A Review of Screening, Assessment and Outcome Measures for Drug and Alcohol Settings

Written by Mark Deady

For the Network of Alcohol & other Drug Agencies (NADA) as part of the Drug and Alcohol and Mental Health Information Management Project

2009

Table of Contents

Abbreviations

Acronym	Full Title
AATOM(-C; -R)	Australian Alcohol Treatment Outcome Measure (Clinicians; Researchers)
ADHDQ	Attention Deficit Hyperactivity Disorder Questionnaire
ADS	Alcohol Dependence Scale
ASI	Addiction Severity Index
ASSIST	The Alcohol, Smoking and Substance Involvement Screening Test
AUDIT	Alcohol Use Disorders Identification Test
AVON	Avon Mental Health Measure
BAI	Beck Anxiety Inventory
BASIS-32	Behaviour and Symptom Identification Scale
BDDQ	Body Dysmorphic Disorder Questionnaire
BDEPQ	Benzodiazepine Dependence Questionnaire
BDI	Beck Depression Inventory
BHS	Beck Hopelessness Scale
BIDQ	Body Image Disturbance Questionnaire
BPRS	Brief Psychiatric Rating Scale
BSI	Brief Symptom Inventory
BSSI	Beck Scale for Suicidal Ideation
BTOM	Brief Treatment Outcome Measure
CABA	Child and Adolescent Burden Assessment
CAGE / CAGEAID	Cutdown, Annoyed, Guilty, Eye-Opener (Adapted to Include Drugs)
CAN	Camberwell Assessment of Need
CAPE-BRS	Clifton Assessment Procedures for the Elderly
CAPS	Clinician-Administered PTSD Scale
CBLC	Child Behavior Check List
CCQ	Cocaine Craving Questionnaire
CDI	Children's Depression Inventory
(C)GAS	(Children's) Global Assessment Scale
CIBIC+	Clinician's Interview Based Impression of Change Scale
CIDI	Composite International Diagnostic Interview
ComQoL	Comprehensive Quality of Life Scale
CPQ	Cannabis Problems Questionnaire
CUDIT	Cannabis Use Disorders Identification Test
D&A	Drug and Alcohol
DALI	Dartmouth Assessment of Lifestyle Instrument
DAS	Disability Assessment Schedule
DASS	Depression Anxiety Stress Scales
DAST(-A)	Drug Abuse Screening Test (Adolescent)
DHS	Dispositional Hope Scale

DSM(-III; -IV)	Diagnostic and Statistical Manual of Mental Disorders (3rd; 4th Edition)
DUDIT	Drug Use Disorders Identification Test
(CH)EAT	(Children's) Eating Attitudes Test
EDI	Eating Disorder Inventory
GAF	Global Assessment of Functioning Scale
GAIN	Global Appraisal of Individual Needs
GHQ	General Health Questionnaire
GWB	General Wellbeing Scale
HoNOS	Health of a Nation Outcome Scale
ICD(-9; -10)	International Classification of Diseases (9th; 10th Revision)
ICQ	Impaired Control Scale
IES	Impact of Events Scale
IRIS	Indigenous Risk Impact Screen
K10	Kessler Psychological Distress Scale
LDQ	Leeds Dependence Questionnaire
LES	Life Experience Survey
LOCSS	Location of Community Support Scale
LRI	Life Regard Index
LSP	Life Skills Profile
MAP	Maudsley Addiction Profile
MAT	Manchester Audit Tool
(b; S)MAST(-G)	(Brief; Short) Michigan Alcoholism Screening Test (Geriatric)
MCQ	Marijuana Craving Questionnaire
MCS	Mental Component Summary
MHI	Mental Health Inventory
MMSE	Mini-Mental State Examination
MOS-SSS	Medical Outcome Study Social Support Survey
NPV	Negative Predictive Value
OAP	Outcome Assessment Program
OMP	Opioid maintenance pharmacotherapy
OPUS	Outcome of Problems of Users of Services
OTI	Opiate Treatment Index
PACS	Penn Alcohol-Craving Scale
PANAS	Positive and Negative Affect Schedule
PANSS	Positive and Negative Syndrome Scale
PC-PTSD	Primary Care PTSD Screen
PCL-C	PTSD Checklist
PCS	Physical Component Summary
PDS	Posttraumatic Stress Diagnostic Scale
PIL	Purpose in Life Test

POSIT	Problem Oriented Screening Instrument for Teenagers substance use/abuse scale
PPV	Positive Predictive Power Value
PSAS	Psychiatric Symptom Assessment Scale
PSH	Psychosocial History
(M)PSS-SR	(Modified) PTSD Symptom Scale Self-Report
PTSD	Posttraumatic Stress Disorder
QoL	Quality of Life assessment
RAS	Recovery Assessment Scale
RAFLS	Resource Associated Functional Level Scale
RCMAS	Revised Children's Manifest Anxiety Scale
RFS	Role Functioning Scale
SA-45	Symptom Assessment
SADD	Short Alcohol Dependence Data Questionnaire
SADQ(-C)	Severity of Alcohol Dependence Questionnaire
SANS	Scale for Assessment of Negative Symptoms
SAS	Social Adjustment Scale
SBS	Social Behaviour Scale
SCAN	Schedules for Clinical Assessment in Neuropsychiatry
SCID	Structured Clinical Interview for DSM-IV
SCL-90-R	Symptom Checklist-90-Revised
SDQ	Strengths & Difficulties Questionnaire
SDS	Severity of Dependence Scale
SDSS	Substance Dependence Severity Scale
SF-36; SF-12	Short-Form Health Survey
SOC	Sense of Coherence Scale
SODQ	Severity of Opiate Dependence Scale
SPWB	Scales of Psychological Well-being
SRQ	Self Reporting Questionnaire
STAI (STAI-S; STAI-T)	Spielberger State Trait Anxiety Inventory
STORI	Stages of Recovery Instrument
T-ACE	Tolerance; Annoyed; Cut down; Eye Opener
TLEQ	Traumatic Life Events Questionnaire
TLFB	Timeline Followback Method
TOP	Treatment Outcomes Profile
TSQ	Trauma Screening Questionnaire
TWEAK	Tolerance; Worried; Eye Opener; Amnesia; (K) Cut down
WHODAS	World Health Organisation Disability Assessment Schedule
WHOQoL	World Health Organisation Quality of Life
YSR	Youth Self Report

Scope of this Review (Aims and Limitations)

Standardised tools cover a range of areas which may be relevant to Drug and Alcohol (D&A) services. This review provides an overview of some useful standardised tools that can be used to measure treatment outcomes and to screen and assess for mental health symptoms and conditions, drug and alcohol use and disorders and general functioning. **Focus has been given to tools that require limited training to use and are freely available**. It should be noted that some of these tools require specialist training, or else mislabelling, misinterpretation, or inappropriate use may occur (Groth-Marnat, 2003; Roche & Pollard, 2006). Some tools are copyright protected and need to be purchased, and/or require the user to have specific qualifications. It is important that workers are aware of what they are, and are not, trained to use, and seek training where required.

Screening is designed only to highlight the existence of symptoms, not to diagnose clients. Most of the measures described are completed as a self-report (i.e., they are completed by the client). Others, however, need to be administered by a worker. It should be noted that, unfortunately, there are no brief measures with established reliability and validity for the identification of possible personality disorders. The possible presence of these disorders needs to be assessed by a health professional that is qualified and trained to do so (e.g., a registered or clinical psychologist, or psychiatrist).

There is a general lack of a standardised approach to screening, assessment and outcome measurement in the D&A sector. A variety of different tools are used, some of which are empirically established instruments whilst others are purpose-built, internally designed tools with increased practicality and utility but unknown validity and reliability (Roche & Pollard, 2006). This review focuses solely on the former. Similarly, it is important to note that this review, in and of itself is not exhaustive, as the number of available instruments is vast. Nevertheless, all attempts have been made to include the most relevant and useful measures.

Methodology

This review will focus on both domestic and international tools with particular consideration given to those tools widely used in NSW. The review includes tools and measures related to substance misuse, mental health symptomology and wider general health and social functioning issues. Information on the psychometric properties, availability, applicability and accessibility of each measure/tool will be presented. To obtain such information, a thorough search of the literature was conducted, using a range of databases (including: Medline, PsychInfo, Cochrane Database of Systemic Reviews, Pubmed, the National Clearinghouse for Alcohol and Drug Information, Evidence Based Medicine Reviews, industry magazines such as 'Of Substance', the Alcohol and Other Drug Council of Australia's (ADCA) Drug database, and associated reference lists), in conjunction with both other published literature and a grey literature search.

Introduction

Given the high rates of co-occurring mental health conditions among clients of D&A treatment services, it is essential that routine screening and assessment be undertaken for these conditions as part of case formulation. Case formulation involves the gathering of information regarding factors that may be relevant to treatment planning, and formulating a hypothesis as to how these factors fit together to form the current presentation of the client's symptoms (Sim, Gwee, & Bateman, 2005). This information is the first step to devising (and later revising) the client's treatment plan. There is no standardised approach to case formulation (Roche & Pollard, 2006), but it is crucial that a range of different dimensions be considered, including history of present illness, D&A use history (amount and frequency, abuse or dependency syndrome), physical/medical conditions, mental state, psychiatric history, trauma history, suicidal or violent ideation, readiness to change, family history, criminal history, social and cultural issues. Consideration also needs to be given to the client's age, sex, sexual orientation, ethnicity, spirituality, socio-economic status, and cognitive abilities. Screening is the initial step in the process of identifying possible conditions (Croton, 2007). This process is not diagnostic (i.e., it cannot establish whether a disorder actually exists); but rather, it identifies the presence of symptoms which may indicate the presence of a disorder. Thus, screening helps to identify individuals with symptoms that may require further investigation and treatment.

Identifying the needs of clients is fundamental to the case formulation process. It is important that whatever needs the client might have are recognised as they will undoubtedly impact upon treatment. Early diagnosis and treatment of conditions can improve client treatment outcomes (Myrick & Brady, 2003). Identification does not necessarily mean that the worker has to personally treat the difficulty the client is experiencing; however, it allows workers to consider the impact of these needs, manage them accordingly, and engage other services where necessary.

Early identification of all conditions present allows for early intervention, which may lead to better prognosis, more comprehensive treatment, and the prevention of secondary disorders (Chan, Dennis, & Funk, 2008). Diagnostic assessment should ideally occur subsequent to a period of abstinence (Hasin, Trautman, & Endicott, 1998; Quello, Brady, & Sonne, 2005), or at least when the person is not intoxicated or in withdrawal. While the length of this period is not well established, a stabilisation period of between two to four weeks is recommended (Strain, 2002). If symptoms persist after this period, they can be viewed as independent rather than substance-induced. Realistically, such a period of abstinence is a luxury rarely afforded in D&A treatment settings and, therefore, to avoid possible misdiagnosis, it has been recommended that multiple assessments be conducted over time. This process allows the D&A worker to formalise a diagnosis and develop a tailored case plan that is reviewed and modified, allowing for greater accuracy and flexibility in assessment and treatment.

Screening forms one of the first parts of the assessment process. Unlike screening, assessment is a process rather than a one-off event, which involves the ongoing monitoring of clients' mental health and substance use disorder symptoms along with their psychosocial functioning and wellbeing. Ongoing

assessment is important because clients' symptoms and functioning may change throughout treatment. For example, a person may present with symptoms of anxiety and/or depression upon treatment entry; however, these symptoms may subside with abstinence. Alternatively, a person may enter treatment with no mental health symptoms, but symptoms may develop after a period of reduced use or abstinence, particularly if the person has been using substances to self-medicate these symptoms.

Groth-Marnat (2003) suggests that a combination of both informal and standardised assessment techniques is the best way to develop a case formulation. In addition to these assessments, with the client's consent, it may be useful to talk with family members or carers; they can provide invaluable information regarding the client's condition which the client may not recognise or may not want to divulge (Australian Government Department of Health and Ageing, 2006).

An equally important component of a treatment program is regular review and assessment of the client's progress in relation to their treatment goals, that is, outcome assessment (Treatment Protocol Project (TPP), 2000). Again, both informal (non-standardised) and formal procedures are necessary. For instance, the measurement of quantity/frequency of substance use, mental health symptoms, and any other relevant outcomes as required by the initial assessment, and/or the service environment within which the individual worker is operating (APA, 2006; NSW Health, 2008).

For the purpose of this paper, only standardised tools for the screening, assessment and outcome process are to be considered, however, it is important for workers to be aware of informal procedures (e.g., the semi-structured interview, available from the NSW Department of Health (2007)), as these procedures may capture individual client factors specifically relevant to their situation.

Standardised tools can be a useful means of gathering data by providing an objective (reliable and valid) view of the client's difficulties and current life situation (Ries, 1995; Winters, 1999). Furthermore, when conducted appropriately the process of standardised assessment can be a source of rapport building.

Groth-Marnat (2003) suggests that when conducting standardised assessment, it is important to:

- Provide the client with the reasons for assessment and the purpose of each instrument.
- Explain that it is a standard procedure.
- Explain how standardised assessment can be useful in helping clients achieve their goals (e.g., by providing an objective measure).
- Provide appropriate and timely feedback of the results of the assessment.

Standardised assessment should be completed upon entry into and exit from treatment, as well as at follow-up (Mattick & Hall, 1993; Winters, 1999). Specific instruments are also often useful and recommended for periodic completion to monitor client condition (e.g., it is recommended that the BTOM be completed every 3 months). Test results can provide useful clinical information (for both the client and D&A worker) on the client's case and an evaluation of how effective treatment has been. A variety of different tools are used, some of which are empirically established instruments, whilst others

are purpose-built, internally designed tools with increased practicality and utility but unknown validity and reliability (Roche & Pollard, 2006). There are no definite rules guiding the frequency with which outcome measurement and monitoring should occur, and therefore is often a matter of judgement by the individual worker (Treatment Protocol Project (TPP), 2000). Several authors have suggested that most interventions are likely to show results within one month (if the client is going to respond at all), depending on the frequency of sessions (Kay-Lambkin, Baker, & Lewin, 2004; Scogin, 2003; Scogin, Hanson, & Welsh, 2003). However, in some service settings (specifically opioid maintenance pharmacotherapy (OMP)), 3-monthly assessment/review sessions are recommended (NSW Health, 2008).

Psychometric Properties

In assessing screening instruments and assessment tools in general, particular attention should be paid to the **sensitivity** and **specificity** of the instrument as well as its **reliability** and **validity**. The term **sensitivity** describes the proportion of individuals with the condition who test positive on that test, while the **specificity** of a test is the proportion of those without the condition who test negative. These have values between 0 and 1 and research in this area is devoted to ascertaining the best cut-off for a test or screen that produces the highest combination of these two variables.

A test needs to be valid in that it is measuring what it is intended to measure. There are a range of different forms of **validity**. **Content validity** refers to the instrument's comprehensiveness (i.e., how adequately the sampling of items reflects its aims). It is often assessed by feedback from consumers and clinicians. **Construct validity** involves conceptually defining the construct to be measured by the instrument, and assessing the internal structure of its components and the theoretical relationship of its item and subscale scores. **Convergent validity** is the degree to which an instrument is similar to (converges on) other measures that it theoretically should also be similar to, while **discriminant validity** describes the degree to which the instrument is not similar to (diverges from) other measures that it theoretically should not be similar to. **Criterion validity** assesses the extent to which the instrument correlates with other established measures (benchmark standard or "gold standard" measures). **Concurrent validity** measures an instrument's similarity to comparative measures, at the same point in time, while **predictive validity** assesses its ability to predict a future outcome.

A test also needs to be **reliable** in that it will give high positively-correlated results when re-administered by the same rater at two different points in time, given nothing else has changed (**test-retest reliability**). A good scale also has good **inter-rater reliability** which relates to the degree of agreement when the same instrument is applied to the same client by different raters at the same point in time. Clearly a scale or test cannot be valid if it is not reliable. A good scale will also be **internally consistent** in that all items make a significant contribution to the final score or rating. The level of reliability of an instrument is traditionally measured by a kappa value. Kappas of ≤ 0.20 are regarded as poor, 0.21-0.40 as fair, 0.41-0.60 as moderate, 0.61-0.80 as good, and ≥0.81 as very good.

Layout of Current Review

Many tools used in the case formulation process do not fit cleanly into one particular category (i.e., a depression measure) but rather contain items that measure a range of factors (e.g., social functioning and general psychological distress). Similarly, they do not necessarily function only as an –outcome tool or a screening tool but can be used as both. This is discussed on an individual basis for each measure. Therefore this review is broken down into several categories:

1. Global measures – tools that measure a range of client factors (e.g., substance use, psychological and physical health, social functioning).

2. General health and functioning measures – tools that rate an individual's functioning abilities and limitations.

3. General mental health measures – tools that measure a range of psychological symptoms (e.g. distress).

4. Specific mental health measures – tools that measure the symptoms of one disorder class only.

5. Positive mental health measures – an emerging area for outcome measurement in mental health has come from the philosophies of recovery, wellbeing, empowerment and rehabilitation.

6. General substance misuse measures – brief tools to ascertain the existence/nature of the substance problem

7. Severity of substance misuse measures – more specific tools to measure the severity of the substance use problem

8. Craving measures – this section provides an outline of some potentially useful drug craving measures

For each tool, information has been included on its psychometric properties (according to available research), its suitability for particular client groups, availability/cost and scoring administration and expertise required.

Summary Table

Instrument	Utility/Measures	Administration	Strengths	Limitations	Cost
GLOBAL INSTRUMENTS					
Addiction Severity Index (ASI)	Assessment and outcome measurement. 30 day & lifetime alcohol use, drug use, medical problems, psychiatric problems, family/social problems, employment, legal problems.	Interview or self-report	Widely used across a range of population groups.	Psychometric and interpretation concerns. Less extensive Australian use. Lengthy	No
Brief Treatment Outcome Measure (BTOM) / Australian Alcohol Treatment Outcome Measure (AATOM)	Outcome measurement. Bloodborne virus risk, drug use, social and psychological functioning, health.	Interview	Adequate reliability and validity. Australian. No training required. Previous use within D&A sector in NSW. Public domain.	Limited testing across populations	No
Global Appraisal of Individual Needs (GAIN)	Assessment and outcome measure. Background information, substance use, physical health, risk behaviors, mental health, environment, legal, and vocational information.	Interview (can also be self-administered)	DSM-IV diagnosis. Includes satisfaction index. Range of short versions. Comprehensive. Good psychometrics.	Copyrighted/cost. Lengthy. American terminology. Not widely used outside USA.	Yes
Health of the Nation Outcome Scale (HoNOS)	Assessment and outcome measurement. Severity of aggression, self harm, D&A use, memory/orientation, physical problems, mood disturbance, hallucination and delusions, other mental, social relationships/environment.	Interview	Generally adequate validity and reliability. Thoroughly evaluated and extensively used across a range of populations (incl. Indigenous Australians). Public domain.	Inter-rater reliability concerns. Training required.	No
Indigenous Risk Impact Screen (IRIS)	Screening. D&A problems and mental health risks.	Interview	Aboriginal and Torres Strait Islander specific. Adequate reliability and validity. Brief. Public domain.	Aboriginal and Torres Strait Islander specific. Limited empirical validation	No
Maudsley Addiction Profile (MAP)	Outcome measurement. Substance use, health risk behaviour, physical and psychological health, social functioning.	Self-report or interview	Adequate reliability and validity. Used widely across different cultural groups. Public domain	Limited validation in specific population groups and outside of Europe.	No
Opiate Treatment Index (OTI)	Assessment and outcome measurement. D&A use, risk taking, social functioning, criminality, health status, psychological adjustment.	Interview	Good reliability and validity. Australian. Public domain.	Training required. Only moderate validation in different populations. Lengthy. Predominantly a research instrument.	No

Instrument	Utility/Measures	Administration	Strengths	Limitations	Cost
GENERAL HEALTH & FUNCTIONING INSTRUMENTS					
Camberwell Assessment of Need (CAN)	Assessment and outcome measurement. Comprehensive needs in 22 areas.	Interview (or self-report)	Widely used. Validated in Australia. Specific versions for different client groups. No formal training required.	Descrepancies between client and clinician ratings of need.	Yes
Children's Global Assessment Scale (CGAS)	Assessment and outcome measurement. Functioning and psychiatric dysfunction for children and adolescents	Interview	Specifically children and adolescents but adult version also exists. Good validity. Use in variety of populations. Freely available.	Concerns over inter-rater reliability. Clinical expertise required.	No
Global Assessment of Functioning Scale (GAF)	Assessment and outcome measurement. Overall level of psychological, social and occupational client functioning.	Self-report	Used in both psychiatric and D&A populations. Freely available. Based on DSM criteria.	Training essential to interpretation. Some validity and reliability concerns.	No
Life Skills Profile (LSP)	Outcome measurement and assessment. Range of functioning areas (Self-care, anti-social, withdrawal, compliance).	Interview	Positive mental health philosophy. Generally good psychometrics. Validated in a range of populations (incl. D&A, psychiatric and Indigenous). Brief and easy to use.	Some inter-rater issues.	No
Short-Form 36 Health Survey (SF-36; SF-12)	Screening, assessment and outcome measurement. Physical/emotional role limitation and functioning, bodily pain, mental health, social functioning, vitality/ general health perceptions.	Self-report	Brief. Very good psychometrics. Widely validated in an Australian context and across a range of populations (incl. D&A users and mentally ill).	Licence fees apply.	Yes
Strengths and Difficulties Questionnaire (SDQ)	Screening, assessment and outcome measurement. Conduct problems, emotional symptoms, hyperactivity, peer relationships and prosocial behaviour.	Variety of collection methods	Moderate-good psychometrics. Widely used in psychiatric and D&A populations and cross culturally. Although copyrighted it is available without cost. Australian norms. Brief.	Limited to children/adolescents	Not for non-profit organi-sations
World Health Organisation Disability Assessment Schedule II (WHODAS II)	Assessment and outcome measurement Assess the activity limitations and participation restrictions experienced by an individual (understanding and communicating, getting around, self-care, getting along with people, life activities, participation in society).	Variety of collection methods	Brief. Public domain. Adequate psychometrics. Cross cultural and Australian validation.	Limited use in D&A populations.	No

Instrument	Utility/Measures	Administration	Strengths	Limitations	Cost
World Health Organisation Quality of Life-BREF (WHOQoL-BREF)	Outcome measurement and assessment. Physical health, psychological health, social relationships, and environment.	Self-report	Brief. Cross cultural and Australian validation. Used widely across a range of populations (incl. D&A users and mentally ill). Adequate psychometrics. Public domain.	Requires competent reading ability.	No
GENERAL MENTAL HEALTH INSTRUMENTS					
The Behaviour and Symptom Identification Scale 32 BASIS-32®	Screening and Outcome measure. Impulsive and addictive behaviour, Psychosis, Relation to self and others, Depression and anxiety and Daily living and role functioning	Variety of collection methods	Good psychometrics. Australian version available. Used in D&A samples	Copyrighted/cost. High reading level required for self-report version.	Yes
Brief Psychiatric Rating Scale (BPRS)	Screening. Thought disorder, withdrawal, anxiety/depression, hostility and activity.	Interview	Successfully used in D&A populations and across age groups. Adequate validity	Inter-rater reliability concerns. Training and qualifications required and intensive ongoing supervision.	No
Depression Anxiety Stress Scale (DASS)	Screening and outcome measurement. Depression, anxiety and stress (general psychological distress).	Self-report	Good reliability and validity. Fairly widely used across cultures and age groups.	Interpretation requires expertise.	No
General Health Questionnaire (GHQ)	Screening. Somatic symptoms, anxiety/insomnia, social dysfunction, severe depression.	Self-report	Good reliability and validity. Widely used across a range of populations (incl. D&A users).	Lower reliability in general population.	Yes but GHQ-28 is part of the OTI
Kessler psychological distress scale (K10)	Screening. Psychological distress.	Self-report (or interview if required)	Good reliability and validity. Widely used across a range of populations (incl. D&A users). Australian norms. Public domain. Brief.	Only screens for distress broadly.	No
Mental Health Inventory (MHI)	Screening and outcome measurement. General psychological distress and wellbeing.	Self-report or interview	Adequate psychometrics. Easy to use.	Only limited use across population groups.	No
PsyCheck	Screening. General mental health, suicidality/self-harm, depression and anxiety symptoms.	Mixed administration	Adequate reliability and validity. Australian. Designed for use in D&A treatment settings.	Few empirical studies.	No

Instrument	Utility/Measures	Administration	Strengths	Limitations	Cost
Symptom Checklist-90-Revised (SCL-90-R)	Screening and outcome measurement. Somatisation, obsessive compulsive, interpersonal sensitivity, depression, anxiety, hostility, phobic anxiety, paranoid ideation, psychoticism.	Self-report	Good reliability and validity. Used in D&A populations. Long and shorter versions available.	Costs involved and qualifications required.	Yes
SPECIFIC MENTAL HEALTH INSTRUMENTS					
Beck Inventories	Screening, assessment and outcome measurement. Symptoms of depression and anxiety, hopelessness and suicidal ideation.	Self-report or interview	Generally well researched, with good psychometrics. Brief.	Cost involved in use and only available to those with psychiatric or psychological qualifications.	Yes
Body Dysmorphic Disorder Questionnaire (BDDQ)	Screening. Disturbed body image	Self-report	Good sensitivity/specificity. Brief.	Lack of reliability/validition studies.	Yes
Eating Attitudes Test (EAT)	Screening and outcome measurement. Disturbed eating patterns.	Self-report	Brief. No special training required. Freely available online. Good psychometrics.	Does not assess more general dysfunctional attitudes and related psychopathology. Limited validation in men and D&A users	No
Impact of Event Scale (IES)	Screening and outcome measurement. Current degree of subjective stress (PTSD symptoms) experienced as a result of a specific event.	Self-report	Extensive testing. Good psychometrics. Freely available online. No special training required. Brief.	Requires competent reading ability.	No
Primary Care PTSD Screen (PC-PTSD)	Screening. PTSD symptoms.	Self-report	Very brief. Public domain. Used in D&A settings, adolescents and veterans.	Limited empirical testing.	No
Psychosis Screener (PS)	Screening. Psychotic symptoms.	Interview	Adequate reliability and validity. Used in homeless and D&A using populations.	Mixed success in prison populations. Sensitivity/specificity varies.	No
PTSD Checklist (PCL)	Screening. PTSD symptoms.	Self-report	Excellent psychometrics. Different versions available. Public domain. Used in a variety of populations (incl. D&A users).	Cut-off points and scoring methods may vary.	No

Instrument	Utility/Measures	Administration	Strengths	Limitations	Cost
PTSD Symptom Scale Self-Report (PSS-SR) / Posttraumatic Stress Diagnostic Scale (PDS)	Screening and outcome measurement. PTSD symptoms.	Self-report	Good psychometrics. Used in wide range of populations (incl. D&A users).	Copyrighted/cost.	Yes
Spielberger State Trait Anxiety Inventory (STAI)	Screening and outcome measurement. Transitory and enduring anxiety.	Self-report	Widely used (incl. D&A users). Moderate-good psychometrics.	Cost involved in use and qualifications required.	Yes
Traumatic Life Events Questionnaire (TLEQ)	Screening. Intense fear, helplessness and horror symptoms of trauma and frequency trauma.	Self-report or interview	Adequate reliability and validity. Used in adolescents, D&A user and prison populations.	Copyrighted/cost involved. Limited empirical testing.	Yes
Trauma Screening Questionnaire (TSQ)	Screening. Re-experiencing and arousal PTSD symptoms.	Self-report	Simple. Adequate psychometrics in preliminary studies. Freely available.	Limited empirical studies.	No
POSITIVE MENTAL HEALTH INSTRUMENTS					
Dispositional Hope Scale (DHS)	Outcome measurement. Hope.	Self-report	Used across cultures. No training required. Freely available. Brief.	Limited empirical evidence.	No
Recovery Assessment Scale (RAS)	Outcome measurement. Empowerment, coping ability, and quality of life.	Self-report	Simple and effective. Freely available online.	Not widely used.	No
Social & Emotional Wellbeing & Empowerment Tool	Outcome measurement Empowerment.	Self-report	Aboriginal and Torres Strait Islander specific.	Still being finalised	-
Scales of Psychological Well-being (SPWB)	Outcome measurement. Psychological well-being: autonomy, positive relationships, environmental mastery, personal growth, purpose in life, and self-acceptance.	Self-report	Used across cultures. No training required. Long and shorter versions are available.	Small cost in obtaining the tool. Moderate and mixed reliability and validity.	Minimal
Stages of Recovery Instrument (STORI)	Outcome measurement. Recovery as the concept is described by mental health consumers.	Self-report	Freely available. Brief. Good preliminary findings.	Limited empirical evidence.	No
GENERAL D&A INSTRUMENTS					
Alcohol, Smoking and Substance Involvement Screening Test (ASSIST)	Screening. D&A use and risk (lifetime/recent substance use, specific substance involvement, frequency, dependence, abuse, intravenous drug use).	Interview	Good psychometrics across a range of cultures. Brief and simple to administer. Includes brief intervention strategies. Public domain.	Limited empirical evidence for sub-populations.	No

Instrument	Utility/Measures	Administration	Strengths	Limitations	Cost
Alcohol Use Disorders Identification Test (AUDIT)	Screening and outcome measurement. Alcohol use: consumption, dependence, and related-problems.	Self-report or interview	Freely available. Brief. Good psychometrics across a vast range of populations (incl. mentally ill). No training required. Australian version.	Concerns about utility in females, Indigenous and older populations. Intended for general population.	No
CAGE/CAGEAID	Screening. Identify problem alcohol use.	Self-report or interview	Very brief. Moderate-good psychometrics. Used in a variety of populations (adapted for Indigenous Australians). Freely available.	Test-retest concerns. Concerns about utility in females and mentally ill populations.	No
Dartmouth Assessment of Lifestyle Instrument (DALI)	Screening. Substance use disorders use with people with severe mental illness.	Interview	Brief and simple. No special training required. Adequate psychometrics	Limited studies in different populations.	No
Drug Abuse Screening Test (DAST)	Screening and assessment. Identify problem drug use.	Self-report or interview	Brief. Freely available. Good psychometrics in range of populations (incl. mentally ill).	Concerns over applicability to women and across cultures. Does not discriminate between past and present use.	No
Michigan Alcoholism Screening Test (MAST)	Screening and assessment. Identify problem alcohol use.	Self-report or interview	Brief. Public domain. Good psychometrics across a range of populations (incl. mentally ill). No training required for use.	Does not discriminate between past and present drinking. Concerns over applicability to women and across cultures.	No
T-ACE/TWEAK	Screening. Specifically designed to identify at-risk drinking pregnant women (but has some utility in other groups).	Interview	Available online without cost. Very brief. Moderate psychometrics. No training required.	Does not provide a picture of pattern of use. Debate over suitable cut-off scores.	No
Timeline Followback Method (TLFB)	Assessment and outcome measurement. Information on the amount/duration of D&A use over a specified period of time.	Variety of collection methods	Widely used (incl. in mentally ill populations). Simple. Adequate psychometrics.	Training required.	Minimal
D&A SEVERITY INSTRUMENTS					
Alcohol Dependence Scale (ADS)	Assessment and outcome measurement. Identify and assess alcohol abuse and dependence.	Self-report	Adequate psychometrics. Brief. Fairly widely used in a variety of populations	Copyrighted/cost.	Yes

Instrument	Utility/Measures	Administration	Strengths	Limitations	Cost
Cannabis Problems Questionnaire (CPQ)	Screening. Problematic cannabis use.	Self-report	Australian developed. Brief. Good psychometrics. Public domain.	Limited empirical studies across different population groups.	No
Leeds Dependence Questionnaire (LDQ)	Assessment and outcome measurement. Severity of D&A dependence	Self-report	Can be used across D&A dependence. Good psychometrics. Brief. Freely available. Incorporates psychological dependence rather than just consumption and physical dependence.	Limited use in all types of substance dependence. Limited use across cultures.	No
Short Alcohol Dependence Data Questionnaire (SADD)	Assessment and outcome measurement. Less severe alcohol dependence, behavioural and subjective changes.	Self-report	Freely available. Adequate psychometrics. Brief.	Limited empirical studies. Lack of use across cultures.	No
Severity of Alcohol Dependence Questionnaire (SADQ)	Assessment and outcome measurement. Severity of dependence on alcohol, withdrawal symptoms etc.	Interview	Freely available. No special training required. Brief. Good psychometrics.	Not widely used across all groups (e.g. psychiatric). Concerns about use in older people and women.	No
Severity of Dependence Scale (SDS)	Screening and outcome measurement. Psychological dependence on a given substance.	Self-report	Very brief. Public domain. Good psychometrics. Widely validated in an Australian context and across drug types.	Very brief.	No
Substance Dependence Severity Scale (SDSS)	Assessment and outcome measurement. Severity of dependence on a variety of substances.	Interview	Generally adequate psychometrics. Range of substances. Based on DSM-IV validation.	Lengthy. Some concern about the cannabis and sedative subscales.	Minimal
CRAVING MEASURES					
Cocaine Craving Questionnaire (CCQ)	Screening. Cocaine craving	Self-report	Brief. Easy to use. Adequate psychometrics. Used in psychiatric populations.	Limited Australian empirical validation.	Minimal
Cocaine Craving Questionnaire (Weiss et al)	Screening. Cocaine craving	Self-report	Brief. Easy to use. Adequate psychometrics.	Limited empirical studies across different groups.	No
Marijuana Craving Questionnaire (MCQ)	Screening. Marijuana craving	Self-report	Brief. Easy to use. Adequate psychometrics.	Limited empirical studies across different groups.	Minimal
Penn Alcohol-Craving Scale (PACS)	Screening. Alcohol craving	Self-report	Brief. Easy to use. Adequate psychometrics. Used in different settings in an Australian context.	Limited empirical studies.	Minimal

Global Screening, Assessment and Outcome Measures

Opiate Treatment Index (OTI)

The Opiate Treatment Index (OTI) (Darke, Ward, Hall, Heather, & Wodak, 1991) is a structured interview primarily developed to allow comparability between research findings. It is an appropriate measure for users of all different drugs, including alcohol (Darke, Hall, Wodak, Heather, & Ward, 1992). However, there is some dispute the utility of the OTI for alcohol as it does not assess patterns of use (Copeland, 2009; personal communication). Mattick and Hall (1993) recommend the use of the OTI as a general initial assessment and outcome measure for clients using both licit and illicit drugs. The OTI measures six independent outcome domains: drug use, HIV risk-taking behaviour, social functioning, criminality, health status and psychological adjustment as measured by the GHQ-28. While it was not intended to be an alternative to a clinical assessment, it is a useful clinical tool providing information from which to evaluate a treatment program.

The OTI was originally validated in an Australian opiate treatment setting in 1992 and demonstrated high levels of test-retest reliability (.77-.99) on all scales regardless of whether the same or a different interviewer administered the test (Darke et al., 1992). These findings were supported by a later UK study (Adelekan, Green et al., 1996). It also demonstrated generally high levels of internal consistency ranging from Cronbach's alpha = .38-.83 in the Australian study and Cronbach's alpha = .34-.93 in the UK study (Adelekan, Green et al., 1996; Adelekan, Metrebian, Tallack, Stimson, & Shanahan, 1996; Darke et al., 1992).

The OTI was found to have two main factors – 'drug using lifestyle' and 'health and well being' – which accounted for 60% and 55.4% of the total variance in the Australian and London studies respectively (Adelekan, Green et al., 1996; Darke et al., 1992). The instrument was also shown to have good convergent validity with the relevant ASI subscales (Darke et al., 1992). Scores on individual scales of the OTI were also compared with other relevant measures and found to significantly correlate with scores on the Health Status scale with independent medical examinations, between drug use scores and urinalysis results and between reported recent behaviours and partner collateral reports (Darke et al., 1992). These findings were again replicated in the UK study, providing support for the convergent validity and cross-cultural validity of the measure (Adelekan, Green et al., 1996).

Despite some concern over standardisation across interviewers, studies have found good levels of inter-rater reliability, with few or no significant differences between information given to clinicians compared with research assistants (Adelekan, Metrebian et al., 1996; Deering & Sellman, 1996). The OTI has also been found to be sensitive to changes in D&A behaviour over time (Baker, Boggs, & Lewin, 2001; Baker et al., 2002a; Padaiga, Subata, & Vanagas, 2007; Shearer et al., 2001; Verthein et al., 2008).

Client groups

Although initially developed for opioid use (Darke et al., 1992) the OTI is effective when used across licit (including alcohol; although concern has been raised regards its applicability to alcohol) and illicit drugs (Barrowcliff, Champney-Smith, & McBride, 1999; Mattick & Hall, 1993).It has been translated into a range of languages (e.g., Liu et al., 2000; Ruz, Gonzalez, & Ruiz, 1998; Swift, Maher, & Sunjic, 1999) and has been used effectively with people with serious mental illness and substance use problems in inner city Sydney (Teesson & Gallagher, 1999) and more widely across Australia (Baker et al., 2002b).

The OTI is also an effective instrument when used in adolescent populations (Mills, Teesson, Darke, Ross, & Lynskey, 2004; Spooner, Mattick, & Noffs, 2001)

Availability/cost

The OTI is in the public domain and may be used without cost but with due acknowledgment of the source. It can be found at:

- http://ndarc.med.unsw.edu.au/NDARCWeb.nsf/resources/TR_29/$file/TR.011.pdf

Scoring, administration and expertise required

The manual clearly describes the administration procedure and scoring system for each of the scales, although scoring does require the calculation of formulas. The OTI can be administered without any specific training and takes approximately 20-40 minutes to complete.

Brief Treatment Outcome Measure (BTOM); Australian Alcohol Treatment Outcome Measure (AATOM)

The Brief Treatment Outcome Measure (BTOM) (Lawrinson, Copeland, & Indig, 2003) is a tool developed specifically to routinely assess outcomes of treatment for clients receiving OMP services and for use in treatment evaluation research. Treatment outcome is measured by scales developed or adapted from other instruments across the domains of drug dependence (measured by SDS), blood borne virus exposure risk, drug use, health, psychological functioning and social functioning (Lawrinson et al., 2003). It was adopted for routine D&A treatment outcome monitoring in NSW. It includes the National Minimum Data Set data items nested within it so does not duplicate existing data collection requirements.

The evaluation of the BTOM took place in 37 metropolitan, rural and prison OMP services in NSW, among 160 OMP clients. The internal reliability of the BTOM was found to be satisfactory. Test-retest reliabilities for the scales and drug use scores are good to excellent indicating their consistency with multiple measurements across time and different interviewers and concurrent validation of BTOM scales with analogous scales from similar instruments (i.e., the OTI) yielded acceptable agreement (Lawrinson, Copeland, & Indig, 2005, 2006).

Preliminary reviews were conducted on the predictors of outcome monitoring and existing outcome measures (Copeland, Rush, Reid, Clement, & Conroy, 2000; Teesson, Clement, Copeland, Conroy, & Reid, 2000). Similarly, an advisory group and on-going appraisal by clinicians and program administrators informed the development of the BTOM content. In this way, the BTOM has good content validity.

A shorter version of the BTOM, the Brief Treatment Outcome Measure – Concise (BTOM-C) is also available. In this version the SDS, the Psychological Functioning Scale, days client spent in hospital and psychiatric medication questions have been dropped (NSW Department of Health, 2004).

The recently designed Australian Alcohol Treatment Outcome Measure (AATOM) was developed as an equivalent alcohol tool. Two versions of the AATOM exist, one intended for use amongst clinicians for the purpose of routine treatment outcome monitoring (AATOM-C) and one for use amongst researchers (AATOM-R). In the initial sample of clients from D&A treatment agencies, the inter-rater and test-retest reliability of the AATOM-C was good to excellent for most scales. Internal consistency was considered acceptable (Cronbach's alpha ≥ .73), as was concurrent validity (correlation coefficients between .55-.93). Content validity was achieved in the same way as the BTOM. The results of the study demonstrate that the AATOM-C is, overall, a valid and reliable instrument, taking on average 10-15 minutes to administer (Simpson, Lawrinson, Copeland, & Gates, 2007). In line with its purpose, the instrument demonstrated the ability to measure change in client functioning over time (Simpson, Lawrinson, Copeland, & Gates, 2009).

Client groups

Both the BTOM and the AATOM were designed to be administered to all new treatment clients and are suitable for all clients who can understand spoken English. Simpson and colleagues (2007) emphasise the importance of comprehensive and consistent training of those who are to administer the AATOM-C, as it was found one interviewer repeatedly produced poorer test-retest reliability. This is fundamental to each of these measures.

Availability/cost

The BTOM is in the public domain and is freely available from:

- http://www.health.nsw.gov.au/pubs/2004/pdf/btomclinicansguide.pdf
- http://ndarc.med.unsw.edu.au/NDARCWeb.nsf/resources/TR_11/$file/TR.156.PDF

The AATOM is published in Simpson and colleagues (2007), NDARC technical report 288. It is available for a minimal fee at:

- http://ndarc.med.unsw.edu.au/NDARCWeb.nsf/page/NDARC%20Technical%20Reports

Scoring, administration and expertise required

Both the BTOM and the AATOM are brief, easy to administer and can be easily scored, requiring no special training. Each take approximately 10-20 minutes to complete and it is recommended that they be completed once every three months. Concise versions are shorter and therefore take less time to complete.

Health of the Nation Outcome Scale (HoNOS)

Health of the Nation Outcome Scale (HoNOS) (Wing et al., 1998) is a measurement tool designed to assess general health and social functioning of mentally ill people. It is a 12-item questionnaire that measures the severity of aggression, self harm, alcohol and drug use, memory/orientation, physical problems, mood disturbance, hallucination and delusions, other mental, social relationships, social environment. It has been found to be a useful measure of treatment effectiveness/client change over time if conducted at set intervals (Gallagher & Teesson, 2000b; Sharma, Wilkinson, & Fear, 1999; Teesson et al., 2000). However, some concerns have been raised by other authors who have recommended that it not be implemented as a major outcome tool, due to concerns regarding its reliability and validity (Brooks, 2000).

Generally studies looking at the inter-rater reliability of the HoNOS have found that the overall agreement between raters is only moderate at best (Amin et al., 1999; Bebbington, Brugha, Hill, Marsden, & Window, 1999; Brooks, 2000; Hope, Trauer, & Keks, 1998; Orrell, Yard, Handysides, & Schapira, 1999; Shergill, Shankar, Seneviratna, & Orrell, 1999; Wing et al., 1998) and that agreement is even poorer on particular items. Factors associated with improved inter-rater reliability include training, familiarity with consumers and setting (Audin, Margison, Clark, & Barkham, 2001; Brooks, 2000). Relatively few studies have examined the test-retest reliability of the HoNOS, but those that have, generally report fair to moderate overall reliability scores (r = .4-.8) (Brooks, 2000; Orrell et al., 1999; Shergill et al., 1999; Teesson et al., 2000).

Several studies have suggested that the HoNOS has moderately high level of internal consistency (Cronbach's alpha ranging from .59 to .76) and low levels of item redundancy, supporting its use as a meaningful summary of severity of symptoms (McClelland, Trimble, Fox, Stevenson, & Bell, 2000; Orrell et al., 1999; Page, Hooke, & Rutherford, 2001; Shergill et al., 1999; Trauer, 1999; Wing et al., 1998). It has been suggested, however, that the HoNOS should not be regarded as unidimensional, measuring a single, underlying construct of mental health status, but rather a tool to assess a broad range of problems typically experienced by consumers of mental health services and some factorial studies have supported this assertion (McClelland et al., 2000; Preston, 2000a; Trauer, 1999).

The content validity of the HoNOS has generally been found to be good (Teesson et al., 2000). A number of studies have reported that consumer/carer advocacy groups and mental health professionals felt the HoNOS was appropriate, well-designed and thorough, and highlights consumers' problems quickly, indicating changes in their mental health status over time. However, not all items were regarded equally. A number of items and terminology was seen to be problematic and this raised concerns about misinterpretation of the scoring (McClelland et al., 2000; Orrell et al., 1999; Shergill et al., 1999).

Preliminary studies found the scores on the HoNOS correlated well with clinician-rated instruments such as the Social Behaviour Scale (SBS), Schedules for Clinical Assessment in Neuropsychiatry (SCAN) (Amin

et al., 1999; Bebbington et al., 1999). Similarly, it has been found to perform well against the Role Functioning Scale (RFS), the BPRS, the GAS (Amin et al., 1999; Browne, Doran, & McGauran, 2000; McClelland et al., 2000; Orrell et al., 1999; Shergill et al., 1999; Wing et al., 1998), the LSP (Parker, O'Donnell, Hadzi-Pavlovic, & Proberts, 2002), the Manchester Audit Tool (MAT) (Rees, Richards, & Shapiro, 2004), the Disability Assessment Schedule (DAS) (Amin et al., 1999), Location of Community Support Scale (LOCSS) (Amin et al., 1999; Orrell et al., 1999). Shergill and colleagues (1999) found the HoNOS to correlate well with a range a measures including the Clifton Assessment Procedures for the Elderly (CAPE-BRS), the Clinical Dementia Rating, the Mini-Mental State Examination (MMSE) and the Scale for Assessment of Negative Symptoms (SANS).

By contrast, the HoNOS has shown poor or mixed performance against consumer-rated instruments such as the Social Adjustment Scale (SAS), SCL-90-R, the SF-36 (Hope et al., 1998; Parker et al., 2002), the Quality of Life assessment (QoL) (Hunter et al., 2004), the Avon Mental Health Measure (AVON), the Outcome of Problems of Users of Services (OPUS) (Audin et al., 2001; McClelland et al., 2000) and even a self-rating version of the HoNOS with a similar question structure (Rees et al., 2004). Brooks (2000), for instance, in a series of studies found the HoNOS to have poor validity in relation to the SF-36 and the SCL-90-R. Other studies have reported correlations between the HoNOS and consumer-rated measures – e.g., the SF-36, the GHQ, the SCL-90-R, SF-36 and the Comprehensive Quality of Life scale (ComQoL) (McClelland et al., 2000; Orrell et al., 1999). However, they tend to vary across domains and be lower than those between the HoNOS and clinician-rated measures. This kind of difference in scores between measures which use self-report compared with informant-rated is not uncommon.

The HoNOS has also been shown to discriminate between individuals with and without mental health diagnoses and symptoms (e.g., high scores on relative scales to be associated with drug and alcohol, psychotic and bipolar disorders, aggressive behaviour, anxiety, depression) (Bech et al., 2003; Browne et al., 2000; McClelland et al., 2000; Rees et al., 2004). However, a recent Italian study was less positive, suggesting the instrument lacks sufficient discriminatory ability in a sample of patients with psychotic disorders (Gigantesco, Picardi, de Girolamo, & Morosini, 2007). High scores on the HoNOS have also been found to predict a significant proportion of variance in treatment outcome and resource use (Ashaye, Seneviratna, Shergill, & Orrell, 1999; Broadbent, 2001; Schneider, Wooff, Carpenter, Brandon, & McNiven, 2002) despite these findings, however, other authors have found little association between the HoNOS and resource use (Boot, Hall, & Andrews, 1997; Goldney, Fisher, & Walmsley, 1998).

The ability of the HoNOS to detect genuine improvement, deterioration or stability in symptoms has been found to generally be greater in inpatient settings compared with community settings (Audin et al., 2001; Goldney et al., 1998; Teesson et al., 2000). Although others have suggested this relationship may be more complex as aspects like severity and diagnosis may be pertinent (Andrews, 2003a). HoNOS scores have also been found to correlate with changes in consumers' and clinical judgements about whether they had improved, remained stable or deteriorated (Gallagher & Teesson, 2000b; Hunter et al., 2004; Taylor & Wilkinson, 1997). Scores on the HoNOS have also been found to correlate with

existing established measures of outcome, the GAS and the BPRS (Ashaye et al., 1999; Hope et al., 1998; Page et al., 2001). Finally, Bech and colleagues (2003) found that HoNOS scores for consumers who received evidence-based therapies (e.g., lithium and/or electro-convulsive therapy) showed greater improvement on the HoNOS than consumers who did not, most notably on the Behaviour and Symptoms subscales.

Overall, psychometric reports are mixed regarding the HoNOS, with some recent concerns raised about the reliability, specificity and the overall clinical utility of the tool as an outcome measure (Audin et al., 2001; Bebbington et al., 1999; Preston, 2000a; Sharma et al., 1999; Trauer et al., 1999). Nevertheless, others believe it is a promising contender for routine use (Gallagher & Teesson, 2000b; McClelland et al., 2000).

Client groups

Since its original development, the HoNOS has been translated into a number of languages (Bech et al., 2003; Bonsack, Borgeat, & Lesage, 2002; Lauzon et al., 2001; Lora et al., 2001; Morosini, Gigantesco, Mazzarda, & Gibaldi, 2003). Different versions of the HoNOS have been established for different age groups (HoNOS for adults, HoNOSCA for children & adolescents (15 items), and HoNOS65+ for older people). Each has been widely used in Australia.

The construct validity for the HoNOSCA was found to be adequate but that subscale/section scores should be treated with caution (Gowers, Bailey-Rogers, Shore, & Levine, 2000; Gowers et al., 1999). Gowers and colleagues (2000) also found the measure to have adequate concurrent validity and reasonably good inter-rater reliability (although some items performed poorly). These findings were supported in subsequent studies (Bilenberg, 2003; Brann, Coleman, & Luk, 2001; Yates, Garralda, & Higginson, 1999). Garralda and colleagues (2000) reported good correlations between scores overtime (test-retest reliability; between 0.69-0.80 over various time spans). A number of studies have reported significant associations between change (or lack of change) recorded on the HoNOSCA and clinician's global judgements of outcome (Bilenberg, 2003; Brann et al., 2001; Garralda et al., 2000; Gowers et al., 2000; Gowers et al., 1999)

A number of authors have found the concurrent validity of the older person's version of the HoNOS (HoNOS65+) to be adequate (Bagley et al., 2000; Burns et al., 1999; Mozley et al., 1999; Spear, Chawla, O'Reilly, & Rock, 2002)and the limited research available on the construct and content validity of the tool suggest this version has adequate validity in this sense (Burns et al., 1999). The reliability studies that exist using the HoNOS65+ offer mixed findings with Burns and colleagues(1999) and Spear and colleagues (2002) both concluding the inter-rater reliability to be good to very good for most items, while others have shown less positive results (Allen et al., 1999). Spear and colleagues (2002) also found the HoNOS65+ to be moderately sensitive to change as measured by assessment and discharge from inpatient and community services and scores on the Clinician's Interview Based Impression of Change Scale (CIBIC+).

In a recent unpublished report the HoNOS was found to show similar levels of internal consistency of the subscales when applied with Indigenous consumers as that observed in non-Indigenous populations (Haswell-Elkins, 2006). As well as good correlations with scores and indicators of wellness and illness identified by Indigenous consumers and carers in in-depth interviews, indicating content validity of the measure. The author recommends the continued use of HoNOS with Indigenous consumers, together with a set of guiding principles, as a tool that has the capacity to capture important information for the consumer, carer and clinician, but that the development and validation of consumer-rated tools that capture the more fundamental and culturally determined aspects of Indigenous mental health needs to occur (Haswell-Elkins, 2006).

Availability/cost

The HoNOS is in the public domain, and available free of charge, however, training costs may apply. It is available at:

- http://www.crufad.com/phc/honos.htm
- http://www.ergpa.com.au/cms/resources/HoNOS_Score.pdf

Scoring, administration and expertise required

The HoNOS is a clinician-rated tool whereby the clinician rates the consumer on each of the items over a recent period of time (previous 2 weeks and symptoms are scored on a scale of one to five. The tool takes approximately 15-30 minutes to complete and must be administered by a trained clinician. HoNOS is recommended for use by qualified mental health care practitioners, called Raters. However, any experienced mental health worker who has been trained in the use of HoNOS and achieves similar scores to other qualified health practitioners can use the HoNOS. One day training is recommended initially, with a half day retraining every two years. Various training packages and resources have been developed (Morris-Yates, Barber, Harris, & Zapart, 1999; Wing, Lelliott, & Beevor, 2000).

It has been suggested that although no instrument will fulfil all needs, the HoNOS is a comprehensive, easy to use tool that is likely to be effective in routine outcome measurement (Ashaye, Mathew, & Dhadphale, 1997; Gallagher & Teesson, 2000b; McClelland et al., 2000; Miles et al., 2003). Others claim that while it is useful as a routinely administered outcome measure, its major use is in research (Gilbody, House, & Sheldon, 2002; Stafrace, 2002; Stein, 1999).

In the United Kingdom, several studies have reported that clinicians were relatively positive about the HoNOS, viewing it as potentially useful, but insisting that its ongoing use would depend on appropriate resourcing, adequate infrastructure, regular feedback and ongoing training (Broadbent, 2001; James & Kehoe, 1999; Milne, Reichelt, & Wood, 2001). In field trials conducted at five sites in Victoria, Trauer (1998) found that clinicians at one site were extremely positive about the HoNOS, whereas those at the

other four were more ambivalent, believing that it contributed only minimally to their treatment practices.

Overall it appears likely that the HoNOS alone would not be expected to guide day-to-day clinical practice, but it may complement other pieces of evidence that normally form the basis for clinical judgements (Wing et al., 2000).

The Maudsley Addiction Profile (MAP)

The Maudsley Addiction Profile (MAP) (Marsden et al., 1998) was developed as a brief, multi-dimensional instrument for assessing treatment outcome in people with drug and/or alcohol problems. It is an easily administered at intake, during and after an index treatment episode. It covers four domains of substance use, health risk behaviour, physical and psychological health and personal social functioning. The MAP, however, does not include the collection of demographic data so it needs to be linked to other client data collection systems.

In the initial sample of 60 drug users and 80 alcohol users, internal reliability and concurrent validity assessments of the scales (compared with the Life Stressors and Social Resources Inventory, the ASI, and self-reported behaviours) and items were found to be in the range of moderate to good. Test-retest reliability was found to be good after three days. However, this is an inadequate timeframe for measuring such reliability (Copeland, 2009; personal communication). MAP subscores have also been found to correlate well with the relevant ASSIST scores (Marsden et al., 1998; Sharma et al., 1999).

More recently, a new tool, based partially on the MAP, has been devised and is being rolled out in some substance misuse treatment centres in the UK. This instrument, the Treatment Outcomes Profile (TOP), was developed as a brief alternative to the MAP and has a total of 20 items across four sections (offending, health and social functioning, substance use and health risk behaviour). The preliminary report on the TOP has found the measure to have good reliability and validity and to show adequate sensitivity to client change (Marsden et al., 2008).

Client groups

The MAP has been translated and validated in a number of different European countries (Bacskai, Rozsa, & Gerevich, 2005; Hernández & Gómez, 2004; József, Erika, & Sándo, 2004; Mandersen et al., 2001; Marsden, Nizzoli et al., 2000). In a sample of 124 subjects in Italy, Spain and Portugal, the internal and test-retest reliabilities of the MAP were satisfactory (Marsden, Nizzoli et al., 2000).

The tool has also been used successfully in psychiatric/comorbid populations (Marsden, Gossop, Stewart, Rolfe, & Farrell, 2000; Miles et al., 2003). One study has used the MAP on a population of adolescents aged between 14 and 16 years (Best, Manning, Gossop, Gross, & Strang, 2006).

However, further studies are required to evaluate the instrument for other population groups and larger samples.

Availability/cost

The MAP is in the public domain and therefore can be used without cost but with due acknowledgement of its source. The tool itself, along with scoring information can be found at:

- http://www.iop.kcl.ac.uk/iopweb/blob/downloads/locator/l_346_MAP.pdf
- http://eib.emcdda.europa.eu/html.cfm/index3667EN.html

The TOP is available at:

- http://www.nta.nhs.uk/areas/outcomes_monitoring/rollout_products.aspx

Scoring, administration and expertise required

The MAP consists of 60 questions and preliminary studies found the interviewer version to take approximately 12-25 minutes to complete. Scoring procedures are specified in the User's Manual and scoring time is approximately 5 minutes.

Recently a self-report version of the MAP has been developed, assessing and monitoring the functioning of opioid-dependent patients. Correlation coefficients between interview and self-completion version for alcohol, drug, psychiatric, family and legal problems were in excess of 0.7 for the majority of the 20 items that were compared. The authors concluded that this version was a practical alternative to the interview (Luty, Perry, Umoh, & Gormer, 2006).

The Addiction Severity Index (ASI)

The Addiction Severity Index (ASI) (McLellan, Luborsky, Woody, & O'Brien, 1980; McLellan et al., 1992) is one of the most commonly used standardized assessment (rather than basic screening) instruments in the field of substance use disorders. The ASI is a 155-item multidimensional structured interview for assessing alcohol and drug dependence. It assesses frequency of drug and alcohol use as well as other psychosocial areas affected by substance use (e.g., a psychiatric subscale is included). The ASI consists of 7 sub-scales assessing past 30 day and lifetime alcohol use, drug use, medical problems, psychiatric problems, family/social problems, employment and legal problems. The scoring of the ASI takes into account both subjective ratings of problems by clients, and objective tests of use (such as laboratory tests) across each scale to provide an overall severity rating. Various changes have been made since the tool's development and it is currently in its fifth-revised edition. A modified short version the ASI has also been developed and has shown promising results (Cacciola, Alterman, McLellan, Lin, & Lynch, 2007).

The ASI was initially developed as an outcome measure to evaluate treatment across a 6-program treatment network and in follow-up studies to assess treatment-related change (McLellan et al., 1980). The ASI assesses *frequency* of use, without addressing *quantity* of use, a marked difference to other instruments. This is because, firstly, quantity correlates with frequency, and furthermore, frequency is easier to recall than quantity, secondly, because drug use has a lack of standardisation (i.e., there is no "standard drink" equivalent) and similarly there is likely to be a disparity between what respondents believe they have consumed, and what they actually consumed (McLellan et al., 1992). It has been widely used as a tool to measure outcome of treatment (Ahmadi, Kampman, & Dackis, 2006; Craig & Olson, 2004; Ghitza, Epstein, & Preston, 2008).

