shepherd SM, Luebbers S, Ferguson M, Ogloff JRP and Dolan M, "The utility of the SAVRY across ethnicity in Australian young offenders" (2014) *Psychology, Public Policy, and Law*, 20(1), pp 31-45.
- Shepherd S and Strand S, "The Utility of the Psychopathy Checklist: Youth Version (PCL: YV) and the Youth Psychopathic Trait Inventory (YPI) – Is it Meaningful to Measure Psychopathy in Young Offenders?" (2016) *Psychological Assessment*, 2(4), pp 405-415.
- Shipley KG and McAfee JG, *Assessment in speech-language pathology: A resource manual* (2015) Nelson Education.
- Shum D, O’Gorman J, Creed P and Myers B, *Psychological Testing and Assessment – 3<sup>rd</sup> ed* (2017) South Melbourne, Victoria: Oxford University Press.
- Siefert CJ, Sinclair SJ, Kehl-Fei KA and Blais MA, "An item-level psychometric analysis of the Personality Assessment Inventory: Clinical Scales in a psychiatric inpatient unit" (2009) *Assessment*, 16(4), pp 373-383.
- Silverman W and Albano A, *The Anxiety Disorder Interview Schedule for Children –IV – Child and Parent Versions* (1996) Boulder, CO: Graywind.
- Silverman WK, Saavedra LM and Pina AA, "Test-retest reliability of anxiety symptoms and diagnoses with the anxiety disorder interview schedule for DSM-IV: child and parent versions" (2001) *Journal of the American Academy of Child and Adolescent Psychiatry*, 40(8), pp 937-944.

Silvetti M, Wiersema JR, Sonuga-Barke E and Verguts T, “Deficient reinforcement learning in medial frontal cortex as a model of dopamine-related motivational deficits in ADHD” (2013) *Neural Networks*, 46, pp 199-209.

Simon E and Bögels SM, “Screening for anxiety disorders in children”, (2009) *European Child & Adolescent Psychiatry*, 18(10), pp 625–634.

Skeem JL and Cooke DJ, “Is criminal behavior a central component of psychopathy? Conceptual directions for resolving the debate” (2010) *Psychological Assessment*, 22, pp 433–445.

Smith GP, “Brief screening measures for the detection of feigned psychopathology”, in R Rogers (ed) (2008) *Clinical Assessment of Malingering and Deception, 3<sup>rd</sup> Edition*, Guilford Press, pp 323-339.

Sonuga-Barke EJ, Cortese S, Fairchild G and Stringaris A, “Annual Research Review: Transdiagnostic neuroscience of child and adolescent mental disorders—differentiating decision making in attention-deficit/hyperactivity disorder, conduct disorder, depression, and anxiety” (2016) *Journal of Child Psychology and Psychiatry* 57(3), pp 321-349.

Sparrow SS, Cicchetti DV and Balla DA, *Vineland Adaptive Behavior Scales – Second Edition (Vineland-II)* (2005) PsychCorp: Pearson.

Spielberger CD, *STAXI-2: Professional manual* (1999) Psychological Assessment Resources, Lutz, FL.

Spielberger CD, Gorsuch RL, Lushene R, Vagg PR and Jacobs GA, *Manual for the State-Trait Anxiety Inventory* (1983) Palo Alto, CA: Consulting Psychology Press.

Spitzer RA and Endicott J, *Schedule for Affective Disorders and Schizophrenia- Change version* (1978) New York: Biometrics Research, State Psychiatric Institute.

Stanford MS, Mathias CW, Dougherty DM, Lake SL, Anderson NE and Patton JH “Fifty years of the Barratt Impulsiveness Scale: An update and Review” (2009) *Personality and Individual Differences*, 47, pp 385-395.

Sulkowski ML, Storch EA, Geffken GR, Ricketts E, Murphy TK and Goodman WK, “Concurrent validity of the Yale-Brown Obsessive-Compulsive Scale-Symptom Checklist” (2008) *Journal of Clinical Psychology*, 64(12), pp 1338-1351.

Taylor JW, “Psychological assessment”, in Ian Freckleton and Hugh Selby (eds) *Expert evidence* (Chapter 60) (2015) Australia: Thomson Reuters.

Tellegen A and Ben-Porath YS, *MMPI-2-RF (Minnesota Multiphasic Personality Inventory-2 Restructured Form): Technical manual* (2008) Minneapolis: University of Minnesota Press.

Thompson RA and Raikes HA, “Toward the next quarter-century: conceptual and methodological challenges for attachment theory” (2003) *Developmental Psychopathology*, 15(3), pp 691-718.

Van Izendoorn MH “Adult attachment representations, parental responsiveness and intact attachment: A meta-analysis on the predictive validity of the Adult Attachment Interview”. (1995) *Psychological Bulletin*, 117(3), pp 387-403.

Van Voorhees EE, Hardy KK and Kollins SH, “Reliability and validity of self- and other ratings of symptoms of ADHD in adults” (2011) *Journal of Attention Disorders*, 15(3), pp 224-234.