While there have been mixed findings across populations, the ASI has been generally found to have good test-retest and inter-rater reliability as well as good content, construct and criterion validity for a range of substance abusing populations in dozens of studies (e.g., Alterman, Brown, Zaballero, & McKay, 1994; Alterman et al., 1998; Appleby, Dyson, Altman, & Luchins, 1997; Argeriou, McCarty, Mulvey, & Daley, 1994; De Jong, Willems, Schippers, & Hendricks, 1995; Drake, McHugo, & Biesanz, 1995; Hendricks, Kaplan, Van Limbeek, & Geerlings, 1989; Hodgins & el-Guebaly, 1992; Joyner, Wright, & Devine, 1996; McCusker, Bigelow, Servignon, & Zorn, 1994; McLellan et al., 1985; Rogalski, 1987; Wertz, Cleaveland, & Stephens, 1995; Zanis, McLellan, Cnaan, & Randall, 1994). Drake and colleagues (1995) found scale scores to have test-retest reliability coefficients of ≥.60 in a sample of homeless people, with higher coefficients for younger, female, less mentally ill respondents. Similarly, Stoffelmayr, Mavis and Kasim (1994) found the ASI to have high inter-rater reliability, as measured by intra-class correlation coefficients for composite scores (ranging from .83 to 1.00 across domains), but considerably lower reliability for severity scores (ranging from .40 to .87 across domains), which remained stable over a two year period. Internal consistency has been shown to vary across scales, ranging from .65 for the

employment problems scale to .89 for medical problems (Leonhard, Mulvey, Gastfriend, & Shwartz, 2000).

Calsyn and colleagues (2004) found the ASI composite scores provided effective initial screening for patients with impaired functional status as measured by the corresponding SF-36 component summary scores, indicating both good concurrent and discriminant validity.

Svikis and colleagues (1996) found the measure to be highly sensitive (.96) and specific (.94) to DSM-III-R diagnoses of alcohol abuse or dependence, among a sample of drug-abusing women when using a definition of heavy drinking as ≥3 drinks per occasion on ≥3 days/week over at least the previous 12 months on the ASI as the predictor. In a more recent study, the ASI was found to have 85% sensitivity and 80% specificity (Rikoon, Cacciola, Carise, Alterman, & McLellan, 2006). Cut-off scores for men ranged from 0.17-0.19 for alcohol dependence and 0.11-0.16 for drug dependence, and 0.13 and 0.12-0.18 respectively for women.

However, a recent review of thirty-seven studies, cast significant doubt on the psychometric performance of the ASI. Claiming inter-rater and test-retest reliabilities of the severity ratings and composites scores vary from excellent to unsatisfactory and high internal consistencies were reported regularly for only three of the seven composite scores (medical status, alcohol use, psychiatric status). The remaining four composite scores (employment status, drug use, legal status, family/social relations) were found to have low consistencies in at least four different studies. Coefficients of criterion validity were also found to be consistently low (Mäkelä, 2004). However, the omission of a number of studies may have lead to drastic – and somewhat biased – conclusions (McLellan, Cacciola, & Alterman, 2004). Nonetheless, other authors have raised similar concerns with both the psychometric properties of the measure and its interpretation (Melberg, 2004).

Client groups

The items of the ASI appear to assess relevant aspects of alcohol and drug use and associated problems, indicating good content validity across populations, although several researchers have noted that some items may be less relevant to some populations (e.g., Carey, Cocco, & Correia, 1997; Corse, Hirschinger, & Zanis, 1995; Wertz et al., 1995; Zanis et al., 1994; Zanis, McLellan, & Corse, 1997). Nevertheless, the tool has been validated and is frequently used across a variety of substance abusing populations, including psychiatric patients, homeless people, pregnant women and incarcerated prisoners, and has been used to assess treatment outcome across a range of substances, including opiates, cocaine and alcohol (Joyner et al., 1996; McLellan et al., 1992).

The ASI has been translated into a range of languages and evaluated across a range of countries (Gerevich, Bácskai, Kó, & Rózsa, 2005; Krenz et al., 2004; Schmidt et al., 2007). However, there have been no validation studies in Australia, nor have there been any reports of its use with Indigenous Australians. Nevertheless, Teesson and colleagues (2000) claim that the ASI does meet requirements for

use in routine assessment and outcome measurement in Australia. It should be noted, however, that a number of items on the ASI use American terminology and such items would need to be altered to take into account Australian-specific terminology and culture.

Although the ASI has been shown to be a reliable and valid measure in female drug abusing populations, some authors have suggested it fails to assess some important female-specific aspects of substance dependence (e.g., pregnancy-related issues, care-giving, violence) (Comfort, Zanis, Whiteley, Kelly Tyler, & Kaltenbach, 1999). As a result, the Psychosocial History (PSH; female-specific 2-hour, 300-item version) was developed by Comfort and Kaltenbach (1996). Preliminary validation research suggests the PSH has adequate validity and reliability in adult women and appears to outperform the ASI, particularly in those specifically targeted areas (Comfort & Kaltenbach, 1996). The measure, however, is considerably lengthier and more time consuming than the ASI.

Similarly, an adolescent version (the Teen-ASI) of the ASI is available, as the standard version neglects adolescent-related issues. Preliminary studies show it to be a promising measure of adolescent drug abuse within teenagers (Kaminer, 2008; Kaminer, Bukstein, & Tarter, 1991; Kaminer, Wagner, Plummer, & Seifer, 1993).

In psychiatric populations the general conclusion drawn from most individual studies and research summaries, is that many of the sub-scales perform poorly with people who have severe mental illness (Carey et al., 1997). Studies validating the ASI with drug dependent clients with severe and persistent psychiatric disorders have consistently found weak reliability and validity (Carey et al., 1997; Corse et al., 1995; Zanis et al., 1994; Zanis et al., 1997). Accordingly, the use of the ASI in this population is not recommended.

Availability/cost

The ASI interview is in the public domain and can be used without cost but with due acknowledgement of the authors. A copy of the ASI can be found at:

- http://www.tresearch.org/resources/instruments.htm
- http://www.densonline.org/DENSASI.pdf
- http://faculty.ugf.edu/jgretch/syllabi/adc_asi_form.pdf

The computerised version is available from:

- Biomedical Computer Research Institute, 9743 Redd Rambler Place, Philadelphia, PA 19115, USA, Tel: 1-215-676-9743.

Copies of the self-administered questionnaire are available from:

- Craig S. Rosen, Center for Health Care Evaluation, VA Palo Alto Health Care System (152-MPD), 795 Willow Road, Menlo Park, CA, 94025, USA. Tel: 1-650-493-5000 ext 22846, fax: 1-650-617-2736, or at crosen@stanford.edu.

The Teen-ASI is not under copyright but may not be reproduced without permission from the author:

- Yifrah Kaminer, Director, Bradley Substance Abuse Intervention Center (BASIC), Bradley Hospital, 1011 Veterans Memorial Parkway, East Providence, RI, 02915.

Scoring, administration and expertise required

The ASI is an interviewer-administered assessment tool. Administration of the interview form is somewhat lengthy and takes between 30-60 minutes. Administration of the ASI does not required any specialist educational pre-requisites and may be administered by physicians, drug treatment personnel, research technicians and other interested persons who have been trained in its use. Scoring takes approximately 10-20 minutes and instructions for the original interviewer-administered format are included in the manual.

The ASI's authors warn against its use as a self-administered instrument due to the high level of literacy required. However, a recently developed self-administered pen and paper version was found to have good internal consistency (Cronbach's alpha >.70) for the alcohol, drug, psychiatric and medical problems scales, with somewhat lower consistency in the remaining scales in a predominately male inpatient sample (Rosen, Henson, Finney, & Moos, 2000). Composite scores on the alcohol and drug use scales were similar across the two formats (r = .87 and .73, respectively). However, further validation of this questionnaire is required in other populations.

A computer program was also developed to assist with administration and scoring (McLellan et al., 1992) and recently a computerised multimedia CD-ROM has been developed (ASI-MV) and has shown promising psychometric characteristics, leading the authors to suggest that this form may be a viable alternative to potentially unreliable interviewer format (Butler et al., 2001). Furthermore, an internet and automated phone administration method has also been developed (Brodey et al., 2004).

Global Appraisal of Individual Needs (GAIN)

The Global Appraisal of Individual Needs (GAIN) (Dennis, White, Titus, & Unsicker, 2006) was developed as a comprehensive biopsychosocial assessment tool in 1993. It is a progressive and integrated series of measures and computer applications designed to support a number of treatment practices, including initial screenings; brief interventions; referrals; standardised clinical assessments for diagnosis and treatment planning; monitoring of changes in clinical status, service utilisation, and costs to society; and specific needs assessment and evaluation (Dennis, White et al., 2006). A number of different versions of the GAIN exist and are outlined below.

The GAIN includes over 100 scales and indices. It can be summarised into eight broad sections background information, substance use, physical health, risk behaviors, mental health, environment, legal, and vocational information (Dennis, Chan, & Funk, 2006; Titus, Dennis, Lennox, & Scott, 2008). The GAIN's main scales have good internal consistency (alpha over 0.90 on main scales, 0.70 on subscales) and test-retest reliability (kappa≥ 0.60 on categorical measures). The scales are also highly correlated with measures of use from timeline follow-back measures, urine tests, collateral reports, treatment records, and blind psychiatric diagnosis (kappa≥ 0.60) (Dennis, Godley et al., 2004; Dennis et al., 2002; Dennis, Funk, Godley, Godley, & Waldron, 2004; Dennis, Scott, & Funk, 2003; Dennis, White et al., 2006; Lennox, Dennis, Ives, & White, 2006; Lennox, Dennis, Scott, & Funk, 2006).

Confirmatory factor analyses suggest that the GAIN's collection of psychiatric and behavioural problems items vary largely along four dimensions: (1) substance use problems (e.g., abuse, dependence, induced), (2) internalising problems (e.g., depression, anxiety, trauma, suicide), (3) externalising problems (e.g., attention deficit, hyperactivity/impulsivity, conduct disorders), and (4) crime and violence (e.g., interpersonal, verbal, and physical violence, property crime, drug related crime, violent crime) (Dennis, Chan et al., 2006)

Using discriminant analysis, the GAIN scales could also reliably predict independent and blind staff psychiatric diagnoses of co-occurring psychiatric disorders including attention deficit hyperactivity disorder (kappa = 1.00), mood disorders (kappa = .85), conduct disorder/oppositional defiant disorder (kappa = .82), adjustment disorder (kappa = .69), or the lack of a non-substance use diagnosis (kappa = .91) (Shane, Jasiukaitis, & Green, 2003).

Interpretive cut-off scores for the full GAIN comorbidity scales have been found to vary by scale. To simplify interpretation, a low severity level is assigned when 0-24% of the items are endorsed, moderate is assigned when 25-74% of the items are endorsed, and a high severity is designated by endorsement of 75-100% of the items on each scale. Using an interpretive range of moderate or above, Titus and colleagues (2008) found the GAIN to have 98% sensitivity for any internalising and any externalising problems for both adults and adolescents, with sensitivity for many common specific problems in the 95 to 99% range. Using this same cut-off point to determine specificity resulted in correctly ruling out 42 to 73% of non-cases. Thus, the moderate/high cut-off point errs on the side of inclusion, with over

identification of potential cases. Using the high severity cut-off point improved specificity to 97% or better, but reduced sensitivity to 49 to 74% (erring on the side of exclusion) (Titus et al., 2008).

Client groups

The psychometrics of the GAIN and the scale norms have been established for both adults and adolescents and is used widely in substance disorder treatment settings and across a range of substance use disorders (Chan, Godley, Godley, & Dennis, 2009; Harris, Griffin, McCaffrey, & Morral, 2008; Liddle, Rowe, Dakof, Henderson, & Greenbaum, 2009; Martin & Copeland, 2008). It has also been used successfully in prison populations (Friedmann, Melnick, Jiang, & Hamilton, 2008)

However, there are a number of cultural limitations (e.g., American phrases and words) which may not be understood or applicable in an Australian context. Empirical findings relating to the GAIN are limited outside of the United States.

Availability/cost

The GAIN is copyrighted by Chestnut Health Systems. Licensing to use any of the GAIN family of instruments is $100 per agency for 5 years of use. Information can be found at:

- http://www.chestnut.org/LI/GAIN

Scoring, administration and expertise required

Training in the use of the GAIN is required. The GAIN can be orally administered to a client by a clinician or self-administered by the client. While self-administration can be efficient and reliable, it typically leads to more missing data and may have less validity (e.g., because questions may be incorrectly interpreted). Therefore, careful monitoring is recommended (Dennis, White et al., 2006).

The GAIN-Initial takes 60-120 minutes to complete and is designed to provide a standardised biopsychosocial assessment for people presenting to a substance abuse treatment using DSM-IV for diagnosis, it also is useful in treatment planning, performance/outcome monitoring, economic analysis and to support referral/communications with other systems. GAIN-Monitoring 90-Day is a quarterly follow-up version of this instrument for evaluating change over time. The GAIN-Quick is a briefer version of the initial GAIN and takes 15-20 minutes to complete and is useful in basic assessment and need for referral to specialty health, mental health, and/or substance use systems, and/or to support motivational interviewing related to substance use. The GAIN Shorter Screener is a 5 minute version used for screening in a general population to quickly identify who is likely to have an internal, external, or substance use diagnosis. Finally, the GAIN Treatment Satisfaction Index is a 3-minute measure of working alliance and engagement.

The Indigenous Risk Impact Screen (IRIS)

The Indigenous Risk Impact Screen (IRIS) (Ober & Schlesinger, 2005) was developed by an expert group of Indigenous and non-Indigenous researchers in Queensland to assist with the early identification of D&A problems and mental health risks. The tool consists of two sections, screening for both mental health problems and D&A misuse. In a validation study of 175 Aboriginal and Torres Strait Islander people from urban, rural, regional and remote locations in Queensland, the IRIS was found to have two factors corresponding to these two sections (Schlesinger, Ober, McCarthy, Watson, & Seinen, 2007).

Preliminary findings found the IRIS alcohol and drug and mental health subscales to demonstrate good convergent validity (correlation coefficients between .55 and .74) with other well-established screening instruments (DASS, LDQ, SDS, AUDIT) and self-reports of mental health symptoms and D&A use. Both of the subscales showed high internal consistency for each factor (alcohol and drug factor: Cronbach's alpha = .84; mental health factor: Cronbach's alpha = .81). Similarly, the IRIS was found to show good temporal stability, as the test-retest reliability of both the alcohol and drug subscale (r = .79) and the mental health subscale (r = 0.81) was excellent. Finally, using a cut-off score of 10 the alcohol and drug subscale had optimal sensitivity (65%) and specificity (86%). While a score of 11 was found to be the best balance of sensitivity (83%) and specificity (84%) for the mental health subscale (Schlesinger et al., 2007).

Client groups

IRIS is suitable for people who have self-identified as being from an Aboriginal or Torres Strait Islander background and are 18 years of age or older. The client must have a basic understanding English and the measure must not be used when the client is in acute withdrawal or the acute phase of physical/mental illness or is intoxicated (Australian Government Department of Health and Ageing, 2007).

Availability/cost

The IRIS is freely available online in the Australian Alcohol Treatment Guidelines for Indigenous Australians (Schlesinger et al., 2007):

- http://www.alcohol.gov.au/internet/alcohol/publishing.nsf/Content/AGI02

Scoring, administration and expertise required

The IRIS is made up of two sets of questions, with items 1 through 7 forming the 'D&A risk' component and items 8 through 13 forming the 'mental health and emotional well-being risk' component. The items assessing mental health and emotional well-being focus on symptoms of anxiety and depression. The client chooses the answer from a list of response options which best describes his/her current situation.

After tallying up the corresponding numbers, a score of 10 or greater on the D&A component indicates problematic use of D&A is likely, while a score of 11 or greater indicates the need for further assessment or brief intervention regarding mental health and emotional well-being (Australian Government Department of Health and Ageing, 2007). Any worker is able to administer/score the screening tool.

General Health and Functioning Measures

The Short-Form Health Survey (SF-36; SF-12)

The Short-Form 36 Health Survey (SF-36) (Ware, Kosinski, & Keller, 1996; Ware & Sherbourne, 1992) is a self-administered, 36-item questionnaire assessing the client's functioning status, symptoms/well-being and overall health. The SF-36 has 8 separate scales including physical functioning, physical role limitation, emotional role limitation, bodily pain, mental health, social functioning, vitality and general health perceptions. The SF-12 includes 12 questions from the SF-36. These include: 2 questions concerning physical functioning; 2 questions on role limitations because of physical health problems; 1 question on bodily pain; 1 question on general health perceptions; 1 question on vitality (energy/fatigue); 1 question on social functioning; 2 questions on role limitations because of emotional problems; and 2 questions on general mental health (psychological distress and psychological well-being). The SF-36 has been found to a particularly useful as an outcome tool (Gandek et al., 1998; Havard, Teesson, Darke, & Ross, 2006; McCallum, 1995; McHorney, Ware, Lu, & Sherbourne, 1994; Oslin, Slaymaker, Blow, Owen, & Colleran, 2005; Sannibale, Fucito, O'Connor, & Curry, 2005; Sanson-Fisher & Perkins, 1998); however, the SF-12 has not been used extensively in this sense, although preliminary findings suggest it too may be useful when used in this sense. Lenert and colleagues (2000), for instance, found SF-12 scores to be associated with a clinical change (remission) in depression, indicating the measure is sensitive to health outcomes. The main limitation in using these measures as outcome tools is that neither version registers quantity or frequency data on D&A use.

Across a range of different population groups the internal reliability of the SF-36 is acceptable to good (consistently ≥ 0.70) (Hopman et al., 2004; Kagee, 2001; Sanson-Fisher & Perkins, 1998). Gandek and colleagues (2004) for instance, found internal consistency was 0.83 to 0.93 for the eight scales and 0.94 (Physical Component Summary; PCS) and 0.89 (Mental Component Summary; MCS) for component summary measures. The test-retest reliability of the SF-36 has also been reported to be good (Beaton, Hogg-Johnson, & Bombardier, 1997; Calsyn et al., 2004; Essink-Bot, Krabbe, Bonsel, & Aaronson, 1997; Prieto, Alonson, Ferrer, & Antò, 1997; Stewart et al., 2003).

The SF-36 has also been found to correlate adequately with other measures of functioning (Brazier et al., 1992; Elliott, Renier, & Palcher, 2003; Jenkinson, Wright, & Coulter, 1994) and has generally been found to have good criterion validity (Brazier et al., 1992; Jenkinson, 1999; Keller et al., 1998; McHorney et al., 1994; McHorney, Ware, & Raczek, 1993; Sanson-Fisher & Perkins, 1998; Ware et al., 1998). The SF-36 has also been found to have strong content and construct validity in a number of studies (Cuthberston, Scott, Strachan, Kilonzo, & Vale, 2005; Ferguson, Robinson, & Splaine, 2002; Jenkinson et al., 1997; Jenkinson, Lawrence, McWhinnie, & Gordon, 1995; Jenkinson, Peto, & Coulter, 1994; Rost, Smith, Burnam, & Burns, 1992; Sharples, Todd, Caine, & Tait, 2000).

Finally, as previously mentioned, the SF-36 has been shown to be adequately sensitive to change in functioning and overall health over time (Salyers, Bosworth, Swanson, Lamb-Pagone, & Osher, 2000; Ware et al., 1996).

On the whole, the SF-12 has been found to have equally good psychometric properties (Ware et al., 1996). Test-retest reliability, for instance, was found to be good for both summary measures (PCS = 0.89, MCS = 0.76). Similarly, Resnick and Parker (2001) found the test retest reliability of the SF-12 to be in a similar range (*r* = 0.73-0.86) and the internal consistency of the revised SF-12 to be in line with the original (Cronbach's alpha coefficients of 0.72 to 0.89). This version too, was found to correlate well with other measures (Burdine, Felix, Abel, Wiltraut, & Musselman, 2000; Johnson & Coons, 1998; Lundberg, Johannesson, Isacson, & Borgquist, 1999; Macran, Weatherly, & Kind, 2003) and both construct and criterion validity have been shown to be strong (Gandek et al., 1998; Jenkinson & Layte, 1997). In Australia, the Hunter Heart and Stroke study (Lim & Fisher, 1999), the Longitudinal Study on Women's Health (Schofield & Mishra, 1998) and two mental health studies (Sanderson & Andrews, 2002a; Sanderson, Andrews, & Jelsma, 2001) have supported the construct validity of the SF-12.

Although international data suggests both forms of the survey are generally comparable (Gandek et al., 1998), one Australian study reported a much poorer degree of reproducibility, with the 12 items explained only 56% of the variance in the MCS and 82% in the PCS (McCallum, 1997). The author concluded that the standard SF-12 was not suitable for use in Australia, and that an Australian short form with a different subset of items may be more appropriate. However, Sanderson and Andrews (2002b) cross-validated the selection of the questionnaire items for the SF-12 in an Australian sample and found the SF-12 items predicted at least 90% of the variance in both the physical and mental summary scales of the SF-36, whether they were scored with Australian or United States normative data, concluding that the SF-12 *was* an appropriate substitute for the SF-36 when a briefer instrument is required (Sanderson & Andrews, 2002b). Consequently, population health data using the SF-12 can be found in the 1997 Australian National Survey of Mental Health and Well-Being (Andrews, Henderson, & Hall, 2001), the 2000 Mental Health Status of South Australian Population Study (Taylor et al., 2000), the 2002 Longitudinal Investigation of Depression Outcomes Study (Herrman et al., 2002) and the 2003 Australian Gulf War Veteran's Health Study (Sims et al., 2003).

In choosing between the SF-12 and the SF-36, users should consider the trade-off between test taker burden (i.e. number of questions, time to complete) and the precision of scores (i.e. how reliable the obtained score needs to be). Ware and colleagues (1996) reports that there is a 10% loss in the SF-12's ability to distinguish between different disease groups as compared to the SF-36.

Client groups

The SF Health Surveys have been translated into a number of languages (Bullinger et al., 1998; Pernegerv, Lepledge, & Etter, 1999). The validity of the SF-12 items in reproducing the SF-36 summary

scales was compared across nine European countries, with the 12 items explaining 89-92% of the variance in PCS scores and 88-94% of the variance in MCS scores (Gandek et al., 1998).

In other cross cultural studies, authors have found translations to be culturally appropriate and comparable in their content, but may prove problematic in those instances where respondents complete the questionnaire via an untrained translator, such as a friend or family member, translating the English version (Jenkinson, Chandola, Coulter, & Bruster, 2001; Wagner et al., 1998). In Australia, the SF-36 has been utilised for people from a non-English speaking background in Western Sydney (Cardona, Jorm, Williamson, & Chey, 1995). Watkins and colleagues (2000) translated the SF-36 and surveyed 1,610 Vietnamese migrants to Australia. The authors found all but two SF-36 items had good discriminant validity, and all eight scales of the Vietnamese version of the SF-36 had good discriminant validity, which supports the use of SF-36 constructs to assess self-reported health status among Vietnamese migrants. However, the mental health, vitality and bodily pain scales demonstrated low internal consistency. In Australia, little research has been reported on the use of SF-12 with people from a non-English speaking background. Neither version has been extensively applied to the Australian Indigenous population. Scott and colleagues (2000), however, reported some difficulties in applying the SF Health Surveys to the Maori and Pacific ethnic groups in New Zealand.

In a longitudinal study of women in 3,500 households across Australia the validity of both forms of the SF Health Survey were examined. The SF-12 PCS discriminated between women with poor versus good physical health, and MCS discriminated between groups who were or were not psychologically distressed on GHQ-12. The SF-36, relative to the SF-12, however, was found to be a more reliable measure for examining changes in health status over time and between groups (Schofield & Mishra, 1998).

The SF Health Surveys have been used successfully in countless subpopulations, for instance, in substance misusing populations (McGregor, Machin, & White, 2003; Morgan, Morgenstern, Blanchard, Labouvie, & Bux, 2003; Ryan & White, 1996; Stein, Herman, & Anderson, 2009), including older (Rosen, Smith, & Reynolds III, 2008) and adolescent users (Goldstein, Asarnow, Jaycox, Shoptaw, & Murray, 2007).

Both versions have also been used extensively in psychiatric samples (Feld, Colantonio, Yoshida, & Odette, 2003; Goldney, Fisher, Wilson, & Cheok, 2001; Sciolla, Patterson, Wetherell, McAdams, & Jeste, 2003; Sherbourne, Wells, & Ludd, 1996). Sanderson and Andrews (Andrews, 2002; Sanderson & Andrews, 2002a; Sanderson & Andrews, 2002b; Sanderson et al., 2001) have consistently found the SF-12 to be a good measure of disability in individuals suffering mild mental disorders (anxiety and depression). Whilst Salyers and colleagues (2000) have found the SF-12 to be a useful tool in a severe mental illness population.

The measures have also been used in studies of elderly populations (Resnick & Nahm, 2001; Sciolla et al., 2003) and homeless persons (Larson, 2002; Mares, Greenberg, & Rosenheck, 2008; Wong, Nath, & Solomon, 2007).

Availability/cost

Both forms of the SF are copyrighted and an annual license fee applies for the use of the instruments. Survey users are required to register and obtain a quote for the annual license fee that applies to their project. The license charge will depend upon whether users require a commercial or research license. Further information is available at:

- www.qualitymetric.com.
- http://www.sf-36.com/

An online version is also available on the CRUfAD website:

- http://www.crufad.com/phc/sf-12.htm

Scoring, administration and expertise required

The SF-36 takes approximately 5- 15 minutes to complete (SF-12 takes approximately two minutes). The scores are summed for each item, but an overall score is not calculated. Rather a profile for scores on the different dimensions is used. No training is required for those professionals with qualifications and experience in psychometrics and statistics. For those professionals without these qualifications, basic training is required in survey administration and the characteristics of the SF Health Surveys.

The traditional scoring method for the SF-12 is somewhat difficult when scored by hand, however, in a recent community and clinical-based sample, scores generated by the standard scoring method were found to correlate with a simplified scoring method. Means and standard deviations were similar and no individual scores deviated by more than 2.89 in the community sample or by 3.06 in the clinical sample. Thus scoring can be simplified where hand scoring is an advantage (Andrews, 2002).

The SF Health Survey elicited mostly positive responses when compared to other measures in a sample of consumers (Stedman et al., 2000). Furthermore, an electronic version of the SF-36 has been found to be slightly quicker to complete, equivalent in performance and more effective than the paper version (Ryan, Corry, Attewell, & Smithson, 2002).

The Children's Global Assessment Scale (CGAS)

The Global Assessment Scale (GAS) (Endicott, Spitzer, Fliess, & Cohen, 1976) is a rating scale for evaluating the overall functioning of a subject during a specified time period on a continuum from psychological/psychiatric sickness to health. This was adapted to form a version for children and adolescents aged 4 to 16, known as the Children's Global Assessment Scale (CGAS) (Schaffer et al., 1983). The original GAS is used only on occasion but the CGAS is currently in routine use in Australia and therefore this review will focus on the latter version.

The CGAS offers a means of establishing levels of dysfunction for children and adolescents. It is designed to be used in conjunction with diagnostic measures, and aims to provide more detailed information upon which to develop management plans and to evaluate improvement or deterioration in functioning following treatment or over time (Dyrborg et al., 2000; Steinhausen, 1987).

Overall the CGAS has performed well on various tests of its psychometric properties. However, the tool has been criticised for its vulnerability to rater manipulation (and therefore accuracy) and the global nature of the scoring has been criticised for failing to consider different domains of functioning in any organised manner (Hodges & Gust, 1995). Others have found the CGAS inter-rater reliability to be moderate (Hanssen-Bauer, Aalen, Ruud, & Heyerdahl, 2007). Nonetheless, Shaffer and colleagues (1983) found both inter-rater and test-retest reliability of the CGAS to be excellent and subsequent studies have supported these findings (Bird, Canino, Rubio-Stipec, & Ribera, 1987; Dyrborg et al., 2000; Green, Shirk, Hanze, & Wanstrath, 1994; Weissman, Warner, & Fendrich, 1990). Weissman and colleagues (1990) also assessed the sensitivity of the CGAS as an outcome measure and observed higher scores in the subgroup of children who developed disorders with those who remained disorder-free. Other studies have supported the use of this tool as an outcome measure (Ginieri-Coccossis, Liappas, Tzavellas, Triantafillou, & Soldatos, 2007; Hintikka et al., 2006; McShane, Bazzano, Walter, & Barton, 2007; Remschmidt et al., 2007).

The development process of the CGAS demonstrated the content validity of the tool (Schaffer et al., 1983) and the concurrent validity has been generally found to be strong with the CGAS demonstrating high correlations with independent measures of competence, intellectual and social functioning, and problem solving, and only moderate correlations with measures of symptomatology (Bird et al., 1987; Green et al., 1994; Schaffer et al., 1983; Sourander & Piha, 1997; Steinhausen & Metzke, 2001; Weissman et al., 1990). Lower average scores have also been demonstrated for inpatients when compared with outpatients (Schaffer et al., 1983), psychiatric services users when compared with non-users (Steinhausen, 1987), referred clients when compared with non-referred clients, and cases meeting diagnostic criteria when compared with those that do not (Bird et al., 1987). These findings give further support to the validity of the measure.

Finally, Sourander and colleagues (1996; 1996) have found that CGAS ratings at admission to inpatient care were predictive of functioning and residential status at follow-up.

Overall, a recent review found the CGAS to have adequate psychometric properties for routine use (Schorre & Vandvik, 2004).

Client groups

The CGAS is specifically designed for young people aged between 4 and 16 but the standard version is available for adults and elderly individuals (Fisher & Copenhaver, 2006). It has been validated as an appropriate tool for cross-cultural comparison and has been translated into a range of languages (Canino et al., 2004; Hanssen-Bauer, Gowers et al., 2007; Petersen, Bilenberg, Hoerder, & Gillberg, 2006; Roberts, Roberts, & Xing, 2006; Szobot, Ketzer, Parente, Biederman, & Rohde, 2004).

The measure has also been used in prison populations (Abram, Paskar, Washburn, & Teplin, 2008) and routinely within various psychiatric populations (Green et al., 2007; McShane et al., 2007).

Availability/cost

The CGAS is freely available at:

- http://depts.washington.edu/washinst/Training/CGAS/CGAS%20Index.htm

Scoring, administration and expertise required

The CGAS is a clinician-administered tool and provides a single global rating of a child or adolescent's lowest level of functioning over the previous two weeks. Ratings range from 1 (severe dysfunction) to 100 (superior functioning), and the threshold of psychopathology is suggested to sit between 61 and 71 (Bird et al., 1990; Weissman et al., 1990). No formal training materials are provided with the CGAS, although the Australian Mental Health Outcomes and Classification Network have developed training materials as part of its standard package. However, in making their rating, clinicians rely heavily on clinical judgement and are therefore, presumed to have clinical expertise and training in the use of psychometric measures. Clinicians are guided by anchor points at every 10th degree on the scale; these offer an indication of the type of behavioural functioning displayed by consumers at that level.

The CGAS is generally regarded as a useful measure of child and adolescent functioning, providing more detailed information for guiding treatment decisions than diagnosis- or symptom-based measures alone (Bird et al., 1990; Dyrborg et al., 2000; Weissman et al., 1990).

The instructions accompanying CGAS are brief and easy to understand. When the information required is available and the raters are experienced in using the scales, scoring is said to take only a minute or two (Schorre & Vandvik, 2004).

Global Assessment of Functioning Scale (GAF)

The Global Assessment of Functioning Scale (GAF) (Endicott et al., 1976) is a slightly modified form of the Global Assessment Scale and it is a tool used to measure overall level of psychological, social and occupational client functioning on a scale ranging from 1 to 100. The GAF can be completed with reference to varying time periods (e.g., currently, highest level of past year) and it constitutes the operationalisation of Axis V of the DSM-IV mutiaxial assessment (American Psychiatric Association, 1994). It has been used as an assessment and an outcome tool in various populations (Fridell & Hesse, 2006; Goldman, Skodol, & Lave, 1992).

An early study by Shanks (1994) reported that the potential utility of GAF in clinical practice was limited. Using a sample of 103 severely mentally ill patients, Jones and colleagues (1995) found the measure to be a reliable and valid measure of psychiatric disturbance. The different subscales showed good internal consistency (coefficients between 0.71 and 0.76). However, rater effects were found to be a significant source of variation in assessments. Having said this, raters had only one brief training session on the instrument. All GAF scores were associated with current support needs of patients. Symptom and disability scores were associated with changes in antipsychotic medication in the previous month indicating good validity. Similar findings were reported in a subsequent study (Coffey, Jones, & Thornicroft, 1996) which concluded that although the GAF in itself appears to be unsuitable for making individual treatment decisions it could be used as a preliminary step prior to more sophisticated scales.

A recent study by Schwartz (2007) investigated the construct validity of the GAF in clients with schizophrenia. This study compared point-specific GAF scores with concurrent ratings of symptomatology and social and occupational impairments using the Functional Assessment Rating Scale, and found work and school-related problems, danger to others, and psychotic symptoms predicted lower GAF scores, indicating good validity. This study showed similar findings to an earlier study (Patterson & Lee, 1995). Similarly, in a sample of 398 individuals with schizophrenia or schizoaffective disorder, a modified GAF exhibited very high levels of reliability, while the occupational and symptom subscales showed good convergent and discriminant validity (compared to the PANAS and the QoL interview). Further support for the validity of the GAF comes from a recent study, in a sample of individuals suffering from schizophrenia, GAF ratings were highly correlated with ratings of symptoms (as measured by the PANAS) and social behaviour (as measured by the Social Behaviour Schedule) at both follow-ups, but not at initial assessment (Startup, Jackson, & Bendix, 2002).

Roy-Byrne and colleagues (1996) found the GAF to correlate strongly with ratings of clinical symptoms, not functioning and concluded that reliance on the GAF as the only tool to assess patients' functioning may be problematic. Similarly, although Dufton and Siddique (1992) found the GAF to be of use in detecting change from admission to discharge, they suggest changes in GAF scores may lack clinical significance, and concurrent validity. However, this study used a very early version which has undergone revisions to correct for these early shortcomings. The original GAF confounded two areas of functioning: symptomatology and social functioning. Goldman and colleagues (1992) suggested that the GAF be

administered as both an overall scale and as two separate measures assessing symptoms and disability. The modified GAF scale has more detailed criteria and a more structured scoring system than the original GAF. This modified version was found to have improved inter-rater reliability (correlation coefficients of 0.81, compared with 0.62 for the original). Validity studies showed a high correlation (0.80) between the two sets of scores. The authors suggest this modified GAF may be particularly useful when persons with varying skills and employment background – and with little GAF training – must rate patients (Hall, 1995).

Client groups

The tool was developed for use in both children and adults and a recent review has generally confirmed this utility (Schorre & Vandvik, 2004), however, Piersma and Boes (1997) conducted a study on three psychiatric populations and found a significant difference between mean GAF for adult and adolescent inpatients both at admission and at discharge. They concluded that validity and reliability established for adults cannot easily be transferred to children and adolescents. It has also been recommended that a modified GAF scale specifically reflecting an older adult's activities during this period of life be created (Mossbarger, 2005; Whitney, Kunik, Molinari, Lopez, & Karner, 2004).

As indicated above the GAF has been successfully used in a range of psychiatric populations (Jones et al., 1995; Patterson & Lee, 1995; Startup et al., 2002), along with a variety of substance using (often comorbid) populations (Batki, Leontieva, Dimmock, & Ploutz-Snyder, 2008; Maremmani et al., 2008; Mátyássy, Kelemen, Sárközi, Janka, & Kéri, 2006; McCarty et al., 2008; Schwartz, Hilscher, & Hayhow, 2007; Wobrock et al., 2007).

Availability/cost

The GAF is freely available online:

* http://psyweb.com/Mdisord/DSM_IV/jsp/Axis_V.jsp

Scoring, administration and expertise required

Training in the use of the scale is fundamentally important to the reliability of the instrument (Bates, Lyons, & Shaw, 2002; Rey, Starling, Wever, Dossetor, & Plapp, 1995; Yamauchi, Ono, Baba, & Ikegami, 2001). A self-report version of the GAF was found to be a valid and reliable unidimensional instrument measuring psychological, social and occupational functioning (Bodlund, Kullgren, Ekselius, Lindström, & von Knorring, 1994).

Life Skills Profile (LSP)

The Life Skills Profile (LSP) (Rosen, Hadzi-Pavlovic, & Parker, 1989) was originally developed as a measure of aspects of functioning which affected how successfully people with schizophrenia lived in the community and has since been applied to other major psychiatric disorders and to a broad range of other diagnoses (Rosen, Hadzi-Pavlovic, Parker, & Trauer, 2006). The LSP has a positive mental health philosophy whereby it aims to emphasize a person's "life skills" rather than their "lack of life skills" (Rosen et al., 2006).

There are three versions of the LSP (the LSP-39 and two abbreviated versions: the LSP-20 and the LSP-16). The subscales of each version are displayed in Table 1 below:

Table 1. LSP versions and respective subscales

Subscale	LSP-39	LSP-20	LSP-16
1	Self-care	Self-care	Self-care
2	Non-turbulence	Anti-social	Anti-social
3	Social contact	Withdrawal	Withdrawal
4	Communication	Bizarre	—
5	Responsibility	Compliance	Compliance

Early studies of the LSP-39 reported good psychometric properties, principally when completed by case workers, residential carers and parents. The internal consistency of the LSP has been reported as moderately high with subscale correlations ranging from 0.64 to 0.88 and total score correlations ranging from 0.93 to 0.94 (Dickinson & Coursey, 2002; Stedman, Yellowlees, Mellsop, Clarke, & Drake, 1997; Trauer, Duckmanton, & Chiu, 1995). The Communication subscale has been shown to have the poorest internal consistency (and was later dropped in the LSP-16) (Parker, Rosen, Emdur, & Hadzi-Pavlov, 1991; Trauer et al., 1995).

Rosen and colleagues (1989) found the measure to have good predictive validity, with worse LSP predicting more changes in accommodation. While the mean total measure inter-rater reliability was moderately high (correlation coefficient was 0.68). In a subsequent study, test-retest reliability was also found to be high (overall correlation coefficients of 0.89 for total scores and 0.78-0.90 for the 5 scales), although interrater reliability was found to differ between the different types of scorers (e.g., carers compared to parents). The authors suggest this may be due to differing rater relationships with the subjects (Parker et al., 1991). Concurrent validity of the LSP was also good. LSP scores were compared with those generated by the Katz Adjustment Scales (Katz & Lyerly, 1963) scales. The total scores on both scales were moderately correlated (r = 0.65). Further support has been found for the concurrent validity of the LSP in that it has been shown to perform well against the HoNOS, the GAF, the RFS, the QoL, the SBS, the Resource Associated Functional Level Scale (RAFLS) and the GAS (Dickinson & Coursey, 2002; Norman et al., 2000; Parker et al., 2002; Stedman et al., 1997; Trauer et al., 1995; Trauer, Eagar,

Gaines, & Bower, 2004; Wooff, Schneider, Carpenter, & Brandon, 2003). However, the LSP has performed less well and more inconsistently against the BASIS-32, the MHI and the SF-36 (Stedman et al., 1997), the General Well-Being Scale (GWB) (Norman et al., 2000; Trauer, Duckmanton, & Chiu, 1998), and the BPRS (Wooff et al., 2003). Although it has been suggested that this may be a result of the self-report style of many of these measures, and that a number assess symptomatology rather than functioning (Pirkis, Burgess, Kirk, Dodson, & Coombs, 2005).

Trauer and colleagues (1995) suggested modifications to the original version based on confirmatory factor analyses among 200 severely mentally ill patients. Again internal consistencies were generally good, but inter-rater reliabilities were only marginally acceptable (some scores were found to vary with how well and how long the rater had known the patient). The Communication subscale had the poorest psychometric properties. The LSP was compared to locus of care (i.e. community or hospital), BPRS ratings and RAFLS ratings in order to attain concurrent validity, which was found to be good.

The LSP-20 was found to have good internal consistency (Cronbach's alpha ranging from .71-.90) while the subscales of the LSP-20 and total score correlated at 0.92 or higher with the original. Test-retest and inter-rater reliabilities are also comparable with the original, and validity was demonstrated with expected correlations with items of the Positive and Negative Syndrome Scale (PANSS). Compared with the LSP-39, scores on the LSP-20 tend to reflect lower functioning, meaning these levels or thresholds pertaining to one instrument cannot be directly compared with those of the other (Rosen, Trauer, Hadzi-Pavlovic, & Parker, 2001).

The LSP-16, has also been found to correlate well with the original measure. Buckingham and colleagues (1998) found an 85% to 90% concordance between the top four items of each subscale with the full subscale of the LSP-39. Likewise, Rosen and colleagues (2001) found the LSP-16 performed well against its parent instrument. Trauer (2003) also found the LSP-16 correlated well with the HoNOS, but that – like the LSP-39 – it showed poor or mixed performance against the BASIS-32.

Parker and colleagues (2002) suggest that the LSP-39 performs less well in predicting client outcomes when compared to other outcome measures. However, other studies have suggested the LSP does show good validity in predicting outcomes relating to retention in the community (Preston, 2000b), hospital readmission (Andrews, Teesson, Stewart, & Hoult, 1990; Parker & Hadzi-Pavlovic, 1995), change in locus of care (Trauer, Duckmanton, & Chiu, 1997), length of inpatient stay (Ballesteros, Martinez, Martin, Ibarra, & Bulbena, 2002; Kisely, Preston, & Rooney, 2000) and overall costs of care (Kisely et al., 2000; Trauer et al., 1998). Similarly, the LSP has been found to demonstrate greater levels of improvement in those who participate in more intensive treatment and care than in those who undergo routine case management (Hambridge & Rosen, 1994; Hamernik & Pakenham, 1999; Johnston et al., 1998; Rosen & Teesson, 2001).

Overall, the LSP has been found to be a useful measure of client treatment outcome (Eagar, Trauer, & Mellsop, 2005; Horner & Asher, 2005).

Stedman and colleagues (1997) found further support for the predictive potential of the LSP when scores were compared with client self-reported improvement or deterioration over time. LSP scores were worse in the group who reported a decline in their levels of functioning. However, this finding requires further study as there was no association between LSP score and self-reported change for any other group.

The LSP-16 is completely contained within the LSP-20, with the major difference between the two, being the presence of a "Bizarre" subscale in the LSP-20. Therefore, the LSP-16 would be a suitable instrument to use with populations in which psychosis was absent or uncommon, but the LSP-20 when psychosis is more prevalent. The LSP-20 may be more suited to routine service disability and aggregated outcome assessments, but less suited than the LSP-39 to detailed research, or to interactive use as part of service user's individual care planning and review.

Client groups

Studies have generally elicited a positive responses from public sector mental health service providers regarding the LSP-39 as a measure of disability (Andrews, Peters, & Teesson, 1994; Stedman et al., 1997). However, the LSP has been criticised for lacking relevance in community setting and long-term inpatient settings because of the restrictiveness of the response options or a lack of opportunity to exhibit some of the skills (Eu, Lee, Parker, & Loh, 2001; Hadzi-Pavlovic, Rosen, & Parker, 1992; Stedman et al., 1997).

In a recent unpublished report the LSP-16 was found to show similar levels of internal consistency for the subscales when applied with Australian Indigenous consumers as that observed in non-Indigenous populations (Haswell-Elkins, 2006). Good correlations with scores and indicators of wellness and illness identified by Indigenous consumers and carers in in-depth interviews, indicating content validity of the measure. The author recommended the continued use of LSP-16 with Indigenous consumers, together with a set of guiding principles, as a tool that has the capacity to capture important information for the consumer, carer and clinician. They further suggested that the development and validation of consumer-rated tools that capture the more fundamental and culturally determined aspects of Indigenous mental health, needs to occur (Haswell-Elkins, 2006).

The LSP has been translated and validated in a wide range of different countries, languages and ethnic groups (Afuwape et al., 2006; Bulbena Vilarrasa, Fernández de Larrinoa Palacios, & Domínguez Panchón, 1992; Burgés, Fernández, Autonell, Melloni, & Bulbena, 2007; Fernández de Larrinoa Palacios, Bulbena Vilarrasa, & Domínguez Panchón, 1992; Hasegawa et al., 1997; Huang, Sousa, Tsai, & Hwang, 2008; Mohr, Simon, Favrod, Fokianos, & Ferrero, 2004; Uys & Zulu, 1996; Zizolfi, 1997). The tool has also been used as an outcome measure in substance misusing populations (Bradley, Baker, & Lewin, 2007; Moller & Linaker, 2006) and homeless populations (Goldfinger, Schutt, Seidman, Turner, & et al., 1996).

The LSP has been used in a range of settings amongst various populations with generally high levels of success. It has been suggested, however, that different groups may show particular patterns of scores on the LSP. For instance, Rosen and colleagues (1989) found that younger people more frequently showed low scores on the Responsibility and Non-turbulence subscales of the original LSP-39. As mentioned above, the tool has been used with homeless populations and those in long-term residential care settings as well as those in more stable accommodation and total scores on the LSP have also been found to distinguish between consumers (i.e., more preferable scores for more stable accommodation) (Andrews et al., 1990; Browne & Courtney, 2004; Keller & Hayes, 1998; Kirkby, Daniels, Jones, & McInnes, 1997; Trauer et al., 1997; Trauer et al., 1998).

Availability/cost

The LSP is in the public domain and may be used without cost, but with due acknowledgment of the source. It is available at:

- http://www.blackdoginstitute.org.au/search/index.cfm
- http://www.sswahs.nsw.gov.au/MHealth/mhoat/LSP16forCSAHS16Mar04.pdf

Scoring, administration and expertise required

The original LSP-36 is scored in line with its positive mental health philosophy so that high scores indicate high levels of life skills. The LSP-16, however, was scored in the direction of impairments (high ratings indicate poorer functioning). Since the LSP-20 represents an abbreviation of the LSP-39 as well as an extension of the LSP-16, either a 'strengths' or 'impairments' oriented scoring template could be used (Rosen et al., 2001). No training is required to use the scale.

Stedman and colleagues(1997) surveyed service providers about their experience with using the LSP, and found the majority of service providers felt the measure was easily understood and viewed the questions as relevant, useful and effective in measuring outcomes for consumers. Respondents in public sector psychiatric settings, in particular, rated the LSP more highly than other observer-rated measures (Stedman et al., 2000; Stedman et al., 1997).

Camberwell Assessment of Need (CAN)

The Camberwell Assessment of Need (CAN) (Phelan et al., 1993) is a comprehensive needs assessment and treatment outcome tool. The CAN is one of the most widely used instruments for this purpose. It assesses need in 22 areas: accommodation, food, self-care, looking after the home, daytime activities, physical health, psychotic symptoms, information about condition and treatment, psychological distress, safety to self and others, alcohol, (abuse of) drugs, company, intimate relationships, sexual expression, child care, access to a telephone, education, transport, budgeting and benefits (Phelan et al., 1995). Both research and clinical versions of the tool exist, while a shorter version, the CANSAS (Camberwell Assessment of Need Short Appraisal Schedule) also exists and uses one item from the 22 areas in the original CAN.

Correlations of the inter-rater and test-retest reliability of the total number of needs identified by staff were 0.99 and 0.78 respectively in the preliminary study of the instrument. The percentage of complete agreement on individual items ranged from 100-81.6% (inter-rater) and 100-58.1% (test-retest) (Phelan et al., 1995). Generally though, concerns have been raised about the relationship between staff-rated needs and client-rated needs and the inter-rater reliability generally. For instance, Slade and colleagues (Slade, Phelan, Thornicroft, & Parkman, 1996) reported staff and clients rated a similar number of needs, but not in the same areas. There was better agreement between staff and clients regarding needs that have a specific service intervention. The authors concluded that needs are very often assessed differently by staff and clients, which has implications for how needs are assessed in clinical practice (Slade et al., 1996).

In an Australian context, the CAN has been validated (Gallagher & Teesson, 2000a; Issakidis & Teesson, 1999). However, although Issakidis and Teesson (1999) found the tool to be applicable to an Australian setting there was poor agreement between clinicians and clients in identification of the 22 need areas. These findings highlight the fact that staff and clients differ in their assessment of need. Similar findings have been made in relation to the CANSAS in an Australian context (Andresen, Caputi, & Oades, 2000). This has been reported, particularly concerning unmet needs, in a range of studies (Hansson et al., 2001; Macpherson, Collins-Atkins, Gregory, Slade, & Lerescu, 2008; Trauer & Tobias, 2004). Test-retest reliability is generally strong (Arvidsson, 2003; Trauer & Tobias, 2004). However, Wennström and Wiesel (2006) found that summary scores conceal changes in need on the underlying items and thus is recommended not to be used as dependent measures when comparisons among populations or between points in time are of interest.

In the initial development process a draft version of the instrument was sent for comments to 50 experienced professionals in the fields of social work, psychiatry, psychology, psychiatric nursing and occupational therapy. The consensus was that there was a requirement for a needs assessment instrument, and that the CAN would be useful and relevant (Phelan et al., 1995). Further evidence of content validity was found through a parallel survey that was conducted of 59 people with severe mental illness who were either current inpatients or attending a psychiatric day-hospital. All topics were

rated as being at least moderately important, indicating that the instrument is free from item bias (Phelan et al., 1995).

The total scores of the CANSAS were found to moderately correlate with the GAF and the HoNOS (Salvi, Leese, & Slade, 2005). Some have suggested, however, that although the CANSAS was found to be useful in assessing when treatment should be commenced or continued, it may be insufficiently sensitive to be used as an outcome measure (Slade, Beck, Bindman, Thornicroft, & Wright, 1999).

In response to the reliability concerns a newly developed patient-rated, short form (CANSAS-P) has been evaluated. The CANSAS-P exhibited comparable detection of needs with its predecessor, better identification of domains that are problematic for patients to respond to, good test-retest reliability, especially for unmet needs, and generally positive evaluations by patients. The CANSAS-P has been recommended as the needs assessment measure of choice for completion by patients (Trauer, Tobias, & Slade, 2008).

Client groups

The CAN has been translated and validated in a range of countries, particularly across Europe (Ericson, Hansson, & Teike, 1997; Heinze, Taylor, Priebe, & Thornicroft, 1997; Knudsen et al., 2000; McCrone et al., 2005; McCrone et al., 2000; Nielsen et al., 1999; Ruggeri, Lasalvia, Nicolaou, & Tansella, 1999; Wiersma, Nienhuis, Giel, & Slooff, 1998; Yeh, Luh, Liu, Lee, & Slade, 2006). The tool is also widely used in Australia (Gallagher & Teesson, 2000a; Issakidis & Teesson, 1999), but has not been specifically used in Indigenous populations.

The tool is also used among D&A populations (Nielsen, Petersen, Werdelin, Hou, & Lindhardt, 2000). Furthermore, there are a number of specialised versions of the CAN that have been developed and validated including a version for adults with developmental and intellectual disabilities (CANDID) (Xenitidis et al., 2000) and the elderly (CANE) (Reynolds et al., 2000) as well as a version for pregnant mothers (CAN-M) (Howard et al., 2007) and the CANFOR for forensic populations (Thomas et al., 2008).

Availability/cost

To attain the full versions of the CAN instruments, the manual must be purchased:

- http://www.rcpsych.ac.uk/publications/howtoorder.aspx

Scoring, administration and expertise required

Generally the CAN is administered in an interview format, although a newly developed self-report also exists (see above). It was developed to be easily learned and used, without formal training and usable by a wide range of professionals (Phelan et al., 1995). For each topic examples are given in the CAN for what constitutes a need, and the presence of a need is rated on a 3-point scale: 0 = "no problem", 1 =

"no/moderate problem because of continuing intervention" and 2 = "current serious problem". If there is a need (need rating = 1 or 2) then further questions on the topic are asked. A 4-point scale (from 0 = "no help" to 3 = "high help") assesses the level of help received from informal sources (friends and relatives) and services (e.g. paid staff) and the level of help the interviewee (patient or staff member) thinks is needed from services. A 2-point scale (0 = "no" and 1 = "yes") is used to rate whether the interviewee thinks the patient is getting the right type of help. For each question a rating of 9 is used for "not known" (Phelan et al., 1995).

Strengths & Difficulties Questionnaire (SDQ)

The Strengths and Difficulties Questionnaire (SDQ) (Goodman, 1997) was developed as a brief screening tool that describes children and adolescents' behaviours, emotions and relationships. It aims to assess both negative and positive attributes of behaviour and to equally cover five dimensions (namely conduct problems, emotional symptoms, hyperactivity, peer relationships and prosocial behaviour). The author suggests that the SDQ can be used for screening, as part of a clinical assessment, as a treatment-outcome measure, and as a research tool (Goodman, 2001).

Various studies have considered the psychometric properties of the SDQ (Pirkis et al., 2005). Although there have been limited studies examining the measure's content validity, confirmatory factor analysis studies have generally found five-factor solutions that corresponded with the original scales proposed by Goodman (Goodman, 2001; Hawes & Dadds, 2004; Koskelainen, Sourander, & Vauras, 2001; Muris, Meesters, & van den Berg, 2003; Roy, Veenstra, & Clench-Aas, 2008; Smedje, Broman, Hetta, & von Knorring, 1999; Thabet, Stretch, & Vostanis, 2000), suggesting good construct validity. However, some recent findings have raised concern about the factors structure (Mellor & Stokes, 2007; Percy, McCrystal, & Higgins, 2008)

The internal consistency of the SDQ has also been found to be strong (Cronbach's alpha ≥ 0.70 for total difficulties) (Goodman, 2001; Goodman, Meltzer, & Bailey, 2003; Koskelainen et al., 2001; Malmberg, Rydell, & Smedje, 2003; Mellor, 2004; Muris et al., 2003; Ronning, Handegaard, Sourander, & Morch, 2004; Smedje et al., 1999; van Widenfelt, Goedhart, Treffers, & Goodman, 2003). However, the parent-rated and teacher-rated scales tend to have better internal consistency than the self-report scales and the Conduct Problems and Peer Problems subscales have been found to have comparatively poorer internal consistency (relative to the good consistency reported in other scales) (Goodman, 2001; Muris et al., 2003). For instance, Bourbon and colleagues (2005) found a Cronbach's alpha coefficient of only 0.46 for the Peer Problems subscale which was in contrast to the coefficients of between 0.63-0.77 for the other subscales and 0.83 for Total Difficulties scale.

A recent study found father and mother ratings to correlate well overall, however fathers reported higher mean scores than mothers for externalising behaviours. Higher reporting by fathers was related to alcohol misuse, the couple relationship, parenting strategies, and father employment. The authors suggest using combined parental reports in clinical settings for improved sensitivity when identifying problem behaviours (Davé, Nazareth, Senior, & Sherr, 2008).

The concurrent validity of the SDQ has also been supported by findings that scores on the tool correlate well with those on the Rutter questionnaires, the Child Behavior Check List (CBLC), the Children's Depression Inventory (CDI), the Revised Children's Manifest Anxiety Scale (RCMAS), the Attention Deficit Hyperactivity Disorder Questionnaire (ADHDQ), an abbreviated version of the Child and Adolescent Burden Assessment (CABA) and the Youth Self Report (YSR) (Goodman, 1997; Goodman & Scott, 1999;

Klasen et al., 2000; Koskelainen, Sourander, & Kalijonen, 2000; Muris et al., 2003; van Widenfelt et al., 2003).

Similarly, the SDQ has been found to distinguish between those children/adolescents receiving treatment and those who are not, and could do so at least as well as other, more established instruments like the Rutter questionnaires, the CBCL, and the YSR (Goodman, 1997, 1999; Goodman, Meltzer, & Bailey, 1998; Goodman et al., 2003; Hawes & Dadds, 2004; Malmberg et al., 2003; Mellor, 2004; Mullick & Goodman, 2001).

Furthermore, within clinic-based samples, the SDQ has been shown to be able to discriminate between particular diagnoses or problematic behaviour as well or better than other instruments like the CBCL and the YSR (Goodman, Ford, Simmons, Gatward, & Meltzer, 2000; Goodman, Renfrew, & Mullick, 2000; Goodman & Scott, 1999; Hawes & Dadds, 2004; Klasen et al., 2000; Koskelainen et al., 2000; Mathai, Anderson, & Bourne, 2004; Mullick & Goodman, 2001). Others have claimed that although the measure has good convergent validity, its discriminant validity is less strong (Hill & Hughes, 2007).

The measure has shown to be adequately sensitive and specific and the Australian normative data for appropriate cut-off scores is presented in Table 2 below.

Table 2. Abnormal/of concern SDQ Cut-off scores for subscales and total difficulties (boys and girls)

Subscale									
	7- 10 years			11-13 years			14-17 years		
	Self report	Parent report	Teacher report	Self report	Parent report	Teacher report	Self report	Parent report	Teacher report
Boys									
Emotional symptoms	≥6	≥5	≥4	≥5	≥5	≥3	≥5	≥5	≥3
Conduct problems	≥5	≥4	≥4	≥5	≥4	≥3	≥5	≥4	≥3
Hyper-activity	≥7	≥8	≥9	≥6	≥7	≥7	≥7	≥7	≥6
Peer problems	≥5	≥5	≥4	≥4	≥4	≥4	≥3	≥6	≥4
Prosocial	≤5	≤6	≤4	≤5	≤5	≤4	≤5	≤5	≤4
Total difficulties score	≥20	≥19	≥17	≥16	≥17	≥15	≥19	≥18	≥15
Girls									
Emotional symptoms	≥6	≥5	≥4	≥6	≥5	≥4	≥6	≥5	≥3
Conduct problems	≥4	≥3	≥3	≥4	≥4	≥3	≥4	≥4	≥3
Hyper-activity	≥6	≥6	≥5	≥6	≥6	≥5	≥6	≥6	≥5
Peer problems	≥4	≥5	≥4	≥4	≥4	≥4	≥3	≥4	≥4
Prosocial	≤6	≤6	≤5	≤6	≤6	≤6	≤6	≤6	≤6
Total difficulties score	≥17	≥17	≥15	≥17	≥17	≥14	≥15	≥17	≥14

Adapted from Mellor (2005)

The stability of the SDQ (test-retest reliability) was found to be good across a range of time intervals (ranging from 3-4 weeks to 12 months), even for young children. However, non-self-report ratings were more stable than those using self-report (Goodman, 1999, 2001; Hawes & Dadds, 2004; Muris et al., 2003).

The majority of studies examining the inter-rater reliability of the SDQ have found positive correlations between different raters on the individual scale scores, the total difficulties score and the impact score (Goodman, 1997, 2001; Goodman et al., 1998; Koskelainen et al., 2000; Mellor, 2004; Muris et al., 2003; van Widenfelt et al., 2003), but there have been some exceptions (Cury & Golfeto, 2003; Thabet et al., 2000).

Again, in general, correlations between the parent and teacher versions of the SDQ have been found to be higher than correlations between either of these versions and the self-report version (Pirkis et al., 2005). In a recent Victorian clinical sample, Mathai and colleagues (2003) found the SDQ to be sensitive to client outcome change over time (as measured by the HoNOSCA).

Client groups

The SDQ has been translated into more than 40 languages (Goodman, 2001) and used widely both internationally and within Australia (Capron, Thérond, & Duyme, 2007; d'Acremont & Linden, 2008; Du, Kou, & Coghill, 2008; Leeuwen, Meerschaert, Bosmans, De Medts, & Braet, 2006; Matsuishi et al., 2008; Mellor, 2005; Rothenberger, Becker, Erhart, Wille, & Ravens-Sieberer, 2008). However, the majority of studies come from European populations. Separate versions for different aged children have also been created (4-10 years and 11-17 years). The major differences are wording modifications to ensure relevance to the age group in question. Similarly, several different informant versions of the SDQ also exist (teacher, parent, self-report).

The SDQ has also been widely used in psychiatric populations (Mathai et al., 2008), and has also been found to be useful within intellectually disabled children (Emerson, 2005; Kaptein, Jansen, Vogels, & Reijneveld, 2008). It has also been used in substance abusing youth (Christie et al., 2007; Cosden, Panteleakos, Gutierrez, Barazani, & Gottheil, 2004).

Availability/cost

The SDQ is copyrighted, however paper versions may be downloaded and subsequently photocopied without charge by individuals or non-profit organisations provided they do not charge clients:

- http://www.sdqinfo.com/

Scoring, administration and expertise required

In each version of the SDQ, either the parent, teacher or child/adolescent is asked to consider the child/adolescent's behaviour over the past six months (or over the last one month in the case of follow-up administrations), and rate each statement. It can be completed in 5-10 minutes and contains 25 items.

For the majority of items, a response of 'not true' is scored 0, a response of 'somewhat true' is scored 1 and a response of 'certainly true' is scored 2. The exceptions are items 7, 11, 14, 21 and 25, where the reverse scoring order applies. This scoring method yields a score of 0-10 for each subscale, and a total difficulties score of 0-40, generated by summing the scores from all of the scales except the Prosocial Behaviour subscale (Goodman, 2001, 2003).

The limited data that exists suggests it is a simple tool to use, particularly compared to other measures by mothers (Goodman & Scott, 1999). A recent review concluded that the SDQ is easy to complete, user-friendly because of its positive attributes items, allows comparisons to be made between different populations and is sensitive to change (Vostanis, 2006).

World Health Organisation Disability Assessment Schedule II (WHODAS II)

The World Health Organisation Disability Assessment Schedule II (WHODAS II) was developed to assess the activity limitations and participation restrictions experienced by an individual, irrespective of medical diagnosis. WHODAS II provides a profile of functioning across six activity domains, as well as a general disability score. The domains included in the instrument are:

- Understanding and communicating
- Getting around
- Self care
- Getting along with people
- Life activities
- Participation in society

The intra-class correlations within each domain on the 36-item questionnaire were good (0.71-0.92) and the six domains loaded on a general disability factor. A short 12-item form was derived and accounted for 85% of the variance of the full 36-item version (Andrews, 2008). All questions have proven psychometric qualities in terms of sensitivity, specificity, reliability, validity and cross-population comparability shown in general population surveys and in clinical sensitivity to change studies and cost-effectiveness studies (Chisolm, Abrams, McArdle, Wilson, & Doyle, 2005; World Health Organization, 2009). One recent study using a modified version reported good internal consistency for most scales. The total coefficient alpha score was reported as 0.86 (Buist-Bouwman et al., 2006). This was supported by a later study reporting the coefficient alpha ranged from 0.70 to 0.97 for the different subscales of WHODAS II (Pösl, Cieza, & Stucki, 2007).

The tool has been reported as ideal for epidemiology and for routine outcome measurement (Andrews, 2008). It was found to be particularly sensitive to change in anxiety symptom change (Perini, Slade, & Andrews, 2006).

In a small sample of patients with long-term psychotic disorders, the WHODAS II was experienced as relatively complex and at times difficult to administer with full co-operation in this clinical sample. The test-retest reliability was low. However, it gave valuable insights into patients' experiences, which should be acknowledged in order to assess treatment outcomes, as patients tended to report fewer activity limitations and impairments in mental functions than reported by clinicians. It was also useful in assessing restrictions in social functioning and environmental barriers (Chopra, Couper, & Herrman, 2004).

Client groups

Currently, the WHODAS II is available in eleven versions and eighteen languages. The instrument was cross-culturally developed and is applicable across the spectrum of cultural and educational backgrounds. The interviewer-administered version can be used where literacy is low.

The WHODAS II has been widely used with a range of different psychiatric populations (Chopra et al., 2004; Morrow-Howell et al., 2008; Mubarak & Barber, 2003; Norton, de Roquefeuil, Benjamins, Boulenger, & Mann, 2004; Perini et al., 2006; Pösl et al., 2007). Similarly, it has been used in a range of substance abusing populations (e.g. alcohol, opiate, cannabis users; both inpatient and outpatient services) (McEvoy & Shand, 2008; Philip, Greg, & Khelifa, 2007). However, some limitations exist in severely mentally ill clients (Chopra et al., 2004). The tool has also been found to be satisfactory for use in older populations (McKibbin, Patterson, & Jeste, 2004).

Availability/cost

It is necessary to register as a user to access the training manual, but the tool itself is on the WHO website at:

- http://www.who.int/icidh/whodas/index.html
- http://www.who.int/classifications/icf/whodasii/en/index.html

Scoring, administration and expertise required

The 36-item version is available in self-administered, interviewer-administered (recommended) and proxy (i.e. to relatives) versions. It is probably best utilised for the determination of status and outcome in an individual. The 12-item version can be self-administered or interviewer-administered. Administration time for the 12-item version is approximately 5 minutes and 15 minutes for the 36-item version (but will vary depending upon administration method).

World Health Organisation Quality of Life-BREF (WHOQoL-BREF)

The World Health Organisation Quality of Life-BREF (WHOQoL-BREF) instrument comprises 26 items, which measure the following broad domains: physical health, psychological health, social relationships, and environment (Murphy, Herrman, Hawthorne, Pinzone, & Evert, 2000). The WHOQoL-BREF is a shorter version of the original instrument that may be more convenient for use in large research studies or clinical trials. Domain scores produced by the WHOQoL-BREF have been shown to correlate at around 0.9 with the WHOQoL-100 domain scores.

The tool was devised for a range of uses including medical practice and research as an outcome and assessment tool (World Health Organization, 1994, 1995, 1998). Both the original WHOQOL-100 and the WHOQoL-BREF have been shown to display good discriminant validity, convergent and criterion validity, sensitivity to change and test-retest reliability (Castro, Oliveira, Miguel, & Araujo, 2007; Nelson & Lotfy, 1999; O'Carroll, Smith, Cousten, Cossar, & Hayes, 2000; World Health Organization, 1997).

In a study of 533 Dutch psychiatric patients, Trompenaars and colleagues (2005) found the WHOQoL-BREF to load on four factors (with internal consistency ranging from 0.66 to 0.80) and correlate at 0.92 with the original version of the WHOQoL. Content and construct validity was also strong, significant correlations were found between the Perceived Social Support Scale (PSSS) and subscales of the SCL-90-R and equivalent domain scores on the WHOQoL-BREF.

In a recent Australian study, Hawthorne and colleagues reported that the WHOQoL-BREF had good validity and reliability and particular utility as an outcome measure. Population norms and psychometric properties which aid interpretation were provided (Hawthorne, Herrman, & Murphy, 2006).

Client groups

The core WHOQoL instruments can assess quality of life in a variety of situations and population groups and were developed collaboratively in a number of centres worldwide. Since the initial development process, the WHOQoL instruments have also been translated and validated in a vast range of countries and cultures (Carpiniello, Pinna, Carta, & Orrù, 2006; Hao, Fang, & Power, 2006; Saxena, Carlson, Billington, & Orley, 2001; Skevington, Sartorius, & Amir, 2004).

The tools have been used successfully in adolescent (Chen, Wu, & Yao, 2006; Mugno, Ruta, D'Arrigo, & Mazzone, 2007) and older populations (Kalfoss, Low, & Molzahn, 2008; Peel, Bartlett, & Marshall, 2007). A specialised version for the older population has also been developed (Power, Quinn, & Schmidt, 2005)

The WHOQoL instruments have been used extensively with psychiatric populations (Berlim, McGirr, & Fleck, 2008; Hsu, Hwang, Lee, & Chen, 2006; Kauer-Sant'Anna et al., 2007; Masskulpan, Riewthong, Dajpratham, & Kuptniratsaikul, 2008; Stengler-Wenzke, Kroll, Riedel-Heller, Matschinger, & Angermeyer, 2007; Yen et al., 2008) and have also been used widely in substance misusing populations (Gunther,

Roick, Angermeyer, & Konig, 2008; Padaiga et al., 2007). In a sample of alcohol dependent males (da Silva Lima, Fleck, Pechansky, de Boni, & Sukop, 2005), the WHOQoL-BREF was found to have good internal consistency (Cronbach's alpha ranging from .78-.89) and criterion validity. The measure correlated significantly with both the SF-36 domains and SCL-90-R total scores. Furthermore, the WHOQoL-BREF discriminated between low/moderate dependent individuals and highly dependent alcohol abusers (as measured by the SADD).

Availability/cost

The appropriate language version of the WHOQoL-BREF, permission for using it and manuals can be obtained from:

- http://www.who.int/substance_abuse/research_tools/whoqolbref/en/

Scoring, administration and expertise required

The WHOQoL-BREF produces domain scores, but not individual facet scores, unlike the longer, original version. It is a self-report measure consisting of 26 items scored on a five point scale, which add to four separate domain scores.

General Mental Health Screening, Assessment and Outcome Measures

Kessler Psychological Distress Scale (K10; K6)

The Kessler psychological distress scale (K10) (Kessler, 1996) is a widely used, simple self-report measure of psychological distress which can be used to identify those in need of further assessment for anxiety and depression. This measure was designed for use in the general population to detect high-prevalence mental health disorders; however, it may also serve as a useful clinical tool, and scores may be an indicator of mental health disorders with lower population prevalence (e.g. schizophrenia) (Croton, 2007). In addition, the tool can also be used as an outcome measure. The K10 comprises 10 questions that are answered using a five-point scale (where 5 = all of the time, and 1 = none of the time). For all questions, the client circles the answer best describing them in the past four weeks. Scores are then summed with the maximum score of 50 indicating severe distress, and the minimum score of 10 indicating no distress. A guide to interpreting K10 scores is provided in Table 3.

Table 3. Severity of psychological distress according to K10 score

K10 score	Level of psychological distress
10-15	Low
16-21	Moderate
22-29	High
30-50	Very High

Adapted from Andrews and Slade(2001)

People who score 0-15 have one quarter the population risk of meeting criteria for an anxiety or depressive disorder as identified by the Composite International Diagnostic Interview (CIDI), and are unlikely to make a suicidal attempt in their lifetime. The moderate risk group (scores 16-30) have three times the population risk of having a current anxiety or depressive disorder and three times the population risk of ever having made a suicide attempt (Andrews, 2003b). Those who score 30 or above have ten times the population risk of meeting criteria for an anxiety or depressive disorder and 20 times the population risk of ever having made a suicide attempt (Andrews, 2003b; Andrews & Slade, 2001). Andrews and Slade (2001) found 68% of the 1997 National Survey of Mental Health and Well Being respondents fell into the first (low risk) group, while 3% scored 30 or above.

The brief questionnaire has been shown to have good construct and criterion validity (Kessler et al., 2002), being significantly associated with measures of mental health symptoms and disability as well as the frequency on consultations for a mental health problem in the previous 12 month period. The process of selecting and refining the potential items on the instrument was assisted by an expert advisory panel of survey researchers, which rated each potential item for clarity and wording. Only

those items that were consistently rated as clear were included in the pool of items from which the final 10 were chosen (Kessler et al., 2002). Scores on the K10 were also significantly associated with current CIDI diagnosis of anxiety and affective disorders, other mental disorder categories and the presence of any current mental disorder. The tool was also found to outperform other screening scales (e.g. GHQ-12) in discriminatory power in detecting DSM-IV anxiety and depressive disorders (Furukawa, Kessler, Slade, & Andrews, 2003). This finding was supported by a recent study which found the K10 and K6 to outperform a number of competing instruments (Gill, Butterworth, Rodgers, & Mackinnon, 2007).

The sensitivity and specificity scores are outlined in Table 4 and indicate a high level of both over the various K10 scores. Similarly strong ratios were found by Kessler and colleagues (2002; 2003), who also found the Cronbach's alpha for the K10 (a measure of internal consistency reliability) to be high (.93). Data from the Australian National Survey of Mental Health and Wellbeing (NSMHWB) yielded a Cronbach's alpha of 0.92 (Andrews & Slade, 2001).

Table 4. Sensitivity and specificity of the K10 in identifying people who met CIDI criteria for any current anxiety or affective disorder (prevalence 7.1%).

K10 Score	Hit Rate (Sensitivity)	Correct Rejection Rate (Specificity)
>=14	0.94	0.63
15	0.90	0.72
16	0.86	0.78
17	0.81	0.83
18	0.77	0.87
19	0.71	0.90
20	0.66	0.92
21	0.60	0.94
22	0.55	0.95
23	0.50	0.97
24	0.45	0.97
25	0.41	0.98
26	0.36	0.98
27	0.33	0.99
28	0.31	0.99
29	0.27	0.99
30	0.24	0.99
31	0.21	1.00
32	0.18	1.00

Andrews and Slade (2001)

Modified versions of the K10 exist and are useful for specific assessment circumstances. For instance, the K10-LM is intended for use mainly at initial assessment and contains the ten distress items plus four

supplementary questions, and uses a rating period of the "last four weeks". While the K10-L3D contains only the ten distress items and has a rating period of the "last three days". It is designed for use to assess distress during care or prior to discharge. A six-item version of this scale (the K6) is also available and has been found to be psychometrically sound.

Client groups

The simplicity of the K10 to administer and score makes it a versatile tool for a wide range of clients, services types and staff. Clients can complete the scale individually in less than ten minutes. The K10 has also been translated into many languages, including Arabic, Bosnian, Chinese, Croatian, Farsi, Greek, Hindi, Italian, Korean, Macedonian, Serbian, Spanish, Tagalog, Turkish and Vietnamese (Australian Government Department of Health and Ageing, 2002). The K10 and K6 have been validated with a number of different cultural groups (Baggaley et al., 2007; Furukawa et al., 2008). Furthermore no clinically significant biases were found on the K10 based on gender or education (Baillie, 2005).

In recent years the K10 has been used successfully in a range of populations. Kilkkinen and colleagues (2007) found the K10 to have good internal reliability (Cronbach's alpha = .87) across three rural Australian communities. Recent studies have also applied the K10 to drug and alcohol populations. In one sample of 103 injecting drug users in Victoria the K10 was found to have high internal consistency (Cronbach's alpha = .84). While a cut-off score of 27 was found to demonstrate the highest levels of sensitivity and specificity in detecting affective disorders in this population. The best trade-off across sensitivity and specificity correctly identified 78% of cases and 74% of non-cases (Hides, Lubman, Devlin et al., 2007). The K10 was found to significantly predict the presence of a current affective disorder (with individuals screening positive being 10 times more likely to have an affective disorder) and had an overall predictive accuracy of 76.7%. Similarly, in a sample of two public youth D&A services in Victoria, the K10's internal consistency was again high (α = .89; Hides, Lubman, Elkins, Catania, & Rogers, 2007). In older populations it has been argued that short screening tools (such as the K10) may give more accurate results than the CIDI anxiety and depressive questions as older people may have difficulty attending to lengthy, complex questionnaires (O'Connor & Parslow, 2008).

In a recent sample of 129 healthy pregnant women, the K10 showed agreeable sensitivity and specificity in detecting depression, posttraumatic stress disorder, panic disorder, and social phobia. The authors concluded that the K10 was a useful screening measure for mood and anxiety disorders in pregnant women (Spies et al., 2009).

Availability/cost

The K10 is in the public domain and therefore freely available with acknowledgment of the source. It can be found at:

- http://www.crufad.com/K10/printk10.html

- http://www.beyondblue.org.au/index.aspx?link_id=89.678
- http://www.sswahs.nsw.gov.au/MHealth/mhoat/kessler10_lm/Module_SR1_v2.pdf
- http://www.sswahs.nsw.gov.au/MHealth/mhoat/kessler10_l3d/Module_SR2_v2.pdf

Scoring, administration and expertise required

As previously mentioned scoring and administration of the K10 is simple. Clients can self complete or it can be interviewer-administered for those with poor literacy. Despite the tool's strengths (e.g. easy to administer/score, brief, accessibility) the K10 is limited by its non-specific content (i.e., a measure of distress only) and clinical judgement is required for further client treatment. Therefore, although no expertise is required in the initial administration or scoring, workers may require the skills or resources to appropriately interpret scores above the low end on the scale.

The General Health Questionnaire (GHQ)

The General Health Questionnaire (GHQ) (Goldberg & Williams, 1988) is a self-report screening instrument which detects the presence of psychological symptoms. It was originally a 60-item measure on which a client rates each statement on a four-point scale, however, shorter (12- and 28-item) forms have demonstrated adequate reliability and validity and are currently used (Goldberg & Williams, 1988; Politi, Piccinelli, & Wilkinson, 1994). The GHQ-28 provides four specific subscales: somatic symptoms, anxiety and insomnia, social dysfunction, and severe depression. In recent years the 12 item GHQ-12 has also been used extensively.

The GHQ has reasonable test-retest reliability in the range of .85 to .90 in clinical populations (Goldberg & Williams, 1988) but tends to be substantially lower in the general population (Layton, 1986). The internal reliability of the GHQ-12 (the most commonly used) has been found to be strong (Cronbach's alphas ranging from .82 to .86) (Politi et al., 1994; Sriram, Chandrashekar, Isaac, & Shanmugham, 1989; Winefield, Goldney, Winefield, & Tiggemann, 1989).

Content and construct validity of the tool is also generally good and GHQ scores have been found to correlate well with standardised psychiatric assessment, the median correlation was .76 for the GHQ-28 and .70 for the GHQ-12 in studies examining this property (Dawe, Loxton, Hides, Kavanagh, & Mattick, 2002). The GHQ has been found to have good sensitivity and specificity, with a review finding the GHQ-12 had an average sensitivity of 89% and specificity of 80%; the GHQ-28 had an average sensitivity of 84% and specificity of 82% and the GHQ-60 had a sensitivity of 78% and a specificity of 85% (Goldberg & Williams, 1988). However, in substance misuse populations this tends to be lower (Ross & Glaser, 1989). Furthermore, the GHQ has been found to correlate well with other similar measures (e.g., $r \geq 0.8$ when compared to the SCL-90-R) (Vallejo, Mañanes, Comeche, & Díaz, 2008).

In an extensive, international comparison of the validity of the GHQ-28 and GHQ-12, the GHQ-12 was found to have the ability to accurately detect CIDI diagnoses, with an overall sensitivity of 83.4% and specificity of 76.3% (a score of 1-2 was optimal cut-off for most sites) (Goldberg et al., 1997). Similarly, the GHQ-28 exhibited an overall high level of sensitivity (79.7%) and specificity (79.2%). However, these were no higher than the GHQ-12, although there tended to be less variation in optimal cut-off scores across centres. The use of simple versus complex scoring methods, translation of the instruments, use in developing countries, and gender, age and educational variables had no significant effect on the sensitivity or specificity of the instruments (Goldberg et al., 1997). These findings (related to mental health disorder detection) was supported by a study of 572 German primary care outpatients (Schmitz, Kruse, Heckrath, Alberti, & Trees, 1999). It has also been suggested the removal of 20 non-significant individual items from the GHQ-28 may slightly increase the predictive power of the tool from 81.6% to 84.4%, but nonetheless showed good predictive validity (Willmott, Boardman, Henshaw, & Jones, 2008).

Well-being is an important determinant of health and social outcomes. Measures of positive mental health states are needed for population-based research and a number have been included in this

review. There is also evidence to suggest that the GHQ-12 measures not only negative aspects of mental health (e.g., symptoms) but also positive aspects. These areas do correlate but are also said to be independent dimensions adding further utility to the measure (Hu, Stewart-Brown, Twigg, & Weich, 2007).

Client groups

The GHQ has been translated into almost 50 languages and has been used in diverse cultural groups and the items appear to have cross-cultural relevance despite cultural variations in the expression of mental illness (Dawe et al., 2002). Indeed, a recent body of work has focused on the psychometric properties of the GHQ-12 and GHQ-28 across a range of cultural groups (e.g., elderly Iranians, Turkish primary care attendees, Mexican, Japanese, and Namibian adults, German and Chinese primary care patients, ethnic Indians living in the UK) (Caraveo Anduaga, Martinez, Saldivar, Lopez, & Saltijeral, 1998; Doi & Minowa, 2003; Haidula, Shino, Plattner, & Feinstein, 2003; Jacob, Bhugra, & Mann, 1997; Kilic et al., 1997; Malakouti, Fatollahi, Mirabzadeh, & Zandi, 2007; Pan & Goldberg, 1990; Schmitz, Kruse, & Tress, 1999) and have found the structure of the GHQ to be robust. Recent reports have provided support for the validity of the GHQ-12 with young adolescents and suggest that adolescents interpret the GHQ-12 in a similar manner to adults (French & Tait, 2004; Tait, French, & Hulse, 2003). However, there have been some concerns raised about the tool's use in antenatal populations (Swallow, Lindow, Masson, & Hay, 2003), and postnatal population findings are mixed (Ip & Martin, 2006; Navarro et al., 2007).

It has been suggested that the GHQ would be relevant to Indigenous Australians and there is some preliminary, related evidence to support this view (Hunter, 1993), however, further research is required.

Availability/cost

The GHQ are copyrighted instruments, however, the GHQ-28 has been incorporated into the OTI and can be used freely via this tool. Otherwise the tools are available for purchase from:

- http://shop.acer.edu.au/acer-shop/group/SD;jsessionid=5468B3660925CF57CFA34E70A2353C16

Scoring, administration and expertise required

The GHQ is easy to administer and score and can be used by a range of health professionals. It takes 5-15 minutes to administer depending on which version is used. It was specifically developed for use by a broad range of clinicians working in community and non-psychiatric settings, and requires no special training. Recent studies have found both computerised and internet versions of this tool to have comparative psychometric properties to the pen and paper version (Vallejo et al., 2008; Wijndaele et al., 2007). Generally a score of 10 or more on the GHQ is considered indicative of significant psychological distress and the presence of an underlying psychological disorder. However, it has been suggested that

approximately 75% of drug users could be expected to obtain scores of 10 or more upon entering treatment. This suggests that high levels of psychological distress may be typical among drug users who enter treatment. Therefore, clients need to be reassessed four weeks after entering treatment (Darke, 2008; personal communication). Findings from prison populations suggest this to be true of this population group as well (Andersen, Sestoft, Lillebæk, Gabrielsen, & Hemmingsen, 2002). If the client continues to score 10 or more, a more in-depth psychological assessment should be conducted.

The Symptom Checklist-90-Revised (SCL-90-R)

The Symptom Checklist-90-Revised (SCL-90-R) (Derogatis, 1994) is a 90-item self-report questionnaire measuring symptoms of somatisation, obsessive compulsive, interpersonal sensitivity, depression, anxiety, hostility, phobic anxiety, paranoid ideation, and psychoticism. It has also been found to be a useful tool in measuring treatment outcome (Ronan, Dreer, & Dollard, 2000).

The Global Severity Index (GSI) is a composite score obtained by summing the scores on the nine symptom dimensions and dividing this score by the total number of items (usually 90 if there are no missing responses). According to Derogatis (1994) the GSI is the best single indicator of severity of disorder and should be used in most instances where a single summary measure is required. The scale provides scores for severity, intensity and extensiveness of symptoms and has been shown to have superior sensitivity to competing scales (Mattick, Oliphant, Bell, & Hall, 1996). Shorter forms of the SCL-90-R have been developed, including the Brief Symptom Inventory (BSI) with 53 items and the Symptom Assessment (SA-45), each of which show adequate reliability and validity. Correlations of the BSI dimensions with the SCL-90-R range from .92 to .99. The BSI has demonstrated adequate internal consistency ranging from .70 (phobic anxiety) to .89 (depression) and test-retest reliability ranging from .68 (somatisation) to .91 (phobic anxiety) over a two-week period (Boulet & Boss, 1991; Derogatis & Meilisaratos, 1983). The SA-45 (Davison et al., 1997) consists of nine 5-item scales assessing the same symptom domains as the SCL-90. The majority of scales display adequate internal consistency reliabilities (.7-.8) across different age and patient status variables. It has also demonstrated discriminant validity between controls and adolescent and adult patients and between patients at intake and follow up.

The majority of factor analytic studies, indicate that both the SCL-90-R and the BSI tap a single predominant factor reflecting general psychological distress (Bonynge, 1993; Rauter, Leonard, & Swett, 1996). The SCL-90-R has shown good test-retest reliability as well as very good internal reliability (alpha coefficients ranging from .79 to .90 for the different dimension scales) in clinical and non-clinical population (Horowitz, Rosenberg, Baer, Ureno, & Villasenor, 1988).

The SCL-90-R has also been found to perform at least as well (if not better than) a range of other instruments in both assessment and in the measurement of change following treatment, suggesting validity and outcome utility are good (Derogatis, 1994).

The SCL-90-R was found to have good overall accuracy (and sensitivity/specificity) at optimum GSI cut-off points of .68 (sensitivity 72%, specificity 87%) in a diabetic sample and 1.00 (sensitivity 77%, specificity 91%) in a bulimic sample (Peveler & Fairburn, 1990). The SCL-90-R has been compared with the GHQ-12 and both performed equally well at detecting psychopathology (Schmitz, Kruse, Heckrath et al., 1999). Similarly, in a study of substance abusing clients the SCL-90-R was found to be a better predictor of CIDI diagnosed anxiety and mood disorders than the ASI-psychiatric problems scale, displaying a moderate degree of specificity and high sensitivity (Franken & Hendriks, 1999).

Client groups

The SCL-90-R is suitable for both male and female respondents and adolescent populations (Hart, Bryer, & Martines, 1991) and has been translated into a number of languages and subsequently has been validated in a range of diverse populations (e.g., Hispanic, Finish, Dutch, Maori, German populations) (Arrindell et al., 2003; Barker-Collo, 2003; Holi, Marttunen, & Aalberg, 2003; Martinez, Stillerman, & Waldo, 2005). There is a lack of research using the tool in the Australian indigenous population but it has been suggested (based on related findings) that it would also be able to provide helpful diagnostic information (Dawe et al., 2002; Hunter, 1993). A Year 8 reading age is required.

Some studies, however, have advised that its use in inpatient and outpatient samples has demonstrated a more variable factor structure, and lower specificity/sensitivity than in normative populations (Brophy, Norvell, & Kiluk, 1988; Cyr, McKenna Foley, & Peacock, 1985; Pariente, Lépine, Boulenger, Zarifian, & et al., 1989; Rauter et al., 1996; Starcevic, Bogojevic, & Marinkovic, 2000). Nevertheless, the tool has been used successfully in substance abusing populations (Franken & Hendriks, 2001).

Availability/cost

The SCL-90-R is a copyrighted instrument and therefore cannot be reproduced. It is published by NCA Assessments and distributed in Australia by Psychological Assessments Australia. All test purchasers must be Registered Psychologists with post-graduate qualifications in Psychology.

Scoring, administration and expertise required

The SCL-90-R is designed for adolescents over the age of 13 years and for adults. There is both a pen and paper and computerised version of the SCL-90-R. The former takes 15-25 minutes to complete. A recent study has found an internet version of this tool to have comparative psychometric properties to the pen and paper version (Vallejo et al., 2008). However, the measure is complex to score and requires special qualifications outlined above.

The Brief Psychiatric Rating Scale (BPRS)

The Brief Psychiatric Rating Scale (BPRS) (Overall & Gorham, 1962) is a clinician administered scale measuring a broad range of psychiatric symptoms. It was initially devised as an instrument to assess the symptoms of schizophrenia on five subscales of thought disorder, withdrawal, anxiety/depression, hostility and activity (Hedlund & Vieweg, 1980; Overall & Gorham, 1962). Factor structure has been fairly consistent across a range of studies (Burger, Calsyn, Morse, Klinkenberg, & Trusty, 1997; Dingemans, Linszen, Lenior, & Smeets, 1995; Long & Brekke, 1999; Morlan & Tan, 1998; Mueser, Curran, & McHugo, 1997; Ownby & Seibel, 1994).

The BPRS has been shown to be effective in various substance use populations (Steer & Schut, 1979; Westermeyer, Tucker, & Nugent, 1995). It has also shown adequate validity (Lykke, Hesse, Austin, & Oestrich, 2008; Morlan & Tan, 1998) and has been found to be an effective measure both of psychopathology and of treatment-related symptom changes (Burlingame et al., 2006; Hedlund & Vieweg, 1980; Varner, Chen, Swann, & Moeller, 2000). Inter-rater reliability for the BPRS has had mixed results but recent studies have tended to find it to be satisfactory (Flemenbach & Zimmerman, 1973; Hafkenscheid, 1991; Ligon & Thyer, 2000; Mueser et al., 1997).

Suggested cut-off scores for the BPRS are included in Table 5 below.

Table 5. BPRS cut-off points

BPRS Cut-off score	Level of psychological distress
0-30	No notable illness
31-40	Minimally ill
41-53	Moderately ill
53+	Markedly ill

Adapted from Leucht and colleagues (2005)

Client groups

The attitudes and beliefs held by the clinician will have some bearing on the rating of particular diagnostic categories and although this is unlikely to be a major issue, in some ethnic or cultural populations it may impact upon clinical judgment. However, as this scale is clinician rated, it does not require the client to read or write, which may be helpful when assessing individuals who are either illiterate or who are unable to read English. It is also suitable for both male and female respondents.

The BPRS has also been used in various research studies in which the relationship between major psychiatric illness and substance misuse has been investigated (Dawe et al., 2002; Lykke et al., 2008). Similarly, as mentioned above, it has been used successfully with individuals with opiate and alcohol disorders (Steer & Schut, 1979; Westermeyer et al., 1995). It has been successfully used on homeless persons as well as elderly and incarcerated populations (Burger, Yonker, Calsyn, Morse, & Klinkenberg,

2003; Greenwood & Burt, 2001; Ownby & Seibel, 1994; Panos, 2004). A children's version has also been developed (Hughes, Rintelmann, Emslie, Lopez, & MacCabe, 2001; Lachar et al., 2001).

The BPRS has also been found to demonstrate significant change during brief stays of 1 week or less in acute inpatient care. Therefore, the BPRS may be an appropriate outcome for aiding clinical decisions regarding suitability for discharge (Varner et al., 2000).

Availability/cost

The BPRS has been reproduced online at:

- http://priory.com/psych/bprs.htm
- www.cnsforum.com/streamfile.aspx?filename=BPRS.pdf&path=pdf

Scoring, administration and expertise required

The BPRS is a clinician rated instrument. Two main versions exist, the original 18-item version and the more recent 24-item version. Ratings are made after a brief (15-20 minutes) unstructured interview with the patient. Each item is rated on a 7-point scale ranging from "not present" to "extremely severe". The tool is limited because it can only be used by adequately trained mental health professionals (e.g., clinical psychologists or psychiatrists), although specific training is available to clinicians with a background in psychology or psychiatric nursing.

PsyCheck

The Australian PsyCheck screening tool was recently developed and has been demonstrated to be a scientifically valid and useful resource for clinicians (Lee et al., 2007). The screening tool has three sections:

- A General Mental Health Screen, including history of treatment.
- Suicide/Self-Harm Risk Assessment.
- The Self Reporting Questionnaire (SRQ) (Beusenberg & Orley, 1994), that assesses current symptoms of depression and anxiety.

PsyCheck screens for the presence of mental health symptoms that may be addressed within specialist services, suicide risk, psychosis and mental health history are also assessed. Scores on the PsyCheck General Mental Health Screen were compared with scores on the diagnostic interview (CIDI-auto) and was found to significantly correlate with the presence of an affective or anxiety disorder in the previous month and was also predictive of a current affective or anxiety disorder within the last month. These findings suggest that this section of the PsyCheck screening tool could be used as an initial indicator of mental health status (Lee et al., 2007). The SRQ was compared to the commonly used the GHQ and both were validated against the CIDI-Auto (WHO, 1997). The SRQ showed good predictive ability, and was superior in terms of sensitivity (0.81) and specificity (0.84) when compared with the GHQ (Lee et al., 2007).

Client groups

PsyCheck has only recently been developed and evidence is lacking on its applicability to a diverse range of populations. However, the SRQ (Beusenberg & Orley, 1994), was developed by the World Health Organisation (WHO) as a screening tool for common mental disorders in primary health care settings and has strong psychometric properties and has been adapted into a range of culturally specific versions. Cut-off points vary considerably depending on setting (community, primary care, hospital) and culture but with a cut off of 5-6, sensitivity and specificity have been found to be consistently between 70-80%. In some cultures the cut-off needs to be raised, sometimes as high as 13 (Alsuwaida & Alwahhabi, 2006). Evidence for content validity is limited as it is merely a screening tool for mental distress and therefore cannot measure specific disorders. But criterion validity and reliability is generally good (Beusenberg & Orley, 1994).

Availability/cost

PsyCheck is in the public domain and therefore can be used without cost but simply acknowledgement of its source. The tool itself, along with further information, training and supervision, treatment and implementation guidelines can be found at:

- www.psycheck.org.au

Scoring, administration and expertise required

The PsyCheck manual (Lee et al., 2007) includes training on how to administer, score and interpret the results of each section, and the subsequent steps to take according to the screening results. If the results of the screening tool indicate a high presence of symptomatology, further assessment may be warranted.

The instrument was designed for use by non-mental health specialists but cannot be self-administered by the client. A degree of training is required for use of the tool. The PsyCheck Training and Clinical Supervision Guidelines consist of four modules, which are self-contained and delivered in four separate sessions via a PowerPoint Training Presentation (approximately one day). Trainer training is provided to clinicians/educators through a two day course. An advantage of the intensive training is that workers are also trained in management strategies for addressing mental health issues.

The Depression Anxiety Stress Scales (DASS)

The Depression Anxiety Stress Scale (DASS) (Lovibond & Lovibond, 1995) has been shown to be a valid and reliable measure of the dimensions of depression, anxiety and stress separately but also tap into a more general dimension of psychological distress (Henry & Crawford, 2005). The DASS is available in two forms: the DASS-21 and the DASS-42. The use of either test in isolation is sufficient in screening for symptoms.

Early findings reported the DASS-42 had excellent internal consistency and test-retest reliability and can distinguish between features of depression, physical arousal, and psychological tension and agitation better than other existing measures (Brown, Chorpita, Korotitsch, & Barlow, 1997; Lovibond & Lovibond, 1995). A later study successfully confirmed these finding and extended it in order to apply to the DASS-21 (Antony, Bieling, Cox, Enns, & Swinson, 1998). Cronbach's alphas for the DASS-42 Depression, Anxiety, and Stress subscales were .97, .92, and .95, respectively. Similarly, Cronbach's alphas for the DASS-21 subscales were .94 for Depression, .87 for Anxiety, and .91 for Stress. The relevant DASS-42 and DASS-21 scales both correlated highly with the BDI, BAI, and the STAI-T (Antony et al., 1998). In a more recent study the reliability of the DASS was excellent, and the measure was again found to possess adequate convergent and discriminant validity (Crawford & Henry, 2003). Furthermore, the DASS-21 has also been found to be a valid routine clinical outcome measure in the private in-patient setting (Ng et al., 2007).

Client groups

The DASS has been validated in a range of cultures including Turkey, Iran, Spain and South America and the Netherlands (Akin & Çetin, 2007; Daza, Novy, Stanley, & Averill, 2002; de Beurs, Van Dyck, Marquenie, Lange, & Blonk, 2001; Sahebi, Asghari, & Salari, 2004). It has also been used in D&A treatment settings (Morley et al., 2006; Sannibale et al., 2005)

The DASS has also been used successfully in adolescents (Duffy, Cunningham, & Moore, 2005) and the structure of the DASS in older samples is consistent with younger age groups, demonstrating good internal consistency, excellent convergent validity, and good discriminative validity (Gloster et al., 2008). It has also been used among Aboriginal and Torres Strait Islanders in research studies (Schlesinger et al., 2007) but its applicability to these populations has not been sufficiently tested.

Availability/cost

The DASS is in the public domain and therefore can be used without cost but with acknowledgement of its source. The tool itself, along with further information and the manual can be found at:

- http://www2.psy.unsw.edu.au/groups/dass/

Scoring, administration and expertise required

The DASS is a self-report instrument, and no special skills are required to administer or score it. However, decisions based on particular score profiles should be made only by experienced clinicians who have carried out an appropriate clinical examination (Lovibond & Lovibond, 1995). The two forms have 21 and 42 items respectively and are each rated on a four-point scale of how much each particular statement applies to the individual. A guide to interpreting DASS scores is provided in Table 6.

Table 6. Interpreting DASS scale scores

DASS scale score	Level of psychological distress
0 – 78	Normal
78 – 87	Mild
87 – 95	Moderate
95 – 98	Severe
98-100	Extremely severe

Adapted from Lovibond and Lovibond (1995)

Mental Health Inventory (MHI)

The Mental Health Inventory (MHI) (Veit & Ware, 1983) was adapted from another mental health instrument, the General Wellbeing Schedule (GWB), and was designed to assess general psychological distress and wellbeing in a non-patient population. The MHI has been found to have some utility as an outcome measure, as consumers who rated themselves as having improved have been found to show corresponding improvements on the majority of MHI subscales, while those who rated themselves as stable have been found to show no change on the MHI, and those who rated themselves as having deteriorated have been found to show a decline on the MHI (Stedman et al., 1997).

Generally the test-retest reliabilities reported have been adequately high for subscale, global scale and total scores on the MHI (correlations ranging from 0.56 to 0.97) (Downe-Wamboldt & Melanson, 1995; Stedman et al., 1997; Veit & Ware, 1983). While other studies have found the inter-rater reliability to be generally good (although mixed for some subscales), significant correlations between the self-report and interviewer-administered versions of the MHI also strengthen reliability (Pomeroy, Clark, & Philp, 2001).

A number of studies have found the internal consistency of the MHI to be good (Cronbach's alpha has ranged from 0.63 to 0.93 for the subscales, 0.90-0.97 for the global scales and 0.93 to 0.97 for the total score) (Brunier, Graydon, Rothman, Sherman, & Liadsky, 2002; Cohen, Hack, de Moor, Katz, & Goss, 2000; Downe-Wamboldt & Melanson, 1995; Florian & Drory, 1990; Kravetz, Faust, & Shalit, 2001; Ostroff, Woolverton, Berry, & Lesko, 1996; Stedman et al., 1997; Veit & Ware, 1983).

Generally the content validity of the MHI is good, however, concerns have been raised with the validity of specific items in certain population groups (e.g. youth) (Huebeck & Neill, 2000; Pirkis et al., 2005). While results from a Health Insurance Study (Ware, Manning, Duan, Wells, & Newhouse, 1984) into service use have shown support for the predictive validity of the instrument. This study indicated that persons with low total scores on the MHI (i.e., high psychological distress and low psychological wellbeing) were more likely to receive mental health care than those with higher scores.

Scores on the MHI have been found to correlate well with comparable constructs on other standardised measures suggesting the measure is concurrently valid. For instance, Manne and Schnoll (2001) found the scores to correlate with those on the Positive and Negative Affect Schedule (PANAS), and the Dyadic Adjustment Scale. While Zika and Chamberlain (1992) found correlations between scores on the Purpose in Life Test (PIL), the Life Regard Index (LRI) and the Sense of Coherence Scale (SOC), and the MHI. Correlations with the GAS have also been observed (Wichstrom, Anderson, Holte, & Wynne, 1996).

Stedman and colleagues (1997) found the MHI to correlate well with the SF-36 and the Behaviour and Symptom Identification Scale (BASIS-32), but less well with the LSP, Role Functioning Scale (RFS), and Health of the Nation outcome scales (HoNOS).

Kornblith and colleagues (2001) found the MHI correlated well with the Medical Outcome Study Social Support Survey (MOS-SSS), the Life Experience Survey (LES), and the Systems of Belief Inventory (SBI), furthermore, the authors found the MHI could discriminate between those who have experienced stressful life events and those who have not, those with low level of social support and those with good social networks, and those with poor physical health and those with no medical problems (Kornblith et al., 2001). This discriminant validity was replicated in similar studies (Cohen et al., 2000; McCabe, Thomas, Brazier, & Coleman, 1996; Ware et al., 1984; Wyshak, 2001). Furthermore, MHI scores have been shown to relate to mental health service use (McCabe et al., 1996) and to discriminate between certain clinical and non-clinical samples (Cassileth, Lusk, & Strouse, 1984; Rosenthal, Downs, Arheart, Deal, & Rosenthal, 1991; Smith, Egert, Winkel, & Jacobson, 2002)

Client groups

The MHI has been translated into several other languages and validated in a number of different cultures and service settings (Florian & Drory, 1990; Hirini, Flett, Long, & Millar, 2005; Liang, Wu, Krause, Chiang, & Wu, 1992; Wu, Xu, & Li, 2002). The instrument has also been employed in a wide range of studies, particularly in the medical field. For example, it has been used to assess the mental health of consumers with cancer and human immunodeficiency virus (HIV) (Cohen et al., 2000; Kornblith et al., 2001; Manne & Schnoll, 2001; Siegel, Gluhoski, & Karus, 1997; Smith et al., 2002).

It has also been used as a screening tool for depression in older people and in rural communities (Dorfman et al., 1995; Kennedy & Yellowlees, 2003; Pomeroy et al., 2001) and as a general measure of mental illness for primary care consumers (Berwick et al., 1991; Weinstein, Berwick, Goldman, Murphy, & Barsky, 1989).

Availability/cost

The MHI is freely available at:

- http://www.rand.org/health/surveys_tools/mos/mos_mentalhealth.html

Scoring, administration and expertise required

The MHI comprises of 38 items and can be completed either as a self-report measure or as part of a structured-interview; either way it takes approximately 10-15 minutes to complete. Each item includes a description of a particular state of mind which is scored on either a five- or six-point scale. Item scores are summed to give six subscale scores; global scale scores and a total score.

Stedman and colleagues (2000) found that consumers rated the MHI higher than either the BASIS-32 or the SF-36 in terms of perceived relevance, effectiveness and usefulness. Overall the MHI was seen as comprehensive, easy to understand, user-friendly, acceptable and appropriate. However, the tool was

criticised for certain aspects that restricted its utility (e.g., particular wording of items, response options and applicability to certain populations) (Stedman et al., 2000).

The MHI elicited the most positive responses when compared to other measures in a sample of consumers (Stedman et al., 2000). A much shorter version (MHI-5) has also recently been developed and has shown adequate psychometric properties (Berwick et al., 1991).

Behaviour and Symptom Identification Scale 32 (BASIS-32®)

The Behaviour and Symptom Identification Scale 32 (BASIS-32®) (Eisen, Dill, & Grob, 1994; Eisen, Grob, & Klein, 1986) was developed as a 32-item, consumer-oriented, self-report measure of symptoms and behavioural distress. The scale consists of five subscales measuring Impulsive and addictive behaviour, Psychosis, Relation to self and others, Depression and anxiety and Daily living and role functioning. The instrument was originally developed and validated among inpatients, but subsequent studies have supported its use in outpatient and residential settings (Eisen, Leff, & Schaefer, 1999; Eisen, Wilcox, Leff, Schaefer, & Culhane, 1999; Hoffmann, Capelli, & Mastrianni, 1997; Klinkenberg, Cho, & Vieweg, 1998; Russo et al., 1997). It is widely used as a measure of mental health outcomes (Dickey et al., 2003; Doerfler, Addis, & Moran, 2002; Hawthorne, Green, Lohr, Hough, & Smith, 1999; Uttaro & Gonzalez, 2002) and has been found to be highly sensitive to change following treatment in both inpatient and outpatient samples (Eisen et al., 1994; Eisen, Grob, & D.L., 1989; Eisen, Leff et al., 1999; Eisen, Wilcox et al., 1999; Eisen, Wilcox, & Schaefer, 1997; Hoffmann et al., 1997; Jerrell, 2005; Klinkenberg et al., 1998; Russo et al., 1997; Sederer, 1992; Stedman et al., 1997; Trauer & Tobias, 2004). However, the Impulsive and addictive behaviours subscale and the Psychosis subscale showed poorer sensitivity to change than the other three subscales (Russo et al., 1997; Trauer & Tobias, 2004).

The BASIS-32® has been found to show consistently good test-retest reliability (Eisen et al., 1994; Eisen et al., 1986; Klinkenberg et al., 1998). Similarly, when the self-report version of the instrument is compared with the interview version, or with reports of a close informant, good overall inter-rater reliability has been found (Eisen et al., 1986; Klinkenberg et al., 1998).

There is some concern about the internal structure of the measure. The developers themselves have noted that there is 'unnecessary redundancy in the instrument', and as a result a shorter version (BASIS-24®) has been developed (Eisen, Normand, Belanger, Spiro, & Esch, 2004). Nevertheless, the internal consistency of the BASIS-32®, subscales has generally been found to be high ranging from 0.6 to 0.9, across settings (inpatients and outpatients) and types of administration (structured-interview and self-administration) (Chow, Snowden, & McConnell, 2001; Eisen, Leff et al., 1999; Eisen, Wilcox et al., 1999; Hoffmann et al., 1997; Klinkenberg et al., 1998; Russo et al., 1997; Stedman et al., 1997; Uttaro & Gonzalez, 2002).

In assessing the content validity of the BASIS-32®, Eisen and colleagues (1999; 1997) reported that the instrument was comprehensive, however, participants in the other reports had mixed responses. Concerns identified in the work by other studies related to ambiguous and complex language, an exclusive focus on difficulties and issues with content areas. Recommendations from consumers included the addition of items to cover the outcome domains of greater relevance to them (Cameron et al., 2001; Graham et al., 2001).

Numerous studies have considered the concurrent validity of the BASIS-32® and the tool has been found to correlate well with a range of comparable scales including: the Client Assessment of Strengths

Interests and Goals (Lecomte, Wallace, Caron, Perreault, & Lecomte, 2004; Wallace, Lecomte, Wilde, & Liberman, 2001) the Outcome Questionnaire (Doerfler et al., 2002), the SF-36 (Eisen, Wilcox et al., 1999; Eisen et al., 1997; Stedman et al., 1997), the SCL-90 (Eisen et al., 1986; Hoffmann et al., 1997), the Child and Adolescent Functional Assessment Scale (CAFAS) (Hoffmann et al., 1997) the Hopkins Symptom Checklist and the BPRS (Klinkenberg et al., 1998), the MHI, and the HoNOS (Stedman et al., 1997), the CAN (Trauer & Tobias, 2004), the Psychiatric Symptom Assessment Scale (PSAS), Lehman's Quality of Life Interview (Russo et al., 1997) and the Outcome Assessment Program (OAP) Questionnaire (Dornelas, Botticello, Goethe, & Fischer, 2001). However, scores on the BASIS-32®, have been found to be unrelated to performance on the Social and Occupational Functioning Assessment Scale and the MMSE (Russo et al., 1997), the Levels of Recovery from Psychotic Disorders Scale (Sousa & Frazier, 2004), the Global Assessment of Relational Functioning Scale (Wilkins & White, 2001), the LSP and the RFS (Stedman et al., 1997).

The BASIS-32® has been shown to correlate well with objective indicators of functioning, effectively discriminating between consumers who were rehospitalised (Eisen et al., 1994; Eisen et al., 1986; Hoffmann et al., 1997) and those who were currently employed or undertaking studies (Eisen et al., 1994; Eisen et al., 1986) and between inpatients and outpatients (Eisen, Wilcox et al., 1999). Furthermore, consumers with depression, psychotic disorders and substance abuse problems have been shown to score highly on the Depression and anxiety, Psychosis and Impulsive and addictive behaviours subscales, respectively (Doerfler et al., 2002; Eisen, 1992; Eisen, Leff et al., 1999; Eisen, Wilcox et al., 1999; Hoffmann et al., 1997; Klinkenberg et al., 1998; Russo et al., 1997). Likewise, consumers with comorbid mental health and substance abuse problems have been shown to score significantly higher on all five subscales than those with fewer disorders of less complexity (Johnson, Brems, & Burke, 2002; Pollack & Cramer, 2000; Pollack, Cramer, & Varner, 2000). Other studies, however, have found no link between BASIS-32® score and diagnosis (Doerfler et al., 2002; Klinkenberg et al., 1998).

Finally, predictive studies have observed that, when assessed at discharge, consumers who subsequently require rehospitalisation score higher than those who do not relapse (Eisen et al., 1986; Eisen et al., 1997; Hoffmann et al., 1997).

Client groups

An Australian version of the BASIS-32® exists (Eisen et al., 2004) and the tool has also been translated into many languages including Spanish, French, Japanese, Chinese, Korean, Cambodian, Vietnamese and Tagalog (Eisen & Culhane, 1999). Some authors have criticised the BASIS-32®, claiming the reading level required is too high for individuals with limited literacy skills, and that it is difficult to complete for individuals who are acutely psychotic or intoxicated, or have dementia (Eisen et al., 2004; Hawthorne et al., 1999; Hoffmann et al., 1997; Russo et al., 1997). However, few tools are consistently effective for use in such populations. Nevertheless, overcoming these concerns was one of the reasons for the development of the BASIS-24® (Eisen et al., 2004).

The BASIS-32® has also been widely used in both younger and older populations, along with D&A/comorbid samples and homeless people (Gamst et al., 2006; Johnson, Brems, Mills, & Freemon, 2005; Oslin et al., 2005; Smelson et al., 2007).

Availability/cost

The BASIS-32® is a commercial instrument and is not available in the public domain. Copyright is held by the McLean Hospital, and there is an annual fee and licence requirement.

- http://basissurvey.org/basis32/

Scoring, administration and expertise required

The BASIS-32® (BASIS-24®) is a consumer-rated instrument comprising 32 (24) items. Each item is rated using a five-point scale (0 = no difficulty; 1 = a little difficulty; 2 = moderate difficulty; 3 = quite a bit of difficulty; 4; extreme difficulty) which are used to calculate subscales and total scores by adding the ratings for each item and dividing by the number of non-omitted items. The exception to this rule is the computation of the Daily living and role functioning subscale, which is determined by an averaging procedure (Victorian Department of Human Services, 2003).

Ratings are based on the difficulty experienced during the preceding period (one-two weeks), ascertained by structured interview (either with a rater present or by telephone) or self-report format (either on-site or through mail-out). The structured interview is generally administered if the consumer is not capable of self-report (e.g. due to illiteracy or an excess of symptoms) in which the rating scale choices are presented to the consumer on large index cards. The interviewer reads each item to the consumer who is then required to indicate their response on the cards (Eisen, 1995). In its interviewer-administered form, the BASIS-32® may be used by both professional and nonprofessional staff due to the absence of training requirements. An instruction manual is available which contains a survey form, a scoring algorithm, a reference list and articles relating to methodology, reliability and validity.

The BASIS-32® takes on average 5-10 minutes to administer in the self-report format, while the structured interview takes approximately 15-10 minutes to complete. Stedman and colleagues (1997), found that the BASIS-32® was rated favourably by consumers in terms of its utility (defined in terms of its perceived relevance, effectiveness and usefulness), although it was not ranked as highly as the MHI. Others have found it to be applicable to a wide range of people receiving mental health treatment, not limited by diagnoses, symptom patterns or treatment setting, user-friendly and adaptable due to its alternative administration modes. It has also been found to place minimal burden on staff due to its brevity, the simplicity of its design and its absence of training requirements (Doerfler et al., 2002; Eisen, Leff et al., 1999; Eisen et al., 1997; Klinkenberg et al., 1998).

Specific Mental Health Screening, Assessment and Outcome Measures

The following tools have been included not as part of general screening/assessment given routinely to all clients but rather, on the basis of high scores on general measures or identifiable symptoms of specific mental health conditions in order to assess symptoms (and occasionally diagnose) specific mental health conditions.

The Psychosis Screener

The Psychosis Screener (PS) (Degenhardt, Hall, Korten, & Jablensky, 2005) is an interview-style questionnaire and was designed specifically for use in epidemiological studies. The PS uses elements of the CIDI to assess the presence of characteristic psychotic symptoms and is comprised of 7 items, three of which (1a, 2a, 3a) were asked only if the respondent endorsed a previous question (1, 2, 3 respectively). The items cover features of psychotic disorders: delusions of control, thought interference and passivity, delusions of reference or persecution, and grandiose delusions. The final item (Question 4) assessed whether a respondent had ever received a diagnosis of schizophrenia.

It was validated against the CIDI, the DSM-III-R and the International Classification of Diseases, 9th and 10th editions (ICD-9; ICD-10). Internal consistency was good in two preliminary studies (Cronbach's alpha ≥ .74), although the items assessing grandiose delusions performed less well (Degenhardt et al., 2005). When using a broad definition of psychosis (schizophrenia, schizoaffective disorder, and affective psychosis as psychotic disorders), the screener performed best at a cut-off score of ≥1. Using a narrow definition of psychosis (a diagnosis of schizophrenia or schizoaffective disorder only) a cut-off score of 3 was found to most accurately include cases and exclude non-cases (Degenhardt et al., 2005). The screener demonstrated moderate-good Positive Predictive Power Value and Negative Predictive Value (PPV; NPV; proportion of patients who are correctly diagnosed with and without the disorder at both definition cut-off points in two separate samples (prison and general population), moderate sensitivity and specificity levels at the cut-off scores were also obtained, however, these were more or less strong depending upon which "gold standard" measure the scores were compared to (Degenhardt et al., 2005). All cut-off points on the screener, using the narrow definition of psychosis, from both samples are given below in Table 7.

Table 7: Sensitivity and specificity of the Psychosis Screener in two samples for narrow definition of psychosis

Cut-off	Sample 1		Sample 2 [1]	
	Sensitivity	Specificity	Sensitivity	Specificity
1	92.8	54.8	99.3	20.6
2	64.3	72.6	91.4	36.4
3	57.1	84.9	81.6	58.9
4	35.7	87.7	52.6	75.7

Cut-off	Sample 1		Sample 2 [1]	
5	7.1	97.3	36.2	86.9
6	0	100	23.0	92.5

[1] Persons categorised as having psychosis if they met either ICD-10 or DSM-III-R criteria for schizophrenia or schizoaffective disorder (as opposed to ICD-9 in sample 1)

Adapted from Degenhardt and colleagues (2005)

The PS has been shown to have a moderate ability to discriminate between those who meet diagnostic criteria for psychotic disorders and those who do not in community and prison samples, however, there is no findings surrounding its applicability to clinical samples (Degenhardt et al., 2005; White & Chant, 2006).

Client groups

The PS has been used successfully to measure psychosis in those suffering D&A disorders (Degenhardt & Hall, 2001; Degenhardt, Hall, & Lynskey, 2001) and the wider population (Scott, Chant, Andrews, & McGrath, 2006). It has been successfully used in a sample of 210 homeless men and women (Teesson, Hodder, & Buhrich, 2004) and has been used with mixed success in prison populations (Allnutt, Wedgwood, Wilhelm, & Butler, 2008; Butler et al., 2006; Nielssen & Misrachi, 2005).

Availability/cost

The PS is freely available in Degenhardt and colleagues (Degenhardt & Hall, 2001; Degenhardt et al., 2005):

- http://ndarc.med.unsw.edu.au/ndarcweb.nsf/resources/TR_37/$file/TR.093.pdf
- http://ndarc.med.unsw.edu.au/NDARCWeb.nsf/resources/TR_2/$file/TR.210.pdf

Scoring, administration and expertise required

The PS consists of seven items; the first six items cover the following features of psychotic disorders: delusions of control, thought interference and passivity, delusions of reference or persecution and grandiose delusions. The final item records whether a respondent reports ever receiving a diagnosis of schizophrenia. Since cut-off points vary depending upon the definition of psychosis, consideration must be given to the nature of the population with which a screening test is to be used before a cut-off point is selected (Degenhardt et al., 2005). Therefore, its current utility for clinicians is limited.

The Traumatic Life Events Questionnaire (TLEQ)

The Traumatic Life Events Questionnaire (TLEQ) (Kubany et al., 2000) has been shown to adequately evaluate both the re-experiencing of symptoms associated with trauma as well as symptoms of intense fear, helplessness and horror following a traumatic episode. It also assesses the frequency of event occurrence.

The development of the TLEQ was based on expert review, indicating good construct validity (Kubany et al., 2000). In the developmental study of the TLEQ, the tool was found to show adequate sensitivity and specificity, and the average test-retest reliability was .83 but was low for particular items (Kubany et al., 2000). The tool was also found to have good convergent validity when compared with the Traumatic Life Events Interview (correlation coefficients averaging .80) Recent studies have suggested that the psychometric properties of this measure are adequate (Read, Bollinger, & Sharkansky, 2003).

Client groups

The TLEQ has been used in substance abusing populations, adolescents and prison populations (Adams, 2004; Conrad, 2004; Gearon, Kaltman, Brown, & Bellack, 2003; Huang, Zhang, Momartin, Cao, & Zhao, 2006; Kubany et al., 2000; Oyefeso, Brown, Chiang, & Clancy, 2008). Similarly, developmental studies were conducted using a range of ethnic groups (Kubany et al., 2000).