- Ventura J, Green MF, Shaner A and Liberman RP, "Training and quality assurance with the Brief Psychiatric Rating Scale: "the drift busters"" (1993a) *International Journal of Methods in Psychiatric Research*, 3, pp 221-224.
- Ventura J, Luckoff D, Neuchterlein KH, Liberman RP, Green MF and Shaner A, "Brief Psychiatric Rating Scale (BPRS) expanded version: Scales, anchor points, and administration Manual" (1993b) *International Journal of Methods in Psychiatric Research*, 3, pp 227-243.
- Viljoen JL, Elkovitch N, Scalora M and Ullman D, "Assessing risk for violence in adolescents who have sexually offended: An examination of the predictive validity of the ERASOR, YLS/CMI, PCL-YV" (2009) *Criminal Justice and Behavior*, 36, pp 981-1000.
- Viljoen JL, Gray AL, Shaffer C, Bhanwer A, Tafreshi D and Douglas KS, "Does reassessment of risk improve predictions? A framework and examination of the SAVRY and YLS/CMI". (2017) *Psychological Assessment*, 29(9), pp 1096-1110.
- Vitacco MJ, Rogers R, Gabel J and Munizza J, "An evaluation of malingering screens with competency to stand trial patients: A known-groups comparison" (2007) *Law and Human Behavior*, 31, pp 249-260.
- Wang YP and Gorenstein C, "Psychometric properties of the Beck Depression Inventory – II: A comprehensive review". (2013). *Revista Brasileira de Psiquiatria*, 35(4), pp 416-431.
- Weathers FW, Bovin MJ, Lee DJ, Sloan DM, Schnurr PP, Marx BP, "The clinician-administered PTSD scale for DSM-5 (CAPS-5): Development and initial psychometric evaluation in military veterans" (2017) *Psychological Assessment*. Advance online publication.
- Weathers FW, Keane TM and Davidson JR, "Clinician-administered PTSD scale: A review of the first ten years of research" (2001) *Depression and Anxiety*, 13, pp 132-156.
- Weathers FW, Blake DD, Schnurr PP, Kaloupek DG, Marx BP and Keane TM, *The Clinician-Administered PTSD Scale for DSM-5 (CAPS-5)* (2013). Interview available from the National Center for PTSD at <http://www.ptsd.va.gov> (accessed 6 February 2018).
- Weber-Rouget B, Gervasoni N, Dubuis V, Gex-Fabry M, Bondolfi G and Aubry J, "Screening for bipolar disorders using the French version of the Mood Disorder Questionnaire (MDQ)" (2005) *Journal of Affective Disorders*, 88, pp 103–108.
- Wechsler D, "Intelligence defined and undefined: A relativistic appraisal". (1975) *American Psychologist*, 30(2), pp 135.
- Wechsler D and Naglieri JA, *Wechsler Nonverbal Scale of Ability (WNV)* (2006). PsychCorp: Pearson.
- Weiss LG, Saklofske D, Holdnack JA and Prifitera A, *WISC-V: Assessment and Interpretation* (2015) USA, Academic Press.
- Widows M and Smith GP, *Structured Inventory of Malingered Symptomology (SIMS) and Professional Manual* (2005) Odessa, FL: Psychological Assessment Resources.
- Williams JB, "A structured interview guide for the Hamilton Depression Rating Scale" (1988) *Archives of General Psychiatry*, 45(8), pp 742-747.
- Williams J, Gibbon M, First MB, Spitzer RL, Davies M, Wittchen H, "The structured clinical interview for DSM-II-R (SCID): Multi-site test retest reliability" (1992) *Archives of General Psychiatry*, 49(8), pp 630-636.

Wing L, Leekam SR, Libby SJ, Gould J and Larcombe M, "The Diagnostic Interview for Social and Communication Disorders: Background, inter-rater reliability and clinical use". (2002) *Journals of Child Psychology and Psychiatry and Allied Disciplines*, 43(3), pp 307-325.

Wong SCP and Gordon A, "The validity and reliability of the Violence Risk Scale: A treatment-friendly violence risk assessment tool" (2006) *Psychology, Public Policy, and Law*, 12(3), pp 279-309.

Wong SCP, Olver ME, Nicholaichuk T and Gordon A, *Violence Risk Scale-Sex Offender Version* (2004-2006). Available online at <http://www.psynergy.ca> (accessed 5 February 2018).

Wood JJ, Piacentini JC, Bergman RL, McCracken J and Barrios V, "Concurrent validity of the anxiety disorders section of the Anxiety Disorders Interview Schedule for DSM-IV: child and parent versions" (2002) *Journal of Clinical Child and Adolescent Psychology*, 31(3), pp 335-342.

World Health Organisation *Mental health: a state of well-being* (2014). Retrieved from: [http://www.who.int/features/factfiles/mental\\_health/en](http://www.who.int/features/factfiles/mental_health/en) (accessed 2 February 2018).

Worling JR, "The Estimate of Risk of Adolescent Sexual Offence Recidivism (ERASOR): Preliminary psychometric data" (2004) *Sexual Abuse*, 16(3), pp 235-254.

Young RL, *Autism Detection in Early Childhood: ADEC* (2007) Camberwell, Victoria, Australia: Australian Council of Educational Research.

Young RC, Biggs JT, Ziegler VE and Meyer DA, "A rating scale for mania: reliability, validity, and sensitivity" (1978) *British Journal of Psychiatry*, 133, pp 429-435.

Zanarini MC, Skodol AE, Bender D et al, "The Collaborative Longitudinal Personality Disorder Study: reliability of Axis I and II diagnoses" (2000) *Journal of Personality Disorders*, 14(4), pp 291-299.

Zeanah CH, Berline LJ and Boris NW "Practitioner review: clinical applications of attachment theory and research for infants and young children" (2011) *Journal of Child Psychology and Psychiatry*, 52(8), pp819-833.

[The next text page is 61-1]