Availability/cost

The TLEQ is copyrighted and must be purchased. It is available at:

- http://portal.wpspublish.com/portal/page?_pageid=53,70508&_dad=portal&_schema=PORTAL

Scoring, administration and expertise required

The TLEQ consists of 21 items (an early version had only 16 items) asking about traumatic events and the fear/helplessness associated with the event. Both self-report and interview-style forms exist. For each event, respondents are asked to provide the number of times it occurred (ranging from "never" to "more than 5 times") and whether fear, helplessness or horror was present. Some events include a question about presence of injury, and for victimisation questions, characteristics of the perpetrator. The last question asks respondents to identify the one event that "causes you the most distress" among those endorsed. Respondents are also asked about their age upon first occurrence, date of last occurrence, and amount of distress the event causes ("no distress" to "extreme distress"). It takes approximately 10-15 minutes to complete.

The Primary Care PTSD Screen (PC-PTSD)

The Primary Care PTSD Screen (PC-PTSD) (Prins et al., 2003) is a very brief, four-item screen that was designed for use in primary care and other medical settings to screen for Posttraumatic Stress Disorder (PTSD) (Prins et al., 2003).

The preliminary study of the PC-PTSD concluded that the tool was a psychometrically sound screen for PTSD with comparable operating characteristics to other screens for mental disorders. The PC-PTSD outperformed the PCL in terms of overall quality, sensitivity, specificity, efficiency, and quality of efficiency. The authors suggest a PC-PTSD cut-off score of 3 for male and female patients was optimal, but recommend a cut-off score of 2 when high sensitivity is required, but this comes at a cost to specificity (Prins et al., 2003).

A recent comparison of the PC-PTSD and the GHQ-12 found the PC-PTSD had a higher positive predictive value than the GHQ for detecting PTSD, indicating that disorder specific screens are important to use in primary care settings. But a combination of results on the two measures captured more cases than either individually (Ouimette, Wade, Prins, & Schohn, 2008).

Client groups

The PC-PTSD has been used successfully in substance use populations (Ford, Hawke, Alessi, Ledgerwood, & Petry, 2007; Goldstein et al., 2007). Among patients with D&A use disorders the screen has been shown to identify 91% of PTSD cases and has demonstrated good test-retest reliability (r = .80) and yielded a sensitivity of .91 and specificity of .80 using a cut-off score of 3 (Kimerling et al., 2006).

It has also been used in adolescent populations (Asarnow et al., 2008), and regularly used among military and veteran populations (Friedman, 2006; Hoge, Auchterlonie, & Milliken, 2006).

Availability/cost

The PC-PTSD screen is in the public domain and therefore can be used without cost but simply acknowledgement of its source. It is available at:

- http://www.ncptsd.va.gov/ncmain/ncdocs/assmnts/the_primary_care_ptsd_screen_pcptsd.html

Scoring, administration and expertise required

The PC-PTSD is very brief (4-items – one concerning each PTSD symptom cluster) and includes an introductory sentence to cue respondents to traumatic events; however, it does not include a list of potentially traumatic events. It can be used by any worker in need of a general screen for PTSD.

Impact of Events Scale (IES)

The Impact of Event Scale (IES) is a 15-20 item self-report inventory of the current degree of subjective stress experienced as a result of a specific event (Horowitz, Wilner, & Alvarez, 1979). A number of revisions have been made since its development. The most recent version is the IES-R (Weiss & Marmar, 1997). In its final form, the IES-R contains eight intrusion and eight avoidance items, derived from the original IES, and adds to this six items assessing hyperarousal (22 items total). The authors of the IES-R intended for the scale to be comparable with the original scale and made only minor changes to the intrusion and avoidance subscales.

Early psychometric findings using the original IES were strong, reporting high internal consistency (Cronbach's alpha = 0.79 – 0.92) and demonstrating sensitivity in significantly discriminating between patient and non-patient samples across time and detecting change over time, thereby supporting its suitability as a measure of treatment outcome (Zilberg, Weiss, & Horowitz, 1982).

The IES has demonstrated concurrent validity with the GHQ-28 in measuring emotional distress (Hodgkinson & Joseph, 1995). In a recent study sensitivity was found to range between 0.93 and 1.00 and specificity between 0.78 and 0.84 with optimal cut-offs depending upon which diagnostic tool was used (Wohlfarth, van den Brink, Winkel, & ter Smitten, 2003). Although others have reported much lower findings (Witteveen, Bramsen, Hovens, & van der Ploeg, 2005).

Although there have been some concerns raised about the factor structure of the IES, recent reviews have found factor structure to be stable over different types of events, that it can discriminate between stress reactions at different times after the event, and that it has convergent validity with observer-diagnosed PTSD; concluding that the IES was a useful screening and treatment outcome tool (Joseph, 2000; Sundin & Horowitz, 2002).

Although the IES-R was developed in 1997 it has generated surprisingly little psychometric evaluation. Those studies, however, t indicate that it too has good psychometric properties (Creamer, Bell, & Failla, 2003; Ljubotina & Muslic, 2003). For instance, Creamer (2003) reported internal consistency to be excellent (Cronbach's alpha = 0.96) and significant correlation with the PCL (0.84). This study found a cut-off score of 33 to provide the best diagnostic accuracy.

In a recent study, Beck and colleagues (2008) found the IES-R seems to be a solid measure of post-trauma phenomena. Internal consistency was high for both the total score (0.95) and subscales (re-experiencing/intrusion = .90, avoidance = .86 and hyperarousal = .85). Concurrent validity was also supported with the measure scoring well against both the earlier version of the IES and a range of comparative tools (Clinician-Administered PTSD Scale (CAPS), STAI, PSS-SR, BDI-II). Furthermore, although the IES-R was not developed as a diagnostic tool, examination of its discriminative validity found that the measure could differentiate between individuals with and without PTSD.

An earlier review found the IES-R to be among the most favoured tools for trauma exposure and posttraumatic assessment (Elhai, Gray, Kashdan, & Franklin, 2005). Furthermore, in a review by Brewin (2005) the IES consistently performed well and was one (of only two) instruments, that both had been validated on independent samples and had been tested within one year of a traumatic event.

Client groups

Both versions have been used successfully with a range of trauma (Cassiday, McNally, VandeCreek, Knapp, & Jackson, 1992; Joseph, 2000; Robbins & Hunt, 1996; Schwarzwald, Solomon, Weisenberg, & Mikulincer, 1987).

Similarly, both versions of the IES have also been translated and validated in dozens of languages and are widely used in a range of cultural groups (Asukai et al., 2002; Báguena et al., 2001; Brunet, St-Hilaire, Jehel, & King, 2003; Chengzhi, Xiangdong, & Lianxi, 2003; Çorapçioglu, Yargiç, Geyran, & Kocabasoglu, 2006; Ferring & Filipp, 1994; Guo-Ping, Ya-Lin, & Hui, 2006; Guo, Xin, & Geng, 2007; Kazlauskas, Gailiene, Domanskaite-Gota, & Trofimova, 2006; Maercker & Schützwohl, 1998; Mystakidou, Tsilika, Parpa, Galanos, & Vlahos, 2007; Olde, Kleber, van der Hart, & Pop, 2006; Perera-Diltz, 2007; Pietrantonio, Gennaro, Di Paolo, & Solano, 2003; van der Ploeg, Mooren, Kleber, van der Velden, & Brom, 2004; Wu & Chan, 2003; Wu & Chan, 2004).

Acceptable internal consistency and convergent validity of the IES-R have been established among substance dependent samples (Rash, Coffey, Baschnagel, Drobes, & Saladin, 2008). The authors of this study suggest a cut-off value of 22 as optimal for a substance using population, for adequate classification accuracy, sensitivity, and specificity.

The IES has been successfully used among adolescents (Sack, Seeley, Him, & Clarke, 1998; Yule, Ten Bruggencate, & Joseph, 1994) and a version (CRIES) has also been established for children (Giannopoulou et al., 2006; Perrin, Meiser-Stedman, & Smith, 2005)

Availability/cost

The IES instruments are freely available online and can be used without cost but with due acknowledgement of the source by people in non-profit research or clinical work. They are available at:

- http://www.mardihorowitz.com/impact_of_events_scale
- www.tracsa.org.au/file.php?f=48Xeky.LjTmqy.1050
- http://members.iinet.net.au/~gmt/IES-R-Scales.pdf
- http://www.atft.org/research/Impact%20of%20Event%20Scale%20-%20Revised.htm

Scoring, administration and expertise required

The IES is a self-report measure and takes between 10-20 minutes to complete. Information is only collected on one specific life event. The respondent indicates whether or not each a range of symptoms had been experienced within the past 7 days on a 4-point (not at all, rarely, sometimes, often) frequency scale. The subscale scoring for the 22-item IES is as follows: Intrusion (1, 4, 5, 6, 10, 11, 14) and Avoidance (2, 3, 7, 8, 9, 12, 13, 15). While the scoring for the IES-R differs: Avoidance (5, 7, 8, 11, 12, 13, 17, 22), Intrusion (1, 2, 3, 6, 9, 14, 16, 20), Hyperarousal (4, 10, 15, 18, 19, 21). Subscales are calculated as the mean of the items that form them, while the overall total is a sum of these subscales. The IES instruments require no special training.

The PTSD Symptom Scale Self-Report (PSS-SR)

The PTSD Symptom Scale Self-Report (PSS-SR) is a screening tool for PTSD which consists of 17 items, each item corresponding to one of the 17 DSM-III-R diagnostic Criteria B, C and D for PTSD (Foa, Riggs, Dancu, & Rothbaum, 1993).

The PSS-SR was found to have excellent internally consistent for total score (Cronbach's alpha = .91) and individual subscales (0.78, 0.80, and 0.82 for re-experiencing, avoidance, and arousal scales, respectively). The tool was also found to have good test-retest reliability over a period of 1 month (r = .74) (Foa et al., 1993). Furthermore, the PSS-SR was found to significantly correlate with a range of other equivalent instruments (Rape Aftermath Symptom Test, IES, BDI, STAI), indicating good concurrent validity. Finally, convergent validity was also supported with the PSS-SR correctly identifying the PTSD status of 86% of the subjects (sensitivity of 62%, specificity of 100%, PPV of 100% and NPV of 82%) (Foa et al., 1993). These findings were supported by a recent study which found sensitivity to range between .80 and .90 and specificity to range between .84 and .88, with optimal cut-offs dependent upon which diagnostic criteria was used (i.e. 15 for DSM-IV and 8-9 for ICD-10) (Wohlfarth et al., 2003).

Recent findings have also suggested the PSS-SR may have particular utility as a treatment outcome tool (Foa et al., 2005; Foa, Zoellner, & Feeny, 2006; Kindt, Buck, Arntz, & Soeter, 2007; Stalker, Gebotys, & Harper, 2005; van Minnen & Foa, 2006).

The PSS-SR was modified (MPSS-SR) by Falsetti and her colleagues (1993), who reported good overall internal consistency and good concurrent validity. The major modifications are that the items are not keyed to any particular traumatic event and that the MPSS-SR includes severity ratings and frequency ratings for each item. Thus, items are rated on 4-point frequency (ranging from 0 = "not at all" to 3 = "5 or more times per week") and intensity scales (ranging from A = "not at all upsetting" to D = "extremely upsetting"). In addition, for each item, respondents are asked to identify, if they can, which event each symptom is linked to.

An updated version, the Posttraumatic Stress Diagnostic Scale (PDS) (Foa, 1995) also exists to correspond with the current DSM-IV. This tool has 49 items clustering around the PTSD symptom structures. Consistent with DSM-IV, the measure assesses frequency of symptoms over the past month. It also provides information on the nature of the event that produced the symptoms, and also inquires about other DSM-IV criteria, such as interference with daily functioning, few other PTSD scales do this (Foa, Cashman, Jaycox, & Perry, 1997). It has been shown to have satisfactory test-retest reliability, internal consistency, and convergent and concurrent validity (Adkins, Weathers, McDevitt-Murphy, & Daniels, 2008; Foa, 1995).

Foa and colleagues (1997) reported Total Symptom Severity internal consistency to be excellent (Cronbach's alpha = 0.92) and good for the separate subscales (Re-experiencing = 0.78, Avoidance = 0.84, Arousal = 0.84). Test-retest reliability was also high (coefficients of .74 for total score and

coefficients ranging between .77 and .85 for the individual subscales). The sensitivity of the PDS was .89, whereas its specificity was .75, yielding correct diagnosis 82% of the time. Significant correlations were also found between the PDS and other equivalent instruments (IES-R, BDI, STAI), indicating good concurrent validity (Foa et al., 1997). In a sample of psychiatric outpatients, performance was maximised with a cut-off of 27, but this method resulted in much lower sensitivity and higher specificity than had previously been reported (Sheeran & Zimmerman, 2002).

A recent review found the PDS to be among the most favoured tools for measuring trauma exposure and posttraumatic assessment (Elhai et al., 2005). Furthermore, in a review by Brewin (2005) both measures were found to be highly promising instruments.

Client groups

Both the PSS-SR and the PDS have been found to be consistent and accurate in a range of trauma populations (Coffey, Gudmundsdottir, Beck, Palyo, & Miller, 2006; Dunmore, Clark, & Ehlers, 1999; Ehring, Kleim, Clark, Foa, & Ehlers, 2007; Engstrom, El-Bassel, Go, & Gilbert, 2008; Kuwert, Spitzer, Träder, Freyberger, & Ermann, 2007; Scher, McCreary, Asmundson, & Resick, 2008; Spasojevic, Heffer, & Snyder, 2000; Thompson et al., 2003; Tolin & Foa, 1999; Ullman & Long, 2008; Weidmann, Fehm, & Fydrich, 2008). The tools have also been found to be valid and reliable in D&A populations (Bonn-Miller, Vujanovic, Feldner, Bernstein, & Zvolensky, 2007; Coffey, Dansky, Falsetti, Saladin, & Brady, 1998; Coffey, Schumacher, Brady, & Cotton, 2007; Engstrom et al., 2008; Kaysen et al., 2008; Plotzker, Metzger, & Holmes, 2007; Sullivan & Holt, 2008). For instance, using the MPSS-SR, Coffey and colleagues (Coffey et al., 1998) reported good internal consistency reliability (Cronbach's alpha = 0.97) for the total score, and the severity (0.95) and the frequency subscales (0.94). It was found to show adequate convergent validity with the SCL-90-R PTSD scale and the IES. At a cut-off point of 28, sensitivity was 89%, while specificity rate was 65%; overall correct classification rate was 74% (compared to PTSD interview), indicating good concurrent validity.

These tools have also been used in young (Goodman, Morgan, Juriga, & Brown, 2004; Landolt, Vollrath, Timm, Gnehm, & Sennhauser, 2005; Self-Brown et al., 2006) as well as older populations (Chung, Berger, Jones, & Rudd, 2006) and validated in a different countries and cultural groups (Douglas, Jimenez, Lin, & Frisman, 2008; Griesel, Wessa, & Flor, 2006; Jobson & O'Kearney, 2008; Mirzamani, Mohammadi, Mahmoudi-Gharaei, & Mirzamani, 2007; Stieglitz, Frommberger, Foa, & Berger, 2001).

Availability/cost

The PSS-SR and the PDS are copyrighted instruments and enquiries into purchasing/obtaining the measures can be made from:

- Edna B. Foa, Center for the Treatment and Study of Anxiety, Medical College of Pennsylvania Division, Allegheny University of the Health Sciences, 3200 Henry Avenue, Philadelphia, Pennsylvania,19129, Email: foa@AUHS.edu
- http://www.pearsonassessments.com/pds.aspx
- Sherry Falsetti, Medical University of South Carolina, Crime Victims Research and Treatment Center, 171 Ashley Avenue, Charleston, SC 29425-0742

Scoring, administration and expertise required

The PSS-SR consists of 17 items corresponding to the 17 DSM-III-R criteria which are rated on a four-point scale of symptom presence. All versions of the scale take less than 15 minutes to administer (Coffey et al., 1998; Foa et al., 1993).

The PDS is a longer 49-item self-report measure that assesses trauma history and all DSM-IV criteria for the diagnosis of PTSD. Respondents rate the frequency of each symptom item on a scale from 0 to 3, with higher scores indicating greater frequency of symptoms.

Trauma Screening Questionnaire (TSQ)

The Trauma Screening Questionnaire (TSQ) (Brewin et al., 2002) is a recently developed 10-item screening tool consisting of the 10 re-experiencing and arousal items from the PSS-SR (Foa, 1995), modified to provide only two response options. Respondents indicate whether or not they have experienced each symptom at least twice in the past week.

Preliminary investigations have shown promising results. In a sample of rail crash survivors, the TSQ was found to have excellent sensitivity (0.86) and specificity (0.93), as well as excellent PPV (0.86) and NPV (0.93) and an overall efficiency score of 0.90 when endorsing at least six re-experiencing or arousal symptoms in any combination. While similar scores were obtained in a second study using victims of crime (sensitivity = 0.76; specificity = 0.97, PPP = 0.91, NPV = 0.92 and overall efficiency = 0.92) (Brewin et al., 2002). The authors conclude that these levels of overall efficiency are superior to equivalent – interviewer-administered or self-administered – instruments (Brewin et al., 2002).

Based on these preliminary results, a subsequent review (Brewin, 2005) reported that the measure performed better than most longer instruments, despite having a simple yes/no response scale.

In a recent study, the TSQ was found to be a useful tool in a sample of assault victims. This sample replicated the strong sensitivity (0.85), specificity 0.89, NPV (0.98) and overall efficiency (0.90) findings. However, the PPV was lower (0.48), but was likely to be the result of a low overall prevalence of PTSD (Walters, Bisson, & Shepherd, 2007). This finding lends support to the use of the TSQ as an effective means of predicting future PTSD.

Client groups

Limited research exists surrounding the use of the TSQ in different population groups. It was recently used in samples of over 2,000 young adults from the Australian Capital Territory after a major bushfire had occurred in the region (Parslow & Jorm, 2006, 2007; Parslow, Jorm, & Christensen, 2006).

Further studies are required with a variety of population groups.

Availability/cost

The TSQ is reproduced in Brewin and colleagues (2002) it is available online at:

- http://bjp.rcpsych.org/cgi/content/full/181/2/158

Scoring, administration and expertise required

The TSQ is a 10-item self-report and takes less than 5 minutes to complete. It is split into two halves, the first of which screens for re-experiencing symptoms and the second of which screens for arousal symptoms and can be easily administered without training. Respondents answer on a dichotomous "yes/no" scale and the scale therefore, is easily scored.

The PTSD Checklist (PCL)

The PTSD Checklist (PCL) (Weathers, Litz, Herman, Huska, & Keane, 1993) is a self-report scale where respondents rate the extent to which they experience each of the DSM-III-R PTSD symptoms. It consists of 17-items corresponding to the 17 DSM-III-R criteria which are rated on a five-point severity scale.

Early studies found the PCL to have good test-retest reliability (0.96) and moderate concurrent validity as indicated by a kappa of 0.64 for diagnosis of PTSD from the Structured Clinical Interview for DSM (SCID). Three symptom clusters were said to comprise the PCL with internal consistency (alpha coefficient) values ranging from 0.89 to 0.92 (Weathers et al., 1993).

In a subsequent study the internal consistency coefficient (Cronbach's alpha) for the total scale was 0.94 and ranging between 0.82 and 0.94, for the different subscales (Blanchard, Jones-Alexander, Buckley, & Forneris, 1996). At a cut-off score of 50, the PCL yielded sensitivity of 0.78, and specificity of 0.86 and an overall diagnostic efficiency of 0.83, but this improved at the lower cut-off score of 44, (overall diagnostic efficiency = 0.90, sensitivity = 0.94 and specificity = 0.86). Overall, the PCL correlated well with the CAPS (0.93) and diagnostic efficiency was 0.90 compared with the CAPS (Blanchard et al., 1996). However, examination of the individual items showed wide ranging values of individual item correlations.

In a recent study, Ruggiero and colleagues (2003) found Cronbach's alpha coefficients (.94, .85, .85, and .87 for the PCL total, re-experiencing, avoidance, and hyperarousal scores, respectively) were indicative of high internal consistency. While high correlations (i.e., r >.75) were found between the PCL total scores and scores obtained on two well-established measures for PTSD: the IES and Mississippi Scale for PTSD indicating good convergent validity. Statistical analyses also yielded support for discriminant validity. Furthermore, test-retest reliability was strong, ranging from 0.92 (immediate) to 0.68 (2-week interval). The scale was most efficient at cut-off scores of 44, 45 or 50 (with an item score of 3-4 for symptom criteria). These findings supported work by Forbes and colleagues (2001) and Mueser and colleagues (2001).

There are military, civilian and specific versions. An abbreviated version has also recently been created for use in primary care (Lang & Stein, 2005).

Client groups

The PCL has been shown to have strong psychometric properties across a range of trauma populations (Andrykowski, Cordova, Studts, & Miller, 1998; Dobie et al., 2002; Manne, Du Hamel, Gallelli, Sorgen, & Redd, 1998; Mueser et al., 2001; Walker, Newman, Dobie, Ciechanowski, & Katon, 2002).

It has been found to be useful in D&A populations (Najavits et al., 1998). For instance, one study has yielded an optimum cut-off score of 52 to determine the presence of PTSD as measured by the CAPS.

However, the cut-off score of 50 provided the optimal balance between sensitivity (.86) and specificity (.79) (Bollinger, Cuevas, Vielhauer, Morgan, & Keane, 2008).

The PCL has also been used in elderly populations, however, a lower cut-off score may be necessary in older populations (Cook, Elhai, & Areán, 2005; Schinka, Brown, Borenstein, & Mortimer, 2007), while a higher cut-off score may be required in severely mentally ill populations (Grubaugh, Elhai, Cusack, Wells, & Frueh, 2007).

The PCL has been translated and is used in different cultural groups internationally (Kocabasoglu, Özdemir, Yargiç, & Geyran, 2005; Miles, Marshall, & Schell, 2008).

Availability/cost

This checklist is in the public domain. It may be used without cost, but with due acknowledgement of its authors. It can be found at:

- http://idacc.healthbase.info/questionnaires.html

Scoring, administration and expertise required

The PCL is a self-report scale which asks the respondent how often they have been bothered by each symptom in the last month on a 5-point scale. The PCL can be scored in several different ways. A total score (range 17-85) can be obtained by summing the scores from each of the 17 items. A second way to score the PCL is to follow the DSM-IV criteria. It has been suggested that a combination of these two approaches (i.e., the requisite number of symptoms are endorsed within each cluster AND the total score is above the specified cut point for a specific population) may be best. Separate scores can also be obtained for Criteria B, C, and D.

The Beck Inventories

A range of measures for the assessment of various mental health symptoms have been developed by Aaron Beck and colleagues. The Beck Depression Inventory (BDI or BDI-II) is a 21-item self-report instrument intended to assess the existence and severity of symptoms of depression (Beck & Steer, 1987; Beck, Steer, & Brown, 1996). The Beck Hopelessness Scale (BHS) is a 20-item scale designed to detect negative feelings about the future and has been found to be a good predictor of suicide attempts (Beck & Steer, 1988). This can be helpful in ongoing treatment where particular thoughts can continue to be monitored. The Beck Scale for Suicidal Ideation (BSSI) is a 21-item scale assessing intention to commit suicide (Beck & Steer, 1991). The Beck Anxiety Inventory (BAI) (Beck & Steer, 1990) consists of 21 items, each describing a common symptom of anxiety.

The BDI is one of the most widely used self-report measures of depression. The items (of the BDI-II) correspond to the DSM-IV criteria of depression and assess severity of these symptoms (Beck et al., 1996), which include pessimism, sense of failure, self dissatisfaction, guilt, self dislike, suicidal ideas, social withdrawal, indecisiveness, agitation, concentration difficulties, worthlessness, insomnia, fatigue, loss of energy, and loss of libido. Importantly, the BDI appears sensitive to change and can be used to evaluate treatment outcome (Richter, Werner, Heerlein, Kraus, & Sauer, 1998).

The BDI has been found to have good internal consistency reliability with Cronbach's alpha ranging from .76 to .95 in psychiatric samples and from .73 to .92 in non-psychiatric samples (Beck, Steer, & Garbin, 1988). The BDI-II was found to have similarly good internal consistency (Cronbach's alpha ranging from .89 to .93) (Dozois, Dobson, & Ahnberg, 1998; Steer & Clark, 1997; Whisman, Perez, & Ramel, 2000). The test-retest reliability of the BDI is also moderate-strong (correlations ranging from .48 to .86 with psychiatric patients and from .60 to .83 with non-psychiatric groups (Beck, Steer, & Garbin, 1988). Of particular pertinence is that these findings have been replicated in substance misusing populations (Buckley, 2001; Kleinman et al., 1990).

Beck and colleagues (1988) found significant correlations between clinical ratings of depression and scores on the BDI, indicating good construct validity. Similarly, high correlations have been found between the BDI and other depression scales (e.g., the depression subscale of the SCL-90-R). It has also been found to accurately discriminate between depressed and non-depressed patients, although its ability to differentiate anxiety from depression has been criticised (Richter et al., 1998). The validity of the BDI-II has been found to be equally strong (Dozois et al., 1998; Riskind, Beck, Brown, & Steer, 1987; Steer & Clark, 1997).

In a student sample, the BDI-II scores correlated strongly (r = .83) with number of SCID depressed mood symptoms reported by students. A BDI-II cut-off score of 16 yielded a sensitivity rate of 84%, test-retest reliability was 0.96 (Sprinkle et al., 2002). While in a psychiatric population of adolescents, internal reliability estimates were good (ranging from .72 to .91) for the BDI-II total and subscale scores.

Evidence for the concurrent, convergent, and discriminant validity of the BDI-II was also established (Osman, Kopper, Barrios, Gutierrez, & Bagge, 2004).

The BDI was reported to have the best combination of sensitivity and specificity compared with competing scales in screening for depression in cocaine addicts (Weiss, Griffin, & Mirin, 1989). However, specificity was generally low. Nevertheless, it has been reported that a cut-off score of 18 has been found to provide good sensitivity (73%) and specificity (73%) for depression (Dawe et al., 2002). Cut-off scores for the BDI-II are reported in Table 8 and have been found to yield a sensitivity of 81% and specificity of 92% (Beck et al., 1996). However, different populations have been found to have varying optimal cut-off scores. For instance, in a medical sample, high sensitivity/specificity and NPV were obtained with a cut-off score of 9/10 (sensitivity = 100%, specificity = 83.1%, NPV = 100%). While high sensitivity/specificity and PPV were obtained with a cut-off score of 13/14 (sensitivity = 93.5%, specificity = 96%, PPV = 85.3%) (Furlanetto, Mendlowicz, & Bueno, 2005).

Table 8. BDI-II cut-off scores

Score	Classification
0-13	Minimal
14-19	Mild depression
20-28	Moderate depression
29–63	Severe depression

Adapted from Beck and colleagues (1996)

The BHS was developed to assess hopelessness and, in particular, negative attitudes about the future. Across seven clinical groups, Beck and Steer (1988) found the instrument to have high internal consistency (Cronbach's alpha = .93). While the test-retest reliability has also been found to be moderate to good (r = .69).

The BHS has been found to be highly correlated with clinician ratings of hopelessness (r = .64-.74) among general outpatients and suicide attempters and was significantly correlated with related items on the BDI (Beck, Weissman, Lester, & Trexler, 1974). The scale correlates highly with seriousness of suicide attempt and measures of suicidal intent (Beck, Kovacs, & Weissman, 1979; Ellis & Ratliff, 1986).

A large scale prospective follow-up study of suicidal inpatients and outpatients found a cut-off score of 9 yielded high sensitivity (94.1%) in the prediction of completed suicide, although moderate to poor specificity (41.0%) (Beck, Brown, Berchick, & Stewart, 1990). Similarly, psychiatric outpatients with a score of 9 or more were found to be at 11 times the relative risk of committing suicide (Beck et al., 1990). However, some have questioned its utility in predicting eventual suicide (Nimeus, Traskman Bendz, & Alsen, 1997).

The BSSI aims to assess a person's thoughts, plans and intent to commit suicide, as well as the number and seriousness of previous suicide attempts. The BSSI demonstrated a high level of internal consistency

in early study samples (Cronbach's alpha = .87 - .90). However, in a subsample of 60 inpatients the test-retest reliability was moderate to poor (*r* = .54). The BSSI has demonstrated concurrent validity with the BDI and BHS (Beck & Steer, 1991; Beck, Steer, & Ranieri, 1988). The BSSI was among the best predictors of admission to hospital in a review by Cochrane-Brink and colleagues (2000). At a cut-off score of 24 the BSSI had a both excellent sensitivity (100%) and specificity (90%) for predicting hospital admission among suicidal patients and an overall PPV of 71% (Cochrane Brink et al., 2000).

The BAI is a measure that assesses the severity of anxiety in adults and adolescents. The psychometric properties have been extensively studied in a range of inpatient and outpatient populations and among those suffering various anxiety disorders (Beck, Epstein, Brown, & Steer, 1988; Beck & Steer, 1990; Creamer, Foran, & Bell, 1995; de Beurs, Wilson, Chambless, Goldstein, & Feske, 1997; Fydrich, Dowdall, & Chambless, 1992; Kabacoff, Segal, Hersen, & Van Hasselt, 1997; Kumar, Steer, & Beck, 1993; Steer, Ranieri, Beck, & Clark, 1993; Steer, Rissmiller, Ranieri, & Beck, 1993).

Both the internal consistency (Cronbach's alpha = 92-.93) and the test-retest reliability (.75-.83) has been found to be consistently high (Beck & Steer, 1990; de Beurs et al., 1997; Fydrich et al., 1992).

The concurrent validity of the BAI has been demonstrated by correlations between BAI scores and anxiety diaries, clinically rated anxiety and other self-report measures of anxiety across a number of studies (Beck, Epstein et al., 1988; Fydrich et al., 1992; Kabacoff et al., 1997; Steer, Ranieri et al., 1993). However, the tool has been criticised for failing to discriminate between anxiety and depression (Beck, Epstein et al., 1988; Fydrich et al., 1992). Nevertheless, these correlations tend to be lower than those with anxiety measures and are likely to be the result of shared symptoms between the disorders (de Beurs et al., 1997). De-Beurs and colleagues (1997) also found the BAI to show good sensitivity to change, indicating its potential as a treatment outcome tool. The greatest utility of the BAI is reported to be its ability to assess panic symptomatology, as opposed to other forms of anxiety (Leyfer, Ruberg, & Woodruff-Borden, 2006).

Limited studies into the sensitivity and specificity of the measure prevent the designation of optimal cut-off scores (Dawe et al., 2002). These points also tend to differ on the specific anxiety concern. One study has suggested the cut-off score for optimal sensitivity (but moderate specificity) for "any anxiety disorder" is 3.5 and 5.5, for optimal specificity (but moderate sensitivity) (Leyfer et al., 2006).

Client groups

The use of the Beck inventories has been extensive across different population groups. Generally, however, a reading level of eighth or ninth grade education is required (Beckman & Lueger, 1997). The BDI has been used and validated in a range of cultural settings and ethnic groups (e.g., Spanish, German, Arabic, Bulgarian, Swedish, Japanese and Chinese) (Alansari, 2006; Bonicatto, Dew, & Soria, 1998; Byrne, Baron, & Balev, 1996; Byrne, Baron, Larsson, & Melin, 1996; Byrne, Stewart, & Lee, 2004; Carmody, 2005; Chang, 2005; Gorenstein, Andrade, Filho, Tung, & Artes, 1999; Kapci, Uslu, Turkcapar, &

Karaoglan, 2008; Kojima et al., 2002a; Kojima et al., 2002b; Uslu, Kapci, Oncu, Ugurlu, & Turkcapar, 2008; Wiebe & Penley, 2005; Xu, 1991).

It has been validated in adolescent and older populations (Jo, Park, Jo, Ryu, & Han, 2007; Osman et al., 2004; Simith, Schwartz, George, & Panke, 2004; Sprinkle et al., 2002).

The BHS, BAI and the BDI have been used in a diverse sample of clinical groups including substance misusers (Buckley, 2001; Hesse, 2006; Husband et al., 1996; Johnson, Neal, Brems, & Fisher, 2006; Kleinman et al., 1990; Luty & O'Gara, 2006; Lykke et al., 2008; Sumnall, Wagstaff, & Cole, 2004).

The BAI has demonstrated validity for use with adult, adolescent and older psychiatric patients and non-clinical samples (Osman et al., 2002; Wetherell & Gatz, 2005). Translated versions of the BAI have demonstrated adequate psychometric properties (Cheng et al., 2002; Freeston, Ladouceur, Thibodeau, Gagnon, & Rheaume, 1994; Sica, Ghisi, & Lange, 2007; Ulusoy, Sahin, & Erkmen, 1998). However, BAI scores have been found to be significantly related to age and gender. Women with anxiety disorders have been found to score on average 4 points higher than males with anxiety disorders in clinical samples (Beck & Steer, 1990) and also score higher among non-clinical samples (Osman, Kopper, Barrios, Osman, & Wade, 1997). Similarly, BAI scores were also found to be inversely related to age, with younger patients reporting more anxiety than older patients with the same anxiety disorders (Beck & Steer, 1990).

Less work has been conducted on the BHS and BSSI than the other Beck scales, however, both have been used in different cultural groups (Aguilar, Hidalgo, Cano, & Lopez 1995; Tanaka, Sakamoto, Ono, & Fujihara, 1996; Wu & Chan, 2007) and elderly and adolescent populations (Boyd, 2007; Chellappa & Araújo, 2007; Kong, Zhang, Jia, & et al., 2007; Rajpal, 2006).

Availability/cost

The BSSI and BHS may be used by a range of mental health professionals. However, the scoring and interpretation of the measure should be supervised by a Registered Psychologist. The BAI and BDI can only be purchased by a Registered Psychologist. All the Beck scales are copyright protected and may not be reproduced without permission. Copies may be purchased from:

- http://www.psychcorp.com

Scoring, administration and expertise required

Each Beck scale takes between 5 and 10 minutes to administer. Most can be either self-completed or administered orally by the clinician. Full administration and scoring guidelines are provided in the respective manuals. Most Beck scales can be administered in paper and pencil or computer-based formats (Steer, Rissmiller, Ranieri, & Beck, 1995).

The Beck scales are quite simple to administer but scoring and interpretation (and in some cases administration) must be supervised by a Registered Psychologist and the instruments must be purchased by a Registered Psychologist.

The Spielberger State Trait Anxiety Inventory (STAI)

The Spielberger State Trait Anxiety Inventory (STAI) (Spielberger, 1983) measures two separate forms of anxiety. *State* anxiety refers to the transient/situational status of stress, while *trait* anxiety refers to a more lasting predisposition to feel anxious (enduring personality characteristic) associated with anxiety. The former is entirely bound to a stressful situation and is measure on the STAI-S subscale, while the later is a more general personality attribute and is measured on the STAI-T subscale.

The preliminary testing process reported the test-retest reliability of the STAI-T scale to be moderately high for both college students (.76) and high school students (.69) (Spielberger, 1983), although lower for the STAI-S (a median reliability coefficient of .33). However, one might assume that this transitory anxious state would be less consistent over time. Internal consistency was found to good (Cronbach's alpha = .90). A recent review found average reliability coefficients for the STAI were acceptable for both internal consistency and test-retest, but variation was present among the estimates. However, STAI-S test-retest coefficients were again much lower than internal consistency coefficients (Barnes, Harp, & Jung, 2002).

The STAI-T and STAI-S scales have demonstrated discriminant validity in distinguishing between psychiatric and non-psychiatric patients. While stressful situations have been found to elicit significantly higher scores on the STAI-S scale (e.g., immediately before an exam, during military training) compared with non-stressful settings (e.g., after a relaxation class) (Spielberger, 1983). Furthermore, a number of experimental studies indicate the STAI (and the STAI-S scale particularly) may have some utility as an outcome measure (Fisher & Durham, 1999; Muris, Mayer, & Merckelbach, 1998; O'Leary et al., 2000; van Balkom et al., 2008).

Client groups

The STAI has been widely used and there are no special issues or concerns regarding its appropriateness as a measure of anxiety in women unlike the BAI (Dawe et al., 2002). A children's version of the scale is available – the State-Trait Anxiety Inventory for Children (STAIC). However, there are some concerns about the validity of the Trait subscale in this population (Carey, Faulstich, & Carey, 1994).

The STAI has also been validated for use with older adults with and without anxiety disorders (Stanley, Beck, & Zebb, 1996). A recent study on older individuals, found the optimal cut-off score on the STAI-S was 55/54 corresponding with sensitivity of 0.82 and specificity of 0.88 (Kvaal, Ulstein, Nordhus, & Engedal, 2005).

The tool has also been used in a variety of substance misusing populations (Cahill, Adinoff, Hosig, Muller, & Pulliam, 2003; de Almeida & Silva, 2005; Drummond & Phillips, 2002; Hoshi et al., 2007; Ilhan, Demirbas, & Dogan, 2007; Nagel, Schweinsburg, Phan, & Tapert, 2005).

A range of translations of the STAI have also been developed (Gauthier & Bouchard, 1993; Gorenstein, Pompeia, & Andrade, 1995; Mote, Natalicio, & Rivas, 1971; Van der Ploeg, 1980). However, further validation with various cultural groups is required (e.g., Indigenous Australians).

Availability/cost

The STAI is copyright protected and may not be reproduced without permission. A copy may be purchased from:

- ACER Press: Australian Council for Educational Research, ACER Customer Service, 347 Camberwell Road (Private Bag 55),Camberwell Victoria Australia 3124, Tel: (03) 9835 7447 Fax: (03) 9835 7499 via e-mail: sales@acer.edu.au or http://www.acerpress.com.au.

Scoring, administration and expertise required

The scale consists of 40-items, rated on a four-point scale and takes approximately 10 minutes to complete. The authors suggest a reading level of 4th or 5th grade is required. Half of each scale is scored normally and the other 10 items on each scale are reverse scored. This test can only be purchased by Registered Psychologists with post graduate training. Full administration and scoring guidelines are provided in the manual.

The Eating Attitudes Test (EAT)

The Eating Attitudes Test (EAT) (Garner, Olmstead, Bohr, & Garfinkel, 1982) is a screening test that detects disturbed eating patterns. It may also prove a useful outcome tool for anorexia and bulimia interventions (Garner & Garfinkel, 1979; Williamson et al., 1989). However, no version of the EAT assesses more general dysfunctional attitudes and related psychopathology. The Eating Disorder Inventory (EDI) is a larger tool (not discussed here) which is useful for more thorough assessment. The 40 items of the EAT were obtained by a series of administrations and inclusion of only those items that reliably discriminated anorexic patients from normal controls (Garner & Garfinkel, 1979). Shortened versions (EAT-26; EAT-16) of the EAT have been developed and shown to have similar properties to the original. The 26-item version was found to reflect the three main facets (Dieting and avoidance; bulimia and food; and oral control) (Garner et al., 1982). Research on the 16-item version has found four factors (Self-perception of body shape, dieting, awareness of food contents, and food preoccupation) (Ocker, Lam, Jensen, & Zhang, 2007).

In the preliminary study the EAT-40 demonstrated good internal consistency for both a clinical sample (Cronbach's alpha = .79) and for a pooled sample of both clinical patients and normal controls (Cronbach's alpha = .94) (Garner & Garfinkel, 1979). The EAT has also demonstrated good test-retest reliability (r = .84) (Williamson, Anderson, Jackman, & Jackson, 1995).

The construct validity has been supported in studies in which the EAT has been found to correlate highly with the subscales from the EDI, for instance the drive for thinness subscale (r = .81) (Gross, Rosen, Leitenberg, & Willmuth, 1986). The EAT also correlates moderately with self-monitoring records of frequency of bingeing (r = .66) and purging (r = .54) and other established measures of bulimia ($r \geq$ .67) (Garner & Garfinkel, 1979). The instrument has demonstrated the ability to discriminate between anorexic and non-anorexic women and between binge-eating women and non-binge-eating women (Prather & Williamson, 1988), but could not distinguish between anorexic and bulimic individuals (Williamson et al., 1995).

Garner and colleagues (1982) reported that a cut-off score of 20 on the EAT-26 or 30 for the EAT-40 correctly identified 84% and 85% of the subjects as either anorexic or controls women respectively. Others authors have suggested different cut-off scores. For instance, according to ICD-10 criteria, the cut-off of 25 was found to be most sensitive (87.5%) and specific (93.9%) (Canals, Carbajo, & Fernández-Ballart, 2002).

Overall, The EAT has been found to have good psychometric properties of reliability and validity, and reasonable sensitivity and specificity (Engelsen & Laberg, 2001; Garfinkel & Newman, 2001; Mintz & O'Halloran, 2000). PPV is generally very low because eating disorders are relatively uncommon. It is used to screen for eating disturbances in general as the first part of a two-part diagnostic screen, it has the ability to compare groups and to measure change between groups and over time (Garfinkel & Newman, 2001).

Client groups

Eating disorders are common in D&A treatment services and are commonly not detected (Holderness, Brooks Gunn, & Warren, 1994). But although the EAT has been used in studies of substance misusing women, the robustness of the EAT's psychometric properties has not been systematically investigated in this population (Dawe et al., 2002). Similarly, while the EAT has been used with male participants, little validation within this population has occurred.

The EAT has been translated into a range of languages, however, when applied outside a western context some items and concepts are lost or do not translate accurately, this is likely to be due to the fact that eating disorders generally are a particularly Western concept (Fedoroff & McFarlane, 1998). Nevertheless, the EAT has been used in a variety of cultures and ethnic groups (Alvarez-Rayón et al., 2004; Choudry & Mumford, 1992; Ko & Cohen, 1998; Lee, 1993; Lee, Kwok, Liau, & Leung, 2002; Leichner, Steiger, Puentes-Neuman, Perreault, & et al., 1994; Nakai, 2003; Neumärker, Bettle, Bettle, Dudeck, & Neumärker, 1998; Nunes, Camey, Olinto, & Mari, 2005).

A modified version, the CHEAT (Children's EAT) has also been developed to assess disordered eating attitudes in younger girls (Maloney, McGuire, & Daniels, 1988).

Availability/cost

This instrument is copyright protected, but may be used free of charge, with due acknowledgement of the source. It is available from:

- http://psychcentral.com/quizzes/eat.htm
- http://medical.state.gov/index.cfm?fuseaction=public.display&id=1cb40add-dda0-4ba1-be22-b3bdbf81ddab
- http://www.stuaff.niu.edu/csdc/EAT26.HTM
- Garner, D.M. (1997). Psychoeducational principles in treatment. In D.M. Garner & P.E. Garfinkel (Eds.) Handbook of treatment for eating disorders, New York: Guilford Press.

Scoring, administration and expertise required

The EAT is a simple, brief tool requiring no specific training to administer/score. It takes less than 10 minutes to administer, however, as it is a self-report measure, a 5th grade reading level is required by the respondent (Williamson et al., 1995). Items are scored on a 6-point scale ranging from "always" to "never."

Body Dysmorphic Disorder Questionnaire (BDDQ)

The Body Dysmorphic Disorder Questionnaire (BDDQ) (Phillips, 1996) was originally developed as a 9-item self-report to screen for disturbed body image. Items include concern (dissatisfaction) with "the appearance of some part(s)" of one's body, whether or not these are weight-related concerns, preoccupation with these concerns, ensuing distress, interference with "social life" and with school work, job, or role functioning, effects on avoidance and on other people in one's life, and the amount of time spent thinking about the perceived physical defect. A 7-item modification of the BDDQ has also found to screen acceptably for body dysmorphic disorder (Dufresne, Phillips, Vittorio, & Wilkel, 2001).

The is based on the criteria outlined in DSM-IV and has been found to be highly correlated with clinician diagnoses of body dysmorphic disorder (Phillips, 1996). In an inpatient psychiatric setting, the BDDQ had a high sensitivity (100%) and acceptable specificity (89%) (Phillips, Atala, & Pope, 1995). Whilst in a dermatology setting, it had a sensitivity of 100% and a specificity of 93% (Dufresne et al., 2001), and in a sample of 105 adult and 17 adolescent psychiatric inpatients, it had a sensitivity of 100% and a specificity of 93% (Grant, Kim, & Crow, 2001).

A recently developed tool derived from the BDDQ is the Body Image Disturbance Questionnaire (BIDQ) (Cash, Phillips, Santos, & Hrabosky, 2004). The advantage this tool has is that it measures a continuum of body image disturbance rather than a categorical (yes/no) measure. This tool was found to be internally consistent and free of impression-management response bias. The measure also converged appropriately with other body image indices (evaluation, affect, investment, and impact), was positively correlated with depression, social anxiety, and eating disturbances. Scores on this assessment also predicted psychosocial functioning above and beyond body dissatisfaction (Cash et al., 2004).

Availability/cost

The BDDQ can be found in Phillips (1996) with copyright held by the publisher:

- Phillips, K. A. (1996). *The broken mirror: Understanding and treating body dysmorphic disorder*. New York: Oxford University Press.

The BIDQ is printed in Cash and colleagues (2004), with copyright held by the Journal. However, scoring and interpretation requires the purchasing of the scale. It can be purchased for $15 at:

- http://www.body-images.com/assessments/order.html

Client groups

The tool has been used in an array on inpatent and outpatient populations (Bartsch, 2007; Conroy et al., 2008; Dufresne et al., 2001; Dyl, Kittler, Phillips, & Hunt, 2006; Grant et al., 2001; Grant, Kim, & Eckert,

2002; Phillips et al., 1995). A slightly modified version of the BDDQ (with language more appropriate for adolescents) has been developed but not thoroughly evaluated (Dyl et al., 2006).

Although not extensively evaluated in exclusively D&A populations, higher scores on the BDDQ have been found to correlate with increased self-reported illicit substance use histories and comorbidity (Stewart, Stack, & Wilhelm, 2008).

A German version of the scale also exists (Daig, Burkert, Albani, Martin, & Brähler, 2008), but to the author's knowledge few other cross cultural studies have been conducted.

Scoring, administration and expertise required

Items on the BDDQ use a yes/no response format, and scoring is not continuous, while on the BIDQ it is continuous. The scales consist of 7-9 and 12 items respectively. No special qualifications are required to administer the scales, and the scales take approximately 5-10 minutes to complete.

Positive Mental Health Outcome Measures

The Recovery Assessment Scale

The Recovery Assessment Scale (RAS) (Giffort, Schmook, Woody, Vollendorf, & Gervain, 1995) is an outcome tool which tests for empowerment, coping ability, and quality of life. The RAS is a 41-item survey rated on a 5 point scale. A shorter, 24-item version the RAS also exists.

In a study of thirty-five inpatients, the RAS was found to have good test-retest reliability ($r = 0.88$) along with good internal consistency (Cronbach's alpha = 0.93). The scale showed recovery to be positively associated with self-esteem, empowerment, social support, and quality of life, indicating good concurrent validity. It was inversely associated with psychiatric symptoms suggesting discriminant validity (Corrigan, Giffort, Rashid, Leary, & Okeke, 1999).

In a recent sample of 1,824 persons with serious mental illness, the RAS yielded five factors: personal confidence and hope, willingness to ask for help, goal and success orientation, reliance on others, and no domination by symptoms. These findings suggest the RAS has good construct validity for assessing the recovery processes (Corrigan, Salzer, Ralph, Sangster, & Keck, 2004). In this study hope was found to be an essential element of recovery. The combination of variables accounted for a substantial amount of variance in the recovery factors (offering some evidence of convergent validity) but also the factors were found to be distinct elements of recovery.

This finding was replicated in a sample of 168 individuals with severe and persistent psychiatric disability in which the RAS was found to load on five factors, each of which had satisfactory internal reliability (Cronbach's alpha range = .73-.91). The factors displayed convergent validity with positive and significant correlations with other recovery measures. Concurrent validity was demonstrated with significant but lower correlations with symptoms and clinician-rated measures of psychiatric functioning (completed self-report recovery and other mental health measures and their case workers completed the HoNOS) (McNaught, Caputi, Oades, & Deane, 2007).

Client groups

Preliminary studies suggest the RAS to be useful tool as an outcome measure for recovery. However, it has not been empirically tested in a range of populations. Nevertheless, the RAS has proved a valid measure in psychiatric populations (Corrigan et al., 2004; McNaught et al., 2007). In addition, it has been used successfully in a war veteran population (Flinn, Ventura, & Bonder, 2005).

Availability/cost

The RAS is freely available at:

- http://chester.nami.org/Job/RAS.pdf
- http://www.stigmaresearch.org/publications/measures/measure.cfm?mdes=RAS

Scoring, administration and expertise required

The RAS is a 24- or 41-item patient self-report measure. Each item is rated on a 5 point scale and takes approximately 5-15 minutes to complete. No special qualifications are required for its use.

Scales of Psychological Well-being (SPWB)

The Scales of Psychological Well-being (SPWB) (Ryff, 1989a, 1989b) consist of a series of statements reflecting the six areas of psychological well-being: autonomy, environmental mastery, personal growth, positive relations with others, purpose in life, and self-acceptance.

Recent studies have questioned the construct validity of the SPWB and have revealed fewer factors than the six proposed by Ryff in the original studies (Kafka & Kozma, 2002; Springer & Hauser, 2006; van Dierendonck, 2004). Ryff's own studies (Ryff, 1989b; Ryff & Keyes, 1995) have reported high correlations among scores for the constructs that were proposed as independent. Abbott and colleagues (2006) suggest that the measures may not, in practice, adequately operationalise the originally proposed constructs. However, the authors did report good predictive validity (Abbott et al., 2006). Concurrent validity findings are limited and correlations between the SPWB subscales and other measures are only moderate (Ryff & Keyes, 1995).

Client groups

This tool has been used in a range of populations including the elderly, and adolescents (Burton, 2006; Clayman, 2005; Lawler-Row & Piferi, 2006; Schanowitz & Nicassio, 2006; Vleioras & Bosma, 2005) and has been translated into a number of languages (Chang, 2006; Cheng & Chan, 2005; Laukka, 2007; Pulkkinen, Feldt, & Kokko, 2006; Vleioras & Bosma, 2005).

Availability/cost

There is no charge to use the SPWB; however, institutions must pay for the cost of reproducing it from the electronic master file, which is sent upon request to:

- Dr. Carol Ryff; University of Wisconsin; Institute on Aging; 2245 Medical Sciences Center; 1300 University Avenue; Madison, WI 53706; Phone: (608) 262-1818; Fax: (608) 263-6211; email: cryff@wisc.edu.

Scoring, administration and expertise required

The original instrument included 120 items (20 per dimension) but shorter versions comprising 84 items (14 per dimension), 54 items (9 per dimension), 42 items (7 per dimension) and 18 items (3 per dimension) are now widely used. A 14-item scale also exists but concerns have been raised about its reliability.

Respondents rate statements on a scale of 1 to 6, with 1 indicating strong disagreement and 6 indicating strong agreement. No testing supervisors are required. Responses are totalled for each of the six

categories (about half of the responses are reverse scored). Higher scores indicate mastery in that area, while low scores indicate a lack of comfort in that area.

Dispositional Hope Scale (DHS)

The Dispositional Hope Scale (DHS) (Snyder et al., 1991) consists of two subsets of items measuring Agency and Pathways, these two factors form the definition of Hope as Snyder and colleagues (1991) defined it: The positive motivational state that is based on an interaction between successful agency (goal-directed energy) and pathways (planning to meet goals). However, recent findings suggest the tool may only have a single factor (Brouwer, Meijer, Weekers, & Baneke, 2008).

Snyder and colleagues (1991) reported that the scale was internally consistent (alphas in the range of .75-.85 for several studies) and evidenced satisfactory construct and discriminant validity in a number of studies (Snyder et al., 1991).

A modified 6-item State Hope Scale has also been developed and found to be internally consistent, reflecting the theorised agency and pathways components. Evidence supporting the concurrent and discriminant validity for this version has also been reported (Snyder et al., 1996).

Client groups

The DHS has been translated and validated in a variety of different countries (Kato & Snyder, 2005; Lee, Lee, & Choi, 2008). In a large multiethnic sample found the structure of the DHS to be relatively stable and robust, however, some gender differences were found (Roesch & Vaughn, 2006).

A children's version has also been developed and shown to have good internal consistency (Cronbach's alpha = .84) along with good convergent and discriminant validity (Snyder et al., 1997).

Availability/cost

The DHS is printed in Snyder and colleagues (1991):

- Snyder, C. R., Harris, C., Anderson, J. R., Holleran, S. A., Irving, L. M., Sigmon, S. T., et al. (1991). The will and the ways: Development and validation of an individual-differences measure of hope. *Journal of Personality and Social Psychology, 60,* 570-588.

The State Hope Scale is printed in Snyder and colleagues (1996), which is freely available online:

- Snyder, C. R., Sympson, S. C., Ybasco, F. C., Tyrone F. Borders, Babyak, M. A., & Higgins, R. L. (1996). Development and Validation of the State Hope Scale. *Journal of Personality and Social Psychology, 70(2),* 321-335.

Scoring, administration and expertise required

The DHS consists of 12 items (including four distracters), with 4-items tapping the agency factor, and 4-items tapping the pathways thinking factor. Participants respond on an 8-point continuum (1 = definitely false, to 8 = definitely true), so that scores can range from a low of 8 to a high of 64.

Stages of Recovery Instrument (STORI)

The Stages of Recovery Instrument (STORI) (Andresen, Caputi, & Oades, 2006) was developed as a method of measuring recovery as the concept is described by mental health consumers. The development of the STORI was based on the Stage Model of Recovery (Andresen, Oades, & Caputi, 2003), which reflects the experiences of people who have recovered from mental illnesses, such as schizophrenia and the scales focus on mental health, psychological well-being, hope, resilience and recovery. It is intended to measure constructs that are more meaningful to consumers than conventional outcome measures. The STORI was developed in response to consumer criticisms of traditional clinical measures, which tend to focus on illness and disability. In contrast, the STORI focuses on *psychological recovery* and personal growth.

Preliminary data suggest the STORI correlates well with all of the psychological health variables, and the five stage subscales were found to be internally consistent (Cronbach's alpha ranging from 0.88 to 0.94). Despite requiring some minor refinement and further testing, the authors concluded that these findings indicated the STORI was a measure of the consumer definition of recovery (Andresen et al., 2006).

A shorter version of the tool, called the Self-Identified Stage of Recovery is also currently being refined and evaluated.

Client groups

Only limited data exists on the STORI but it has been translated into French and Spanish versions.

Availability/cost

Provided no profit is made, the STORI may be freely downloaded with due acknowledgement of the authors:

- http://www.uow.edu.au/health/iimh/stori/

Scoring, administration and expertise required

The STORI is a self-report measure consisting of fifty items, each rated from 0 to 5. The items are presented in 10 groups of five (subscales), with one item in each group representing a different stage of recovery. The highest subscale score indicates the stage of recovery that the person is experiencing. The subscales are as follows:

- **Moratorium** – A stage of hopelessness and self-protective withdrawal.
- **Awareness** – The realisation that recovery and a fulfilling life is possible
- **Preparation** – The search for personal resources and external sources of help.

- **Rebuilding** – Taking positive steps towards meaningful goals.
- **Growth** – A sense of having control over life and looking forward to the future.

Social and Emotional Wellbeing and Empowerment Tool

Empowerment programs seek to encourage people to take control of their own lives. There is considerable international and national evidence that interventions which empower socially excluded populations across psychological, organisational and community levels have achieved improved health outcomes and quality of life of disadvantaged groups (Wallerstein, 1992; Wallerstein, 2006). Indigenous people of Australia are one such population who have experienced severe and systematic disempowerment with devastating health and social impacts.

Work is currently being done on a quantitative tool to measure empowerment outcomes for individual, organisational and structural levels in Australian Indigenous peoples/communities. Preliminary data has shown positive results (Haswell-Elkins, 2009; personal communication).

This tool will enable cost benefit and sustainability analysis of empowerment interventions, based on aspects of empowerment as defined by Australian Indigenous people, and may eventually be used across health promotion and community development activities more broadly. The development of the measure had a strong Indigenous focus and involved the analysis of individual interviews on the process of empowerment (Haswell-Elkins, 2009; personal communication). Work by Tsey and colleagues (Tsey & Every, 2000; Tsey et al., 2005; Tsey et al., 2007) on the Family Well-being Tool and the Family Well Being Empowerment Program formed the context for the development of this empowerment tool.

The tool has three main components: K10, and two newly developed instruments: a 13-item Emotional Empowerment Scale (EES) and set of 12 scenarios (12S). It was piloted in eight small group settings involving 90 participants working or volunteering in Indigenous social health activities. Both the EES (self-capacity and inner peace) and 12S (psychological and social empowerment) have been found to have adequate reliability and validity (CRCAH, 2008).

Client groups

This tool was specifically designed for the Aboriginal and Torres Strait Islander community.

Availability/cost

The empowerment tool is still being finalised and publication of data and the tool itself is scheduled for 2009.

Scoring, administration and expertise required

The K10 component is scored as normal, the two new scales consist of 13 items and 12 scenarios, scoring for these items is explained in the tool itself.

General Drug and Alcohol Screening, Assessment and Outcome Measures

The Alcohol Use Disorders Identification Test (AUDIT)

The Alcohol Use Disorders Identification Test (AUDIT) is a 10-item screening instrument developed by a WHO collaborative designed to screen for a range of drinking problems (Saunders, Aasland, Babor, Fuente, & Grant, 1993). It is generally considered the 'gold standard' for providing an indication of current alcohol use disorders (e.g. harmful use, abuse, dependence), and aims to measure three aspects of alcohol misuse: consumption, dependence, and related-problems. Although originally found to load on three factors (and often used as a one-factor screening instrument with a single cut-off score) recent findings support a two-factor solution for the AUDIT (alcohol consumption and alcohol-related consequences) (Doyle, Donovan, & Kivlahan, 2007).

The psychometric properties of the AUDIT have been explored in a number of populations, including inpatient care, rural and urban communities and emergency room patients, the unemployed and college students (Reinert & Allen, 2002). Internal reliability has been consistently strong, with Cronbach's alpha scores in the range of.80-.94 (Allen, Litten, Fertig, & Babor, 1997; Bohn, Babor, & Kranzler, 1995; Shields & Caruso, 2003). While test-retest reliability has been has also shown good temporal stability (r = .88) (Daeppen, Yersin, Landry, Pecoud, & Decrey, 2000). AUDIT scores have been used to predict alcohol-related physical disorders and social problems (Conigrave, Hall, & Saunders, 1995; Conigrave, Saunders, & Reznik, 1995) and also the likelihood of remaining unemployed after a two-year period (Claussen & Aasland, 1993). Similarly, the AUDIT score was also shown to be a better predictor of subsequent alcohol-related medical and social problems than standard biochemical markers (Conigrave, Saunders et al., 1995). These findings indicate the AUDIT has high levels of predictive validity. The construct validity of the AUDIT has also been studied, with AUDIT scores showing moderate to high correlations with other self-report alcohol screening tools (e.g. the MAST and the CAGE), but lower correlations with biochemical measures (Allen et al., 1997; MacKenzie, Langa, & Brown, 1996), although it has been suggested that this may be due to a lack of sensitivity from such measures (Aertgeerts, Buntinx, Ansoms, & Fevery, 2001).

Saunders and colleagues (1993) suggested cut-off points of 8 and 10 for maximal sensitivity and specificity. Of those individuals who scored 8 or more on the AUDIT, 95-100% were classified in the hazardous alcohol consumption group; 93-100% were classified as having abnormal drinking behaviour; 100% were alcohol dependent (Dawe et al., 2002). Other authors have suggested≥8 as the cut -off for harmful consumption,≥10 for hazardous consumption and ≥19 for abuse and dependence diagnosis (MacKenzie et al., 1996). In a clinical sense, Babor and colleagues (2001) suggest individuals with scores between 8 and 15 should receive simple advice focused on the reduction of hazardous drinking. Those with scores between 16 and 19 should receive brief counselling and continued monitoring and those with scores of 20 or above clearly warrant further diagnostic evaluation for alcohol dependence. However, these cut-off scores may need to be modified depending upon the characteristics of the client

group (Cherpitel, 1995). The cut-off values of 7 and 16 were observed to have the best trade-offs between sensitivity, specificity, the ratio of positive likelihood to negative likelihood ratios, and PPV in a recent male Sri Lankan sample (De Silva, Jayawardana, & Pathmeswaran, 2008). While Bradley and colleagues (1998) suggest that lowering the recommended cut-off point from 8 to 4 may enhance the sensitivity of the AUDIT in adult women. This was supported by a recent review (Reinert & Allen, 2007), which otherwise supported the use of the AUDIT in most populations.

A recent meta-analysis of 19 relevant studies, at a cut-off point of 8, sensitivity ranged from .31 to .89 and specificity ranged from .83 to .96 across the eight studies conducted in primary care. A single trial in general hospital inpatients found a sensitivity of .93 and a specificity of .94; another trial in emergency-department patients found a sensitivity of .72 and a specificity of .88. A study in university students found a sensitivity of .82 and a specificity of .78. Three studies in elderly patients found sensitivities between .55 and .83 at a pooled specificity of .96. The authors concluded the large heterogeneity between results, could only partly be explained by setting diversity (Berner, Kriston, Bentele, & Härter, 2007).

Client groups

As mentioned above the AUDIT has been used successfully in a number of different populations (Berner et al., 2007; Cherpitel, 1998; Gómez, Conde, Santana, & Jorrín, 2005; Kokotailo et al., 2004; Reinert & Allen, 2002, 2007). It was purposely devised to be literally translated to different languages (Saunders, Aasland, Amundsen, & Grant, 1993) and has been validated in a number of cultures (Adewuya, 2005; Bergman & Källmén, 2002; Carey, Carey, & Chandra, 2003; Dybek et al., 2006; Gache et al., 2005; Giang, Spak, Dzung, & Allebeck, 2005; Kim, Gulick, Nam, & Kim, 2008; Lima et al., 2005; Tsai, Tsai, Chen, & Liu, 2005).

A modified version of the AUDIT has been developed for use with the Australian population. The AusAUDIT (Conigrave & Elvy, 1998; Degenhardt, Conigrave, Wutzke, & Saunders, 2001) retains the 10 items of the original AUDIT but makes some changes to detect lower, but still harmful drinking. The AusAUDIT can be found in Degenhardt and colleagues (2001). The AusAUDIT was found to show good sensitivity but somewhat limited specificity (Degenhardt, Conigrave et al., 2001). To improve specificity, the authors recommend increasing the cut-off points (from 6 for women and 7 for men) to 7 for women and to 10 for men. This resulted in sensitivity of .85 and specificity of .70 for both men and women. Using a cut-off point of 7 for both men and women, the AusAUDIT was somewhat less sensitive to detecting those participants who met ICD-10 diagnoses of dependence and/or harmful use, with sensitivities ranging from 85.7% to 87.2%.

The AUDIT has been shown to be an equally valid or superior measure for adolescents (compared to the TWEAK, the Problem Oriented Screening Instrument for Teenagers substance use/abuse scale (POSIT) and the CAGE) (Chung et al., 2000; Knight, Sherritt, Harris, Gates, & Chang, 2003). However, there has been some concern that the instrument lacks sensitivity in females and older populations (Dawe et al.,

2002). For instance Powell and McInness (1994) found the AUDIT had low sensitivity (57%) to alcohol abuse in a large sample of hospitalised Australian inpatients over 65 years of age. Morton and colleagues (1996) found the AUDIT sensitivity to be even lower in American male war veterans over 65 years (33%). Dawe and colleagues (2002) suggests the poor performance of the AUDIT in the elderly may be due to the emphasis on the actual consumption of alcohol, which may be less relevant to alcohol misuse and related problems in this age group (Conigliaro, Kraemer, & McNeil, 2000). Nonetheless, recent findings have tended to find the AUDIT to be an adequate tool in ageing populations (Berner et al., 2007; Gómez et al., 2006).

The AUDIT has been used among the Aboriginal and Torres Strait Islanders in research studies (Brady, Sibthorpe, Bailie, Ball, & Sumnerdodd, 2002; Kelly & Kowalyszyn, 2003; Schlesinger et al., 2007), however; there has been a dearth of specific field trials within this population so the use of the instrument and interpretation of scores should only proceed with caution. Despite the items for the AUDIT being derived from a cultural diverse cross national data set (Saunders, Aasland, Babor et al., 1993), one study has reported the AUDIT was felt to be intrusive and some questions were poorly understood by an Australian Indigenous population (Brady et al., 2002).

In a sample of Australian male psychiatric in-patients with a primary diagnosis of schizophrenia, the AUDIT showed good sensitivity (87%) and specificity (90%) in detecting past 12 month CIDI-diagnosed alcohol disorders, when the standard cut-off of 8 was used (Dawe, Seinen, & Kavanagh, 2000). Similarly, Kavanagh and colleagues (1999) found the tool had a sensitivity of 100% and specificity of 77% in detecting current alcohol disorder in a sample of young inpatients suffering psychotic episodes. Preliminary research in this area has shown the AUDIT to be an appropriate and valuable instrument. Equally strong results were found in various psychiatric outpatient populations (Hill & Chang, 2007; Maisto, Carey, Carey, Gordon, & Gleason, 2000). Similarly, the AUDIT has been found to have good sensitivity and moderate to good specificity for hazardous drinking, and alcohol use disorders in drug dependent samples (Skipsey, Burleson, & Kranzler, 1997).

O'Hare and colleagues (2004) found the AUDIT to be a reliable screening tool in a severely mentally ill population, with good concurrent validity with other measures of alcohol abuse and psychosocial difficulties. A lower cut-off score (3) however, may lead to more accurate detection of alcohol use disorders than the traditional cut-off score. However, in a recent study of a psychosis-sufferers the AUDIT functioned best with a problem drinking cut-off score of 10 (sensitivity, 85%; specificity, 91%) (Cassidy, Schmitz, & Malla, 2008).

Recently a number of new modified versions of the AUDIT have been developed for use in other drug using populations (i.e., other than alcohol). The Cannabis Use Disorders Identification Test (CUDIT) was found to have a PPV of 84.6% and sensitivity of 73.3% at a cut-off of 8, making it a superior instrument for measuring frequency of cannabis use compared to simply the number of cannabis use days (Adamson & Sellman, 2003). Internal consistency was good (Cronbach's alpha = 0.84). However, there

has been concerns raised about the validity of particular items on the measure and further studies are required (Piontek, Kraus, & Klempova, 2008).

Similarly, the Drug Use Disorders Identification Test (DUDIT) is an 11-item self-report instrument, which has recently been developed and tested (Berman, Bergman, Palmstierna, & Schlyter, 2005). Psychometric properties were examined in a sample of heavy drug users from prison, probation, and inpatient detoxification settings, and in a general Swedish population sample. In the drug user sample, the DUDIT predicted drug dependence with a sensitivity of 90% for both DSM-IV and ICD-10 and a respective specificity of 78 and 88%. Reliability according to Cronbach's alpha coefficient was 0.80.

Availability/cost

The AUDIT is in the public domain and therefore can be used without cost but with due acknowledgement of the source. This tool is available from a number of websites:

- http://whqlibdoc.who.int/hq/2001/WHO_MSD_MSB_01.6a.pdf
- http://www.nceta.flinders.edu.au/pdf/GP-Project/GP-Resource-Kit_files/B31-Ho5.pdf
- http://www.therightmix.gov.au/pdfs/HealthProviderAUDIT.pdf

Scoring, administration and expertise required

The AUDIT can be self- or clinician-administered and can be administered and scored without specific training. It is scored by simply adding the scores on each of the ten items (items 1 to 8 are scored on a 0 – 4 scale and items 9 and 10 are scored 0, 2, 4). A score of 8 or above (for men and perhaps 4 and above for adolescents and women) is thought to be indicative of alcohol problems. It takes between 2-5 minutes to complete and one minute to score. It has been shown to require a minimum reading level of seventh grade (Hays, Merz, & Nicholas, 1995), which suggests it is suitable for people with low levels of literacy (e.g., those for whom English was a second language).

CAGE / CAGEAID

The CAGE (Ewing, 1984) is a four-item screening tool that is designed to identify problem drinking via four constructs (each its own question). **C**utdown, **A**nnoyed, **G**uilty, **E**ye-Opener. The CAGEAID (**A**dapted to **I**nclude **D**rugs) is an equivalent tool developed to screen for drug use disorders has been.

The items on the CAGE have good internal reliability (Mischke & Venneri, 1987), but test-retest reliability may be less consistent over the lifetime (Green & Whichelow, 1994) and may function better as a screen for recent problems (Watson et al., 1995). In an early UK sample it was reported that using a cut-off point of ≥2 affirmative responses, the CAGE showed good sensitivity (84%) and specificity (95%) for detecting current high-risk drinking (defined as 8 or more standard drinks a day), and had an overall PPV of 45% (King, 1986). MacKenzie and colleagues (1996) recommend using a cut-off point of 1 for detecting hazardous/harmful use and a cut-off point of 3 for identifying those likely to meet a DSM diagnosis. However, using a cut-off point of 1 for harmful use is likely to overly compromise specificity (for the sake of sensitivity) and lead to a high rate of false positives which may render the CAGE somewhat impractical (Watkins, Eisele, & Matthews, 2000). Some authors have suggested that introducing the CAGE questionnaire in a non-judgmental way (e.g., "do you have a drink now and then?") dramatically increases its sensitivity without this compromise to specificity (Steinweg & Worth, 1993).

With a cut-off of ≥1 the CAGE-AID exhibited sensitivity of .79 and specificity of .77. At a cut off ≥2 sensitivity dropped to .70 and specificity increased to .85 (Brown & Rounds, 1995). This was higher in schizophrenic patients with alcohol use disorders, high sensitivity (0.91) and specificity (0.83) at the lower cut-off and the higher cut-off points (sensitivity = 0.82, specificity = 0.94) (Dervaux et al., 2006). In a sample of war veterans a cut-off score of ≥ 1 achieved a sensitivity of 86% and specificity of 93% when using the diagnostic interview as the criterion standard (Liskow, Campbell, Nickel, & Powell, 1995). While in a French hospital inpatient population, at a cut-off of 2 the CAGE had a sensitivity of 77% and a specificity of 94%. The CAGE test was more sensitive for patients diagnosed as alcohol-dependent than for alcohol abusers (61% vs. 84%) with the same specificity (94%) (Malet, Schwan, Boussiron, Aublet-Cuvelier, & Llorca, 2005).

A review found reliability of the CAGE to be varied with median internal consistency reliability across 22 samples reaching .74, but ranged from .52 to .90. Sample age was the only identified sample characteristic that demonstrated a statistically significant relationship with CAGE score reliability (Shields & Caruso, 2004).

Client groups

The CAGE has been used in many cultures worldwide, but often careful translation is required to retain accuracy. One problem that both forms of the CAGE may face is that an increased awareness of the dangers associated with alcohol/drug consumption (for instance Australia's growing public health

campaign) may produce an increase in the number of non-problem drinkers answering in the affirmative to items 1 and 3 (Waterson & Murray-Lyon, 1988).

Unlike many of the other instruments, the CAGE has been used in studies with Australian Indigenous communities. Skowrow and Smith (1986) found that high scorers on the CAGE consumed significantly more alcohol both on the day before interview, and on a typical drinking day and drank more often in a sample of 106 homeless Aboriginal men. Further studies found CAGE scores to correlate with both quantity and frequency of alcohol intake among Indigenous Australians in the Kimberley region (Hunter, Hall, & Spargo, 1991). After community consultation, these authors deemed it necessary to alter the wording for the items. The amended items are shown below (Hunter et al., 1991):

1. Do you sometimes think you shouldn't drink, or maybe drink less?
2. Do you feel angry or upset when other people get on your back about drinking, or tell you to cut down?
3. Do you ever feel shame or guilty about drinking?
4. Do you sometimes take a drink early in the morning for headache or because you feel no good, a reviver?

The CAGE was found to perform poorly in younger age groups (Aertgeerts, Buntinx, Bande-Knops et al., 2000; Chung et al., 2000; O'Hare & Tran, 1997). Some suggest this is because the items may be less relevant to adolescent populations, such as morning drinking, which is indicative of long-term alcohol dependence, rather than more recent problematic drinking (O'Hare & Tran, 1997). To increase the sensitivity of the CAGE in this population, Aertgeerts and colleagues (2000) suggested replacing the "Annoyed" item with "Have you ever been under the influence of alcohol in a situation where it increased your chances of getting hurt, for example, when riding a bicycle or driving a car", referred to as the "Under the influence" item (CUGE). However, validation of this new form of the instrument is required. Dawe and colleagues (2002) suggest, in light of these findings, that instruments that tap psychological symptomatology of alcohol misuse, such as cravings, high consumption and loss of control/memory (i.e., the AUDIT & TWEAK) may be more appropriate to the assessment of young alcohol misusers than those that index physiological dependence and long-term alcohol-related problems (i.e., the MAST & CAGE). In a similar sense, some authors have suggested the CAGE is less sensitive to the discrimination of less severe cases of alcohol misuse in the general population (Bisson, Nadeau, & Demers, 1999; MacKenzie et al., 1996).

However, the CAGE is perhaps the best measure of problematic drinking behaviour in older populations. It appears to be as sensitive as the MAST-G in the United States, with sensitivity for alcohol problems and dependence ranging from .77 to .94 at a cut-off point of ≥1 for persons over 60 years, but its brevity and ease of use make it arguably more efficient (Conigliaro et al., 2000). Buchsbaum and colleagues (1992) found the CAGE effectively discriminated between primary care patients over 60 with a history of alcohol misuse from those who did not. However, like the MAST-G, the CAGE showed extremely low sensitivity (.15 and .13 for excessive alcohol intake and dependence, respectively) in a British sample of

elderly emergency room attendees, suggesting that its utility within older Australian populations requires further investigation.

It has also been argued that the CAGE may lack sensitivity in the female population, especially when the usual cut-off point of 2 was used (Bradley et al., 1998; Cherpitel, 1998; Cherpitel, 1999; O'Hare & Tran, 1997).

Wolford and colleagues (1999) report the CAGE as having a sensitivity of 61% and specificity of 69% in detecting alcohol-related disorders, in inpatients with serious mental disorders. Lending support to early findings casting doubt on its effectiveness in psychiatric populations (Breakey, Calabrese, Rosenblatt, & Crum, 1998; Watson et al., 1995).

Availability/cost

The CAGE and CAGEAID are both in the public domain and therefore can be used without cost but with due acknowledgement of the source. This tool is available from a number of websites:

- http://ajp.psychiatryonline.org/cgi/reprint/131/10/1121
- http://counsellingresource.com/quizzes/alcohol-cage/index.html
- http://www.mentalneurologicalprimarycare.org/downloads/primary_care/11-1_cage_questionnaire.pdf
- http://www.recoveroz.com.au/CAGE
- http://nyc.gov/html/doh/downloads/pdf/csi/alcoholkit-clin-sticker.pdf
- http://www.mqic.org/pdf/CAGE_CAGE_AID_QUESTIONNAIRES.pdf
- https://www.mhn.com/static/pdfs/CAGE-AID.pdf
- http://www.cqaimh.org/pdf/tool_cageaid.pdf

Scoring, administration and expertise required

Both forms of the CAGE are easily administered and scored and even memorised. The tests themselves take only one minute to perform and despite being traditionally interview-style measures no difference was found between the oral and the written versions of the CAGE (Aertgeerts, Buntinx, Fevery, & Ansoms, 2000). They are simply scored by adding the "yes" responses and can be used by any health worker requiring a brief scan for substance problems. Further assessment to attain more specific information is generally required, due to the brevity of the instrument.

Michigan Alcoholism Screening Test (MAST)

The Michigan Alcoholism Screening Test (MAST) is a 24-item screening tool designed to identify and assess alcohol disorders (Selzer, 1971). Early studies showed strong internal consistency (Cronbach's alpha = .95) (Selzer, Vinokur, & van Rooijen, 1975) but more recent studies suggest a number of items are not highly correlated and that the instrument itself might not be measuring one factor but rather several factors related to problem-drinking (Crook, Oei, & Young, 1994; Parsons, WallBrown, & Myers, 1994; Saltstone, Halliwell, & Hayslip, 1994).

Selzer (1971) suggested a cut-off point of 5 to identify harmful or hazardous drinking. However, a cut-off score of 13 (at which the test has sensitivity of .91 and specificity of .76) is suggested for detecting the presence of alcohol abuse and dependence (Ross, Gavin, & Skinner, 1990).

Recent studies have reported the MAST to correlate with the AUDIT moderately well and correlate more highly than the AUDIT or the SADQ with DSM-IV criteria (Conley, 2001; Conley, 2002), indicating excellent construct validity. Internal consistency was also strong (0.86).

However, the MAST does not discriminate between past and present drinking and is therefore more useful in detecting lifetime alcohol issues than those which may be current (Dawe et al., 2002). A number of short versions of the MAST also exist and recently the standard version was shortened to 22 items. The most frequently studied short versions are the 10-item Brief MAST (bMAST) (Pokorny, Miller, & Kaplan, 1972) and the 13-item Short MAST (SMAST) (Selzer et al., 1975).

Connor and colleagues (2007) found the bMAST to have good construct reliability and both single-factor and two-factor scoring were equally effective as the AUDIT in assessing dependence severity. The authors concluded that the decision to use the original or two-factor bMAST should be based on criteria of purpose and efficiency. If the primary purpose is screening, the 10-item bMAST has demonstrated reliability and efficiency. In a recent meta-analysis of the MAST and the SMAST, Shields and colleagues (2007) found that both the MAST and the SMAST observe moderate to good internal consistency reliability estimates. However, in individual assessment and outcome measurement where personal and social costs are considered significant, the MAST and SMAST should be used with caution.

Client groups

Like the AUDIT, the MAST seems more applicable to men and less reliable and sensitive to detecting alcohol problems in females (Cherpitel, 1998; Shields et al., 2007). Several items are specifically aimed at men while others seem to focus on particularly male behavioural issues, which occur less frequently in women. However, it has been shown to have some utility in pregnant females (Chang, 2001). Similarly, the instrument includes a number of American words and phrases and may not be properly understood in the Australian community. However, the limited studies that exist suggest language of administration

and sample ethnicity were found to have very little association with the variation in MAST and SMAST reliability (Shields et al., 2007).

It has been suggested that the MAST may be insensitive to older individuals. However, in an elderly Brazilian population a cut-off score of 4/5 was associated with a sensitivity of 91.4%, specificity of 83.9%, and PPV and NPV of 69.6% and 96.0%, respectively (Hirata, Almeida, Funari, & Klein, 2001), suggesting that the MAST is a good screening test for the detection of alcohol abuse and dependence in an elderly male population. Furthermore, a geriatric version of the MAST (MAST-G) is available, along with an adolescent version. Each have been found to be satisfactory (Blow, 1991; Snow, Thurber, & Hodgson, 2002). The MAST-G was found to have good sensitivity (ranging from .70-.95) and specificity (.65-.84) in an American sample of individuals aged over 65, when a cut-off score of≥5 was used (Fingerhood, 2000), but these findings were not replicated in a British sample suggesting the tool might be limited in some non-American cultures (Conigliaro et al., 2000). The instrument has not been validated on Australian Indigenous populations and requires evaluation in an Australian context.

The MAST has been found to differentiate between non-alcoholic and alcoholic patients with schizophrenia with an overall detection rate of 80% (Searles, Alterman, & Purtill, 1990). It has also been used successfully on an outpatient basis with women who were psychiatric patients (Swett, Cohen, Surrey, & Compaine, 1991) and with clients undergoing methadone treatment programs (Stastny & Potter, 1991). However, in a recent study of in-patients with severe psychiatric disorders, Wolford and colleagues (1999) reported only moderate sensitivity (63%) and specificity (68%). Nevertheless, the psychometric properties of the MAST in psychiatric populations have been the topic of two meta-analyses (Teitelbaum & Mullen, 2000; Teitelbaum & Carey, 1996), both indicating the MAST was a useful tool and generally showed good sensitivity and moderate specificity in most studies. In the most recent meta-analysis, average sensitivity was 87.7%, while average specificity was 68.1% across nine MAST validity studies in psychiatric populations (Teitelbaum & Mullen, 2000). As in the general population, the MAST shows lower sensitivity to subclinical levels of alcohol misuse. The meta-analysis found the psychometric properties of the MAST were unaffected by respondents' psychiatric condition (i.e., psychotic, mood, or anxiety diagnoses).

Availability/cost

The MAST may be reproduced for non-commercial use (clinical, research, training purposes) as long as the author is credited. It has also been reproduced online at:

- http://www.ncadd-sfv.org/symptoms/mast_test.html
- http://www.dorris.com/objects/MichiganAlcoholScreeningTest.pdf

Scoring, administration and expertise required

The MAST is a self-report measure and can be used by any worker needing to screen for alcohol problems. It takes approximately 10 minutes to complete. Scoring is simple and instructions are provided with the test itself but differ slightly depending on what test is used. In addition to the original 22-item test, the bMAST and SMAST are composed of 10 and 13 items, respectively. All tests are dichotomously endorsed as yes or no.

Drug Abuse Screening Test (DAST)

Drug Abuse Screening Test (DAST) (Skinner, 1982) is a screening instrument developed in order to identify drug abuse problems and was based on the same style of questions as the MAST. In contrast to the MAST, the DAST items refer to the past 12-months rather than lifetime. There are several forms including 28-items, 20 items and 10 item tests and are highly correlated with each other (Cocco & Carey, 1998).

The DAST has been shown to have good internal consistency reliability (28-item DAST; Cronbach's alpha = .92; 20-item DAST, Cronbach's alpha = .95) (Skinner, 1982) and criterion validity. It was found to correlate highly with the ASI (Skinner & Goldberg, 1986). Scores have also been found to correlate highly with the frequency of use for a range of drugs including cannabis, barbiturates, amphetamine and opiates. DAST scores also discriminated accurately between alcohol and drug problems (Appleby et al., 1997).

Using a clinical sample of 501 drug/alcohol patients Gavin and colleagues (1989) found the DAST to have good concurrent and discriminant validity. Subjects were classified according to the presence or absence of any current DSM-III drug disorder (excluding alcohol and tobacco). The DAST attained 85% overall accuracy in identifying subjects who met DSM-III diagnosis. The authors suggest a cut-off score of 5/6 for optimum sensitivity and specificity on the 28-item DAST. Similarly, a cut-off score of 3 on the 10-item DAST correctly classified 93% of patients (Bohn, Babor, & Kranzler, 1991).

In a recent meta-analysis, the DAST was found to be an easy to administer, reliable, and valid tool with good sensitivity, and specificity. In general, all versions of the DAST yielded satisfactory levels of reliability and validity for use as clinical or research tools (Yudko, Lozhkina, & Fouts, 2007). Internal reliability was consistently high (.74-.95) for each version of the DAST. Test-retest correlation coefficients of 0.85 was reported for DAST-28, 0.78 for the DAST 20, 0.71 for DAST-10, and 0.89 for an adolescent version (DAST-A). The review also found evidence supporting the construct, criterion and discriminant validity of the DAST. Sensitivity and specificity (and optimal cut-off points) varied depending on DAST version and population studied. The authors recommend that in order "to obtain maximum sensitivity, a lower cut-off score from a possible cut-off score range is recommended when screening for drug abusers, and a higher cut-off score is recommended when screening for non drug abusers. The specificity of the DAST is increased when the cut-off score is high and, as a result, sensitivity decreases. Clinicians should select what cut-off scores to use according to the screening purpose" (Yudko et al., 2007, p.197).

Client groups

Like the MAST, it is likely that the DAST may have limitations when assessing drug use in women due to questions on social and behavioural aspects of drug use which may be more relevant to men, it has however, been used with success in an incarcerated female population (Saltstone et al., 1994).

Furthermore, there is limited information regarding its applicability in a wide range of cultures (Carey et al., 2003), and again like the MAST there are some concerns about the terminology used. The tool also requires validation in Australia's Indigenous population.

Recent research on the DAST with psychiatric outpatient populations has confirmed the internal scale properties with this group and established acceptable test-retest reliability, criterion-related validity, sensitivity and specificity. Maisto and colleagues (2000) for instance, found that a cut-off score of 2 provided good sensitivity and specificity for identifying a current diagnosis of an alcohol or drug use disorder among psychiatric outpatients. The internal consistency (Cronbach's alpha >.85) and test-retest reliability (*r* >.70) has also been found to be acceptable within this population (Cocco & Carey, 1998; Teitelbaum & Carey, 2000). In 250 psychiatric patients drawn from 4 treatment programs, the DAST evidenced high internal consistency reliability and good item-total score correlations. A factor analysis of the DAST revealed 5 factors: (1) self-recognition of a drug problem, (2) serious social consequences of drug use, (3) help-seeking for drug abuse, (4) illegal drug-related activities, and (5) inability to control drug use (Staley & El-Guebaly, 1990). In a first-episode psychosis population the DAST functioned best with a problem drug use cut-off score of 3 (sensitivity, 85%; specificity, 73%) (Cassidy et al., 2008).

The applicability of the DAST to the older population is unknown but an adolescent version (DAST-A) has recently been developed (Martino, Grilo, & Fehon, 2000). Initial validation findings indicate this version also has good internal consistency (Cronbach's alpha = .91), high test-retest reliability (*r* = .89). A cut-off score of 6 yielded 79% sensitivity and 84.5% specificity for identifying DSM-IV diagnoses of drug related disorders.

Availability/cost

There is a small fee to purchase copies of the DAST, which are available at:

- http://www.camh.net/Publications/CAMH_Publications/drug_abuse_screening_test.html

However, the DAST may be reproduced for non-commercial use (clinical, research, training purposes) with appropriate acknowledgement of the authors. It has also been reproduced online at:

- http://counsellingresource.com/quizzes/drug-abuse/index.html
- http://www.drtepp.com/pdf/substance_abuse.pdf
- http://www.veteransoutreachcenter.org/documents/DRUGABUSESCREENINGTEST2.pdf

Scoring, administration and expertise required

The DAST can be either interviewer or self-administered and can be administered without specific training. It takes less than five minutes to complete and is scored by adding the number of "yes" responses (except items 4 and 5 which are scored 1 for "no" responses).

T-ACE and the TWEAK

Both the T-ACE (Sokol, Martier, & Ager, 1989) and the TWEAK (Russell & Bigler, 1979) were developed to specifically identify at-risk drinking pregnant women as alternatives to the MAST. However, both have shown to be useful in assessment of non-pregnant women and men (Chang, McNamara, Orav, & Wilkins-Haug, 2006). Both scales have a similar structure to the CAGE: **T**olerance; **A**nnoyed; **C**ut down; **E**ye Opener (T-ACE) and **T**olerance; **W**orried; **E**ye Opener; **A**mnesia; (**K**) Cut down (TWEAK).

In pregnant women both the T-ACE and the TWEAK show consistently higher sensitivity and specificity than the MAST or CAGE (Chang et al., 1998; Russell et al., 1996). For instance, using a cut-off point of ≥1, the T-ACE had a sensitivity of 76% in predicting risky drinking during pregnancy compared with 59% for the CAGE and 76% for the MAST. Specificities for the T-ACE, CAGE and MAST were 79%, 82% and 76% (Sokol et al., 1989). Russell (1994) reported similarly strong findings, with TWEAK sensitivity of 79% and specificity of 83%, while T-ACE sensitivity was 70% and specificity was 85%, significantly higher than both the CAGE (49%; 93%) and the MAST (49%; 95%). The T-ACE Tolerance question was later changed to "How many drinks can you hold?" This increased the sensitivity and specificity of the T-ACE to 91% and 81% (Russell et al., 1994). Using this version of the Tolerance question, TWEAK sensitivity increased from 79% to 91% using a cut-off point of 2 (Russell et al., 1994). Based on these findings, Russell concluded that the TWEAK appeared to be somewhat more sensitive and less specific than the T-ACE but with both clearly outperforming the MAST and the CAGE test in screening for risk drinking during pregnancy.

Mixed findings exist surrounding the applicability of these measures to the general population. There is some evidence that the TWEAK outperforms the AUDIT in detecting female alcohol abuse and dependence (Bradley et al., 1998) and is more sensitive to detecting current alcohol dependence in men in the general public than the CAGE (Cherpitel, 1999). Some studies of emergency room and primary care patients indicate the TWEAK outperforms the CAGE and the MAST for screening alcohol abuse and dependence in women in the general population (Bradley et al., 1998; Cherpitel, 1997). However, other studies have found the CAGE to be more sensitive to alcohol use disorders than the TWEAK, particularly in certain ethnic groups (Cherpitel, 1999). Overall, it appears that the TWEAK is a sensitive instrument for detecting alcohol problems in both pregnant women and in the general population (Cherpitel, 1998). However, like other instruments, it does not provide a picture of the client's pattern of consumption. Therefore, a positive identification by the TWEAK may be supplemented by the AUDIT.

Chan and colleagues (1993) suggest a cut-off point of three to show optimal sensitivity and specificity for assessing heavy drinking and alcohol dependence in a combined sample of men and women. Other authors have found this cut-off point to be a suitable cut-off for men (sensitivity of between 76%-91% and specificity of between 74%-86%), but to result in unacceptably low sensitivity for women (sensitivity ranging from 57%-80%) in clinical and general populations (Cherpitel, 1998; Cherpitel, 1999). Accordingly, a cut-off point of two has been found to have optimal sensitivity (89%-91%) and specificity (77%- 87%) for detecting alcohol problems in women (Cherpitel, 1995; Russell et al., 1996). In older

women this may be lower again, with one study reporting the "best predictor model" had T-ACE scores of 1 or higher (Stevenson & Masters, 2005).

When the T-ACE was compared with the DSM-IV criteria for lifetime alcohol abuse or dependence (SCID, as Gold Standard), risky drinking was identified with sensitivities and specificities of 0.88 and 0.59 respectively at a cut-off point of 2 in an outpatient psychiatric clinic (Hill & Chang, 2007).

Client groups

As most studies are conducted in US samples, these questionnaires require Australian evaluation to ensure the findings are generalisable across cultures and class in this country, but the international findings are encouraging (Varescon, Gaugue, & Wendland, 2007). In a recent Brazilian sample, both the T-ACE and the TWEAK had excellent inter-rater reliability (coefficients ranging from 0.78 to 0.95). Internal consistency was low and construct validity was moderate (Moraes, Viellas, & Reichenheim, 2005).

Concerns exist surrounding the T-ACE and the TWEAK utility with older persons, as Tolerance (a significant component of these instruments) tends to be a poor indicator of alcohol misuse in this age group (DeHart & Hoffman, 1997). In the adolescent population, Chung and colleagues (2000) compared the utility of the AUDIT, TWEAK and CAGE in detecting DSM-IV diagnoses of alcohol abuse and dependence in a sample of American adolescent emergency room attendees and found the TWEAK to have good sensitivity (84%) and specificity (80%) for this population using a cut-off point of 1 (although the tool was less sensitive than the AUDIT).

The TWEAK has been found to have relatively low sensitivities but adequate specificities among female war veteran outpatients (Bush et al., 2003).

Availability/cost

The T-ACE is a copyrighted instrument, while the TWEAK is in the public domain and may be used without cost but with due acknowledgment of the source. Both are reproduced in Dawe and colleagues (2002), which is available online at:

- http://www.alcohol.gov.au/internet/alcohol/publishing.nsf/Content/DA90

Scoring, administration and expertise required

Both the T-ACE and the TWEAK are easily administered and scored and even memorised. The tests themselves take only one minute to perform and must be administered by the worker. The T-ACE is simply scored by adding the "yes" responses, a score of 2 or more is suggestive of alcohol problems. The TWEAK is scored slightly differently with item 1 scoring 2 points if a woman reports she can hold more

than five drinks, a "yes" response to item 2 scores 2 points and a "yes" response to items 3-5 score 1 point each. Therefore the maximum score obtained is 7. A total score of ≥3 for men and ≥2 for women is indicative of harmful/hazardous alcohol use. Both tools can be used by any health worker requiring a brief screen for alcohol problems. Further assessment to attain more specific information is generally required, due to the brevity of the instruments.

Timeline Followback Method (TLFB)

The Timeline Followback Method (TLFB) (Sobell & Sobell, 1995) is a technique developed to obtain precise information on the amount of alcohol (or other drugs) consumed and duration of each drinking/drug use session over a specified period of time, usually 3 months. This measure is particularly useful for assessing treatment outcomes and overall assessment and is not appropriate for screening. The key feature of this method is a blank calendar on which the client provides an estimate of the amount of alcohol drunk/drugs used on each drinking/use occasion during the time period. To assist in memory recall and to provide a framework for the client to work within, the first task is to note all events that may assist with recall (e.g., national holidays, significant personal events).

The TLFB has been used extensively in the research literature and has been found to have high test-retest reliability, with coefficients ranging from .79 to .96 over 30- to 90-day follow-up periods across a range of drinking populations (Sobell & Sobell, 2000). There is also a high degree of agreement between client self-report and official records such as days in jail or treatment facilities. The more recent computerised version has been found to have good reliability with test-retest correlations exceeding .85 (Fals-Stewart, O'Farrell, Freitas, McFarlin, & Rutigliano, 2000).

Overall consumption, number of heavy drinking days, and number of mean drinks per drinking day have all been found to positively correlate with the instrument (Sobell & Sobell, 2000), along with scores on the ADS and the MAST indicating that the level of alcohol problems or dependence was directly related to drinking behaviour as determined by the TLFB method. There was a similarly high level of agreement between those drinking variables derived from the TLFB method and biochemical indices of alcohol-related liver dysfunction (Sobell & Sobell, 2000).

Some authors have expressed concern that this type of self-monitoring measure may underestimate actual alcohol consumption (e.g., Carney, Tennen, Affleck, del-Boca, & Kranzler, 1998). For instance, one study reported the TLFB consistently and significantly underestimated alcohol consumption compared to the aggregated daily reports (Searles, Helzer, Rose, & Badger, 2002). This underestimate, however, was stable across the three reporting periods. Generally though, it is believed that the disparity between TLFB and actual consumption is small and appears to be influenced by individual differences in reporting (Dawe et al., 2002; Donohue et al., 2004; Vinson, Reidinger, & Wilcosky, 2003). Nevertheless, it is important that data collected by the TLFB be compared to other retrospective measures for accuracy.

Although originally designed for alcohol consumption, the TLFB has more recently been validated for collecting consumption data for other drug use (Carey & Correia, 1998; Fals-Stewart et al., 2000).

Client groups

The TLFB has been validated across a number of countries including Australia, Canada, Mexico, Poland, and Sweden (Annis et al., 1996; Sobell et al., 2001), along with a variety of subpopulations, including

those with severe mental illnesses, those in alcohol treatment facilities, and college students (DeMarce, Burden, Lash, Stephens, & Grambow, 2007). Carey (1997) has assessed the reliability and validity of the TLFB interview among psychiatric outpatients and concluded that it can appropriately be used with this population. Similarly, it has been used successfully in homeless populations and pregnant mothers (Sacks, Drake, Williams, Banks, & Herrell, 2003; Savage, Wray, Ritchey, & Fulmer, 2002). However, there are no reports of the use of the TLFB with Indigenous Australians.

The TLFB has been found to be reliable in adolescent populations (Dennis, Funk et al., 2004; Levy et al., 2004).

Availability/cost

The TLFB is under copyright but the pen and paper version may be obtained free of charge from:

- Linda C. Sobell, Center for Psychological Studies, Nova Southwestern University, 3301 College Avenue, Fort Lauderdale, FL 33314, (phone 954-262-5811; fax 954-262-3895) or via e-mail: sobelll@cps.nova.edu.
- http://www.nova.edu/gsc/online_files.html#time_followback

Scoring, administration and expertise required

The TLFB can be administered by the interviewer as a pen and paper instrument or via a client administered computer program (Sobell & Sobell, 1996). A modified version for telephone interviews has also been developed (Sobell, Brown, Leo, & Sobell, 1996). The time taken to complete the TLFB depends upon the time period covered, an individual drinker's pattern of consumption and the method of administration used. Vakili and colleagues (2008) found that shorter time windows, which are more time and resource efficient, can be used with little to no loss in the accuracy of the data. That is, one month for large samples (e.g., surveys), three months for individual cases (e.g., clinical use) and smaller samples. Dawe and colleagues (2002) suggest that at least 30 minutes should be allowed. Some training is needed for proper administration. Printed instructions are available for pencil-and-paper administration and are included in the computer-assisted program. A training video has also been developed and is available from the below address:

- http://www.camh.net/Publications/CAMH_Publications/timeline_followbk_usersgd.html

The computerised version of the TLFB provides clients with detailed instructions for self-administration and country-specific information regarding standard drinks, and allows them to incorporate their own personal events or holidays into the provided calendar. The computerised version allows measurement of time intervals up to 12 months and takes the same amount of time to administer as the pen and paper version (Dawe et al., 2002). Computerised scoring and interpretation is also available. The

shortened version, though, may have some limitations, which may affect its utility in clinical and research programs.

When treating clients in group settings a group-administered version of the TLFB has also shown promising results (LaBrie, Pedersen, & Earleywine, 2005; Pedersen & LaBrie, 2006)

Dartmouth Assessment of Lifestyle Instrument (DALI)

The Dartmouth Assessment of Lifestyle Instrument (DALI) (Rosenberg et al., 1998) is a screening instrument developed specifically for substance use disorders use with people with severe mental illness. It consists of 18 items derived from various existing screening tools. It was developed to be interviewer-administered. Eight items test for drug use disorders, nine test for alcohol use disorders. Two items overlap alcohol and drug use disorders. The items in the DALI address several dimensions of substance use disorder: Patterns of use, loss of control, the physiological syndrome of dependence, consequences of use, and subjective distress.

In the psychiatric population in which it was devised, the DALI was found to have a sensitivity of 80% and specificity of 85% in identifying alcohol use disorders and a sensitivity of 100% and specificity of 80% for cannabis and cocaine use disorders (Rosenberg et al., 1998). Overall classification accuracy of the measure for alcohol use disorders was 83.1%, while the overall classification accuracy for cannabis or cocaine use disorders was 89.7%. Results suggest that it is reliable over time and across interviewers, and that it is more sensitive and specific than several measures including the MAST, TWEAK, CAGE or DAST. Further replication of these results is required before the scale is recommended for routine use.

A recent study found the diagnostic accuracy of the DALI instrument was 74% for alcohol disorders and 88% for drug disorders (Ford, 2003). Using a cut-off score of 2, the specificity of the DALI alcohol scale was 0.98 and its sensitivity was 0.35. While the specificity of the DALI drug scale was 0.97 and its sensitivity was 0.50 (using a -1 cut-off). It showed good concurrent validity with the LDQ (Cronbach's alpha = 0.92, R^2=0.44).

The Simple Substance Use Screening Scale (SUSS), is a recently developed scale based on the DALI and has been found to correctly identify 86% of participants for problematic alcohol use (sensitivity 88%, specificity 84%) and 84% for problematic drug use (sensitivity 82%, specificity 84%) in a preliminary study (Ley, Jeffery, Shaw, & Weaver, 2007).

Client groups

As outlined above the DALI was specifically designed for use in psychiatric populations and has proved a useful tool in these settings (Swanson et al., 2002). Further studies are required in a range of populations and cultural groups to ensure these preliminary findings generalise across populations.

Availability/cost

The DALI is freely available at:

- http://dms.dartmouth.edu/prc/instruments/dali/

Scoring, administration and expertise required

The DALI is a short and easy to administer interview. It is slightly more difficult to score than some other measures, but requires no special training. It can be used by any worker requiring a brief screening tool for substance use related problems particularly people with severe mental illness

.

The Alcohol, Smoking and Substance Involvement Screening Test (ASSIST)

The Alcohol, Smoking and Substance Involvement Screening Test (ASSIST) was a tool developed through the World Health Organisation (WHO) by an international group of specialist addiction researchers and clinicians in response to the worldwide public health burden associated with problematic substance use (World Health Organization ASSIST Working Group, 2002). The instrument was designed to screen for problem or risky use of tobacco, alcohol, cannabis, cocaine, amphetamine-type stimulants, sedatives, hallucinogens, inhalants, opioids and 'other drugs'. A risk score is obtained for each substance and falls into either a 'low', 'moderate' or 'high' risk category which determines the type of intervention ('none', 'brief intervention', 'brief intervention plus referral').

The ASSIST screens for lifetime/recent substance use, specific substance involvement, frequency, dependence, abuse, intravenous drug use and was found to have excellent concurrent, construct, predictive and discriminative validity (Humeniuk & Ali, 2006). Internal consistency (Cronbach's alpha) was over 0.80 for the majority of domains and ASSIST items. Preliminary reliability studies were conducted at ten collaborating centres chosen for their ability to provide access to culturally diverse samples of individuals with different substance use patterns in Australia (coordinating centre), Brazil, India, Ireland, Israel, United Kingdom, the Palestinian Territories, Puerto Rico, the USA and Zimbabwe (Ali et al., 2002). In general, the test-test reliabilities were in the range of good to excellent, with coefficients ranging from 0.90 (consistency of reporting 'ever' use of substance) to 0.58 (regretted what was done under influence of substance). The average test-retest reliability coefficients for substance classes ranged from 0.61 for sedatives to 0.78 for opioids. (Ali et al., 2002; World Health Organization ASSIST Working Group, 2002).

Concurrent validity of the ASSIST was investigated by statistical comparison with other Gold Standards and standardised measures of parameters comparable with the ASSIST. Significant correlations were demonstrated between the ASSIST and similarly worded items of other questionnaires (the ASI, the SDS, the MINI International Neuropsychiatric Interview (MINI-Plus), the Rating of Injection Site Condition (RISC), the DAST, the Revised Fagerstrom Tolerance Questionnaire (RTQ), the MAP and the AUDIT) (Newcombe, Humeniuk, & Ali, 2005). Correlation coefficients between ASSIST scores and scores from these other measures were generally strong (e.g., r = 0.76-0.88 (ASI), 0.82 (AUDIT) and 0.78 (RTQ)) Construct validity was established by significant correlations between ASSIST scores and measures of risk factors for the development of drug and alcohol problems (r = 0.48-0.76) (Humeniuk & Ali, 2006; Humeniuk, Ali et al., 2008).

Discriminative validity of the ASSIST was investigated by comparison of ASSIST scores as grouped by known standards of dependence, abuse and non-problematic use. The measure was found to discriminate between substance use, abuse and dependence (Humeniuk & Ali, 2006; Newcombe et al., 2005). Suitable specificities (50-96%) and sensitivities (54-97%) for most substances were established and outlined in Table 9.

Table 9: Rounded cut-off scores for all substances for the ASSIST (version 3.0)

Risk	Alcohol	Other Substances
Low risk	0-10	0-3
Moderate risk (Abuse)	11-26	4-26
High risk (dependence)	27+	27+

Adapted from Humeniuk and Ali (2006)

A final process in the development of the ASSIST program was the evaluation of an ASSIST-linked brief intervention. Overall, it was found that this was effective in getting participants to reduce their substance use and risk as measured by their ASSIST score. Similarly, clients also perceived this as the case at the 3-month follow-up interview (Humeniuk, Dennington, & Ali, 2008).

Recently the National Institute on Drug Abuse has adapted this tool into a new measure to aid medicos in assessing whether or not their patients are using and/or "abusing" drugs (it can be found at: http://www.drugabuse.gov/nidamed/index.php).

Client groups

The development of the ASSIST was involved a range of different countries (Australia, Brazil, India, Thailand, UK, USA and Zimbabwe). Similarly, versions exist in a variety of languages, including English, French, Spanish, German, Hindi, and Portuguese.

Availability/cost

The ASSIST is in the public domain and may be used without cost, but with due acknowledgment of the source. It is available at:

- http://www.who.int/substance_abuse/activities/assist/en/index.html

Scoring, administration and expertise required

This tool is clinician administered and can be conducted in 5 to 10 minutes. The ASSIST comprising of eight questions commences with a general screening question, frequency/recency of use in the past 3 months. The later items assess psychological dependence and problems associated with use. At the completion of the interview a number of domains can be derived for each respondent, from their responses to the questions. Scoring is simple and instructions are included within the instrument and may take 10-15 minutes. No expertise is required to administer the ASSIST, however, manuals, guidelines for use and information on brief interventions/self-help can be obtained from the above website.

Drug and Alcohol Severity Screening, Assessment and Outcome Measures

Severity of Alcohol Dependence Questionnaire (SADQ)

Severity of Alcohol Dependence Questionnaire (SADQ) (Stockwell, Hodgson, Edwards, Taylor, & Rankin, 1979) is a 20-item questionnaire designed to measure the severity of dependence on alcohol. It is divided into five subscales: Physical Withdrawal Symptoms, Affective Withdrawal Symptoms, Craving and Withdrawal Relief Drinking, Consumption and Reinstatement. The more recent version of this scale (the 16 item, SADQ-C) has an additional companion scale, the Impaired Control Scale (ICQ) that focuses on the physical and affective aspects of alcohol dependence.

Unlike a number of the instruments outlined above, the SADQ has been widely used in Britain and Australia, and has demonstrated good reliability and validity in in-patient, out-patient and community-based treatment samples. It is useful for predicting dependence, withdrawal severity and the likelihood of achieving a moderate drinking outcome. It is often used as an outcome and assessment instrument (Foster, Peters, & Marshall, 2000; Kavanagh, Sitharthan, Spilsbury, & Vignaendra, 1999) but works less well as a screening tool (Dawe et al., 2002). It is particularly useful in attaining a specific measurement of alcohol dependence with drinkers related to the potentially fatal symptoms of alcohol withdrawal.

The original SADQ is a widely used measure of the severity of alcohol dependence and has been found to have particularly good reliability and validity compared to a number of major self-report questionnaires (Davidson, 1986; Stockwell et al., 1979; Stockwell, Murphy, & Hodgson, 1983). The original study found internal consistently to be strong (Stockwell et al., 1979). Scores on the SADQ also correlated highly with clinicians' case notes and diagnoses of alcohol dependence (82% concordance). This was supported by similar findings in a later study (Meehan, Webb, & Unwin, 1985). A score of 30 on the SADQ is generally taken to indicate severe dependence (Smith, 1986). Scores on the SADQ have also been found to correlate with withdrawal severity (Davidson, 1986; Meehan et al., 1985; Stockwell et al., 1983).

Scores on equivalent items for the SADQ and SADQ-C have been found to correlate highly with each other and with total scores for both scales. Internal consistency was also strong for this more recent version (Cronbach's alpha = .97), the SADQ-C. When the five ICQ items were added, a coefficient of 0.98 was obtained (Stockwell et al., 1983). In a Western Australian community sample, Cronbach's alpha coefficients of 0.86 (for the SADQ-C alone) and 0.87 (when it was combined with the ICQ) were obtained (Stockwell, Sitharthan, McGrath, & Lang, 1994).

Client groups

Only limited findings exist surrounding the SADQ's cross-cultural relevance (e.g., Ee Heok et al., 1990). It has not been widely translated and due to the level of English required to understand the SADQ, some cultural and linguistically diverse groups and those with low levels of literacy may have difficulty with the measure. The measure has not been evaluated in the Indigenous Australian population. Due to the number of consumption style questions it could be suggested that the measure may be less reliable in older population. Similarly, women may require a lower cut-off score, since women sustain alcohol-related damage at lower levels of consumption than men (Dawe et al., 2002).

To the authors knowledge, the tool has not been assessed among psychiatric populations (Singh, Mattoo, Sharan, & Basu, 2005).

Availability/cost

The SADQ-C is in the public domain and can be used without cost but with due acknowledgement of the source. It is available at:

- http://www.prisonmentalhealth.org/downloads/professional_resources/09-5_sadq_alcohol_assessment.pdf
- http://www.alcohol-and-drug-guide.com/support-files/assesspdf.pdf

Scoring, administration and expertise required

The SADQ-C is a client-administered tool which takes approximately five minutes to complete. It does not require specialised training to administer and is easily scored. Items 1, 3 and 4 of the ICQ are scored on a 4-point scale ranging from 0 (never or almost never) to 3 (nearly always). Items 2 and 5 are scored in reverse with a score of 0 (nearly always) to a score of 3 (never or almost never). The twenty items of the SADQ are all scored as follows: 0 = never or almost never, 1 = sometimes, 2 = often, 3 = nearly always.

Short Alcohol Dependence Data Questionnaire (SADD)

The Short Alcohol Dependence Data Questionnaire (SADD) (Raistrick, Dunbar, & Davidson, 1983) is a 15-item self-report based on the same formulation of alcohol dependence syndrome as the SADQ (Davidson & Raistrick, 1986; Edwards & Gross, 1976). It is less focused on the experience of withdrawal symptoms than the SADQ and includes items reflecting behavioural and subjective changes associated with drinking. It has been suggested that these differences make this scale more useful for less severe drinking problems (Davidson & Raistrick, 1986).

Davidson, Bunting and Raistrick (1989) demonstrated that the SADD measured alcohol dependence syndrome. The instrument has also been found to show good test-retest reliability (*r* = .87) and acceptable split-half reliability (a correlation between total score on odd and even numbered questions) (McMurran & Hollins, 1989; Raistrick et al., 1983). The SADD has been correlated with a number of other measures and found to have strong construct validity due to correlations between SADD scores and overall alcohol intake, SADQ scores, clinical assessment (Davidson & Raistrick, 1986; Doherty & Webb, 1989).

Client groups

The authors argue that the SADD is relatively independent of socio-cultural influences but there are only limited studies which have been conducted in order to confirm this. One early study using a Brazilian population does, however, lend support for this position (Jorge & Masur, 1985). Whether it is appropriate for use across different cultural groups in an Australian context or with Indigenous Australians, requires further research. There appears to be no specific studies to have tested the validity and reliability of the SADD with youth, elderly or women.

The tool has also been used successfully in mentally ill populations (Carrigan, Drobes, & Randall, 2004; Stasiewicz et al., 2008).

Both SADQ and SADD are routinely administered to adolescent clients of the NSW Juvenile Justice System and although not empirically tested, they appear useful in this context (Dawe et al., 2002) and the SADD particularly, is easily understood (McMurran & Hollins, 1989).

Availability/cost

The SADD is in the public domain and may be used without cost but with due acknowledgment of the source. It can be found at:

- http://www.dva.gov.au/health/younger/mhealth/alcohol/training/appendo.htm

Scoring, administration and expertise required

The SADD is a self-report measure and takes only five minutes to administer. It can be used by any worker regardless of expertise. A total score is obtained by adding the score from each of the items. Scores between 1 and 9 indicate low dependence, those between 10 and 19 indicate medium dependence and a score of 20 or more indicates high dependence. Each item is scored as follows: never = 0; sometimes = 1; often = 3; nearly always = 4.

Alcohol Dependence Scale (ADS)

The Alcohol Dependence Scale (ADS) (Skinner & Horn, 1984) is a 25-item self-report questionnaire designed to identify and assess alcohol abuse and dependence. Like the SADQ and SADD, it is based upon Edwards and Gross (1976) conceptualisation of the alcohol dependence syndrome. Items forming the ADS were found to have good internal consistency (Cronbach's alpha = .92). The measure was also found to correlate with daily consumption of alcohol and lifetime use of alcohol, social consequences from drinking, prior treatment for alcohol abuse, use of alcohol to change mood and feelings of guilt over drinking (Skinner & Horn, 1984). It has also been found to show good construct validity with the MAST and diagnostic validity regarding DSM diagnoses (Ross et al., 1990; Skinner & Horn, 1984). Ross and colleagues (1990) reported that a cut-off score of 9 was associated with sensitivity of 91% and specificity of 82% in identifying alcohol abuse or dependence disorders in a primarily male sample. The overall accuracy of the ADS at this cut-off point was 89%. A study with homeless women found a cut-off point of 8 to be optimal for mild/moderate dependence, while scores greater than 15 indicated severe dependence (Chantarujikapong, Smith, & Fox, 1997).

The ADS has also shown some utility as a treatment outcome measure (Loeber, Croissant, Heinz, Mann, & Flor, 2006; Ralevski, Ball, Nich, Limoncelli, & Petrakis, 2007; Ray, Hutchison, & Bryan, 2006).

Shorter versions of the ADS have been developed and have been found to correlate highly with the original scale, and with other measures of alcohol involvement (Kahler, Strong, Hayaki, Ramsey, & Brown, 2003; Kahler, Strong, Stuart, Moore, & Ramsey, 2003). Despite generally good findings regarding the ADS's psychometric properties, a recent study found the ADS to lack accuracy in identifying physiological dependence or withdrawal in treatment-seeking individuals with DSM-IV alcohol dependence (Saxon, Kivlahan, Doyle, & Donovan, 2007).

Client groups

The ADS has also been successfully adapted for use with a variety of different cultures and ethnic groups (Fu et al., 2008; Rajendran & Cheridan, 1990; Solís, Cordero, Cordero, & Martínez, 2007). The translated versions of the ADS were found to have high internal reliability.

The ADS was found to correlate well with alcohol-related problems and post-release drinking goals with incarcerated male offenders (Hodgins & Lightfoot, 1989). Similarly, Peters and colleagues (2000) found the ADS to be one of the most effective screening instruments for detecting substance use disorders in a prison population. The ADS (cut-off point of≥14) was combined with the drug use section of the ASI (cut-off point of ≥ 11) and found to have sensitivity of 74% and a specificity of 92% in detecting alcohol or drug dependence. Overall, 83% of the sample was correctly assigned.

The tool has also been used in psychiatric population samples (Bischof, Rumpf, Meyer, Hapke, & John, 2005; Petrakis et al., 2006; Ralevski et al., 2007).

Importantly, the ADS appears to be an equally valid and reliable measure of alcohol dependence in women (Chantarujikapong et al., 1997; Drake & Mercer-McFadden, 1995). The ADS has been used successfully in several studies investigating alcohol dependence in homeless and incarcerated women, and studies of alcohol misuse in Australian female university students (e.g., Biron, Brochu, & Desjardins, 1995; Chantarujikapong et al., 1997; Williams, Connor, & Ricciardelli, 1998). However, there have not been any reports of its use with Indigenous Australians.

Availability/cost

The ADS is copyrighted and the kit (including questionnaires and users guide) can be purchased for $15 - $20 at:

- http://www.camh.net/Publications/CAMH_Publications/alcohol_dependence_scale.html

Scoring, administration and expertise required

The ADS is a brief self-report measure that takes between five and ten minutes to complete. The manual provides scoring instructions and no special training or expertise is required to use or interpret the ADS. A computerised version is also available.

Cannabis Problems Questionnaire (CPQ)

The Cannabis Problems Questionnaire (CPQ) (Copeland, Gilmour, Gates, & Swift, 2005; Copeland, Swift, & Rees, 2001) is a brief, recently developed Australian screen of cannabis-related problems. Originally 53 items long, the CPQ was later cut to 27 and then 22 items.

In its validation study, the 22-item CPQ was found to be a valid, reliable and sensitive measure of cannabis-related problems for use with clinical and research populations of current cannabis users (Copeland et al., 2005). Exploratory factor analyses revealed a three factor solution best described the data accounting for 57% of the variance in the larger item set. Test-retest correlations were between 0.92 and 1.00 and inter-rater reliability correlations between 0.74 and 1.00. The total CPQ score classified DSM-IV cannabis dependence with 84% specificity and sensitivity and daily cannabis use with 83% specificity and 55% sensitivity (Copeland et al., 2005).

Availability/cost

The CPQ is in the public domain, and may be used without cost, but with due acknowledgment of the source. It is available at:

- http://ncpic.org.au/

Client groups

The CPQ is a newly developed tool and hasn't been widely tested in different populations, however, an adolescent version of the CPQ has been recently developed and validated (CPQ-A) (Martin, Copeland, Gilmour, Gates, & Swift, 2006).

Scoring, administration and expertise required

The scale is dichotomously scored as "yes/no" and no special training is required to use the scale. The scale is self completed by the client and takes less than 10 minutes to complete.

Leeds Dependence Questionnaire (LDQ)

The Leeds Dependence Questionnaire (LDQ) (Raistrick et al., 1994) was developed as a 10-item assessment and outcome measure for treatment of substance dependence. In the LDQ, severity of dependence incorporates broader notions of psychological dependence rather than simply measuring consumption and physically dependent symptoms. It has been shown to have good psychometric properties for alcohol and opiates but only limited findings concerning its psychometric properties for measuring the severity of dependence on other illicit substances exist. It has been found to be an effective assessment and outcome measurement tool (McCambridge & Maria, 2008; Raistrick, Tober, Heatherward, & Clark, 2007)

Preliminary findings indicate the measure has good levels of internal consistency (Cronbach's alpha = .94) and test-retest reliability (*r* = .95) as well as concurrent, discriminant and convergent validity for alcohol and opiate dependence. The LDQ was significantly correlated with the DALI, Severity of Opiate Dependence Scale (SODQ) and the SADQ scores (Ford, 2003; Raistrick et al., 1994). However, no specific cut-off score has been established to indicate dependence.

In a recent sample of clinic attendees, the LDQ had high internal consistency. A number of factors were independent predictors of higher LDQ scores, these included age (younger), gender (male), higher score on the GHQ and substance category (opioid or other drugs compared with alcohol) (Heather, Raistrick, Tober, Godfrey, & Parrott, 2001).

Client groups

Lennings (1999; 2003) found the LDQ performed well in measuring drug dependence among tertiary students and juvenile delinquents. Among students, the LDQ was predictive of alcohol use even after controlling for other variables. It has been found to be a comparatively good measure of alcohol dependence in a youth population when administered online, with good internal consistency and test-retest correlation statistics (Thomas & McCambridge, 2008). Similarly, in a sample of psychiatric inpatients, Ford (2003) found the LDQ to have very good psychometric properties and suggests it is a useful measure of substance dependence in this population.

It has also been used among Aboriginal and Torres Strait Islanders in research studies (Schlesinger et al., 2007) but its applicability to these populations has not been sufficiently tested.

Availability/cost

The LDQ is in the public domain, and may be used without cost, but with due acknowledgment of the source. It is available at:

- http://eib.emcdda.europa.eu/attachements.cfm/att_3954_EN_leeds.pdf

Scoring, administration and expertise required

The LDQ is a self-completion questionnaire and no qualifications are required for its use. Respondents are instructed to answer the questions about their substance use in the past week and to tick the relevant response. Each of the items is scored on a "never" (0), "sometimes" (1), "often" (2) and "nearly always" (3) scale, yielding a maximum score of 30.

Substance Dependence Severity Scale (SDSS)

The Substance Dependence Severity Scale (SDSS) (Miele et al., 2000a) is a relatively newly developed semi-structured clinical interview designed to obtain a measure of severity of dependence on a variety of substances, in order to evaluate treatment outcomes. It provides DSM-IV diagnoses of dependence for alcohol, cocaine, heroin, stimulants, illicit opiates, sedatives, methadone, cannabis, hallucinations and 'other' substances (e.g., inhalants). Each substance is rated for symptom severity and frequency in the last 30 days. Preliminary validation studies supported the internal consistency and test-retest reliability of the SDSS for alcohol, heroin, cocaine and sedative use. The SDSS was found to have concurrent and predictive validity in assessing the severity of DSM-IV alcohol, cocaine and heroin dependence. The SDSS relies on psychometric evidence for reliability and validity both as a dependence severity assessment tool and as a treatment outcome indicator in substance abuse patients (González-Saiz, 2007).

The SDSS was validated using a sample of 175 recent admissions to both inpatient and outpatient drug and alcohol rehabilitation programs, dual diagnosis units and methadone maintenance programs (Miele et al., 2000a; Miele et al., 2000b). The internal consistency of the four sub-scales was assessed for alcohol, cocaine, heroin, cannabis and sedative use. Internal consistency estimates on the Usual Severity Subscale (Cronbach's alpha = .79-.91), Worst Severity Subscale (Cronbach's alpha = .67-.89), Total number of days the symptom occurred (Cronbach's alpha = .75-.91) and Total number of days symptom at worst severity (Cronbach's alpha = .66-.82) were mostly in the acceptable range for all substances, with estimates for cannabis tending to fall in the lower end of the acceptable range (Miele et al., 2000a). Test-retest reliability was found to be generally good to excellent for alcohol, heroin, cocaine and sedatives; but moderate for cannabis (Miele et al., 2000a).

Concurrent validity of the SDSS was demonstrated with severity scale scores for all substances, except sedatives, being significantly correlated with clinical dependence severity ratings (DSM-IV). SDSS severity and frequency subscales were significantly correlated with frequency of alcohol, heroin, cocaine and cannabis use (Miele et al., 2000a; Miele et al., 2000b).

Further studies found alcohol scales were significantly related to the ASI alcohol composite score but not to the ASI drug composite score, while the SDSS cocaine and heroin scales were significantly related to the ASI drug composite score but not the ASI alcohol composite score. All SDSS scales were significantly correlated with substance-specific measures of the consequences of substance use. As evidence of the scale's predictive validity, it was found that SDSS scores were significantly related to length of time to first post-treatment alcohol, cocaine and heroin use (Miele et al., 2000b).

Miele and colleagues (2001) again found support for the use of the SDSS for assessing the severity of the ICD-10 dependence and harmful use diagnoses, however, it appears to perform less well in assessing cannabis disorders compared with other drugs. In a sample of 180 substance users, test-retest reliabilities for the ICD-10 dependence scales ranged from good to excellent for alcohol, cocaine, heroin,

and cannabis (ranging from .69 to .90). Test-retest reliabilities for the SDSS' ICD-10 harmful use scales were also adequate for alcohol, cocaine, and heroin (.54-.79) and in the poor to fair range for cannabis (.39-.40). Internal consistency, diagnostic concordance, and concurrent validity results were comparable to the test-retest findings.

Client groups

As explained above, initial verification studies were conducted with a variety of population groups but its utility in an Australian context has not been systematically explored. Some concerns are held relating to the SDSS cannabis scale as it performed more poorly and inconsistently compared to other scales and was not related to other indicators of dependence. This requires further examination, along with instrument's reliability and validity for assessing sedative and other drug dependence.

Availability/cost

A copy of the SDSS may be requested from:

- Gloria Miele, Research Assessment Associates, Inc., 60 Haven Avenue, Suite 4D, New York NY10032, USA, Tel. (001) 212 781 1678, E-mail: gmm23@columbia.edu.
- http://pubs.niaaa.nih.gov/publications/Assesing%20Alcohol/InstrumentPDFs/67_SDSS.pdf

Scoring, administration and expertise required

The interview takes approximately 30 to 45 minutes to administer depending on the number of substances. However, it was designed to be administered by clinicians with a post-graduate degree and clinical experience with patients with substance abuse or mental disorders. A separate score sheet is used for every substance used in the past 30 days. This results in each substance having four sum scores for usual severity, worst severity, total number of days symptom occurred and total number of days symptom at worst severity. Each scale is scored on a slightly different scoring range. Computerised scoring and interpretation is available.

Severity of Dependence Scale (SDS)

The Severity of Dependence Scale (SDS) (Gossop et al., 1995) is a brief, five-item screening measure of psychological aspects of dependence. It was originally included as the final section in the SODQ. The SDS has been validated across a range of drug using groups, including heroin, cannabis, cocaine, amphetamine and benzodiazepine users (de las Cuevas, Sanz, de la Fuente, Padilla, & Berenguer, 2000; Gossop et al., 1995; Martin, Copeland, Gates, & Gilmour, 2006).

Studies among heroin, amphetamine and cocaine users have shown the SDS to be a reliable measure of psychological dependence. The SDS has been found to have good internal consistency (Cronbach's alpha ranging from .80 to .90) and good test-retest reliability (.89) over a one day interval in a sample of heroin users (Gossop, Best, Marsden, & Strang, 1997; Gossop et al., 1995). The construct validity of the SDS has been supported by significant correlations with behavioural indices of dependence including dose, frequency and duration of use (Darke, Ross, & Hall, 1996). Severity of dependence was also influenced by route of drug administration, with heroin smokers having significantly lower dependence scores than those who injected.

Other studies concerning the SDS for cannabis dependence compared with alcohol and opiate dependence, have been weaker, with the SDS demonstrating only a moderate level of internal consistency (Cronbach's alpha = .72) (Swift, Hall, Didcott, & Reilly, 1998). However, the SDS total score was more strongly related to the belief that cannabis use was a problem than the ICD or DSM-III-R scores. At a cut-off score of 3 as indicative of cannabis dependence as identified by the CIDI, the SDS was found to show sensitivity of 64% and specificity of 82% for cannabis users (Swift et al., 1998).

More recent studies however, have found the SDS to be a brief, valid and reliable screen for cannabis dependence among people with psychosis. The scale has been found to demonstrate high levels of internal consistency and strong construct and concurrent validity in this population. Individuals with a score of 2 on the SDS were nearly 30 times more likely to have DSM-IV cannabis dependence (Hides, Dawe, Young, & Kavanagh, 2007). The SDS was the strongest predictor of DSM-IV cannabis dependence after controlling for other predictor variables. Similarly, Martin and colleagues (2006) found the SDS to have good internal consistency (Cronbach's alpha = 0.83) and test-retest reliability (0.88) in adolescent cannabis users. Total SDS score correlated significantly with frequency of cannabis use and a number of DSM-IV dependence criteria, indicating good concurrent validity. These authors suggested, however, that a cut-off score of 4 was optimal for use as an indicator of cannabis dependence.

In a recent study of Sydney cocaine users, the SDS was found to also have optimal sensitivity (67%) and specificity (93%) for both males and females at a cut-off of ≥3. The authors concluded that the SDS to be a test of high diagnostic utility for the measurement of cocaine dependence, when compared with the CIDI (Kaye & Darke, 2002). Others have reported this cut-off score to also be optimal for alcohol dependence (Lawrinson, Copeland, Gerber, & Gilmour, 2007).

In a sample of 100 regular benzodiazepine users attending a mental health service, de las Cuevas and colleagues (2000) examined the utility of the SDS as a screening instrument for benzodiazepine dependence. Reliability analysis using Cronbach's alpha yielded a value of .81. Using a cut-off score of 6 for problematic benzodiazepine use identified by the CIDI, the SDS demonstrated specificity of 94.2% and a sensitivity of 97.9%.

It has also been found to be a useful instrument to measure treatment outcome and forms a part of the BTOM (Gossop, Stewart, & Marsden, 2007; Lawrinson et al., 2003; Sannibale et al., 2005).

Client groups

The SDS has been validated in a range of translations (Chen et al., 2008; Ferri, Marsden, de Araujo, Laranjeira, & Gossop, 2000; Gu et al., 2008; Steiner, Baumeister, & Kraus, 2008) and used in both adolescent and elderly populations (Cook, Biyanova, Thompson, & Coyne, 2007; Lozano et al., 2008; Thomas & McCambridge, 2008). The tool has been translated (and back translated) into Vietnamese in a Sydney study of heroin users (Swift et al., 1999). The wording of the SDS is straightforward and the concepts appear to be understood by a variety of drug users. It has also been used among Aboriginal and Torres Strait Islanders in research studies (Schlesinger et al., 2007), but its applicability to these populations has not been sufficiently tested.

The utility of the SDS as a measure of cannabis dependence was recently examined among a sample of sample of 153 in-patients with a schizophrenia spectrum disorder in a recent Australian study (Hides, Dawe et al., 2007). The SDS had high internal consistency, with Cronbach's alpha of 0.81 and optimal sensitivity (86%) and specificity (83%) at a cut-off point of ≥2. At this cut -off, the SDS significantly predicted the presence of DSM-IV cannabis dependence using the CIDI (Hides, Dawe et al., 2007).

Availability/cost

The SDS is incorporated into the BTOM it is available at:

- http://ndarc.med.unsw.edu.au/NDARCWeb.nsf/resources/TR_11/$file/TR.156.PDF

The version for cannabis is available at:

- http://ncpic.org.au/assets/downloads/workforce/cannabisinfo/assessment-tools/severity-of-dependence-scale

It is also freely available online at:

- http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2636780
- http://eib.emcdda.europa.eu/html.cfm/index7343EN.html

Scoring, administration and expertise required

The SDS self-report contains only five items and takes less than one minute to complete and one minute to score. Each item is scored on a four-point scale, and no specific training is required for use of the scale.

Craving measures

There are limited measures designed specifically to measure craving and fewer still that measure craving generally across all substances. The majority have been aimed at assessing cocaine craving specifically and many have been insufficiently evaluated to warrant strong recommendation across all groups.

Cocaine Craving Questionnaire (CCQ)

The Cocaine Craving Questionnaire (CCQ) (Tiffany, Singleton, Haertzen, & Henningfield, 1993) is a 45-item questionnaire on cocaine craving. Two versions exist: The Now version which measures current craving for cocaine, and the General version measuring average craving over the preceding week. The 45 items of the CCQ are grouped according to the live content categories: desire to use, intention/planning, anticipation of positive outcome, anticipation of relief for withdrawal/dysphoria, lack of control over use.

Internal consistency of the different scales was moderate to high (.70-.93). Examination of item content, correlations of factors across versions, and external correlates of the factors suggested that both versions were represented by the same hierarchical factor structure. Higher total scores on the CCQ-Now were associated significantly with lower confidence in ability to quit using cocaine, greater frequency of use over the past 6 months, current negative mood, and lifetime use of cocaine (Tiffany et al., 1993).

A shorter, 10-item version has also been developed (CCQ-Brief) and was found to be significantly correlated with the original, the Voris Cocaine Craving Scale, the BDI-II, the BAI and recent drug use. The internal consistency of the CCQ-Brief was strong (alpha = .90). The authors concluded that the CCQ-Brief was a valid and reliable instrument that can be easily administered as a measure of current cocaine craving (Sussner et al., 2006).

Client groups

The CCQ has been used in psychiatric populations (Brown, Nejtek, Perantie, Orsulak, & Bobadilla, 2003; Copersino et al., 2004; Smelson et al., 2002) and in different cultural groups (da Silveira, Fernandes, Silveira, & Jorge, 2006). However, to the author's knowledge there are no validation studies in Australian samples.

Availability/cost

The CCQ-Brief is available in Sussner and colleagues (2006).

Scoring, administration and expertise required

The CCQ is scored on a 7-point Likert-type scale on the extent to which the individual agreed or disagreed with each item. The endpoints of the scale are labeled 'strongly disagree' (1) and 'strongly agree' (7). The 45-item General version of the CCQ was constructed from the Now version by rewriting each item in the past tense. The CCQ requires little training and is administered by the client. The CCQ-Brief takes less time to administer than the full version (less than 5 minutes compared to approximately 10-15 minutes).

Weiss Cocaine Craving Questionnaire

An identically named Cocaine Craving Questionnaire created by Weiss and colleagues (1995; 1997) consisted of only 3-5 questions to measure different aspects of cocaine craving: 1) current intensity, 2) intensity during the previous 24 hours, 3) frequency, 4) responsiveness to drug-related conditioned stimuli, and 5) imagined likelihood of use if in a setting with access to drugs. The scale has a high level of internal consistency, with Cronbach's alpha ranging from a low of 0.82 to a high of 0.94 in hospitalised cocaine users and 0.85 to 0.90 in outpatients and each of the five items showed significant decreases in craving over time. However, the scale showed a lack of predictive validity (Weiss et al., 1995).

The scale was shortened to just three items in a later study (Weiss et al., 2003):

1. Please rate how strong your desire was to use cocaine during the last 24 hours.
2. Please imagine yourself in the environment in which you previously used drugs and/or alcohol. If you were in this environment today, what is the likelihood that you would use cocaine?
3. Please rate how strong your urges are for cocaine when something in the environment reminds you of it.

The two items eliminated were those that correlated most highly (0.78–0.91) with the remaining items and that added no predictive validity to the questionnaire. The omitted items were the following:

1. Please rate how strong your desire for cocaine is right now.
2. Please rate how often you had the urge to use cocaine during the past 24 hours.

Cronbach's alpha for the three-item Cocaine Craving Scale was high (0.78) and the tool was found to predict the relative likelihood of cocaine use during the subsequent week (Weiss et al., 2003).

An amphetamine version of the 3-item Cocaine Craving Scale has also been developed (Shearer, in press).

Availability/cost

The Cocaine Craving Questionnaire is outlined above, and scoring below.

Client groups

To this author's knowledge there are no studies examining the questionnaire's applicability to specific populations.

Scoring, administration and expertise required

No training is required for the use of the scale and it takes less than 5 minutes to complete. Response options ranged from 0 for "no desire/likelihood of use" to 9 for "strong desire/likelihood of use." The composite score was a sum of these three items, ranging from 0 to 27.

Marijuana Craving Questionnaire (MCQ)

The Marijuana Craving Questionnaire (MCQ) (Heishman, Singleton, & Liguori, 2001) is a 47-item self-report measure assessing demographics, drug use history, marijuana quit attempts and current mood. It was based on the CCQ and measures four constructs characterising marijuana craving: Compulsivity (an inability to control marijuana use); Emotionality (use of marijuana in anticipation of relief from withdrawal or negative mood); Expectancy (anticipation of positive outcomes from smoking marijuana) and Purposefulness (intention and planning to use marijuana for positive outcomes).

The subscales were found to have adequate internal consistencies, low to moderate positive intercorrelations and were significantly correlated with marijuana use history and a wide range of single-item measures of craving. The authors concluded that the MCQ is a valid and reliable instrument for assessing marijuana craving in individuals not seeking drug abuse treatment and that marijuana craving can be measured in the absence of withdrawal (Heishman et al., 2001). A follow up study using active imagery of auditorily presented scripts verify and extend the reliability and validity of the MCQ as a multidimensional measurement of marijuana craving (Singleton, Trotman, Zavahir, Taylor, & Heishman, 2002).

Availability/cost

The MCQ is available upon request from:

- Stephen J. Heishman, PhD, National Institute on Drug Abuse (NIDA) Ph: 410-550-1547 Fax: 410-550-1849 E-mail: sheish@intra.nida.nih.gov

Client groups

To this author's knowledge there are no studies examining the questionnaire's applicability to specific populations.

Scoring, administration and expertise required

The MCQ takes approximately 7 minutes to compete, which is adequate for both research and clinical purposes (Heishman et al., 2001). Each item is rated on a seven-point Likert-type scale from "strongly disagree" to "strongly agree". A shorter 12-item version of the instrument also exists.

Penn Alcohol-Craving Scale (PACS)

The Penn Alcohol-Craving Scale (PACS) (Flannery, Volpicelli, & Pettinati, 1999) is a 5-item, single-factor scale that is quickly and easily administered. The first three questions are centered on the frequency, intensity, and duration of thoughts about drinking. The fourth question asks the individual to rate his/her ability to resist drinking if alcohol is available. The final question asks the subject to rate his/her overall average craving for alcohol during the previous week.

The PACS demonstrated excellent internal consistency. Construct validity of the PACS was demonstrated via its convergence with two commonly used measures for assessing craving, the Obsessive Compulsive Drinking Scale and the Alcohol Urge Questionnaire. Lack of correlation between PACS scores and several other non-craving, self-report measures indicate that the PACS also had good discriminant validity. Additional analyses revealed that there were significant differences in craving scores during the initial 3 weeks of the trial among those who did and those who did not relapse during weeks 3-12 indicating good predictive validity (Flannery et al., 1999). A later study found that scores on the PACS were the strongest predictor of drinking during treatment when compared with the Obsessive Compulsive Drinking Scale and the Alcohol Urge Questionnaire.

Availability/cost

The PACS is freely available in Flannery and colleagues (1999), with copyright held by the Journal, alternately, copies can be obtained from:

- Barbara Flannery RTI International 3040 Cornwallis Road, PO Box 12194, Research Triangle Park, NC 27709-1294; Phone: 919-316-3457; Fax: 919-541-6683; E-mail: bflannery@rti.org

Client groups

There is a dearth of empirical investigation of the use of the PACS in different populations. However, it has been used in both inpatient and outpatient care and in an Australian context (Flannery et al., 1999; Richardson et al., 2008; Yoon, Kim, Thuras, Grant, & Westermeyer, 2006).

Scoring, administration and expertise required

Each question is scaled from 0 to 6. It can be administered without training and takes less than 5 minutes to administer, score and interpret.

Conclusion

It is important to stress that the list above is not exhaustive but it nevertheless aims to provide a useful review of the available tools which have utility for screening, assessment and outcome measurement across mental health, D&A and general health/social functioning. Although no tool can fit all needs and requirements, some instruments may be more useful than others depending on a clinician's particular work environment and the client they are working with.

The author was asked to review the measures in terms of their psychometric properties, applicability and ease of use across a range of settings by a range of professionals as well as availability and cost. With these criteria in mind, the measures reviewed in this document that may be particularly useful are the AATOM and the BTOM as global measures, the CAN, the SF-12 and the LSP as general social functioning measures, the AUDIT and ASSIST for problematic drug and alcohol use, and the K10 for general mental health.

The contents of this review document will inform NADA's choice of measures to be included in the NGO Drug and Alcohol and Mental Health Information Management Project.

References

Abbott, R. A., Ploubidis, G. B., Huppert, F. A., Kuh, D., Wadsworth, M. E., & Croudace, T. J. (2006). Psychometric evaluation and predictive validity of Ryff's psychological well-being items in a UK birth cohort sample of women. *Health and Quality of Life Outcomes, 76*(4).

Abram, K. M., Paskar, L. D., Washburn, J. J., & Teplin, L. A. (2008). Perceived barriers to mental health services among youths in detention. *Journal of the American Academy of Child & Adolescent Psychiatry, 47*(3), 301-308.

Adams, N. L. (2004). *The relationship between childhood traumatic experiences and gang-involved delinquent behavior in adolescent boys.* ProQuest Information & Learning: US.

Adamson, S. J., & Sellman, J. D. (2003). A prototype screening instrument for cannabis use disorder: the Cannabis Use Disorders Identification Test (CUDIT) in an alcohol-dependent clinical sample. *Drug and Alcohol Review, 22*(3), 309-315.

Adelekan, M., Green, A., DasGupta, N., Tallack, F., Stimson, G. V., & Wells, B. (1996). Reliability and validity of the opiate treatment index among opioid users in the United Kingdom. *Drug and Alcohol Review, 15*, 261-270.

Adelekan, M., Metrebian, N., Tallack, F., Stimson, G. V., & Shanahan, W. (1996). Who should collect Opiate Treatment Index data in opiate treatment outcome monitoring: Clinic staff or researchers? *Drug and Alcohol Review, 15*(1), 65-71.

Adewuya, A. O. (2005). Validation of the Alcohol Use Disorders Identification Test (AUDIT) as a screening tool for alcohol-related problems among Nigerian university students. *Alcohol and Alcoholism, 40*(6), 575-577.

Adkins, J. W., Weathers, F. W., McDevitt-Murphy, M., & Daniels, J. B. (2008). Psychometric properties of seven self-report measures of posttraumatic stress disorder in college students with mixed civilian trauma exposure. *Journal of Anxiety Disorders, 22*(8), 1393-1402.

Aertgeerts, B., Buntinx, F., Ansoms, S., & Fevery, J. (2001). Screening properties of questionnaires and laboratory tests for the detection of alcohol abuse or dependence in a general practice population.[see comment]. *British Journal of General Practice, 51*(464), 206-217.

Aertgeerts, B., Buntinx, F., Bande-Knops, J., Vandermeulen, C., Roelants, M., Ansoms, S., et al. (2000). The value of CAGE, CUGE, and AUDIT in screening for alcohol abuse and dependence among college freshmen. *Alcoholism: Clinical and Experimental Research, 24*(1), 53-57.

Aertgeerts, B., Buntinx, F., Fevery, J., & Ansoms, S. (2000). Is there a difference between CAGE interview and written CAGE questionnaires? *Alcoholism: Clinical and Experimental Research, 24*(5), 733-736.

Afuwape, S. A., Johnson, S., Craig, T. J. K., Miles, H., Leese, M., Mohan, R., et al. (2006). Ethnic differences among a community cohort of individuals with dual diagnosis in South London. *Journal of Mental Health, 15*(5), 551-567.

Aguilar, G. I., Hidalgo, M. D., Cano, G. R., & Lopez , J. C. (1995). Prospective study of hopelessness in first-episode psychotic patients: Psychometric characteristics of the Beck Hopelessness Scale in this group. *Anales de Psiquiatria, 11*(4), 121-125.

Ahmadi, J., Kampman, K., & Dackis, C. (2006). Outcome predictors in cocaine dependence treatment trials. *The American Journal on Addictions, 15*(6), 434-439.

Akin, A., & Çetin, B. (2007). The Depression Anxiety and Stress Scale (DASS): The study of validity and reliability. *Kuram ve Uygulamada Egitim Bilimleri, 7*(1), 260-268.

Alansari, B. M. (2006). Internal consistency of Arabic adaptation of the Beck Depression Inventory-II with college students in eighteen Arab countries *Social Behavior & Personality 34*, 425-430.

Ali, R., Awwad, E., Babor, T., Bradley, F., Butau, T., Farrell, M., et al. (2002). The Alcohol, Smoking and Substance Involvement Screening Test (ASSIST): Development, reliability and feasibility. *Addiction, 97*(9), 1183-1194.

Allen, J. P., Litten, R. Z., Fertig, J. B., & Babor, T. (1997). A review of research on the Alcohol Use Disorders Identification Test (AUDIT). *Alcoholism: Clinical & Experimental Research, 21*(4), 613-619.

Allen, L., Bala, S., Carthew, R., Daley, S., Doyle, E., Driscoll, P., et al. (1999). Experience and application of HoNOS65+. *Psychiatric Bulletin, 23*(4), 203-206.

Allnutt, S., Wedgwood, L., Wilhelm, K., & Butler, T. (2008). Australian and New Zealand Journal of Psychiatry. *Temperament, substance use and psychopathology in a prisoner population: implications for treatment, 42*, 969-975.

Alsuwaida, A., & Alwahhabi, F. (2006). The Diagnostic Utility of Self-Reporting Questionnaire (SRQ) as a Screening Tool for Major Depression in Hemodialysis Patients. *Saudi Journal of Kidney Disorders and Transplantation, 17*, 503-510.

Alterman, A. I., Brown, L. S., Zaballero, A., & McKay, J. R. (1994). Interviewer severity ratings and composite scores of the ASI: a further look. *Drug and Alcohol Dependence, 34*, 201-209.

Alterman, A. I., McDermot, P. A., Metzger, M., Rutherford, J., Cacciola, J. S., & Brown , I. S., Jr. (1998). New scales to assess change in the Addiction Severity Index for the opioid, cocaine and alcohol dependent. *Psychology of Addictive Behaviors, 12*(4), 233-246.

Alvarez-Rayón, G., Mancilla-Díaz, J. M., Vázquez-Arévalo, R., Unikel-Santoncini, C., Caballero-Romo, A., & Mercado-Corona, D. (2004). Validity of the Eating Attitudes Test: A study of Mexican eating disorders patients. *Eating and Weight Disorders, 9*(4), 243-248.

American Psychiatric Association. (1994). *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*. Washington DC: American Psychiatric Association.

Amin, S., Singh, S. P., Croudace, T., Jones, P., Medley, I., & Harrison, G. (1999). Evaluating the Health of the Nation Outcome Scales. Reliability and validity in a three-year follow-up of first-onset psychosis. *British Journal of Psychiatry, 174*, 399-403.

Andersen, H. S., Sestoft, D., Lillebæk, T., Gabrielsen, G., & Hemmingsen, R. (2002). Validity of the General Health Questionnaire (GHQ-28) in a prison population: Data from a randomized sample of prisoners on remand. *International Journal of Law and Psychiatry, 25*(6), 573-580.

Andresen, R., Caputi, P., & Oades, L. (2006). Stages of recovery instrument: development of a measure of recovery from serious mental illness. *Australian & New Zealand Journal of Psychiatry, 40*(11-12), 972-980.

Andresen, R., Caputi, P., & Oades, L. G. (2000). Interrater reliability of the Camberwell Assessment of Need Short Appraisal Schedule. *Australian and New Zealand Journal of Psychiatry, 34*(5), 856-861.

Andresen, R., Oades, L., & Caputi, P. (2003). The experience of recovery from schizophrenia: towards an empirically-validated stage model. *Australian and New Zealand Journal of Psychiatry, 37*, 586-594.

Andrews, G. (2002). A brief integer scorer for the SF-12: validity of the brief scorer in Australian community and clinic settings. *Australian and New Zealand Journal of Public Health, 26*(6), 508-510.

Andrews, G. (2003a). HoNOS.   Retrieved 17 February, 2009, from
    http://www.crufad.com/phc/honos.htm

Andrews, G. (2003b). K10 Symptom Scale.   Retrieved 8th February, 2009, from
    http://www.crufad.unsw.edu.au/K10/k10info.htm

Andrews, G. (2008). WHO DAS II - Using the World Health Organization Disability Assessment Scale to
    operationalize the concepts in the International Classification of Functioning, Disability and
    Health [Electronic Version]. Retrieved 6th March 2009 from
    http://www.aihw.gov.au/eventsdiary/iidf2006/abstracts_pdf_doc/gavin_andrews.pdf.

Andrews, G., Henderson, S., & Hall, W. (2001). Prevalence, comorbidity, disability and service utilisation.
    Overview of the Australian National Mental Health Survey. *British Journal of Psychiatry, 178*,
    145-153.

Andrews, G., Peters, L., & Teesson, M. (1994). *The measurement of consumer outcome in mental health*.
    Canberra, ACT: Australian Government Publishing Service.

Andrews, G., & Slade, T. (2001). Interpreting scores on the Kessler Psychological Distress Scale (K10).
    Australian and New Zealand Journal of Psychiatry, 25(6), 494-497.

Andrews, G., Teesson, M., Stewart, G., & Hoult, J. (1990). Follow-up of community placement of the
    chronic mentally ill in New South Wales. *Hospital & Community Psychiatry., 41*(2), 184-188.

Andrykowski, M. A., Cordova, M. J., Studts, J. L., & Miller, T. W. (1998). Posttraumatic stress disorder
    after treatment for breast cancer: Prevalence of diagnosis and use of the PTSD Checklist--Civilian
    Version (PCL--C) as a screening instrument. *Journal of Consulting and Clinical Psychology, 66*(3),
    586-590.

Annis, H. M., Sobell, L. C., Ayala-Velazquez, H., Rybakowski, J. K., Sandahl, C., Saunders, B., et al. (1996).
    Drinking-related assessment instruments: Cross-cultural studies. *Substance Use and Misuse, 31*,
    1525-1546.

Antony, M. M., Bieling, P. J., Cox, B. J., Enns, M. W., & Swinson, R. P. (1998). Psychometric properties of
    the 42-item and the 21-item versions of the Depression Anxiety Stress Scales in clinical groups
    and a community sample. *Psychological Assessment, 10*(2), 176-181.

APA. (2006). Practice guideline for the treatment of patients with substance use disorders: Second
    edition. *American Journal of Psychiatry, 163*(8), S1-S82.

Appleby, L., Dyson, V., Altman, E., & Luchins, D. J. (1997). Assessing substance use in multiproblem
    patients: reliability and validity of the Addiction Severity Index in a mental hospital population.
    Journal of Nervous & Mental Disease, 185(3), 159-165.

Argeriou, M., McCarty, D., Mulvey, K., & Daley, M. (1994). Use of the Addiction Severity Index with
    homeless substance abusers. *Journal of Substance Abuse Treatment, 11*(4), 359-365.

Arrindell, W., Ettema, H., Groenman, N., Brook, F., Janssen, I., Slaets, J., et al. (2003). De groeiende
    inbedding van de Nederlandse SCL-90-R*: Psychodiagnostisch gereedschap. [Further Dutch
    experiences with the Symptom Checklist-90 Revised.]. *Psycholoog, 38*(11), 576-582.

Arvidsson, H. (2003). Test-retest reliability of the Camberwell Assessment of Need (CAN). *Nordic Journal
    of Psychiatry, 57*(4), 279-283.

Asarnow, J. R., Baraff, L. J., Berk, M., Grob, C., Devich-Navarro, M., Suddath, R., et al. (2008). Pediatric
    emergency department suicidal patients: Two-site evaluation of suicide ideators, single
    attempters, and repeat attempters. *Journal of the American Academy of Child & Adolescent
    Psychiatry, 47*(8), 958-966.

Ashaye, K., Seneviratna, K., Shergill, S. S., & Orrell, M. (1999). Do the Health of the Nation Outcome
    Scales predict outcome in the elderly mentally ill? A 1-year follow-up study. *Journal of Mental
    Health (UK), 8*(6), 615-620.

Ashaye, O., Mathew, G., & Dhadphale, M. (1997). A comparison of older longstay psychiatric and learning disability inpatients using the Health of the Nation Outcome Scales. *International Journal of Geriatric Psychiatry, 12*(5), 548-552.

Asukai, N., Kato, H., Kawamura, N., Kim, Y., Yamamato, K., Kishimoto, J., et al. (2002). Reliability and validity of the Japanese-language version of the Impact of Event Scale-revised (IES-R-J): Four studies of different traumatic events. *Journal of Nervous and Mental Disease, 190*(3), 175-182.

Audin, K., Margison, F. R., Clark, J. M., & Barkham, M. (2001). Value of HoNOS in assessing patient change in NHS psychotherapy and psychological treatment services. *British Journal of Psychiatry, 178*, 561-566.

Australian Government Department of Health and Ageing. (2002). *National Outcomes and Casemix Collection: Overview of Clinical Measures and Data Items*. Canberra: Commonwealth Department of Health and Ageing.

Australian Government Department of Health and Ageing. (2006). *National Alcohol Strategy 2006-2009: Towards Safer Drinking Cultures*. Canberra: Australian Government Department of Health and Ageing.

Australian Government Department of Health and Ageing. (2007). *Alcohol Treatment Guidelines for Indigenous Australians*. Canberra: Australian Government Department of Health and Ageing.

Babor, T. F., Higgins-Biddle, J. C., Saunders, J. B., & Monteiro, M. G. (2001). *AUDIT: The Alcohol Use Disorders Identification Test Guidelines for Use in Primary Care* (Second ed.): World Health Organization.

Bacskai, E., Rozsa, S., & Gerevich, J. (2005). Clinical experiences with the Maudsley Addiction Profile (MAP) among heroin addicts. *Orvosi Hetilap, 146*(31), 1635-1639.

Baggaley, R. F., Ganaba, R., Filippi, V., Kere, M., Marshall, T., Sombié, I., et al. (2007). Short communication: Detecting depression after pregnancy: the validity of the K10 and K6 in Burkina Faso. *Tropical Medicine & International Health, 12*(10).

Bagley, H., Cordingley, L., Burns, A., Mozley, C. G., Sutcliffe, C., Challis, D., et al. (2000). Recognition of depression by staff in nursing and residential homes. *Journal of Clinical Nursing, 9*(3), 445-450.

Báguena, M. J., Villarroya, E., Beleña, Á., Amelia, D., Roldán, C., & Reig, R. (2001). Propiedades psicométricas de la versión española de la Escala Revisada de Impacto del Estresor (EIE-R). [Psychometric properties of the Spanish version of the Impact of Event Scale-Revised (IES-R).]. *Análisis y Modificación de Conducta, 27*(114), 581-604.

Baillie, A. J. (2005). Predictive gender and education bias in Kessler's psychological distress Scale (K10). *Social Psychiatry and Psychiatric Epidemiology, 40*, 743-748.

Baker, A., Boggs, T. G., & Lewin, T. J. (2001). Randomized controlled trial of brief cognitive-behavioural interventions among regular users of amphetamine.[see comment]. *Addiction, 96*(9), 1279-1287.

Baker, A., Lewin, T., Reichler, H., Clancy, R., Carr, V., Garrett, R., et al. (2002a). Evaluation of a motivational interview for substance use within psychiatric in-patient services. *Addiction, 97*(10), 1329-1337.

Baker, A., Lewin, T. J., Reichler, H., Clancy, R., Carr, V. J., Garrett, R., et al. (2002b). Evaluation of a motivational interview for substance use within psychiatric in-patient services. *Addiction, 87*, 1329-1337.

Ballesteros, R. J., Martinez, S. M., Martin, C. M., Ibarra, G. N., & Bulbena, V. A. (2002). Assessment of the Life Skills Profile and the Brief Psychiatric Rating Scale as predictors of the length of psychiatric hospitalization. *Actas Esplanolas de Psiquiatria, 30*, 225-232.

Barker-Collo, S. L. (2003). Culture and validity of the Symptom Checklist-90-Revised and Profile of Mood States in a New Zealand student sample. *Cultural Diversity and Ethnic Minority Psychology, 9*(2), 185-196.

Barnes, L. L. B., Harp, D., & Jung, W. S. (2002). Reliability generalization of scores on the Spielberger State-Trait Anxiety Inventory. *Educational and Psychological Measurement, 62*(4), 603-618.

Barrowcliff, A., Champney-Smith, J., & McBride, A. J. (1999). Use of a modified version of the Opiate Treatment Index with amphetamine users: Validation and pilot evaluation of a prescribing service. *Journal of Substance Abuse 4*, 98-103.

Bartsch, D. (2007). Prevalence of body dysmorphic disorder symptoms and associated clinical features among Australian university students. *Clinical Psychologist, 11*(1), 16-23.

Bates, L. W., Lyons, J. A., & Shaw, J. B. (2002). Effects of brief training on application of the Global Assessment of Functioning Scale. *Psychological Reports, 91*(3 Pt 1), 999-1006.

Batki, S. L., Leontieva, L., Dimmock, J. A., & Ploutz-Snyder, R. (2008). Negative symptoms are associated with less alcohol use, craving, and "high" in alcohol dependent patients with schizophrenia. *Schizophrenia Research, 105*(1-3), 201-207.

Beaton, D. E., Hogg-Johnson, S., & Bombardier, C. (1997). Evaluating changes in health status: Reliability and responsiveness of five generic health status measures in workers with musculoskeletal disorders. *Journal of Clinical Epidemiology, 50*, 73-93.

Bebbington, P., Brugha, T., Hill, T., Marsden, L., & Window, S. (1999). Validation of the Health of the Nation Outcome Scales. *British Journal of Psychiatry, 174*, 389-394.

Bech, P., Bille, J., Schutze, T., Sondergaard, S., Wiese, M., & Waarst, S. (2003). Health of the Nation Outcome Scales (HoNOS): Implementability, subscale structure and responsiveness in the daily psychiatric hospital routine over the first 18 months. *Nordic Journal of Psychiatry, 57*(4), 285-290.

Beck, A., & Steer, R. (1988). *Beck Hopelessness Scale Manual*. San Antonio, Tex: The Psychological Corporation.

Beck, A., & Steer, R. (1991). *Manual for the Beck Scale for Suicidal Ideation*. San Antonio, Tex: Psychological Corporation.

Beck, A. T., Brown, G., Berchick, R. J., & Stewart, B. L. (1990). Relationship between hopelessness and ultimate suicide: A replication with psychiatric outpatients. *American Journal of Psychiatry, 147*(2), 190-195.

Beck, A. T., Epstein, N., Brown, G., & Steer, R. A. (1988). An inventory for measuring clinical anxiety: Psychometric properties. *Journal of Consulting and Clinical Psychology, 56*(6), 893-897.

Beck, A. T., Kovacs, M., & Weissman, A. (1979). Assessment of suicidal intention: The Scale for Suicide Ideation. *Journal of Consulting and Clinical Psychology, 47*(2), 343-352.

Beck, A. T., & Steer, R. A. (1987). *Beck Depression Inventory: Manual*. USA: Harcourt, Brace, Jovanovich.

Beck, A. T., & Steer, R. A. (1990). *Beck Anxiety Inventory: Manual*. USA: Harcourt, Brace, Jovanovich.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the Beck Depression Inventory-II*. San Antonio, TX: Psychological Corporation.

Beck, A. T., Steer, R. A., & Garbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty five years of evaluation. *Clinical Psychology Review, 8*, 77-100.

Beck, A. T., Steer, R. A., & Ranieri, W. F. (1988). Scale for suicide ideation: Psychometric properties of a self-report version. *Journal of Clinical Psychology, 44*(4), 499-505.

Beck, A. T., Weissman, A., Lester, D., & Trexler, L. (1974). The measurement of pessimism: The Hopelessness Scale. *Journal of Consulting and Clinical Psychology, 42*(6), 861-865.

Beck, J. G., Grant, D. M., Read, J. P., Clapp, J. D., Coffey, S. F., Miller, L. M., et al. (2008). The Impact of Event Scale-Revised: Psychometric properties in a sample of motor vehicle accident survivors. *Journal of Anxiety Disorders, 22*(2), 187-198.

Beckman, H. T., & Lueger, R. J. (1997). Readability of self-report clinical outcome measures. *Journal of Clinical Psychology, 53*(8), 785-789.

Bergman, H., & Källmén, H. (2002). Alcohol use among Swedes and a psychometric evaluation of the Alcohol Use Disorders Identification Test. *Alcohol and Alcoholism, 37*(3), 245-251.

Berlim, M. T., McGirr, A., & Fleck, M. P. (2008). Can sociodemographic and clinical variables predict the quality of life of outpatients with major depression? *Psychiatry Research, 160*(3), 364-371.

Berman, A. H., Bergman, H., Palmstierna, T., & Schlyter, F. (2005). Evaluation of the Drug Use Disorders Identification Test (DUDIT) in Criminal Justice and Detoxification Settings and in a Swedish Population Sample. *European Addiction Research, 11*(1).

Berner, M. M., Kriston, L., Bentele, M., & Härter, M. (2007). The Alcohol Use Disorders Identification Test for detecting at-risk drinking: A systematic review and meta-analysis. *Journal of Studies on Alcohol and Drugs, 68*(3), 461-473.

Berwick, D. M., Murphy, J. M., Goldman, P. A., Ware, J. E., Barsky, A. J., & Weinstein, M. C. (1991). Performance of a five-item mental health screening test. *Medical Care, 29*(2), 169-176.

Best, D., Manning, V., Gossop, M., Gross, S., & Strang, J. (2006). Excessive drinking and other problem behaviours among 14-16 year old schoolchildren. *Addictive Behaviors, 31*(8), 1424-1435.

Beusenberg, M., & Orley, J. (1994). *The Self-Reporting Questionnaire*. Geneva: World Health Organisation.

Bilenberg, N. (2003). Health of the Nation Outcome Scales for Children and Adolescents (HoNOSCA): Results of a Danish field trial. *European Child and Adolescent Psychiatry, 12*(6), 298-302.

Bird, H. R., Canino, G. J., Rubio-Stipec, M., & Ribera, J. C. (1987). Further measures of the psychometric properties of the Children's Global Assessment Scale. *Archives of General Psychiatry, 44*(9), 821-824.

Bird, H. R., Yager, T. J., Staghezza, B., Gould, M. S., Canino, G., & Rubio-Stipec, M. (1990). Impairment in the epidemiological measurement of childhood psychopathology in the community. *Journal of the American Academy of Child and Adolescent Psychiatry, 29*(5), 796-803.

Biron, L. L., Brochu, S., & Desjardins, L. (1995). The issue of drugs and crime among a sample of incarcerated women. *Deviant Behavior, 16*(1), 25-43.

Bischof, G., Rumpf, H. J., Meyer, C., Hapke, U., & John, U. (2005). Influence of psychiatric comorbidity in alcohol dependent subjects in a representative population survey on treatment utilization and natural recovery. *Addiction, 100*(3), 405-413.

Bisson, J., Nadeau, L., & Demers, A. (1999). The validity of the CAGE scale to screen for heavy drinking and drinking problems in a general population survey. *Addiction, 94*(5), 715-722.

Blanchard, E. B., Jones-Alexander, J., Buckley, T. C., & Forneris, C. A. (1996). Psychometric properties of the PTSD Checklist (PCL). *Behaviour Research and Therapy, 34*(8), 669-673.

Blow, F. (1991). Michigan Alcoholism Screening Test-Geriatric Version (MAST-G). Baltimore: University of Michigan Alcohol Research Center.

Bodlund, O., Kullgren, G., Ekselius, L., Lindström, E., & von Knorring, L. (1994). Axis V--Global Assessment of Functioning Scale. Evaluation of a self-report version. *Acta Psychiatrica Scandinavica, 90*(5), 342-347.

Bohn, M. J., Babor, T., & Kranzler, H. R. (1991). Validity of the Drug Abuse Screening Test (DAST-10) in inpatient substance abusers. *Problems of Drug Dependence, 119*, 233-235.

Bohn, M. J., Babor, T. F., & Kranzler, H. R. (1995). The Alcohol Use Disorders Identification Test (AUDIT): validation of a screening instrument for use in medical settings. *Journal of Studies on Alcohol, 56*(4), 423-432.

Bollinger, A. R., Cuevas, C. A., Vielhauer, M. J., Morgan, E. E., & Keane, T. M. (2008). The operating characteristics of the PTSD Checklist in detecting PTSD in HIV+ substance abusers. *Journal of Psychological Trauma, 7*(4), 213-234.

Bonicatto, S., Dew, A. M., & Soria, J. J. (1998). Analysis of the psychometric properties of the Spanish version of the Beck Depression Inventory in Argentina. *Psychiatry Research, 79*, 277-288.

Bonn-Miller, M. O., Vujanovic, A. A., Feldner, M. T., Bernstein, A., & Zvolensky, M. J. (2007). Posttraumatic stress symptom severity predicts marijuana use coping motives among traumatic event-exposed marijuana users. *Journal of Traumatic Stress, 20*(4), 577-586.

Bonsack, C., Borgeat, F., & Lesage, A. (2002). Measuring patients' problems severity and outcomes in a psychiatric sector: A field study with the French version of the Health of Nation Outcome Scales (HoNOS-F) *Annales Medico-Psychologiques, 160*(7), 483-488.

Bonynge, E. R. (1993). Unidimensionality of SCL-90-R scales in adult and adolescent crisis samples. *Journal of Clinical Psychology, 49*(2), 212-215.

Boot, B., Hall, W., & Andrews, G. (1997). Disability, outcome and case-mix in acute psychiatric in-patient units. *British Journal of Psychiatry, 171*, 242-246.

Boulet, J., & Boss, M. W. (1991). Reliability and validity of the Brief Symptom Inventory. *Psychological Assessment, 3*(3), 433-437.

Bourdon, K. H., Goodman, R., Rae, D. S., Simpson, G., & Koretz, D. S. (2005). The Strengths and Difficulties Questionnaire: U.S. normative data and psychometric properties. *Journal of the American Academy of Child & Adolescent Psychiatry, 44*(6), 557-564.

Boyd, J. W. (2007). *The effects of verbal abuse on adolescent males: An attachment theory perspective.* ProQuest Information & Learning: US.

Bradley, A. C., Baker, A., & Lewin, T. J. (2007). Group intervention for coexisting psychosis and substance use disorders in rural Australia: Outcomes over 3 years. *Australian and New Zealand Journal of Psychiatry, 41*(6), 501-508.

Bradley, K. A., Boyd-Wickizer, J., Powell, S. H., & Burman, M. L. (1998). Alcohol screening questionnaires in women: a critical review [see comments]. *JAMA, 280*(2), 166-171.

Brady, M., Sibthorpe, B., Bailie, R., Ball, S., & Sumnerdodd, P. (2002). The feasibility and accaptability of introducing brief intervention for alcohol misuse in an urban Aboriginal medical service. *Drug and Alcohol Review, 21*(4), 375-380.

Brann, P., Coleman, G., & Luk, E. (2001). Routine outcome measurement in a child and adolescent mental health service: An evaluation of HoNOSCA. *Australian and New Zealand Journal of Psychiatry, 35*(3), 370-376.

Brazier, J. E., Harper, R., Jones, N. M. B., O'Cathian, A., Thomas, K. J., Usherwood, T., et al. (1992). Validating the SF-36 health survey questionniare:new outcome measure for primary care. *BMJ, 305*, 160-164.

Breakey, W. R., Calabrese, L., Rosenblatt, A., & Crum, R. M. (1998). Detecting alcohol use disorders in the severely mentally ill. *Community Mental Health Journal, 34*(2), 165-174.

Brewin, C. R. (2005). Systematic Review of Screening Instruments for Adults at Risk of PTSD. *Journal of Traumatic Stress, 18*(1), 53-62.

Brewin, C. R., Rose, S., Andrews, B., Green, J., Tata, P., McEvedy, C., et al. (2002). Brief screening instrument for post-traumatic stress disorder [see comment]. *British Journal of Psychiatry, 181*, 158-162.

Broadbent, M. (2001). Reconciling the information needs of clinicians, managers and commissioners: A pilot project. *Psychiatric Bulletin, 25*(11), 423-425.

Brodey, B. B., Rosen, C. S., Brodey, I. S., Sheetz, B. M., Steinfeld, R. R., & Gastfriend, D. R. (2004). Validation of the Addiction Severity Index (ASI) for internet and automated telephone self-report administration. *Journal of Substance Abuse Treatment, 26*(4), 253-259.

Brooks, R. (2000). The reliability and validity of the Health of the Nation Outcome Scales: Validation in relation to patient derived measures. *Australian and New Zealand Journal Psychiatry, 34*(3), 504-511.

Brophy, C. J., Norvell, N. K., & Kiluk, D. J. (1988). An examination of the factor structure and convergent and discriminant validity of the SCL-90R in an outpatient clinic population. *Journal of Personality Assessment, 52*(2), 334-340.

Brouwer, D., Meijer, R. R., Weekers, A. M., & Baneke, J. J. (2008). On the dimensionality of the Dispositional Hope Scale. *Psychological Assessment, 20*(3), 310-315.

Brown, E. S., Nejtek, V. A., Perantie, D. C., Orsulak, P. J., & Bobadilla, L. (2003). Lamotrigine in patients with bipolar disorder and cocaine dependence. *Journal of Clinical Psychiatry, 64*(2), 197-201.

Brown, R. L., & Rounds, L. A. (1995). Conjoint screening questionnaires for alcohol and drug abuse. *Wisconsin Medical Journal, 94*, 135-140.

Brown, T. A., Chorpita, B. F., Korotitsch, W., & Barlow, D. H. (1997). Psychometric properties of the Depression Anxiety Stress Scales (DASS) in clinical samples. *Behaviour Research and Therapy, 35*, 79-89.

Browne, G., & Courtney, M. (2004). Measuring the impact of housing on people with schizophrenia. *Nursing and Health Sciences, 6*(1), 37-44.

Browne, S., Doran, M., & McGauran, S. (2000). Health of the Nation Outcome Scales (HoNOS): Use in an Irish psychiatric outpatient population. *Irish Journal of Psychological Medicine, 17*(1), 17-19.

Brunet, A., St-Hilaire, A., Jehel, L., & King, S. (2003). Validation of a French version of the Impact of Event Scale-Revised. *The Canadian Journal of Psychiatry / La Revue canadienne de psychiatrie, 48*(1), 56-61.

Brunier, G., Graydon, J., Rothman, B., Sherman, C., & Liadsky, R. (2002). The psychological well-being of renal peer support volunteers. *Journal of Advanced Nursing, 38*(1), 40-49.

Buchsbaum, D. G., Buchanan, R. G., Welsh, J., Centor, R. M., & Schnoll, S. H. (1992). Screening for drinking disorders in the elderly using the CAGE questionnaire. *Journal of the American Geriatric Society, 40*(7), 662-665.

Buckingham, W., Burgess, P., Solomon, S., Pirkis, J., & Eagar, K. (1998). *Developing a Casemix Classification for Mental Health Service*. Canberra: Commonwealth Department of Health and Family Services.

Buckley, T. (2001). A psychometric evaluation of the BDI-II in treatment-seeking substance abusers. *Journal of Substance Abuse Treatment, 20*(3), 197-204.

Buist-Bouwman, M. A., Graaf, R. D., Vollebergh, W. A. M., Alonso, J., Bruffaerts, R., & Ormel, J. (2006). Functional disability of mental disorders and comparison with physical disorders: a study among the general population of six European countries. *Acta Psychiatrica Scandinavica, 113*(6), 492-500.

Bulbena Vilarrasa, A., Fernández de Larrinoa Palacios, P., & Domínguez Panchón, A. I. (1992). Adaptación castellana de la escala LSP (Life Skills Profile). Perfil de habilidades de la vida cotidiana. Estructura y composición factorial. [Castillian adaptation of the LSP (Life Skills Profile): Profile of the skills employed in daily living: Structure and factorial composition.]. *Actas Luso-Espanolas de Neurologia, Psiquiatria y Ciencias Afines, 20*(2), 51-60.

Bullinger, M., Alonso, J., Apolone, G., Leplege, A., Sullivan, M., Wood-Dauphinee, S., et al. (1998). Translating health status questionnaires and evaluating their quality: the IQOLA Project approach. International Quality of Life Assessment. *Journal of Clinical Epidemiology, 51*, 913-923.

Burdine, J. N., Felix, M. R., Abel, A. L., Wiltraut, C. J., & Musselman, Y. J. (2000). The SF-12 as a population health measure: an exploratory examination of potential for application. *Health Services Research, 35*, 885-904.

Burger, G., Calsyn, R., Morse, G., Klinkenberg, W., & Trusty, M. (1997). Factor structure of the Expanded Brief Psychiatric Rating Scale. *Journal of Clinical Psychology, 53*(5), 451-454.

Burger, G. K., Yonker, R. D., Calsyn, R. J., Morse, G. A., & Klinkenberg, W. D. (2003). A confirmatory factor analysis of the Brief Psychiatric Rating Scale in a homeless sample. *International Journal of Methods in Psychiatric Research, 12*(4), 192-196.

Burgés, V., Fernández, A., Autonell, J., Melloni, F., & Bulbena, A. (2007). Adaptación y validación Española del Life Skills Profile, forma abreviada (LSD-20): Un instrumento para valorar las habilidades de la vida cotidiana en contextos clínicos reales. [Spanish adaptation and validation of the brief form of the Life Skills Profile-20: An instrument to assess daily living skills in real clinical settings.]. *Actas Españolas de Psiquiatría, 35*(2), 79-88.

Burlingame, G. M., Seaman, S., Johnson, J. E., Whipple, J., Richardson, E., Rees, F., et al. (2006). Sensitivity to change of the Brief Psychiatric Rating Scale-Extended (BPRS-E): An item and subscale analysis. *Psychological Services, 3*(2), 77-87.

Burns, A., Beevor, A. S., Lelliott, P., Wing, J. K., Blakey, A., Orrell, M., et al. (1999). Health of the Nation Outcome Scales for elderly people (HoNOS 65+). *British Journal of Psychiatry, 174*, 424-427.

Burton, L. C. (2006). *ADHD, gender role discrepancy, and well-being in adult women.* ProQuest Information & Learning: US.

Bush, K. R., Kivlahan, D. R., Davis, T. M., Dobie, D. J., Sporleder, J. L., Epler, A. J., et al. (2003). The TWEAK is weak for alcohol screening among female veterans affairs outpatients. *Alcoholism: Clinical and Experimental Research, 27*(12), 1971-1978.

Butler, S. F., Budman, S. H., Goldman, R. J., Newman, F. J., Beckley, K. E., Trottier, D., et al. (2001). Initial validation of a computer-administered Addiction Severity Index: The ASI-MV. *Psychology of Addictive Behaviors, 15*(1), 4-12.

Butler, T., Andrews, G., Allnutt, S., Sakashita, C., Smith, N. E., & Basson, J. (2006). Mental disorders in Australian prisoners: a comparison with a community sample. *Australian and New Zealand Journal of Psychiatry ; :, 40*, 272-276.

Byrne, B. M., Baron, P., & Balev, J. (1996). The Beck Depression Inventory: Testing for its factorial validity and invariance across gender for Bulgarian non-clinical adolescents. *Personality and Individual Differences, 21*(5), 641-651.

Byrne, B. M., Baron, P., Larsson, B., & Melin, L. (1996). Measuring depression for Swedish nonclinical adolescents: Factorial validity and equivalence of the Beck Depression Inventory across gender. *Scandinavian Journal of Psychology, 37*(1), 37-45.

Byrne, B. M., Stewart, S. M., & Lee, P. W. H. (2004). Validating the Beck Depression Inventory II for Hong Kong Community Adolescents. *International Journal of Testing, 4*, 199-216.

Cacciola, J. S., Alterman, A. I., McLellan, A. T., Lin, Y.-T., & Lynch, K. G. (2007). Initial evidence for the reliability and validity of a "Lite" version of the Addiction Severity Index. *Drug and Alcohol Dependence, 87*(2-3), 297-302.

Cahill, M. A., Adinoff, B., Hosig, H., Muller, K., & Pulliam, C. (2003). Motivation for treatment preceding and following a substance abuse program. *Addictive Behaviors, 28*(1), 67-79.

Calsyn, D. A., Saxon, A. J., Bush, K. R., Howell, D. N., Baer, J. S., Sloan, K. L., et al. (2004). The Addiction Severity Index medical and psychiatric composite scores measure similar domains as the SF-36 in substance-dependent veterans: Concurrent and discriminant validity. *Drug & Alcohol Dependence, 76*, 165-171.

Cameron, I., Eagles, J., Howie, F., Andrew, J. E., Crawford, J. R., Kohler, C., et al. (2001). Preliminary validation of a UK-modified version of the BASIS-32. *International Journal of Psychiatry in Clinical Practice, 5*(1), 41-48.

Canals, J., Carbajo, G., & Fernández-Ballart, J. (2002). Discriminant validity of the Eating Attitudes Test according to American Psychiatric Association and World Health Organization criteria of eating disorders. *Psychological Reports, 91*(3,Pt2), 1052-1056.

Canino, G., Shrout, P. E., Rubio-Stipec, M., Bird, H. R., Bravo, M., Ramirez, R., et al. (2004). The DSM-IV rates of child and adolescent disorders in Puerto Rico: prevalence, correlates, service use, and the effects of impairment. *Archives of General Psychiatry, 61*(1), 85-93.

Capron, C., Thérond, C., & Duyme, M. (2007). Psychometric properties of the French version of the self-report and teacher Strengths and Difficulties Questionnaire (SDQ). *European Journal of Psychological Assessment, 23*(2), 79-88.

Caraveo Anduaga, J. J., Martinez, N. A., Saldivar, G., Lopez, J., & Saltijeral, M. T. (1998). Performance of the GHQ-12 in relation to current and lifetime CIDI psychiatric diagnoses:GHQ-12 in relation to CIDI diagnoses. *Salud Mental, 21*(4), 1-11.

Cardona, M., Jorm, L., Williamson, M., & Chey, T. (1995). *The Blacktown Health Survey of People from Non English Speaking Background*. Sydney Western Sector Public Health Unit.

Carey, K. B. (1997). Challenges in assessing substance use patterns in persons with comorbid mental and addictive disorders. *NIDA Research Monograph, 172*, 16-32.

Carey, K. B., Carey, M. P., & Chandra, P. S. (2003). Psychometric evaluation of the Alcohol Use Disorders Identification Test and short Drug Abuse Screening Test with psychiatric patients in India. *Journal of Clinical Psychiatry, 64*(7), 767-774.

Carey, K. B., Cocco, K. M., & Correia, C. J. (1997). Reliability and validity of the Addiction Severity Index among outpatients with severe mental illness. *Psychological Assessment, 9*(4), 422-428.

Carey, K. B., & Correia, C. J. (1998). Severe mental illness and addictions: assessment considerations. *Addictive Behaviors, 23*(6), 735-748.

Carey, M. P., Faulstich, M. E., & Carey, T. C. (1994). Assessment of anxiety in adolescents: Concurrent and factorial validities of the Trait Anxiety scale of Spielberger's State-Trait Anxiety Inventory for Children. *Psychological Reports, 75*(1, Pt 1), 331-338.

Carmody, D. P. (2005). Psychometric characteristics of the Beck Depression Inventory-II with college students of diverse ethnicity. *International Journal of Psychiatry in Clinical Practice, 9*(1), 22-28.

Carney, M. A., Tennen, H., Affleck, G., del-Boca, F. K., & Kranzler, H. R. (1998). Levels and patterns of alcohol consumption using timeline follow-back, daily diaries and real-time "electronic interviews." *Journal of Studies on Alcohol, 59*(4), 447-454.

Carpiniello, B., Pinna, M., Carta, M. G., & Orrù, M. G. (2006). Reliability, validity and acceptability of the WHOQOL-Bref in a sample of Italian psychiatric outpatients. *Epidemiologia e Psichiatria Sociale, 15*(3), 228-232.

Carrigan, M. H., Drobes, D. J., & Randall, C. L. (2004). Attentional Bias and Drinking to Cope With Social Anxiety. *Psychology of Addictive Behaviors, 18*(4), 374-380.

Cash, T. F., Phillips, K. A., Santos, M. T., & Hrabosky, J. I. (2004). Measuring "negative body image": Validation of the Body Image Disturbance Questionnaire in a nonclinical population. *Body Image, 1*(4), 363-372.

Cassiday, K. L., McNally, R. J., VandeCreek, L., Knapp, S., & Jackson, T. L. (1992). The "Impact of Event Scale-Revised" for assessing rape victims with post-traumatic stress disorder. In *Innovations in clinical practice: A source book, Vol. 11.* (pp. 325-333): Professional Resource Press/Professional Resource Exchange: Sarasota.

Cassidy, C. M., Schmitz, N., & Malla, A. (2008). Validation of the alcohol use disorders identification test and the drug abuse screening test in first episode psychosis. *The Canadian Journal of Psychiatry / La Revue canadienne de psychiatrie, 53*(1), 26-33.

Cassileth, B., Lusk, E., & Strouse, T. (1984). Psychosocial status in chronic illness. *New England Journal of Medicine, 311*, 506-511.

Castro, M. d. G., Oliveira, M. d. S., Miguel, A. C., & Araujo, R. B. (2007). WHOQOL-BREF psychometric properties in a sample of smokers. *Revista Brasileira de Psiquiatria, 29*(3), 254-257.

Chan, A. W. K. (1993). Recent developments in detection and biological indicators of alcoholism. *Drugs and Society, 8*(1), 31-67.

Chan, Y.-F., Dennis, M. L., & Funk, R. R. (2008). Prevalence and comorbidity of major internalizing and externalizing problems among adolescents and adults presenting to substance abuse treatment. *Journal of Substance Abuse Treatment, 34*(1), 14-24.

Chan, Y.-F., Godley, M. D., Godley, S. H., & Dennis, M. L. (2009). Utilization of mental health services among adolescents in community-based substance abuse outpatient clinics. *The Journal of Behavioral Health Services & Research, 36*(1), 35-51.

Chang, E. C. (2006). Perfectionism and dimensions of psychological well-being in a college student sample: A test of a stress-mediation model. *Journal of Social & Clinical Psychology, 25*(9), 1001-1022.

Chang, G. (2001). Alcohol-screening instruments for pregnant women. *Alcohol Research & Health, 25*(3), 204-209.

Chang, G., McNamara, T., Orav, E. J., & Wilkins-Haug, L. (2006). Identifying risk drinking in expectant fathers. *Birth: Issues in Perinatal Care, 33*(2), 110-116.

Chang, G., Wilkins Haug, L., Berman, S., Goetz, M. A., Behr, H., & Hiley, A. (1998). Alcohol use and pregnancy: improving identification. *Obstetrics and Gynecology, 91*(6), 892-898.

Chang, H. (2005). Dimensions of the Chinese Beck Depression Inventory-II in a university sample. *Individual Differences Research, 3*(3), 193-199.

Chantarujikapong, S. I., Smith, E. M., & Fox, L. W. (1997). Comparison of the Alcohol Dependence Scale and Diagnostic Interview Schedule in homeless women. *Alcoholism: Clinical and Experimental Research, 21*(4), 586-595.

Chellappa, S. L., & Araújo, J. F. (2007). Sleep disorders and suicidal ideation in patients with depressive disorder. *Psychiatry Research, 153*(2), 131-136.

Chen, K.-H., Wu, C.-H., & Yao, G. (2006). Applicability of the WHOQOL-BREF on early adolescence. *Social Indicators Research, 79*(2), 215-234.

Chen, V. C.-H., Chen, H., Lin, T.-Y., Chou, H.-H., Lai, T.-J., Ferri, C. P., et al. (2008). Severity of heroin dependence in Taiwan: Reliability and validity of the Chinese version of the Severity of Dependence Scale (SDS-super([Ch])). *Addictive Behaviors, 33*(12), 1590-1593.

Cheng, S.-T., & Chan, A. C. M. (2005). Measuring Psychological well-being in the Chinese. *Personality and Individual Differences, 38*(6), 1307-1316.

Cheng, S. K.-W., Wong, C.-W., Wong, K.-C., Chong, G. S.-C., Wong, M. T.-P., Chang, S. S.-Y., et al. (2002). [A study of psychometric properties, normative scores, and factor structure of the Beck Anxiety Inventory--the Chinese version.]. *Chinese Journal of Clinical Psychology, 10*(1), 4-6.

Chengzhi, Z., Xiangdong, W., & Lianxi, C. (2003). [Reliability and Validity of the Impact of Event Scale in Chinese Traumatic Event Suffers.]. *Chinese Mental Health Journal, 17*(10), 679-681.

Cherpitel, C. J. (1995). Analysis of cut points for screening instruments for alcohol problems in the emergency room. *Journal of Studies on Alcohol, 56*(6), 695-700.

Cherpitel, C. J. (1997). Comparison of screening instruments for alcohol problems between black and white emergency room patients from two regions of the country. *Alcoholism: Clinical & Experimental Research, 21*(8), 1391-1397.

Cherpitel, C. J. (1998). Differences in performance of screening instruments for problem drinking among blacks, whites and Hispanics in an emergency room population. *Journal of Studies on Alcohol, 59*(4), 420-426.

Cherpitel, C. J. (1999). Screening for alcohol problems in the U.S. general population: A comparison of the CAGE and TWEAK by gender, ethnicity, and services utilization. *Journal of Studies on Alcohol, 60*(5 Sep), 705-711.

Chisolm, T. H., Abrams, H. B., McArdle, R., Wilson, R. H., & Doyle, P. J. (2005). The WHO-DAS II: Psychometric Properties in the Measurement of Functional Health Status in Adults With Acquired Hearing Loss. *Trends in Amplification, 9*(3), 111-126.

Chopra, P. K., Couper, J. W., & Herrman, H. (2004). The assessment of patients with long-term psychotic disorders: Application of the WHO Disability Assessment Schedule II. *Australian and New Zealand Journal of Psychiatry, 38*(9), 753-759.

Choudry, I. Y., & Mumford, D. B. (1992). A pilot study of eating disorders in Mirpur (Pakistan) using an Urdu version of the Eating Attitudes Test. *International Journal of Eating Disorders, 11*(3), 243-251.

Chow, J. C.-C., Snowden, L. R., & McConnell, W. (2001). A confirmatory factor analysis of the BASIS-32 in racial and ethnic samples. *Journal of Behavioral Health Services and Research, 28*(4), 400-411.

Christie, G., Marsh, R., Sheridan, J., Wheeler, A., Suaalii-Sauni, T., Black, S., et al. (2007). The Substances and Choices Scale (SACS) - The development and testing of a new alcohol and other drug screening and outcome measurement instrument for young people. *Addiction, 102*(9), 1390-1398.

Chung, M. C., Berger, Z., Jones, R., & Rudd, H. (2006). Posttraumatic stress disorder and general health problems following myocardial infarction (Post-MI PTSD) among older patients: The role of personality. *International Journal of Geriatric Psychiatry, 21*(12), 1163-1174.

Chung, T., Colby, S. M., Barnett, N. P., Rohsenow, D. J., Spirito, A., & Monti, P. M. (2000). Screening adolescents for problem drinking: Performance of brief screens against DSM-IV alcohol diagnoses. *Journal of Studies on Alcohol, 61*(4), 579-587.

Claussen, B., & Aasland, O. (1993). The Alcohol Use Disorders Identification Test (AUDIT) in a routine health examination of long-term unemployed. *Addiction, 88*, 363-368.

Clayman, S. L. (2005). *The relationship among disclosure, internalized homophobia, religiosity, and psychological well-being in a lesbian population.* ProQuest Information & Learning: US.

Cocco, K. M., & Carey, K. B. (1998). Psychometric properties of the Drug Abuse Screening Test in psychiatric outpatients. *Addiction, 94*(2), 269-281.

Cochrane Brink, K. A., Lofchy, J. S., & Sakinofsky, I. (2000). Clinical rating scales in suicide risk assessment. *General Hospital Psychiatry, 22*(6), 445-451.

Coffey, M., Jones, S. H., & Thornicroft, G. (1996). A brief mental health outcome scale: relationships between scale scores and diagnostic/sociodemographic variables in the long-term mentally ill. *Journal of Psychiatric and Mental Health Nursing, 3*(2), 89-93.

Coffey, S. F., Dansky, B. S., Falsetti, S. A., Saladin, M. E., & Brady, K. T. (1998). Screening for PTSD in a substance abuse sample: psychometric properties of a modified version of the PTSD Symptom Scale Self-Report. Posttraumatic stress disorder. *Journal of Traumatic Stress, 11*(2), 393-399.

Coffey, S. F., Schumacher, J. A., Brady, K. T., & Cotton, B. D. (2007). Changes in PTSD symptomatology during acute and protracted alcohol and cocaine abstinence. *Drug and Alcohol Dependence, 87*(2-3), 241-248.

Coffey, S. R., Gudmundsdottir, B., Beck, J. G., Palyo, S. A., & Miller, L. (2006). Screening for PTSD in Motor Vehicle Accident Survivors Using the PSS-SR and IES. *Journal of Traumatic Stress, 19*(1), 119-128.

Cohen, L., Hack, T. F., de Moor, C., Katz, J., & Goss, P. E. (2000). The effects of type of surgery and time on psychological adjustment in women after breast cancer treatment. *Annals of Surgical Oncology, 7*(6), 427-434.

Comfort, M., & Kaltenbach, K. A. (1996). The psychosocial history: an interview for pregnant and parenting women in substance abuse treatment and research. In E. R. Rahdert (Ed.), *Treatment for drug-exposed women and their children: Advances in research methodology. NIDA Research Monograph 165* (pp. 123-142). Rockville, MD: U.S. Department of Health and Human Services.

Comfort, M., Zanis, D. A., Whiteley, M. J., Kelly Tyler, A., & Kaltenbach, K. A. (1999). Assessing the needs of substance abusing women: Psychometric data on the psychosocial history. *Journal of Substance Abuse Treatment, 17*, 79-83.

Conigliaro, J., Kraemer, K., & McNeil, M. (2000). Screening and identification of older adults with alcohol problems in primary care. *Journal of Geriatric Psychiatry and Neurology, 13*(3), 106-114.

Conigrave, K. M., & Elvy, G. (1998). An Australian modification of the popular AUDIT questionnaire [Letter]. *Drug and Alcohol Review, 17*(2), 227-228.

Conigrave, K. M., Hall, W. D., & Saunders, J. B. (1995). The AUDIT questionnaire: Choosing a cut-off score. *Addiction, 90*(10), 1349-1356.

Conigrave, K. M., Saunders, J. B., & Reznik, R. B. (1995). Predictive capacity of the AUDIT questionnaire for alcohol-related harm. *Addiction, 90*(11), 1479-1485.

Conley, T. B. (2001). Construct validity of the MAST and AUDIT with multiple offender drunk drivers. *Journal of Substance Abuse Treatment, 20*(4), 287-295.

Conley, T. B. (2002). *Construct validity of the SAD-Q, AUDIT and MAST with multiple-offender drunk drivers.* ProQuest Information & Learning: US.

Connor, J. P., Grier, M., Feeney, G. F. X., & Young, R. M. (2007). The validity of the Brief Michigan Alcohol Screening Test (bMAST) as a problem drinking severity measure. *Journal of Studies on Alcohol and Drugs, 68*(5), 771-779.

Conrad, M. L. I. (2004). *The relation between interpersonal trauma and perceived health in women: A structural model of the mediation of the PTSD symptom clusters and depressive symptoms.* ProQuest Information & Learning: US.

Conroy, M., Menard, W., Fleming-Ives, K., Modha, P., Cerullo, H., & Phillips, K. A. (2008). Prevalence and clinical characteristics of body dysmorphic disorder in an adult inpatient setting. *General Hospital Psychiatry, 30*(1), 67-72.

Cook, J. M., Biyanova, T., Thompson, R., & Coyne, J. C. (2007). Older primary care patients' willingness to consider discontinuation of chronic benzodiazepines. *General Hospital Psychiatry, 29*(5), 396-401.

Cook, J. M., Elhai, J. D., & Areán, P. A. (2005). Psychometric properties of the PTSD Checklist with older primary care patients. *Journal of Traumatic Stress, 18*(4), 371-376.

Copeland, J., Gilmour, S., Gates, P., & Swift, W. (2005). The Cannabis Problems Questionnaire: Factor structure, reliability, and validity. *Drug and Alcohol Dependence, 80*(3), 313-319.

Copeland, J., Rush, B., Reid, A., Clement, N., & Conroy, A. (2000). *Alcohol and Other Drug Treatment: Predictors of Outcome and Routine Monitoring Systems*. Sydney: National Drug and Alcohol Research Centre, Monograph no. 45.

Copeland, J., Swift, W., & Rees, V. (2001). Clinical profile of participants in a brief intervention program for cannabis use disorder. *Journal of Substance Abuse Treatment, 20*(1), 45-52.

Copersino, M. L., Serper, M. R., Vadhan, N., Goldberg, B. R., Richarme, D., Chou, J. C. Y., et al. (2004). Cocaine craving and attentional bias in cocaine-dependent schizophrenic patients. *Psychiatry Research, 128*(3), 209-218.

Çorapçioglu, A., Yargiç, I., Geyran, P., & Kocabasoglu, N. (2006). "Olaylarin etkisi ölçegi" (IES-R) Türkçe versiyonunun geçerlilik ve güvenilirligi. [Validity and reliability of Turkish Version of "Impact of Event Scale-Revised" (IES-R).]. *Yeni Symposium: psikiyatri, nöroloji ve davranis bilimleri dergisi, 44*(1), 14-22.

Corrigan, P. W., Giffort, D., Rashid, E., Leary, M., & Okeke, I. (1999). Recovery as a psychological construct. *Community Mental Health Journal, 35*(3), 231-239.

Corrigan, P. W., Salzer, M., Ralph, R. O., Sangster, Y., & Keck, L. (2004). Examining the factor structure of the recovery assessment scale. *Schizophrenia Bulletin, 30*(4), 1035-1041.

Corse, S. J., Hirschinger, N. B., & Zanis, D. (1995). The use of the Addiction Severity Index with people with severe mental illness. *Psychiatric Rehabilitation Journal, 19*(1), 9-18.

Cosden, M., Panteleakos, F., Gutierrez, L., Barazani, S., & Gottheil, E. (2004). Strength-Based Assessment of Adolescents Who Abuse Drugs: Implications for Helping High-Risk Youth. *California School Psychologist, 9*, 115-126.

Craig, R. J., & Olson, R. E. (2004). Predicting methadone maintenance treatment outcomes using the Addiction Severity Index and the MMPI-2 content scales (negative Treatment indicators and cynism scales). *American Journal of Drug and Alcohol Abuse, 30*(4), 823-839.

Crawford, J. R., & Henry, J. D. (2003). The Depression Anxiety Stress Scales (DASS): Normative data and latent structure in a large non-clinical sample. *British Journal of Clinical Psychology, 42*(2), 111-131.

CRCAH. (2008). *CRC for Aboriginal Health annual report 2007-2008: Aboriginal Solutions to Aboriginal Health*. Darwin: The Cooperative Research Centre for Aboriginal Health (CRCAH).

Creamer, M., Bell, R., & Failla, S. (2003). Psychometric properties of the Impact of Event Scale-Revised. *Behaviour Research and Therapy, 41*(12), 1489-1496.

Creamer, M., Foran, J., & Bell, R. (1995). The Beck Anxiety Inventory in a non-clinical sample. *Behaviour Research and Therapy, 33*(4), 477-485.

Crook, G. M., Oei, T. P. S., & Young, R. M. (1994). Structure of the MAST with an Australian sample of alcoholics. *Drug and Alcohol Review, 13*, 41-46.

Croton, G. (2007). *Screening for and assessment of co-occurring substance use and mental health disorders by Alcohol & Other Drug and Mental Health Services*. Victoria: Victorian Dual Diagnosis Initiative Advisory Group.

Cury, C. R., & Golfeto, J. H. (2003). Strengths and difficulties questionnaire (SDQ): A study of school children in Ribeirao Preto. *Revista brasileira de psiquiatria Sao Paulo, Brazil, 25*(3), 139-145.

Cuthberston, B. H., Scott, J., Strachan, M., Kilonzo, M., & Vale, L. (2005). Quality of life before and after intensive care. *Anaesthesia, 60*, 332-339.

Cyr, J. J., McKenna Foley, J. M., & Peacock, E. (1985). Factor structure of the SCL-90-R: Is there one? *Journal of Personality Assessment, 49*(6), 571-578.

d'Acremont, M., & Linden, M. V. d. (2008). Confirmatory factor analysis of the Strengths and Difficulties Questionnaire in a community sample of French-speaking adolescents. *European Journal of Psychological Assessment, 24*(1), 1-8.

da Silva Lima, A. F. B., Fleck, M., Pechansky, F., de Boni, R., & Sukop, P. (2005). Psychometric properties of the World Health Organization Quality of Life instrument (WHOQoL-BREF) in alcoholic males: A pilot study. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care & Rehabilitation, 14*(2), 473-478.

da Silveira, D. X., Fernandes, M., Silveira, E. D., & Jorge, M. R. (2006). Cocaine craving questionnaire: Assessing craving among cocaine users in Brazil. *Psychiatry Research, 142*(2-3), 257-259.

Daeppen, J. B., Yersin, B., Landry, U., Pecoud, A., & Decrey, H. (2000). Reliability and validity of the Alcohol Use Disorders Identification Test (AUDIT) imbedded within a general health risk screening questionnaire: Results of a survey in 332 primary care patients. *Alcoholism: Clinical and Experimental Research, 24*(5), 659-665.

Daig, I., Burkert, S., Albani, C., Martin, A., & Brähler, E. (2008). Zusammenhang zwischen körperdysmorphen beschwerden, körperbild und selbstaufmerksamkeit an einer repräsentativen stichprobe. [Associations between body dysmorphic symptoms, body image and self-consciousness in a representative population sample.]. *Psychotherapie Psychosomatik Medizinische Psychologie, 58*(1), 16-22.

Darke, S., Hall, W., Wodak, A., Heather, N., & Ward, J. (1992). Development and validation of a multi-dimensional instrument for assessing outcome of treatment among opiate users: The Opiate Treatment Index. *British Journal of Addiction, 87*, 733-742.

Darke, S., Ross, J., & Hall, W. (1996). Overdose among heroin users in Sydney, Australia: I. Prevalence and correlates of non-fatal overdose. *Addiction, 91*(3), 405-411.

Darke, S., Ward, J., Hall, W., Heather, N., & Wodak, A. (1991). *The opiate treatment index (OTI) manual* (Technical Report No. 11). Sydney, Australia: National Drug and Alcohol Research Centre.

Davé, S., Nazareth, I., Senior, R., & Sherr, L. (2008). A comparison of father and mother report of child behaviour on the Strengths and Difficulties Questionnaire. *Child Psychiatry & Human Development, 39*(4), 399-413.

Davidson, R. (1986). Assessment of the alcohol dependence syndrome: A review of self-report screening questionnaires. *British Journal of Addiction, 26*, 243-255.

Davidson, R., Bunting, B., & Raistrick, D. (1989). The homogeneity of the alcohol dependence syndrome: A factorial analysis of the SADD questionnaire. *British Journal of Addiction, 84*, 907-915.

Davidson, R., & Raistrick, D. (1986). The validity of the Short Alcohol Dependence Data (SADD) questionnaire: A short self-report questionnaire for the assessment of alcohol dependence. *British Journal of Addiction, 81*, 217-222.

Davison, M. L., Bershadsky, B., Bieber, J., Silversmith, D., Maruish, M. E., & Kane, R. L. (1997). Development of a brief, multidimensional, self-report instrument for treatment outcomes assessment in psychiatric settings: Preliminary findings Erratum *Assessment, 4*(4), 419.

Dawe, S., Loxton, N. J., Hides, L., Kavanagh, D. J., & Mattick, R. P. (2002). *Review of diagnostic and screening for alcohol and other drug use and other psychiatric disorders*. Canberra: Commonwealth Department of Health and Aged Care.

Dawe, S., Seinen, A., & Kavanagh, D. (2000). An examination of the utility of the AUDIT in people with schizophrenia. *Journal of Studies on Alcohol., 61*(5), 744-750.

Daza, P., Novy, D. M., Stanley, M. A., & Averill, P. (2002). The Depression Anxiety Stress Scale-21: Spanish translation and validation with a Hispanic sample. *Journal of Psychopathology and Behavioral Assessment, 24*(3), 195-205.

de Almeida, S. P., & Silva, M. T. A. (2005). Characteristics of Ecstasy Users in São Paulo, Brazil. *Substance Use & Misuse, 40*(3), 395-404.

de Beurs, E., Van Dyck, R., Marquenie, L. A., Lange, A., & Blonk, R. W. B. (2001). De DASS: Een vragenlijst voor het meten van depressie, angst en stress. [The DASS: A questionnaire for the measurement of depression, anxiety, and stress.]. *Gedragstherapie, 34*(1), 35-53.

de Beurs, E., Wilson, K. A., Chambless, D. l., Goldstein, A. J., & Feske, U. (1997). Convergent and divergent validity of the Beck Anxiety Inventory for patients with panic disorder and agoraphobia. *Depression and Anxiety, 6*(4), 140-146.

De Jong, C., A., Willems, J., Schippers, G., M., & Hendricks, V., M. (1995). The Addiction Severity Index: reliability and validity in a Dutch alcoholic population. *International Journal of the Addictions, 30*(5), 605-616.

de las Cuevas, C., Sanz, E. J., de la Fuente, J. A., Padilla, J., & Berenguer, J. C. (2000). The Severity of Dependence Scale (SDS) as screening test for benzodiazepine dependence: SDS validation study. *Addiction, 95*(2), 245-250.

De Silva, P., Jayawardana, P., & Pathmeswaran, A. (2008). Concurrent validity of the Alcohol Use Disorders Identification Test (AUDIT). *Alcohol and Alcoholism, 43*(1), 49-50.

Deering, D. E. A., & Sellman, J. D. (1996). An inter-rater reliability study of the Opiate Treatment Index. *Drug and Alcohol Review, 15*(1), 57-63.

Degenhardt, L., & Hall, W. (2001). The association between psychosis and problematical drug use among Australian adults: findings from the National Survey of Mental Health and Well-Being. *Psychological Medicine., 31*(4), 659-668.

Degenhardt, L., Hall, W., Korten, A., & Jablensky, A. (2005). *Use of brief screening instrument for psychosis: Results of a ROC analysis. Technical Report No. 210*. Sydney: National Drug and Alcohol Research Centre.

Degenhardt, L., Hall, W., & Lynskey, M. (2001). Alcohol, cannabis and tobacco use among Australians: a comparison of their associations with other drug use and use disorders, affective and anxiety disorders, and psychosis. *Addiction., 96*(11), 1603-1614.

Degenhardt, L. J., Conigrave, K. M., Wutzke, S. E., & Saunders, J. B. (2001). The validity of an Australian modification of the AUDIT questionnaire. *Drug and Alcohol Review, 20*, 143-154.

DeHart, S. S., & Hoffman, N. G. (1997). Screening and diagnosis: Alcohol use disorders in older adults. In A. M. Gurnack (Ed.), *Older adults' misuse of alcohol, medicines, and other drugs: research and practice issues* (pp. 25-53). New York: Springer.

DeMarce, J. M., Burden, J. L., Lash, S. J., Stephens, R. S., & Grambow, S. C. (2007). Convergent validity of the Timeline Followback for persons with comorbid psychiatric disorders engaged in residential substance use treatment. *Addictive Behaviors, 32*(8), 1582-1592.

Dennis, M., Godley, S. H., Diamond, G., Tims, F. M., Babor, T., Donaldson, J., et al. (2004). The Cannabis Youth Treatment (CYT) Study: Main findings from two randomized trials. *Journal of Substance Abuse Treatment, 27*(3), 197-213.

Dennis, M., Titus, J. C., Diamond, G., Donaldson, J., Godley, S. H., Tims, F. M., et al. (2002). The Cannabis Youth Treatment (CYT) experiment: Rationale, study design, and analysis plans. *Addiction, 97*, S16-S34.

Dennis, M. L., Chan, Y.-F., & Funk, R. (2006). Development and validation of the GAIN Short Screener (GSS) for psychopathology and crime/violence among adolescents and adults. *American Journal on Addictions, 15*(Suppl. 1), 80-91.

Dennis, M. L., Funk, R., Godley, S. H., Godley, M. D., & Waldron, H. (2004). Cross-validation of the alcohol and cannabis use measures in the Global Appraisal of Individual Needs (GAIN) and Timeline

Followback (TLFB; Form 90) among adolescents in substance abuse treatment. *Addiction, 99*(Suppl2), 120-128.

Dennis, M. L., Scott, C. K., & Funk, R. (2003). An experimental evaluation of recovery management checkups (RMC) for people with chronic substance use disorders. *Evaluation and Program Planning, 26*, 339-352.

Dennis, M. L., White, M. K., Titus, J. C., & Unsicker, J. I. (2006). Global Appraisal of Individual Needs (GAIN): Administration guide for the GAIN and related measures (Version 5). from http://www.chestnut.org/LI/gain/GAIN_I/GAIN-I_v_5-4/Index.html

Derogatis, L. R. (1994). *Symptom Checklist-90-Revised: Administration, scoring and procedures manual* (3 ed.). Minneapolis, MN: National Computer Systems, Inc.

Derogatis, L. R., & Meilisaratos, N. (1983). The Brief Symptom Inventory: An introductory report. *Psychological medicine, 13*, 595-605.

Dervaux, A., Baylé, F. J., Laqueille, X., Bourdel, M.-C., Leborgne, M., Olié, J.-P., et al. (2006). Validity of the CAGE questionnaire in schizophrenic patients with alcohol abuse and dependence. *Schizophrenia Research, 81*(2-3), 151-155.

Dickey, B., Normand, S.-L., Hermann, R. C., Eisen, S. V., Corte´s, D. E., Cleary, P. D., et al. (2003). Guideline recommendations for treatment of schizophrenia. *Archives of General Psychiatry, 60*(4), 340-348.

Dickinson, D., & Coursey, R. D. (2002). Independence and overlap among neurocognitive correlates of community functioning in schizophrenia. *Schizophrenia Research, 56*(1-2), 161-170.

Dingemans, P. M. A. J., Linszen, D. H., Lenior, M. E., & Smeets, R. M. W. (1995). Component structure of the expanded Brief Psychiatric Rating Scale (BPRS-E). *Psychopharmacology, 122*(3), 263-267.

Dobie, D. J., Kivlahan, D. R., Maynard, C., Bush, K. R., McFall, M., Epler, A. J., et al. (2002). Screening for post-traumatic stress disorder in female Veteran's Affairs patients: Validation of the PTSD Checklist. *General Hospital Psychiatry, 24*(6), 367-374.

Doerfler, L. A., Addis, M. E., & Moran, P. W. (2002). Evaluating mental health outcomes in an inpatient setting: Convergent and divergent validity of the OQ-45 and BASIS-32. *Journal of Behavioral Health Services and Research, 29*(4), 394-403.

Doherty, B., & Webb, M. (1989). The distribution of alcohol dependence severity among in-patient problem drinkers. *British Journal of Addiction, 84*, 917-922.

Doi, Y., & Minowa, M. (2003). Factor structure of the 12-item General Health Questionnaire in the Japanese general adult population. *Psychiatry and Clinical Neurosciences, 57*(4), 379-383.

Donohue, B., Azrin, N. H., Strada, M. J., Silver, N. C., Teichner, G., & Murphy, H. (2004). Psychometric evaluation of self- and collateral Timeline Follow-Back Reports of drug and alcohol use in a sample of drug-abusing and conduct-disordered adolescents and their parents. *Psychology of Addictive Behaviors, 18*(2), 184-189.

Dorfman, R. A., Lubben, J. E., Mayer-Oakes, A., Atchison, K., Schweitzer, S. O., De Jong, F. J., et al. (1995). Screening for depression among a well elderly population. *Social Work, 40*(3), 295-304.

Dornelas, E. A., Botticello, A. L., Goethe, J. W., & Fischer, E. H. (2001). Validity of a brief measure of posthospital adjustment for psychiatric patients. *Comprehensive Psychiatry, 42*(5), 410-415.

Douglas, A. N., Jimenez, S., Lin, H.-J., & Frisman, L. K. (2008). Ethnic differences in the effects of spiritual well-being on long-term psychological and behavioral outcomes within a sample of homeless women. *Cultural Diversity and Ethnic Minority Psychology, 14*(4), 344-352.

Downe-Wamboldt, B. L., & Melanson, P. M. (1995). Emotions, coping, and psychological well-being in elderly people with arthritis. *Western Journal of Nursing Research, 17*(3), 250-265.

Doyle, S. R., Donovan, D. M., & Kivlahan, D. R. (2007). The factor structure of the Alcohol Use Disorders Identification Test (AUDIT). *Journal of Studies on Alcohol and Drugs, 68*(3), 474-479.

Dozois, D. J. A., Dobson, K. S., & Ahnberg, J. L. (1998). A psychometric evaluation of the Beck Depression Inventory-II. *Psychological Assessment, 10*(2), 83-89.

Drake, R. E., McHugo, G. J., & Biesanz, J. C. (1995). The test-retest reliability of standardized instruments among homeless persons with substance use disorders. *Journal of Studies on Alcohol, 56*(2), 161-167.

Drake, R. E., & Mercer-McFadden, C. (1995). Assessment of substance use among persons with chronic mental illness. In A. F. Lehman & L. B. Dixon (Eds.), *Double Jeopardy: Chronic Mental Illness and Substance Use Disorders*. Chur, Switzerland: Harwood Academic Publishers.

Drummond, D. C., & Phillips, T. S. (2002). Alcohol urges in alcohol-dependent drinkers: Further validation of the Alcohol Urge Questionnaire in an untreated community clinical population. *Addiction, 97*(11), 1465-1472.

Du, Y., Kou, J., & Coghill, D. (2008). The validity, reliability and normative scores of the parent, teacher and self report versions of the Strengths and Difficulties Questionnaire in China. *Child and Adolescent Psychiatry and Mental Health, 2*.

Duffy, C. J., Cunningham, E. G., & Moore, S. M. (2005). Brief report: The factor structure of mood states in an early adolescent sample. *Journal of Adolescence, 28*(5), 677-680.

Dufresne, R. G., Phillips, K. A., Vittorio, C. C., & Wilkel, C. S. (2001). A screening questionnaire for body dysmorphic disorder in a cosmetic dermatologic surgery practice. *Dermatologic Surgery, 27*, 457-462.

Dufton, B. D., & Siddique, C. M. (1992). Measures in the day hospital. I. The global assessment of functioning scale. *International Journal of Partial Hospitalization, 8*(1), 41-49.

Dunmore, E., Clark, D. M., & Ehlers, A. (1999). Cognitive factors involved in the onset and maintenance of posttraumatic stress disorder (PTSD) after physical or sexual assault. *Behaviour Research and Therapy, 37*(9), 809-829.

Dybek, I., Bischof, G., Grothues, J., Reinhardt, S., Meyer, C., Hapke, U., et al. (2006). The reliability and validity of the Alcohol Use Disorders Identification Test (AUDIT) in a German general practice population sample. *Journal of Studies on Alcohol, 67*(3), 473-481.

Dyl, J., Kittler, J., Phillips, K. A., & Hunt, J. I. (2006). Body Dysmorphic Disorder and Other Clinically Significant Body Image Concerns in Adolescent Psychiatric Inpatients: Prevalence and Clinical Characteristics. *Child Psychiatry & Human Development, 36*(4), 369-382.

Dyrborg, J., Larsen, F., Nielsen, S., Byman, J., Buhl Nielsen, B., & Gautre-Delay, F. (2000). The Children's Global Assessment Scale (CGAS) and Global Assessment of Psychosocial Disability (GAPD) in clinical practice: Substance and reliability as judged by intraclass correlations. *European Child and Adolescent Psychiatry, 9*(3), 195-201.

Eagar, K., Trauer, T., & Mellsop, G. (2005). Performance of routine outcome measures in adult mental health care. *Australian and New Zealand Journal of Psychiatry, 39*(8), 713-718.

Edwards, G., & Gross, M. (1976). Alcohol dependence: provisional description of a clinical syndrome. *British Medical Journal, 1*, 1058-1061.

Ee Heok, K., Soo Meng, K., Cheong Sung, T., Yeo, B., Halstead, S., & Mahendran, R. (1990). A follow-up study of Asian problem drinkers. *British Journal of Addiction, 85*, 1261-1264.

Ehring, T., Kleim, B., Clark, D. M., Foa, E. B., & Ehlers, A. (2007). Screening for posttraumatic stress disorder: What combination of symptoms predicts best? *Journal of Nervous and Mental Disease, 195*(12), 1004-1012.

Eisen, S. V. (1992). Alcohol, drugs and psychiatric disorders: A current view of hospitalised adolescents. *Journal of Adolescent Research, 7*, 250-275.

Eisen, S. V. (1995). Assessment of subjective distress by patients self-report versus structured interview. *Psychological Reports, 76*, 35-39.

Eisen, S. V., & Culhane, M. A. (1999). Behavior and Symptom Identification Scale (BASIS-32). In M. E. Maruish (Ed.), *The Use of Psychological Testing for Treatment Planning and Outcomes Assessment* (2 ed., pp. 759-790). Mahwah, NJ: Lawrence Erlbaum Associates.

Eisen, S. V., Dill, D. L., & Grob, M. C. (1994). Reliability and validity of a brief patient-report instrument for psychiatric outcome evaluation. *Hospital and Community Psychiatry, 45*(3), 242-247.

Eisen, S. V., Grob, M. C., & D.L., D. (1989). Substance abuse in an inpatient population. *McLean Hospital Journal, 14*, 1-22.

Eisen, S. V., Grob, M. C., & Klein, A. A. (1986). BASIS: The development of a self-report measure for psychiatric inpatient evaluation. *Psychiatric Hospital, 17*(4), 165-171.

Eisen, S. V., Leff, H. S., & Schaefer, E. (1999). Implementing outcome systems: lessons from a test of the BASIS-32 and the SF-36. *Journal of Behavioral Health Services and Research, 26*(1), 18-27.

Eisen, S. V., Normand, S. L., Belanger, A. J., Spiro, A., & Esch, D. (2004). The revised Behavior and Symptom Identification Scale (BASIS-R). *Medical Care, 42*, 1230-1241.

Eisen, S. V., Wilcox, M., Leff, H. S., Schaefer, E., & Culhane, M. A. (1999). Assessing behavioral health outcomes in outpatient programs: reliability and validity of the BASIS-32. *Journal of Behavioral Health Services and Research, 26*(1), 5-17.

Eisen, S. V., Wilcox, M., & Schaefer, E. (1997). *Use of BASIS-32 for Outcome Assessment of Recipients of Outpatient Mental Health Services.* Belmont, MA: McLean Hospital: Evaluation Center at the Human Services Research Institute.

Elhai, J. D., Gray, M. J., Kashdan, T. B., & Franklin, C. L. (2005). Which instruments are most commonly used to assess traumatic event exposure and Posttraumatic effects?: A survey of traumatic stress professionals. *Journal of Traumatic Stress, 18*(5), 541-545.

Elliott, T. E., Renier, C. M., & Palcher, J. (2003). Chronic pain, depression, and quality of life: correlations and predictive value of the SF-36. *Pain Medicine, 4*, 331-339.

Ellis, T. E., & Ratliff, K. G. (1986). Cognitive characteristics of suicidal and nonsuicidal psychiatric inpatients. *Cognitive Therapy and Research, 10*(6), 625-634.

Emerson, E. (2005). Use of the Strengths and Difficulties Questionnaire to assess the mental health needs of children and adolescents with intellectual disabilities. *Journal of Intellectual & Developmental Disability, 30*(1), 14-23.

Endicott, J., Spitzer, R. L., Fliess, J. H., & Cohen, J. (1976). The Global Assessment Scale: A procedure for measuring overall severity of psychiatric disturbance. *Archives of General Psychiatry, 33*, 766-771.

Engelsen, B. K., & Laberg, J. C. (2001). A comparison of three questionnaires (EAT-12, EDI, and EDE-Q) for assessment of eating problems in healthy female adolescents. *Nordic Journal of Psychiatry, 55*(2), 129-135.

Engstrom, M., El-Bassel, N., Go, H., & Gilbert, L. (2008). Childhood sexual abuse and intimate partner violence among women in methadone treatment: A direct or mediated relationship? *Journal of Family Violence, 23*(7), 605-617.

Ericson, B. G., Hansson, J.-H., & Teike, T. (1997). Urgent need for a Need Assessment Scale in monitoring social support: A validation study on the Swedish version of the Camberwell Assessment of Need. *Nordic Journal of Psychiatry, 51*(3), 173-182.

Essink-Bot, M. L., Krabbe, P. F. M., Bonsel, G. J., & Aaronson, N. K. (1997). An empirical comparison of four generic health status measures: The Nottingham Health Profile, the Medical Outcomes Study 36 Item Short Form, the COOP/WONCA Charts, and the EuroQol Instrument. *Medical Care, 35*, 522-537.

Eu, P. W., Lee, C., Parker, G., & Loh, J. (2001). The disability profile of patients with schizophrenia in psychiatric hospital and community settings in Singapore. *Singapore Medical Journal, 42*(12), 559-562.

Ewing, J. A. (1984). Detecting alcoholism. The CAGE questionnaire. *Jama., 252*(14), 1905-1907.

Fals-Stewart, W., O'Farrell, T. J., Freitas, T. T., McFarlin, S. K., & Rutigliano, P. (2000). The Timeline Followback reports of psychoactive substance use by drug-abusing patients: Psychometric properties. *Journal of Consulting and Clinical Psychology, 68*(1), 134-144.

Falsetti, S. A., Resnick, H. S., Resick, P. A., & Kilpatrick, D. G. (1993). The modified PTSD Symptom Scale: A brief self-report measure of posttraumatic stress disorder. *The Behavior Therapist, 16*, 161-162.

Fedoroff, I. C., & McFarlane, T. (1998). Cultural aspects of eating disorders. In S. S. Kazarian & D. R. Evans (Eds.), *Cultural Clinical Psychology: Theory, research and practice* (pp. 152-176). New York: Oxford University Press.

Feld, R., Colantonio, A., Yoshida, K., & Odette, F. (2003). Mental health and vitality among Canadian women with physical disabilities. *Psychological Reports, 93*, 75-83.

Ferguson, R. J., Robinson, A. B., & Splaine, M. (2002). Use of the Reliable Change Index to evaluate clinical significance in SF-36 outcomes. *Quality of Life Research, 11*, 509-516.

Fernández de Larrinoa Palacios, P., Bulbena Vilarrasa, A., & Domínguez Panchón, A. I. (1992). Estudio de fiabilidad, validez y consistencia interna de la escala LSP (Life Skills Profile). Perfil de habilidades de la vida cotidiana. [Study of the reliability, validity, and internal consistency of the LSP (Life Skills Profile) scale.]. *Actas Luso-Espanolas de Neurologia, Psiquiatria y Ciencias Afines, 20*(2), 71-75.

Ferri, C. P., Marsden, J., de Araujo, M., Laranjeira, R. R., & Gossop, M. (2000). Validity and reliability of the Severity of Dependence Scale (SDS) in a Brazilian sample of drug users. *Drug and Alcohol Review, 19*(4), 451-455.

Ferring, D., & Filipp, S.-H. (1994). Teststatistische Überprüfung der Impact of Event-Skala: Befunde zu Reliabilität und Stabilität. [Statistical testing of the Impact of Event Scale: Reliability and stability.]. *Diagnostica, 40*(4), 344-362.

Fingerhood, M. (2000). Substance abuse in older people. *Journal of the American Geriatric Society, 48*(8), 985-995.

Fisher, K. M., & Copenhaver, V. (2006). Assessing the mental health of rural older adults in public housing facilities: a comparison of screening tools. *Journal of Gerontological Nursing, 32*(9), 26-33.

Fisher, P. L., & Durham, R. C. (1999). Recovery rates in generalized anxiety disorder following psychological therapy: An analysis of clinically significant change in the STAI-T across outcome studies since 1990. *Psychological Medicine, 29*(6), 1425-1434.

Flannery, B. A., Volpicelli, J. R., & Pettinati, H. M. (1999). Psychometric properties of the Penn Alcohol Craving Scale. *Alcoholism: Clinical and Experimental Research, 23*(8), 1289-1295.

Flemenbach, A., & Zimmerman, R. L. (1973). Inter- and intra-rater reliability of the Brief Psychiatric Rating Scale. *Psychological Reports, 36*, 783-792.

Flinn, S., Ventura, D., & Bonder, B. (2005). Return to work experiences for veterans with severe mental illness living in rural group home facilities. *Work, 24*(1), 63-70.

Florian, V., & Drory, Y. (1990). Mental Health Inventory (MHI): Psychometric properties and normative data in the Israeli population. *Psychologia: Israel Journal of Psychology, 2*, 26-35.

Foa, E. B. (1995). *Posttraumatic Stress Diagnostic Scale*. Minneapolis, MN: National Computer Systems.

Foa, E. B., Cashman, L., Jaycox, L., & Perry, K. (1997). The validation of a self-report measure of posttraumatic stress disorder: The Posttraumatic Diagnostic Scale. *Psychological Assessment, 9*(4), 445-451.

Foa, E. B., Hembree, E. A., Cahill, S. P., Rauch, S. A. M., Riggs, D. S., Feeny, N. C., et al. (2005). Randomized trial of prolonged exposure for posttraumatic stress disorder with and without cognitive restructuring: Outcome at academic and community clinics. *Journal of Consulting and Clinical Psychology, 73*(5), 953-964.

Foa, E. B., Riggs, D. S., Dancu, C. V., & Rothbaum, B. O. (1993). Reliability and validity of a brief instrument for assessing posttraumatic stress disorder. *Journal of Traumatic Stress, 6*, 459-473.

Foa, E. B., Zoellner, L. A., & Feeny, N. C. (2006). An Evaluation of Three Brief Programs for Facilitating Recovery After Assault. *Journal of Traumatic Stress, 19*(1), 29-43.

Forbes, D., Creamer, M., & Biddle, D. (2001). The validity of the PTSD checklist as a measure of symptomatic change in combat-related PTSD. *Behaviour Research and Therapy, 39*(8), 977-986.

Ford, J. D., Hawke, J., Alessi, S., Ledgerwood, D., & Petry, N. (2007). Psychological trauma and PTSD symptoms as predictors of substance dependence treatment outcomes. *Behaviour Research and Therapy, 45*(10), 2417-2431.

Ford, P. (2003). An evaluation of the Dartmouth Assessment of Lifestyle Inventory and the Leeds Dependence Questionnaire for use among detained psychiatric inpatients. *Addiction, 98*(1), 111-118.

Foster, J. H., Peters, T. J., & Marshall, E. J. (2000). Quality of life measures and outcome in alcohol-dependent men and women. *Alcohol, 22*(1), 45-52.

Franken, I. H., & Hendriks, V. M. (1999). Predicting outcome of inpatient detoxification of substance abusers. *Psychiatric Services., 50*(6), 813-817.

Franken, I. H. A., & Hendriks, V. M. (2001). Screening and diagnosis of anxiety and mood disorders in substance abuse patients. *The American Journal on Addictions, 10*(1), 30-39.

Freeston, M. H., Ladouceur, R., Thibodeau, N., Gagnon, F., & Rheaume, J. (1994). L'inventaire d'anxiete de Beck. Proprietes psychometriques d'une traduction francaise. [The Beck Anxiety Inventory. Psychometric properties of a French translation]. *Encephale, 20*(1), 47-55.

French, D., & Tait, R. J. (2004). Measurement invariance in the general health questionnaire-12 in young Australian adolescents. *European Child & Adolescent Psychiatry, 13*(1), 1-7.

Fridell, M., & Hesse, M. (2006). Psychiatric severity and mortality in substance abusers a 15-year follow-up of drug users. *Addictive Behaviors, 31*(4), 559-565.

Friedman, M. J. (2006). Posttraumatic stress disorder among military returnees from Afghanistan and Iraq. *Am J Psychiatry, 163*(4), 586-593.

Friedmann, P. D., Melnick, G., Jiang, L., & Hamilton, Z. (2008). Violent and disruptive behavior among drug-involved prisoners: Relationship with psychiatric symptoms. *Behavioral Sciences & the Law, 26*(4), 389-401.

Fu, S. S., Kodl, M., Willenbring, M., Nelson, D. B., Nugent, S., Gravely, A. A., et al. (2008). Ethnic differences in alcohol treatment outcomes and the effect of concurrent smoking cessation treatment. *Drug and Alcohol Dependence, 92*(1-3), 61-68.

Furlanetto, L. M., Mendlowicz, M. V., & Bueno, J. R. (2005). The validity of the Beck Depression Inventory-Short Form as a screening and diagnostic instrument for moderate and severe depression in medical inpatients. *Journal of Affective Disorders, 86*(1), 87-91.

Furukawa, T. A., Kawakami, N., Saitoh, M., Ono, Y., Nakane, Y., Nakamura, Y., et al. (2008). The performance of the Japanese version of the K6 and K10 in the World Mental Health Survey Japan. *International Journal of Methods in Psychiatric Research, 17*(3), 152-158.

Furukawa, T. A., Kessler, R. C., Slade, T., & Andrews, G. (2003). The performance of the K6 and K10 screening scales for psychological distress in the Australian National Survey of Mental Health and Well-Being. *Psychological Medicine, 33*, 357-362.

Fydrich, T., Dowdall, D., & Chambless, D. L. (1992). Reliability and validity of the Beck Anxiety Inventory. *Journal of Anxiety Disorders, 6*(1), 55-61.

Gache, P., Michaud, P., Landry, U., Accietto, C., Arfaoul, S., Wenger, O., et al. (2005). The Alcohol Use Disorders Identification Test (AUDIT) as a Screening Tool for Excessive Drinking in Primary Care: Reliability and Validity of a French Version. *Alcoholism: Clinical and Experimental Research, 29*(11), 2001-2007.

Gallagher, J., & Teesson, M. (2000a). Measuring disability, need and outcome in Australian community mental health services. *Australian and New Zealand Journal of Psychiatry, 34*(5), 850-855.

Gallagher, J., & Teesson, M. (2000b). Measuring disability, need and outcome in Australian community mental health services. *Australian & New Zealand Journal of Psychiatry, 34*(5), 850-855.

Gamst, G., Herdina, A., Mondragon, E., Munguia, F., Pleitez, A., Stephens, H., et al. (2006). Relationship Among Respondent Ethnicity, Ethnic Identity, Acculturation, and Homeless Status on a Homeless Population's Functional Status. *Journal of Clinical Psychology, 62*(12), 1485-1501.

Gandek, B., Sinclair, S. J., Kosinski, M., & Ware, J. E. J. (2004). Psychometric evaluation of the SF-36 health survey in Medicare managed care. *Health Care Financing Review, 25*, 5-25.

Gandek, B., Ware, J. E., Aaronson, N. K., Apolone, G., Bjorner, J. B., Brazier, J. E., et al. (1998). Cross-validation of item selection and scoring for the SF-12 Health Survey in nine countries: Results from the IQOLA Project. International Quality of Life Assessment. *Journal of Clinical Epidemiology, 51*(11), 1171-1178.

Garfinkel, P. E., & Newman, A. (2001). The Eating Attitudes Test: Twenty-five years later. *Eating and Weight Disorders, 6*(1), 1-24.

Garner, D. M., & Garfinkel, P. E. (1979). The Eating Attitudes Test: An index of the symptoms of anorexia nervosa. *Psychological Medicine, 9*, 273-279.

Garner, D. M., Olmstead, M. P., Bohr, Y., & Garfinkel, P. E. (1982). The eating attitudes test: Psychometric features and clinical correlates. *Psychological Medicine, 12*, 871-878.

Garralda, M. E., Yates, P., & Higginson, I. (2000). Child and adolescent mental health service use. HoNOSCA as an outcome measure. *British Journal of Psychiatry, 177*, 52-58.

Gauthier, J., & Bouchard, S. (1993). Adaptation canadienne-française de la forme révisée du State-Trait Anxiety Inventory de Spielberger. [A French-Canadian adaptation of the revised version of Spielberger's State-Trait Anxiety Inventory.]. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement, 25*(4), 559-578.

Gavin, G. R., Ross, H. E., & Skinner, A. (1989). Diagnostic validity of the Drug Abuse Screening Test in the assessment of DSM-III drug disorders. *British Journal of Addiction, 84*, 301-307.

Gearon, J. S., Kaltman, S. I., Brown, C., & Bellack, A. S. (2003). Traumatic Life Events and PTSD Among Women With Substance Use Disorders and Schizophrenia. *Psychiatr Serv, 54*(4), 523-528.

Gerevich, J., Bácskai, E., Kó, J., & Rózsa, S. (2005). Reliability and validity of the Hungarian version of the European Addiction Severity Index. *Psychopathology, 38*(6), 301-309.

Ghitza, U. E., Epstein, D. H., & Preston, K. L. (2008). Self-report of illicit benzodiazepine use on the Addiction Severity Index predicts treatment outcome. *Drug and Alcohol Dependence, 97*(1-2), 150-157.

Giang, K. B., Spak, F., Dzung, T. V., & Allebeck, P. (2005). The use of AUDIT to assess level of alcohol problems in rural Vietnam. *Alcohol and Alcoholism, 40*(6), 578-583.

Giannopoulou, I., Smith, P., Ecker, C., Strouthos, M., Dikaiakou, A., & Yule, W. (2006). Factor structure of the Children's Revised Impact of Event Scale (CRIES) with children exposed to earthquake. *Personality and Individual Differences, 40*(5), 1027-1037.

Giffort, D., Schmook, A., Woody, C., Vollendorf, C., & Gervain, M. (1995). *Construction of a Scale to Measure Consumer Recovery*. Springfield, IL: Illinois Office of Mental Health.

Gigantesco, A., Picardi, A., de Girolamo, G., & Morosini, P. (2007). Discriminant Ability and Criterion Validity of the HoNOS in Italian Psychiatric Residential Facilities. *Psychopathology, 40*, 111-115.

Gilbody, S. M., House, A. O., & Sheldon, T. A. (2002). Psychiatrists in the UK do not use outcomes measures: National survey. *British Journal of Psychiatry, 180*(2), 101-103.

Gill, S., Butterworth, P., Rodgers, B., & Mackinnon, A. (2007). Validity of the mental health component scale of the 12-item Short-Form Health Survey (MCS-12) as measure of common mental disorders in the general population. *Psychiatry Research, 152*(1).

Ginieri-Coccossis, M., Liappas, I. A., Tzavellas, E., Triantafillou, E., & Soldatos, C. (2007). Detecting changes in quality of life and psychiatric symptomatology following an in-patient detoxification programme for alcohol-dependent individuals: the use of WHOQOL-100. *In Vivo, 21*(1), 99-106.

Gloster, A. T., Rhoades, H. M., Novy, D., Klotsche, J., Senior, A., Kunik, M., et al. (2008). Psychometric properties of the Depression Anxiety and Stress Scale-21 in older primary care patients. *Journal of Affective Disorders, 110*(3), 248-259.

Goldberg, D., & Williams, P. (1988). *A user's guide to the General Health Questionnaire*. Windsor: NFER-NELSON Publishing Company.

Goldberg, D. P., Gater, R., Sartorius, N., Ustun, T. B., Piccinelli, M., Gureje, O., et al. (1997). The validity of two versions of the GHQ in the WHO study of mental illness in general health care. *Psychological Medicine, 27*(1), 191-197.

Goldfinger, S. M., Schutt, R. K., Seidman, L. J., Turner, W. M., & et al. (1996). Self-report and observer measures of substance abuse among homeless mentally ill persons in the cross-section and over time. *Journal of Nervous and Mental Disease, 184*(11), 667-672.

Goldman, H. H., Skodol, A. E., & Lave, T. R. (1992). Revising axis V for DSM-IV: A review of measures of social functioning. *American Journal of Psychiatry, 149*, 1148-1156.

Goldney, R. D., Fisher, L. J., & Walmsley, S. H. (1998). The Health of the Nation Outcome Scales in psychiatric hospitalisation: A multicentre study examining outcome and prediction of length of stay. *Australian and New Zealand Journal Psychiatry, 32*(2), 199-205.

Goldney, R. D., Fisher, L. J., Wilson, D. H., & Cheok, F. (2001). Suicidal ideation and health related quality of life in the community. *Medical Journal of Australia, 175*, 546-549.

Goldstein, R. B., Asarnow, J. R., Jaycox, L. H., Shoptaw, S., & Murray, P. J. (2007). Correlates of "non-problematic" and "problematic" substance use among depressed adolescents in primary care. *Journal of Addictive Diseases, 26*(3), 39-52.

Gómez, A., Conde, A., Santana, J. M., & Jorrín, A. (2005). Diagnostic usefulness of brief versions of Alcohol Use Disorders Identification Test (AUDIT) for detecting hazardous drinkers in primary care settings. *Journal of Studies on Alcohol, 66*(2), 305-308.

Gómez, A., Conde, A., Santana, J. M., Jorrín, A., Serrano, I. M., & Medina, R. (2006). The diagnostic usefulness of AUDIT and AUDIT-C for detecting hazardous drinkers in the elderly. *Aging & Mental Health, 10*(5), 558-561.

González-Saiz, F. (2007). Some remarks concerning proposals for dimensional criteria for substance use disorders in DSM-V. *Addiction, 102*(5), 827-828.

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

Goodman, R. (1999). The extended version of the Strengths and Difficulties Questionnaire as a guide to child psychiatric caseness and consequent burden. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 40*(5), 791-799.

Goodman, R. (2001). Psychometric properties of the Strengths and Difficulties Questionnaire. *Journal of the American Academy of Child and Adolescent Psychiatry, 40*(11), 1337-1345.

Goodman, R. (2003). The Strengths and Difficulties Questionnaire (SDQ). In L. VandeCreek (Ed.), *Innovations in Clinical Practice: Focus on Children and Adolescents* (pp. 109-111). Sarasota, FL: Ellis Human Development Institute.

Goodman, R., Ford, T., Simmons, H., Gatward, R., & Meltzer, H. (2000). Using the Strengths and Difficulties Questionnaire (SDQ) to screen for child psychiatric disorders in a community sample. *British Journal of Psychiatry, 177*, 534-539.

Goodman, R., Meltzer, H., & Bailey, V. (1998). The Strengths and Difficulties Questionnaire: A pilot study on the validity of the self-report version. *European Child and Adolescent Psychiatry, 7*(3), 125-130.

Goodman, R., Meltzer, H., & Bailey, V. (2003). The Strengths and Difficulties Questionnaire: A pilot study on the validity of the self-report version. *International Review of Psychiatry, 15*(1-2), 173-177.

Goodman, R., Renfrew, D., & Mullick, M. (2000). Predicting type of psychiatric disorder from Strengths and Difficulties Questionnaire (SDQ) scores in child mental health clinics in London and Dhaka. *European Child and Adolescent Psychiatry, 9*(2), 129-134.

Goodman, R., & Scott, S. (1999). Comparing the Strengths and Difficulties Questionnaire and the Child Behavior Checklist: Is small beautiful? *Journal of Abnormal Child Psychology, 27*(1), 17-24.

Goodman, R. F., Morgan, A. V., Juriga, S., & Brown, E. J. (2004). Letting the story unfold: A case study of client-centered therapy for childhood traumatic grief. *Harvard Review of Psychiatry, 12*(4), 199-212.

Gorenstein, C., Andrade, I., Filho, A. H. G. V., Tung, T. C., & Artes, R. (1999). Psychometric properties of the Portuguese version of the Beck Depression Inventory on Brazilian college students. *Journal of Clinical Psychology, 55*, 553-562.

Gorenstein, C., Pompeia, S., & Andrade, L. (1995). Scores of Brazilian University students on the Beck Depression and the State Trait Anxiety Inventories. *Psychological Reports, 77*(2), 635-641.

Gossop, M., Best, D., Marsden, J., & Strang, J. (1997). Test-retest reliability of the Severity of Dependence Scale. *Addiction, 92*(3), 353.

Gossop, M., Darke, S., Griffiths, P., Hando, J., Powis, B., Hall, W., et al. (1995). The severity of dependence scale (SDS): Psychometric properties of the SDS in English and Australian samples of heroin, cocaine and amphetamine users. *Addiction, 90*, 607-614.

Gossop, M., Stewart, D., & Marsden, J. (2007). Readiness for change and drug use outcomes after treatment. *Addiction, 102*(2), 301-308.

Gowers, S. G., Bailey-Rogers, S. J., Shore, A., & Levine, W. (2000). The Health of the Nation Outcome Scales for Child and Adolescent Mental Health (HoNOSCA). *Child Psychology and Psychiatry Review, 5*(2), 50-56.

Gowers, S. G., Harrington, R. C., Whitton, A., Lelliott, P., Beevor, A. S., Wing, J. K., et al. (1999). Brief scale for measuring the outcomes of emotional and behavioural disorders in children. Health of the Nation Outcome Scales for children and Adolescents (HoNOSCA). *British Journal of Psychiatry, 174*, 413-416.

Graham, C., Coombs, T., Buckingham, W., Eagar, K., Trauer, T., & Callaly, T. (2001). *Victorian Mental Health Outcomes Measurement Strategy: Consumer Perspectives on Future Directions for Outcome Self-Assessment: Report of the Consumer Consultation Project*. Victoria: Department of Human Services.

Grant, J. E., Kim, S. W., & Crow, S. J. (2001). Prevalence and clinical features of body dysmorphic disorder in adolescent and adult psychiatric inpatients. *Journal of Clinical Psychiatry, 62*, 517-522.

Grant, J. E., Kim, S. W., & Eckert, E. D. (2002). Body dysmorphic disorder in patients with anorexia nervosa: Prevalence, clinical features and delusionality of body image. *International Journal of Eating Disorders, 32*(3), 291-300.

Green, B., Shirk, S., Hanze, D., & Wanstrath, J. (1994). The Children's Global Assessment Scale in clinical practice: An empirical evaluation. *Journal of the American Academy of Child and Adolescent Psychiatry, 33*(8), 1158-1164.

Green, J., Jacobs, B., Beecham, J., Dunn, G., Kroll, L., Tobias, C., et al. (2007). Inpatient treatment in child and adolescent psychiatry--a prospective study of health gain and costs. *Journal of Child Psychology & Psychiatry & Allied Disciplines, 48*(12), 1259-1267.

Green, S., & Whichelow, M. J. (1994). Longitudinal validity of the CAGE questionnaire. *Addiction Research, 2*, 195-201.

Greenwood, A., & Burt, G. (2001). Validity of the Brief Psychiatric Rating Scale within a forensic inpatient hospital. *Expert Evidence, 8*(1), 15-30.

Griesel, D., Wessa, M., & Flor, H. (2006). Psychometric qualities of the German version of the Posttraumatic Diagnostic Scale (PTDS). *Psychological Assessment, 18*(3), 262-268.

Gross, J., Rosen, J. C., Leitenberg, H., & Willmuth, M. E. (1986). Validity of the Eating Attitudes Test and the Eating Disorders Inventory in bulimia nervosa. *Journal of Consulting and Clinical Psychology, 54*(6), 875-876.

Groth-Marnat, G. (2003). *Handbook of psychological assessment*. USA: John Wiley & Sons.

Grubaugh, A. L., Elhai, J. D., Cusack, K. J., Wells, C., & Frueh, B. C. (2007). Screening for PTSD in public-sector mental health settings: The diagnostic utility of the PTSD Checklist. *Depression and Anxiety, 24*(2), 124-129.

Gu, J., Lau, J. T. F., Chen, H., Liu, Z., Lei, Z., Li, Z., et al. (2008). Validation of the Chinese version of the Opiate Addiction Severity Inventory (OASI) and the Severity of Dependence Scale (SDS) in China. *Addictive Behaviors, 33*(5), 725-741.

Gunther, O. H., Roick, C., Angermeyer, M. C., & Konig, H.-H. (2008). Responsiveness of EQ-5D utility indices in alcohol-dependent patients. *Drug & Alcohol Dependence, 92*(1-3), 291-295.

Guo-Ping, H., Ya-Lin, Z., & Hui, X. (2006). [The Chinese Version of the Impact of Event Scale-Revised: Reliability and Validity.]. *Chinese Mental Health Journal, 20*(1), 28-31.

Guo, S.-R., Xin, Z.-Q., & Geng, L.-N. (2007). [Reliability and validity of Chinese version of the Impact of Event Scale-Revised.]. *Chinese Journal of Clinical Psychology, 15*(1), 15-17.

Hadzi-Pavlovic, D., Rosen, A., & Parker, G. (1992). The relevance and use of life skills assessments. In K. DJ (Ed.), *Schizophrenia: An Overview and Practical Handbook* (pp. 206-220). London: Chapman and Hall.

Hafkenscheid, A. (1991). Psychometric evaluation of a standardized and expanded Brief Psychiatric Rating Scale. *Acta Psychiatrica Scandinavica, 84*(3), 294-300.

Haidula, L., Shino, E., Plattner, l., & Feinstein, A. (2003). A Namibian version of the 28 item General Health Questionnaire. *South African Psychiatry Review, 6*(4), 23-25.

Hall, R. C. (1995). Global assessment of functioning. A modified scale. *Psychosomatics, 36*(4), 416-417.

Hambridge, J. A., & Rosen, A. (1994). Assertive community treatment for the seriously mentally ill in suburban Sydney: a programme description and evaluation. *Australian & New Zealand Journal of Psychiatry, 28*(3), 438-445.

Hamernik, E., & Pakenham, K. I. (1999). Assertive community treatment for persons with severe mental disorders: A controlled treatment outcome study. *Behaviour Change, 16*(4), 259-268.

Hanssen-Bauer, K., Aalen, O. O., Ruud, T., & Heyerdahl, S. (2007). Inter-rater reliability of clinician-rated outcome measures in child and adolescent mental health services. *Administration & Policy in Mental Health, 34*(6), 504-512.

Hanssen-Bauer, K., Gowers, S., Aalen, O. O., Bilenberg, N., Brann, P., Garralda, E., et al. (2007). Cross-national reliability of clinician-rated outcome measures in child and adolescent mental health services. *Administration & Policy in Mental Health, 34*(6), 513-518.

Hansson, L., Vinding, H. R., Mackeprang, T., Sourander, A., Werdelin, G., Bengtsson-Tops, A., et al. (2001). Comparison of key worker and patient assessment of needs in schizophrenic patients living in the community: A Nordic multicentre study. *Acta Psychiatrica Scandinavica, 103*(1), 45-51.

Hao, Y.-T., Fang, J.-Q., & Power, M. J. (2006). [The Equivalence of WHOQOL-BREF among 13 Culture Versions.]. *Chinese Mental Health Journal, 20*(2), 71-75.

Harris, K. M., Griffin, B. A., McCaffrey, D. F., & Morral, A. R. (2008). Inconsistencies in self-reported drug use by adolescents in substance abuse treatment: Implications for outcome and performance measurements. *Journal of Substance Abuse Treatment, 34*(3), 347-355.

Hart, K. M., Bryer, J. B., & Martines, K. A. (1991). SCL-90-R symptom patterns for adolescent psychiatric inpatients. *Journal of Personality Assessment, 56*(1), 168-178.

Hasegawa, K., Ogawa, K., Kondoh, C., Iseda, T., Ikebuchi, E., & Miyake, Y. (1997). [The reliability and validity of the Japanese version of the Life Skills Profile.]. *Seishin Igaku (Clinical Psychiatry), 39*(5), 547-555.

Hasin, D., Trautman, K., & Endicott, J. (1998). Psychiatric research interview for substance and mental disorders: Phenomenologically based diagnosis in patients who abuse alcohol or drugs. *Psychopharmacology Bulletin, 34*, 3-8.

Haswell-Elkins, M. (2006). Indigenous adult mental health outcomes Project in the Cairns network far north Queensland.

Havard, A., Teesson, M., Darke, S., & Ross, J. (2006). Depression among heroin users: 12-Month outcomes from the Australian Treatment Outcome Study (ATOS). *Journal of Substance Abuse Treatment, 30*(4), 355-362.

Hawes, D. J., & Dadds, M. (2004). Australian data and psychometric properties of the Strengths and Difficulties Questionnaire. *Australian New Zealand Journal of Psychiatry, 38*(8), 644-651.

Hawthorne, G., Herrman, H., & Murphy, B. (2006). Interpreting the WHOQOL-Brèf: Preliminary population norms and effect sizes. *Social Indicators Research, 77*(1), 37-59.

Hawthorne, W. B., Green, E. E., Lohr, J. B., Hough, R., & Smith, P. G. (1999). Comparison of outcomes of acute care in short-term residential treatment and psychiatric hospital settings. *Psychiatric Services, 50*(3), 401-406.

Hays, R. D., Merz, J. F., & Nicholas, R. (1995). Response burden, reliability and validity of the CAGE, Short MAST and AUDIT alcohol screening measures. *Behaviour Research Methods, Instruments and Computers, 27*, 277-280.

Heather, N., Raistrick, D., Tober, G., Godfrey, C., & Parrott, S. (2001). Leeds Dependence Questionnaire: New data from a large sample of clinic attenders. *Addiction Research & Theory, 9*(3), 253-269.

Hedlund, J. L., & Vieweg, B. W. (1980). The Brief Psychiatric Rating Scale (BPRS): A comprehensive review. *Journal of Operational Psychiatry, 11*, 48–65.

Heinze, M., Taylor, R. E., Priebe, S., & Thornicroft, G. (1997). The quality of life of patients with paranoid schizophrenia in London and Berlin. *Social Psychiatry and Psychiatric Epidemiology, 32*(5), 292-297.

Heishman, S. J., Singleton, E. G., & Liguori, A. (2001). Marijuana craving questionnaire: Development and initial validation of a self-report instrument. *Addiction, 96*(7), 1023-1034.

Hendricks, V. M., Kaplan, C. D., Van Limbeek, J., & Geerlings, P. (1989). The Addiction Severity index: reliability and validity in a Dutch Addict Population. *Journal of Substance Abuse Treatment, 6*, 133-141.

Henry, J. D., & Crawford, J. R. (2005). The short-form version of the Depression Anxiety Stress Scales (DASS-21): Construct validity and normative data in a large non-clinical sample. *British Journal of Clinical Psychology, 44*, 227-239.

Hernández, M. Á. T., & Gómez, C. F. (2004). Spanish validation of the Maudsley Addiction profile (MAP). *Adicciones, 16*(4), 267-275.

Herrman, H., Patrick, D. L., Diehr, P., Martin, M. L., Fleck, M., Simon, G. E., et al. (2002). Longitudinal investigation of depression outcomes in primary care in six countries: the LIDO study. Functional status, health service use and treatment of people with depressive symptoms. *Psychological Medicine, 32*, 889-902.

Hesse, M. (2006). The Beck Depression Inventory in patients undergoing opiate agonist maintenance treatment. *British Journal of Clinical Psychology, 45*(3), 417-425.

Hides, L., Dawe, S., Young, R. M., & Kavanagh, D. J. (2007). The reliability and validity of the severity of dependence scale for detecting cannabis dependence in psychosis. *Addiction, 102*, 35-40.

Hides, L., Lubman, D. I., Devlin, H., Cotton, S., Aitken, C., Gibbie, T., et al. (2007). Reliability and validity of the Kessler 10 and Patient Health Questionnaire among injecting drug users. *Australian and New Zealand Journal of Psychiatry, 41*(2), 166-168.

Hides, L., Lubman, D. I., Elkins, K., Catania, L. S., & Rogers, N. (2007). Feasibility and acceptability of a mental health screening tool and training programme in the youth alcohol and other drug (AOD) sector. *Drug and Alcohol Review, 26*(5), 509-515.

Hill, C. R., & Hughes, J. N. (2007). An examination of the convergent and discriminant validity of the Strengths and Difficulties Questionnaire. *School Psychology Quarterly, 22*(3), 380-406.

Hill, K. P., & Chang, G. (2007). Brief screening instruments for risky drinking in the outpatient psychiatry clinic. *The American Journal on Addictions, 16*(3), 222-226.

Hintikka, U., Marttunen, M., Pelkonen, M., Laukkanen, E., Viinamaki, H., & Lehtonen, J. (2006). Improvement in cognitive and psychosocial functioning and self image among adolescent inpatient suicide attempters. *BMC Psychiatry, 6*, 58.

Hirata, E. S., Almeida, O. P., Funari, R. R., & Klein, E. L. (2001). Validity of the Michigan Alcoholism Screening Test (MAST) for the detection of alcohol-related problems among male geriatric outpatients. *American Journal of Geriatric Psychiatry, 9*(1), 30-34.

Hirini, P., Flett, R., Long, N., & Millar, M. (2005). Frequency of Traumatic Events, Physical and Psychological Health among Maori. *New Zealand Journal of Psychology, 34*(1), 20-27.

Hodges, K., & Gust, J. (1995). Measures of impairment for children and adolescents. *Journal of Mental Health Administration, 22*(4), 403-413.

Hodgins, D. C., & el-Guebaly, N. (1992). More data on the Addiction Severity Index. Reliability and validity with the mentally ill substance abuser. *Journal of Nervous & Mental Disease, 180*(3), 197-201.

Hodgins, D. C., & Lightfoot, L. O. (1989). The use of the Alcohol Dependence Scale with incarcerated male offenders. *International Journal of Offender Therapy and Comparative Criminology, 33*(1), 59-67.

Hodgkinson, P., & Joseph, S. A. (1995). Factor analysis of the Impact of Events Scale with female bank staff following an armed raid. *Personality and Individual Differences, 19*(5), 773-775.

Hoffmann, F. L., Capelli, K., & Mastrianni, X. (1997). Measuring treatment outcome for adults and adolescents: reliability and validity of BASIS-32. *Journal of Mental Health Administration, 24*(3), 316-331.

Hoge, C. W., Auchterlonie, J. L., & Milliken, C. S. (2006). Mental Health Problems, Use of Mental Health Services, and Attrition From Military Service After Returning From Deployment to Iraq or Afghanistan. *JAMA, 295*(9), 1023-1032.

Holderness, C. C., Brooks Gunn, J., & Warren, M. P. (1994). Co-morbidity of eating disorders and substance abuse review of the literature. *International Journal of Eating Disorders, 16*(1), 1-34.

Holi, M. M., Marttunen, M., & Aalberg, V. (2003). Comparison of the GHQ-36, the GHQ-12 and the SCL-90 as psychiatric screening instruments in the Finnish population. *Nordic Journal of Psychiatry, 57*(3), 233-238.

Hope, J. D., Trauer, T., & Keks, N. A. (1998). Reliability, validity and utility of the Health of the Nation Outcomes Scale (HoNOS) in Australian adult psychiatric services. *Schizophrenia Research, 29*(1,2), 9-10.

Hopman, W. M., Berger, C., Joseph, L., Towheed, T., vandenKerkhof, E., Anastassiades, T., et al. (2004). Stability of normative data for the SF-36: results of a three-year prospective study in middle aged Canadians. *Canadian Journal of Public Health, 95*, 387-391.

Horner, D., & Asher, K. (2005). General practitioners and mental health staff sharing patient care: Working model. *Australasian Psychiatry, 13*(2), 176-180.

Horowitz, L. M., Rosenberg, S. E., Baer, B. A., Ureno, G., & Villasenor, V. S. (1988). Inventory of interpersonal problems: Psychometric properties and clinical applications. *Journal of Consulting and Clinical Psychology, 56*, 885-892.

Horowitz, M., Wilner, N., & Alvarez, W. (1979). Impact of event scale: A measure of subjective stress. *Psychosomatic Medicine, 41*(3), 209-219.

Hoshi, R., Mullins, K., Boundy, C., Brignell, C., Piccini, P., & Curran, H. V. (2007). Neurocognitive function in current and ex-users of ecstasy in comparison to both matched polydrug-using controls and drug-naïve controls. *Psychopharmacology, 194*(3), 371-379.

Howard, L., Hunt, K., Slade, M., O'Keane, V., Senevirante, T., Leese, M., et al. (2007). Assessing the needs of pregnant women and mothers with severe mental illness: The psychometric properties of the Camberwell Assessment of Need - Mothers (CAN-M). *International Journal of Methods in Psychiatric Research, 16*(4), 177-185.

Hsu, H. T., Hwang, S. L., Lee, P. H., & Chen, S. C. (2006). Impact of liver donation on quality of life and physical and psychological distress. *Transplantation Proceedings, 38*(7), 2102-2105.

Hu, Y., Stewart-Brown, S., Twigg, L., & Weich, S. (2007). Can the 12-item General Health Questionnaire be used to measure positive mental health? *Psychological Medicine, 37*(7), 1005-1013.

Huang, C.-Y., Sousa, V. D., Tsai, C.-C., & Hwang, M.-Y. (2008). Social support and adaptation of Taiwanese adults with mental illness. *Journal of Clinical Nursing, 17*(13), 1795-1802.

Huang, G., Zhang, Y., Momartin, S., Cao, Y., & Zhao, L. (2006). Prevalence and characteristics of trauma and posttraumatic stress disorder in female prisoners in china. *Comprehensive Psychiatry, 47*(1), 20-29.

Huebeck, B. G., & Neill, J. T. (2000). Confirmatory factor analysis and reliability of the Mental Health Inventory for Australian adolescents. *Psychological Reports, 87*(2), 431-440.

Hughes, C. W., Rintelmann, J., Emslie, G. J., Lopez, M., & MacCabe, N. (2001). A revised anchored version of the BPRS-C for childhood psychiatric disorders. *Journal of Child and Adolescent Psychopharmacology, 11*(1), 77-93.

Humeniuk, R., & Ali, R. (2006). *Validation of the Alcohol, Smoking and Substance Involvement Screening Test (ASSIST) and pilot brief intervention [electronic resource]: A technical report of phase II findings of the WHO ASSIST Project*. Geneva: World Health Organization.

Humeniuk, R., Ali, R., Babor, T. F., Farrell, M., Formigoni, M. L., Jittiwutikarn, J., et al. (2008). Validation of the Alcohol, Smoking and Substance Involvement Screening Test (ASSIST). *Addiction, 103*(6), 1039-1047.

Humeniuk, R., Dennington, V., & Ali, R. (2008). *The effectiveness of a brief intervention for illicit drugs linked to the Alcohol, Smoking and Substance Involvement Screening Test (ASSIST) in primary health care settings: A technical report of phase III findings of the WHO ASSIST randomized controlled trial [electronic resource]*. Geneva: World Health Organization.

Hunter, E., Hall, W., & Spargo, R. (1991). Alcohol consumption and its correlates in a remote aboriginal population. *Aboriginal Law Bulletin, 2*(51), 8-10.

Hunter, E. M. (1993). *Aboriginal health and history: power and prejudice in remote Australia*. Melbourne and New York: Cambridge University Press.

Hunter, R., McLean, J., Peck, D., Pullen, I., Greenfield, A., McArthur, W., et al. (2004). The Scottish 700 Outcomes Study: A comparative evaluation of the Health of the Nation Outcome Scale (HoNOS), the Avon Mental Health Measure (AVON), and an Idiographic Scale (OPUS) in adult mental health. *Journal of Mental Health, 13*(1), 93-105.

Husband, S. D., Marlowe, D. B., Lamb, R. J., Iguchi, M. Y., Bux, D. A., Kirby, K. C., et al. (1996). Decline in self-reported dysphoria after treatment entry in inner-city cocaine addicts. *Journal of Consulting and Clinical Psychology, 64*(1), 221-224.

Ilhan, I. O., Demirbas, H., & Dogan, Y. B. (2007). Psychosocial factors in alcohol use-related problems of working youth. *Substance Use & Misuse, 42*(10), 1537-1544.

Ip, W. Y., & Martin, C. R. (2006). Psychometric properties of the 12-item General Health Questionnaire (GHQ-12) in Chinese women during pregnancy and in the postnatal period. *Psychology, Health & Medicine, 11*(1), 60-69.

Issakidis, C., & Teesson, M. (1999). Measurement of need for care: A trial of the Camberwell Assessment of Need and the Health of the Nation Outcome Scales. *Australian and New Zealand Journal of Psychiatry, 33*(5), 754-759.

Jacob, K. S., Bhugra, D., & Mann, A. H. (1997). The validation of the 12-item General Health Questionnaire among ethnic Indian women living in the United Kingdom. *Psychological Medicine, 27*(5), 1215-1217.

James, M., & Kehoe, R. (1999). Using the Health of the Nation Outcome Scales in clinical practice. *Psychiatric Bulletin, 23*(9), 536-538.

Jenkinson, C. (1999). Comparison of UK and US methods for weighting and scoring the SF-36 summary measures. *Journal of Public Health Medicine, 21*, 372-376.

Jenkinson, C., Chandola, T., Coulter, A., & Bruster, S. (2001). An assessment of the construct validity of the SF-12 summary scores across ethnic groups. *Journal of Public Health Medicine, 23*, 187-194.

Jenkinson, C., Gray, A., Doll, H., Lawrence, K., Keoghane, S., & Layte, R. (1997). Evaluation of index and profile measures of health status in a randomized controlled trial. Comparison of the Medical

Outcomes Study 36-Item Short Form Health Survey, EuroQol, and disease specific measures. *Medical Care, 35*, 1109-1118.

Jenkinson, C., Lawrence, K., McWhinnie, D., & Gordon, J. (1995). Sensitivity to change of health status measures in a randomized controlled trial: Comparison of the COOP charts and the SF 36. *Quality of Life Research, 4*, 47-52.

Jenkinson, C., & Layte, R. (1997). Development and testing of the UK SF-12 (short form health survey). *Journal of Health Services & Research Policy, 2*, 14-18.

Jenkinson, C., Peto, V., & Coulter, A. (1994). Measuring change over time: a comparison of results from a global single item of health status and the multi dimensional SF 36 health status survey questionnaire in patients presenting with menorrhagia. *Quality of Life Research, 3*, 317-321.

Jenkinson, C., Wright, L., & Coulter, A. (1994). Criterion validity and reliability of the SF 36 in a population sample. *Quality of Life Research, 3*, 7-12.

Jerrell, J. M. (2005). Behavior and Symptom Identification Scale 32: Sensitivity to Change Over Time. *The Journal of Behavioral Health Services & Research, 32*(3), 341-346.

Jo, S. A., Park, M. H., Jo, I., Ryu, S.-H., & Han, C. (2007). Usefulness of Beck Depression Inventory (BDI) in the Korean elderly population. *International Journal of Geriatric Psychiatry, 22*(3), 218-223.

Jobson, L., & O'Kearney, R. (2008). Cultural differences in personal identity in post-traumatic stress disorder. *British Journal of Clinical Psychology, 47*(1), 95-109.

Johnson, J. A., & Coons, S. J. (1998). Comparison of the EQ-5D and SF-12 in an adult US sample. *Quality of Life Research, 7*, 155-166.

Johnson, M. E., Brems, C., & Burke, S. (2002). Recognizing comorbidity among drug users in treatment. *American Journal of Drug & Alcohol Abuse, 28*(2), 243-261.

Johnson, M. E., Brems, C., Mills, M. E., & Freemon, M. (2005). Utility of the Behavior and Symptom Identification Scale (BASIS-32) at an Alcohol Detoxification Unit. *Alcoholism Treatment Quarterly, 23*(1), 17-29.

Johnson, M. E., Neal, D. B., Brems, C., & Fisher, D. G. (2006). Depression as measured by the Beck Depression Inventory-II among injecting drug users. *Assessment, 13*(2), 168-177.

Johnston, S., Salkeld, G., Sanderson, K., Issakidis, C., Teesson, M., & Buhrich, N. (1998). Intensive case management: A cost-effectiveness analysis. *Australian and New Zealand Journal of Psychiatry, 32*(4), 551-559.

Jones, S., Thornicroft, G., Coffey, M., & Dunn, G. (1995). A brief mental health outcome scale-reliability and validity of the Global Assessment of Functioning (GAF). *British Journal of Psychiatry, 166*(5), 654-659.

Jorge, M. R., & Masur, J. (1985). The use of the short-form alcohol dependence data questionnaire (SADD) in Brazilian alcoholic patients. *British Journal of Addiction, 80*, 301-305.

Joseph, S. (2000). Psychometric evaluation of Horowitz's Impact of Event Scale: A review. *Journal of Traumatic Stress, 13*(1), 101-113.

Joyner, L. M., Wright, J. D., & Devine, J. A. (1996). Reliability and validity of the addiction severity index among homeless substance misusers. *Substance Use & Misuse, 31*(6), 729-751.

József, G., Erika, B., & Sándo, R. (2004). Hungarian validation of Maudsley Addiction Profile (MAP). *Psychiatria Hungarica, 19*(2), 123-130.

Kabacoff, R. I., Segal, D. L., Hersen, M., & Van Hasselt, V. B. (1997). Psychometric properties and diagnostic utility of the Beck Anxiety Inventory and the State-Trait Anxiety Inventory with older adult psychiatric outpatients. *Journal of Anxiety Disorders, 11*(1), 33-47.

Kafka, G. J., & Kozma, A. (2002). The Construct Validity of Ryff's Scales of Psychological Well-Being (SPWB) and their Relationship to Measures of Subjective Well-Being *Social Indicators Research, 57*, 171-190.

Kagee, A. (2001). Review of the SF-36 Health Survey. In B. S. Plake & J. C. Impara (Eds.), *The fourteenth mental measurements yearbook*. Lincoln NE: Buros Institute of Mental Measurements.

Kahler, C. W., Strong, D. R., Hayaki, J., Ramsey, S. E., & Brown, R. A. (2003). An item response analysis of the Alcohol Dependence Scale in treatment-seeking alcoholics. *Journal of Studies on Alcohol, 64*(1), 127-136.

Kahler, C. W., Strong, D. R., Stuart, G. L., Moore, T. M., & Ramsey, S. E. (2003). Item functioning of the alcohol dependence scale in a high-risk sample. *Drug and Alcohol Dependence, 72*(2), 183-192.

Kalfoss, M. H., Low, G., & Molzahn, A. E. (2008). The suitability of the WHOQOL-BREF for Canadian and Norwegian older adults. *European Journal of Ageing, 5*(1), 77-89.

Kaminer, Y. (2008). The Teen Addiction Severity Index around the globe: The Tower of Babel revisited. *Substance Abuse, 29*(3), 89-94.

Kaminer, Y., Bukstein, O., & Tarter, R. E. (1991). The Teen-Addiction Severity Index: rationale and reliability. *International Journal on Addictions, 26*(2), 219-226.

Kaminer, Y., Wagner, E., Plummer, B., & Seifer, R. (1993). Validation of the Teen Addiction Severity Index (T-ASI). Preliminary findings. *American Journal on Addictions, 2*(3), 250-254.

Kapci, E. G., Uslu, R., Turkcapar, H., & Karaoglan, A. (2008). Beck Depression Inventory II: Evaluation of the psychometric properties and cut-off points in a Turkish adult population. *Depression and Anxiety, 25*, E104-E110.

Kaptein, S., Jansen, D. E. M. C., Vogels, A. G. C., & Reijneveld, S. A. (2008). Mental health problems in children with intellectual disability: Use of the Strengths and Difficulties Questionnaire. *Journal of Intellectual Disability Research, 52*(2), 125-131.

Kato, T., & Snyder, C. R. (2005). [Relationship between hope and subjective well-being: reliability and validity of the dispositional Hope Scale, Japanese version]. *Shinrigaku Kenkyu - Japanese Journal of Psychology, 76*(3), 227-234.

Katz, M. M., & Lyerly, S. B. (1963). Methods for measuring adjustment and social behavior in the community: I. Rationale, description, discriminative validity and scale development. *Psychological Reports, 13*(2), 503-535.

Kauer-Sant'Anna, M., Frey, B. N., Andreazza, A. C., Cereser, K. M., Gazalle, F. K., Tramontina, J., et al. (2007). Anxiety comorbidity and quality of life in bipolar disorder patients. *Canadian Journal of Psychiatry - Revue Canadienne de Psychiatrie, 52*(3), 175-181.

Kavanagh, D. J., Saunders, J. B., Young, R., White, A., Jenner, L., Clair, A., et al. (1999). *Evaluation of screening and brief intervention for substance abuse in early psychosis. A report to AUSEINET*. Brisbane: Department of Psychiatry, University of Queensland.

Kavanagh, D. J., Sitharthan, T., Spilsbury, G., & Vignaendra, S. (1999). An Evaluation of Brief Correspondence Programs for Problem Drinkers. *Behavior Therapy, 30*(4), 641-656.

Kay-Lambkin, F. J., Baker, A., & Lewin, T. L. (2004). The "co-morbidity roundabout": a framework to guide assessment and intervention strategies and engineer change among people with co-morbid problems. *Drug and Alcohol Review, 23*(4), 407 - 423.

Kaye, S., & Darke, S. (2002). Determining a diagnostic cut-off on the Severity of Dependence Scale (SDS) for cocaine dependence. *Addiction, 97*(6), 727-731.

Kaysen, D., Pantalone, D. W., Chawla, N., Lindgren, K. P., Clum, G. A., Lee, C., et al. (2008). Posttraumatic stress disorder, alcohol use, and physical health concerns. *Journal of Behavioral Medicine, 31*(2), 115-125.

Kazlauskas, E., Gailiene, D., Domanskaite-Gota, V., & Trofimova, J. (2006). Ivykio poveikio skales - revizuotos (IES-R) lietuvi?kos versijos psichometrines savybes. [Psychometric Properties of the Lithuanian Version of the Impact of Event Scale-Revised (IES-R).]. *Psichologija, 33*, 22-30.

Keller, S., & Hayes, R. (1998). The relationship between the Allen Cognitive Level Test and the Life Skills Profile. *American Journal of Occupational Therapy, 52*(10), 851-856.

Keller, S. D., Ware, J. E. J., Bentler, P. M., Aaronson, N. K., Alonso, J., Apolone, G., et al. (1998). Use of structural equation modelling to test the construct validity of the SF-36 Health Survey in ten countries: results from the IQOLA Project. International Quality of Life Assessment. *Journal of Clinical Epidemiology, 51*, 1179-1188.

Kelly, A. B., & Kowalyszyn, M. (2003). The association of alcohol and family problems in a remote indigenous Australian community. *Addictive Behaviors, 28*(4), 761-767.

Kennedy, C., & Yellowlees, P. (2003). The effectiveness of telepsychiatry measured using the Health of the Nation Outcome Scale and the Mental Health Inventory. *Journal of Telemedicine and Telecare, 9*(1), 12-16.

Kessler, R. C. (1996). *Kessler's 10 Psychological Distress Scale*. Boston, MA: Harvard Medical School.

Kessler, R. C., Andrews, G., Colpe, l. J., Hiripi, E., Mroczek, D. K., Normand, S.-l. T., et al. (2002). Short screening scales to monitor population prevalences and trends in non-specific psychological distress. *Psychological Medicine, 32*(6), 959-976.

Kessler, R. C., Barker, P. R., Colpe, L. J., Epstein, J. F., Groerer, J. C., Hiripi, E., et al. (2003). Screening for serious mental illness in the general population. *Archives of General Psychiatry, 60*(2), 184-189.

Kilic, C., Rezaki, M., Rezaki, B., Kaplan, I., Ozgen, G., Sagduyu, A., et al. (1997). General Health Questionnaire (GHQ12 & GHQ28): psychometric properties and factor structure of the scales in a Turkish primary care sample. *Social Psychiatry and Psychiatric Epidemiology, 32*(6), 327-331.

Kilkkinen, A., Kao-Philpot, A., O'Neil, A., Philpot, B., Reddy, P., Bunker, S., et al. (2007). Prevalence of psychological distress, anxiety and depression in rural communities in Australia. *Australian Journal of Rural Health, 15*(2), 114 - 119.

Kim, S. S., Gulick, E. E., Nam, K. A., & Kim, S.-H. (2008). Psychometric properties of the alcohol use disorders identification test: A Korean version. *Archives of Psychiatric Nursing, 22*(4), 190-199.

Kimerling, R., Trafton, J. A., Nguyen, B., Kimerling, R., Trafton, J. A., & Nguyen, B. (2006). Validation of a brief screen for Post-Traumatic Stress Disorder with substance use disorder patients. *Addictive Behaviors, 31*(11), 2074-2079.

Kindt, M., Buck, N., Arntz, A., & Soeter, M. (2007). Perceptual and conceptual processing as predictors of treatment outcome in PTSD. *Journal of Behavior Therapy and Experimental Psychiatry, 38*(4), 491-506.

King, M. (1986). At risk drinking among general practice attendees: Prevalence, characteristics and alcohol-related problems. *British Journal of Psychiatry, 148*, 533-540.

Kirkby, K., Daniels, B., Jones, I., & McInnes, M. (1997). A survey of social outcome in schizophrenia in Tasmania. *Australian and New Zealand Journal of Psychiatry, 31*(3), 405-410.

Kisely, S., Preston, N., & Rooney, M. (2000). Pathways and outcomes of psychiatric care: Does it depend on who you are, or what you've got? *Australian and New Zealand Journal of Psychiatry, 34*(6), 1009-1014.

Klasen, H., Woerner, W., Wolke, D., Meyer, R., Overmeyer, S., Kaschnitz, W., et al. (2000). Comparing the German versions of the Strengths and Difficulties Questionnaire (SDQ-Deu) and the Child Behavior Checklist. *European Child and Adolescent Psychiatry, 9*(4), 271-276.

Kleinman, P. H., Miller, A. B., Millman, R. B., Woody, G. E., Todd, T., Kemp, J., et al. (1990). Psychopathology among cocaine abusers entering treatment. *Journal of Nervous and Mental Disease, 178*, 442-447.

Klinkenberg, W. D., Cho, D. W., & Vieweg, B. (1998). Reliability and validity of the interview and self-report versions of the BASIS-32. *Psychiatric Services, 49*(9), 1229-1231.

Knight, J. R., Sherritt, L., Harris, S. K., Gates, E. C., & Chang, G. (2003). Validity of brief alcohol screening tests among adolescents: A comparison of the AUDIT, POSIT, CAGE, and CRAFFT. *Alcoholism: Clinical and Experimental Research, 27*(1), 67-73.

Knudsen, H. C., Vázquez-Barquero, J. L., Welcher, B., Gaite, L., Becker, T., Chisholm, D., et al. (2000). Translation and cross-cultural adaptation of outcome measurements for schizophrenia: EPSILON Study 2. *British Journal of Psychiatry, 177*(Suppl39), s8-s14.

Ko, C., & Cohen, H. (1998). Intraethnic comparison of eating attitudes in Native Koreans and Korean Americans using a Korean translation of the Eating Attitudes Test. *Journal of Nervous and Mental Disease, 186*(10), 631-636.

Kocabasoglu, N., Özdemir, A. Ç., Yargiç, I., & Geyran, P. (2005). Türkçe "PTSD Checklist - Civilian Version" (PCL-C) ölçeginin geçerlilik ve güvenilirligi. [The validity and safety of Turkish "PTSD Checklist - Civilian Version" (PCL-C) scale.]. *Yeni Symposium: psikiyatri, nöroloji ve davranis bilimleri dergisi, 43*(3), 126-134.

Kojima, M., Furukawa, T. A., Takahashi, H., Kawai, M., Nagaya, T., & Tokudome, S. (2002a). Cross-cultural validation of the Beck Depression Inventory-II in Japan. *Psychiatry Research, 110*(3), 291-299.

Kojima, M., Furukawa, T. A., Takahashi, H., Kawai, M., Nagaya, T., & Tokudome, S. (2002b). Cross-cultural validation of the Beck Depression Inventory II in Japan. *Psychiatry Research, 110*, 291-299.

Kokotailo, P. K., Egan, J., Gangnon, R., Brown, D., Mundt, M., & Fleming, M. (2004). Validity of the Alcohol Use Disorders Identification Test in college students. *Alcoholism: Clinical and Experimental Research, 28*(6), 914-920.

Kong, Y.-Y., Zhang, J., Jia, S.-H., & et al. (2007). [Reliability and validity of the Beck Hopelessness Scale for Adolescents.]. *Chinese Mental Health Journal, 21*(10), 686-689.

Kornblith, A. B., Herndon, J. E., Zuckerman, E., Viscoli, C. M., Horwitz, R. I., Cooper, M. R., et al. (2001). Social support as a buffer to the psychological impact of stressful life events in women with breast cancer. *Cancer, 91*(2), 443-454.

Koskelainen, M., Sourander, A., & Kalijonen, A. (2000). The Strengths and Difficulties Questionnaire among Finnish school-aged children and adolescents. *European Child and Adolescent Psychiatry, 9*(4), 277-284.

Koskelainen, M., Sourander, A., & Vauras, M. (2001). Self-reported strengths and difficulties in a community sample of Finnish adolescents. *European Child and Adolescent Psychiatry, 10*(3), 180-185.

Kravetz, S., Faust, M., & Shalit, M. (2001). The mental health of men with a psychiatric disorder or traumatic brain injury: Wives' labeling and husbands' perceived control. *Rehabilitation Psychology, 46*(4), 450-465.

Krenz, S., Dieckmann, S., Favrat, B., Spagnoli, J., Leutwyler, J., Schnyder, C., et al. (2004). French Version of the Addiction Severity Index (5th Edition): Validity and Reliability among Swiss Opiate-Dependent Patients. *European Addiction Research, 10*(4), 173-179.

Kubany, E. S., Haynes, S. N., Leisen, M. B., Owens, J. A., Kaplan, A. S., Watson, S. B., et al. (2000). Development and preliminary validation of a brief broad-spectrum measure of trauma exposure: the Traumatic Life Events Questionnaire. *Psychological Assessment, 12*(2), 210-224.

Kumar, G., Steer, R. A., & Beck, A. T. (1993). Factor structure of the Beck Anxiety Inventory with adolescent psychiatry inpatients. *Anxiety, Stress and Coping: An International Journal, 6*(2), 125-131.

Kuwert, P., Spitzer, C., Träder, A., Freyberger, H. J., & Ermann, M. (2007). Sixty years later: Post-traumatic stress symptoms and current psychopathology in former German children of World War II. *International Psychogeriatrics, 19*(5), 955-961.

Kvaal, K., Ulstein, I., Nordhus, I. H., & Engedal, K. (2005). The Spielberger State-Trait Anxiety Inventory (STAI): The state scale in detecting mental disorders in geriatric patients. *International Journal of Geriatric Psychiatry, 20*(7), 629-634.

LaBrie, J., Pedersen, E., & Earleywine, M. (2005). A group-administered Timeline Followback Assessment of alcohol use. *Journal of Studies on Alcohol, 66*(5), 693-697.

Lachar, D., Randle, S. L., Harper, R. A., Scott-Gurnell, K. C., Lewis, K. R., Santos, C. W., et al. (2001). The brief psychiatric rating scale for children (BPRS-C): Validity and reliability of an anchored version. *Journal of the American Academy of Child & Adolescent Psychiatry, 40*(3), 333-340.

Landolt, M. A., Vollrath, M., Timm, K., Gnehm, H. E., & Sennhauser, F. H. (2005). Predicting posttraumatic stress symptoms in children after road traffic accidents. *Journal of the American Academy of Child & Adolescent Psychiatry, 44*(12), 1276-1283.

Lang, A. J., & Stein, M. B. (2005). An abbreviated PTSD checklist for use as a screening instrument in primary care. *Behaviour Research and Therapy, 43*(5), 585-594.

Larson, C. O. (2002). Use of the SF-12 instrument for measuring the health of homeless persons. *Health Services Research, 37*, 733-750.

Laukka, P. (2007). Uses of music and psychological well-being among the elderly. *Journal of Happiness Studies, 8*, 215-241.

Lauzon, S., Corbiere, M., Bonin, J.-P., Bonsack, C., Lesage, A., & Ricard, N. (2001). Validation of the French version of the Health of the Nation Outcome Scale (HoNOS). *Canadian Journal of Psychiatry, 46*, 841-846.

Lawler-Row, K. A., & Piferi, R. L. (2006). The forgiving personality: Describing a life well lived? *Personality and Individual Differences, 41*(6), 1009-1020.

Lawrinson, P., Copeland, J., Gerber, S., & Gilmour, S. (2007). Determining a cut-off on the Severity of Dependence Scale (SDS) for alcohol dependence. *Addictive Behaviors, 32*(7), 1474-1479.

Lawrinson, P., Copeland, J., & Indig, D. (2003). *The Brief Treatment Outcome Measure: Opioid Maintenance Pharmacotherapy (BTOM) Manual*. Sydney: NDARC, University of New South Wales, NDARC Technical Report No. 156.

Lawrinson, P., Copeland, J., & Indig, D. (2005). Development and validation of a brief instrument for routine outcome monitoring in opioid maintenance pharmacotherapy services: The brief treatment outcome measure (BTOM). *Drug and Alcohol Dependence, 80*(1), 125-133

Lawrinson, P., Copeland, J., & Indig, D. (2006). Regional differences in injecting practices and other substance use-related behaviour among entrants into opioid maintenance pharmacotherapy treatment in New South Wales, Australia. *Drug and Alcohol Dependence, 82*(Supplement 1), S95-S102.

Layton, C. (1986). Note on test-retest characteristics of the General Health Questionnaire. *Perceptual and Motor Skills, 62*, 221-222.

Lecomte, T., Wallace, C. L., Caron, J., Perreault, M., & Lecomte, J. (2004). Further validation of the Client Assessment of Strengths Interests and Goals. *Schizophrenia Research, 66*(1), 59-70.

Lee, D.-G., Lee, H.-K., & Choi, Y.-H. (2008). *Validation of Dispositional Hope Scale in Korea*. Paper presented at the American Psychological Association 2008 Convention Presentation.

Lee, N., Jenner L., Kay-Lambkin, F., Hall, K., Dann, F., Roeg, S., et al. (2007). *PsyCheck: Responding to mental health issues within alcohol and drug treatment*. Canberra: Commonwealth of Australia.

Lee, S. (1993). How abnormal is the desire for slimness? A survey of eating attitudes and behaviour among Chinese undergraduates in Hong Kong. *Psychological Medicine, 23*(2), 437-451.

Lee, S., Kwok, K., Liau, C., & Leung, T. (2002). Screening Chinese patients with eating disorders using the eating attitudes test in Hong Kong. *International Journal of Eating Disorders, 32*(1), 91-97.

Leeuwen, K. V., Meerschaert, T., Bosmans, G., De Medts, L., & Braet, C. (2006). The strengths and difficulties questionnaire in a community sample of young children in flanders. *European Journal of Psychological Assessment, 22*(3), 189-197.

Leichner, P., Steiger, H., Puentes-Neuman, G., Perreault, M., & et al. (1994). Validation d'une échelle d'attitudes alimentaires auprès d'une population québécoise francophone. [Validation of the Eating Attitudes Test (EAT-26) in a French-speaking population of Quebec.]. *The Canadian Journal of Psychiatry / La Revue canadienne de psychiatrie, 39*(1), 49-54.

Lenert, L. A., Sherbourne, C. D., Sugar, C., & Wells, K. B. (2000). Estimation of utilities for the effects of depression from the SF-12. *Medical Care, 38*, 763-770.

Lennings, C. J. (1999). An evaluation of the Leeds Dependence Questionnaire. *Journal of Child and Adolescent Substance Abuse, 8*(3), 73-87.

Lennings, C. J., Copeland, J., & Howard, J. (2003). Substance use patterns of young offenders and violent crime. *Aggressive Behavior, 29*, 414-422.

Lennox, R., Dennis, M. L., Ives, M., & White, M. K. (2006). The construct and predictive validity of different approaches combining urine and self-reported drug use measures among adolescents in substance abuse treatment. *American Journal on Addictions, 15*(Suppl. 1), 92-101.

Lennox, R. D., Dennis, M. L., Scott, C. K., & Funk, R. R. (2006). Combining psychometric and biometric measures of substance use. *Drug and Alcohol Dependence, 83*, 95-103.

Leonhard, C., Mulvey, K., Gastfriend, D. R., & Shwartz, M. (2000). The Addiction Severity Index: A field study of internal consistence and validity. *Journal of Substance Abuse Treatment, 18*, 129-135.

Leucht, S., Kane, J. M., Kissling, W., Hamann, J., Etschel, E., & Engel, R. (2005). Clinical implications of Brief Psychiatric Rating Scale scores. *British Journal of Psychiatry, 187*(4), 366-371.

Levy, S., Sherritt, L., Harris, S. K., Gates, E. C., Holder, D. W., Kulig, J. W., et al. (2004). Test-Retest Reliability of Adolescents' Self-Report of Substance Use. *Alcoholism: Clinical and Experimental Research, 28*(8), 1236-1241.

Ley, A., Jeffery, D., Shaw, S., & Weaver, T. (2007). Development of a brief screen for substance misuse amongst people with severe mental health problems living in the community. *Journal of Mental Health, 16*(5), 679-690.

Leyfer, O. T., Ruberg, J. L., & Woodruff-Borden, J. (2006). Examination of the utility of the Beck Anxiety Inventory and its factors as a screener for anxiety disorders. *Journal of Anxiety Disorders, 20*(4), 444-458.

Liang, J., Wu, S. C., Krause, N. M., Chiang, T. L., & Wu, H. (1992). The structure of the Mental Health Inventory among Chinese in Taiwan. *Medical Care, 30*(8), 659-676.

Liddle, H. A., Rowe, C. L., Dakof, G. A., Henderson, C. E., & Greenbaum, P. E. (2009). Multidimensional family therapy for young adolescent substance abuse: Twelve-month outcomes of a randomized controlled trial. *Journal of Consulting and Clinical Psychology, 77*(1), 12-25.

Ligon, J., & Thyer, B. A. (2000). Interrater reliability of the Brief Psychiatric Rating Scale used at a community-based inpatient crisis stabilization unit. *Journal of Clinical Psychology, 56*(4), 583-587.

Lim, L. L., & Fisher, J. D. (1999). Use of the 12-item short-form (SF-12) Health Survey in an Australian heart and stroke population. *Quality of Life Research 8*, 1-8.

Lima, C. T., Freire, A. C. C., Silva, A. P. B., Teixeira, R. M., Farrell, M., & Prince, M. (2005). Concurrent and construct validity of the AUDIT in an urban Brazilian sample. *Alcohol and Alcoholism, 40*(6), 584-589.

Liskow, B., Campbell, J., Nickel, E., & Powell, B. J. (1995). Validity of the CAGE questionnaire in screening for alcohol dependence in a walk-in (triage) clinic. *Journal of Studies on Alcohol, 56*(3), 277-281.

Liu, Z., Lian, Z., Zhou, W., Mu, Y., Teng, Y., Zhu, R., et al. (2000). A preliminary study on the severity of drug addiction in drug abusers. *Chinese Mental Health Journal, 14*(4), 231-234.

Ljubotina, D., & Muslic, L. (2003). Convergent validity of four instruments for measuring posttraumatic stress disorder. *Review of Psychology, 10*(1), 11-21.

Loeber, S., Croissant, B., Heinz, A., Mann, K., & Flor, H. (2006). Cue exposure in the treatment of alcohol dependence: Effects on drinking outcome, craving and self-efficacy. *British Journal of Clinical Psychology, 45*(4), 515-529.

Long, J., & Brekke, J. (1999). Longitudinal factor structure of the Brief Psychiatric Rating Scale in schizophrenia. *Psychological Assessment, 11*, 498-506.

Lora, A., Bai, G., Bianchi, S., Bolongaro, G., Civenti, G., Erlicher, A., et al. (2001). The Italian version of HoNOS (Health of the Nation Outcome Scales), a scale for evaluating the outcome and the severity in mental health services. *Epidemiologia e Psichiatria Sociale, 10*, 198-204.

Lovibond, S. H., & Lovibond, P. F. (1995). *Manual for the Depression Anxiety Stress Scales* (2nd ed.). Sydney: Psychology Foundation.

Lozano, O. M., Domingo-Salvany, A., Martinez-Alonso, M., Brugal, M. T., Alonso, J., & de la Fuente, L. (2008). Health-related quality of life in young cocaine users and associated factors. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care & Rehabilitation, 17*(7), 977-985.

Lundberg, L., Johannesson, M., Isacson, D. G., & Borgquist, L. (1999). The relationship between health-state utilities and the SF-12 in a general population. *Medical Decision Making, 19*, 128-140.

Luty, J., & O'Gara, C. (2006). Validation of the 13-Item Beck Depression Inventory in alcohol-dependent people. *International Journal of Psychiatry in Clinical Practice, 10*(1), 45-51.

Luty, J., Perry, V., Umoh, O., & Gormer, D. (2006). Validation and development of a self-report outcome measure (MAP-sc) in opiate addiction. *Psychiatry Bulletin, 30*(4), 134-139.

Lykke, J., Hesse, M., Austin, S. F., & Oestrich, I. (2008). Validity of the BPRS, the BDI and the BAI in dual diagnosis patients. *Addictive Behaviors, 33*(2), 292-300.

MacKenzie, D. M., Langa, A., & Brown, T. M. (1996). Identifying hazardous or harmful alcohol use in medical admissions: A comparison of AUDIT, CAGE and brief MAST. *Alcohol and Alcoholism, 31*(6), 591-599.

Macpherson, R., Collins-Atkins, C., Gregory, N., Slade, M., & Lerescu, T. (2008). The relationships between user, carer and staff perceptions of need in an assertive outreach team. *Journal of Mental Health, 17*(5), 452-461.

Macran, S., Weatherly, H., & Kind, P. (2003). Measuring population health: a comparison of three generic health status measures. *Medical Care, 41*, 218-231.

Maercker, A., & Schützwohl, M. (1998). Erfassung von psychischen Belastungsfolgen: Die Impact of Event Skala-revidierte Version (IES-R). [Assessment of post-traumatic stress reactions: The Impact of Event Scale-Revised (IES-R).]. *Diagnostica, 44*(3), 130-141.

Maisto, S. A., Carey, M. P., Carey, K. B., Gordon, C. M., & Gleason, J. R. (2000). Use of the AUDIT and the DAST-10 to identify alcohol and drug use disorders among adults with a severe and persistent mental illness. *Psychological Assessment., 12*(2), 186-192.

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

Malakouti, S. K., Fatollahi, P., Mirabzadeh, A., & Zandi, T. (2007). Reliability, validity and factor structure of the GHQ-28 used among elderly Iranians. *International Psychogeriatrics, 19*(4), 623-634.

Malet, L., Schwan, R., Boussiron, D., Aublet-Cuvelier, B., & Llorca, P.-M. (2005). Validity of the CAGE questionnaire in hospital. *European Psychiatry, 20*(7), 484-489.

Malmberg, M., Rydell, A. M., & Smedje, H. (2003). Validity of the Swedish version of the Strengths and Difficulties Questionnaire (SDQ-Swe). *Nordic Journal of Psychiatry, 57*(5), 357-363.

Maloney, M. J., McGuire, J., & Daniels, S. R. (1988). Reliability testing of a children's version of the Eating Attitude Test. *Journal of the American Academy of Child and Adolescent Psychiatry, 27*, 541-543.

Mandersen, J., Nizzoli, U., Corbelli, C., Margaron, H., Torres, M. A., Castro del Prado, I., et al. (2001). Reliability of the Maudsley Addiction Profile ERIT Version (MAP) and the Treatment Perceptions Questionnaire (TPQ) in Italy, Spain and Portugal for the evaluation of treatments. *Adicciones, 13*(2), 217-227.

Manne, S., & Schnoll, R. (2001). Measuring cancer patients' psychological distress and well-being: A factor analytic assessment of the Mental Health Inventory. *Psychological Assessment, 13*(1), 99-109.

Manne, S. L., Du Hamel, K., Gallelli, K., Sorgen, K., & Redd, W. H. (1998). Posttraumatic stress disorder among mothers of pediatric cancer survivors: Diagnosis, comorbidity, and utility of the PTSD Checklist as a screening instrument. *Journal of Pediatric Psychology, 23*(6), 357-366.

Maremmani, I., Pacini, M., Lubrano, S., Perugi, G., Tagliamonte, A., Pani, P. P., et al. (2008). Long-term outcomes of treatment-resistant heroin addicts with and without DSM-IV Axis 1 psychiatric comorbidity (dual diagnosis). *European Addiction Research, 14*(3), 134-142.

Mares, A., Greenberg, G., & Rosenheck, R. (2008). Client-level Measures of Services Integration Among Chronically Homeless Adults. *Community Mental Health Journal, 44*(5), 367-376.

Marsden, J., Farrell, M., Bradbury, C., Dale-Perera, A., Eastwood, B., Roxburgh, M., et al. (2008). Development of the treatment outcomes profile. *Addiction, 103*, 1450-1460.

Marsden, J., Gossop, M., Stewart, D., Best, D., Farrell, M., Lehmann, P., et al. (1998). The Maudsley Addiction Profile (MAP): A brief instrument for assessing treatment outcome. *Addiction, 93*(12), 1857-1868.

Marsden, J., Gossop, M., Stewart, D., Rolfe, A., & Farrell, M. (2000). Psychiatric symptoms among clients seeking treatment for drug dependence. Intake data from the National Treatment Outcome Research Study [see comments]. *British Journal of Psychiatry, 176*, 285-289.

Marsden, J., Nizzoli, U., Corbelli, C., Margaron, H., Torres, M., Prada De Castro, I., et al. (2000). New European instruments for treatment outcome research: reliability of the maudsley addiction profile and treatment perceptions questionnaire in Italy, Spain and Portugal. *European Addiction Research, 6*(3), 115-122.

Martin, G., & Copeland, J. (2008). The adolescent cannabis check-up: Randomized trial of a brief intervention for young cannabis users. *Journal of Substance Abuse Treatment, 34*(4), 407-414.

Martin, G., Copeland, J., Gates, P., & Gilmour, S. (2006). The Severity of Dependence Scale (SDS) in an adolescent population of cannabis users: Reliability, validity and diagnostic cut-off. *Drug and Alcohol Dependence, 83*(1), 90-93.

Martin, G., Copeland, J., Gilmour, S., Gates, P., & Swift, W. (2006). The Adolescent Cannabis Problems Questionnaire (CPQ-A): Psychometric properties. *Addictive Behaviors, 31*(12), 2238-2248.

Martinez, S., Stillerman, L., & Waldo, M. (2005). Reliability and Validity of the SCL-90-R With Hispanic College Students. *Hispanic Journal of Behavioral Sciences, 27*(2), 254-264.

Martino, S., Grilo, C. M., & Fehon, D. C. (2000). Development of the Dug Abuse Screening Test for adolescents (DAST-A). *Addictive Behaviors, 25*, 57-70.

Masskulpan, P., Riewthong, K., Dajpratham, P., & Kuptniratsaikul, V. (2008). Anxiety and depressive symptoms after stroke in 9 rehabilitation centers. *Journal of the Medical Association of Thailand, 91*(10), 1595-1602.

Mathai, J., Anderson, P., & Bourne, A. (2003). Use of the Strengths and Difficulties Questionnaire as an outcome measure in a child and adolescent mental health service. *Australasian Psychiatry, 11*, 334-337.

Mathai, J., Anderson, P., & Bourne, A. (2004). Comparing psychiatric diagnoses generated by the Strengths and Difficulties Questionnaire with diagnoses made by clinicians. *Australian New Zealand Journal of Psychiatry, 38*(8), 639-643.

Mathai, J., Jespersen, S., Bourne, A., Donegan, T., Akinbiyi, A., & Gray, K. (2008). Use of the Strengths and Difficulties Questionnaire in identifying emotional and behavioural problems in children of parents with a mental illness in Australia. *Australian e-Journal for the Advancement of Mental Health (AeJAMH), 7*(3), 1446-7984.

Matsuishi, T., Nagano, M., Araki, Y., Tanaka, Y., Iwasaki, M., Yamashita, Y., et al. (2008). Scale properties of the Japanese version of the Strengths and Difficulties Questionnaire (SDQ): A study of infant and school children in community samples. *Brain & Development, 30*(6), 410-415.

Mattick, R., Oliphant, D., Bell, J., & Hall, W. (1996). *Psychiatric morbidity in methadone maintenance patients: prevalence, effect on drug use and detection.* Paper presented at the Substance use and mental illness: Proceedings of the fourth Lingard symposium, Newcastle Hunter Institute of Mental Health.

Mattick, R. P., & Hall, W. (1993). *A treatment outline for approaches to opioid dependence*. Sydney: National Drug and Alcohol Research Centre (NDARC).

Mátyássy, A., Kelemen, O., Sárközi, Z., Janka, Z., & Kéri, S. (2006). Recognition of complex mental states in patients with alcoholism after long-term abstinence. *Alcohol and Alcoholism, 41*(5), 512-514.

McCabe, C. J., Thomas, K. J., Brazier, J. E., & Coleman, P. (1996). Measuring the mental health status of a population: a comparison of the GHQ-12 and the SF-36 (MHI-5). *British Journal of Psychiatry, 169*(4), 516-521.

McCallum, J. (1995). The SF 36 in an Australian sample: validating a new, generic health status measure. *Australian Journal of Public Health, 19*, 160-166.

McCallum, J. (1997). *The shorter short form: Analysis of SF-12 items in Australian data*. Canberra: Australian Institute of Health and Welfare.

McCambridge, J., & Maria, D. (2008). Randomized controlled trial of the effects of completing the Alcohol Use Disorders Identification Test questionnaire on self-reported hazardous drinking. *Addiction, 103*(2), 241-248.

McCarty, D., Fuller, B., Kaskutas, L. A., Wendt, W. W., Nunes, E. V., Miller, M., et al. (2008). Treatment programs in the National Drug Abuse Treatment Clinical Trials Network. *Drug and Alcohol Dependence, 92*(1-3), 200-207.

McClelland, R., Trimble, P., Fox, M. L., Stevenson, M. R., & Bell, B. (2000). Validation of an outcome scale for use in adult psychiatric practice. *Quality in Health Care, 9*(2), 98-105.

McCrone, P., Bhui, K., Craig, T., Mohamud, S., Warfa, N., Stansfeld, S. A., et al. (2005). Mental health needs, service use and costs among Somali refugees in the UK. *Acta Psychiatrica Scandinavica, 111*(5), 351-357.

McCrone, P., Leese, M., Thornicroft, G., Schene, A. H., Knudsen, H. C., Vázquez-Barquero, J. L., et al. (2000). Reliability of the Camberwell Assessment of Need--European Version: EPSILON Study 6. *British Journal of Psychiatry, 177*(Suppl39), s34-s40.

McCusker, J., Bigelow, C., Servignon, C., & Zorn, M. (1994). Test-retest reliability of the Addiction Severity Index Composite Scores among clients in residential treatment. *American Journal on Addictions, 3*, 254-262.

McEvoy, P. M., & Shand, F. (2008). The effect of comorbid substance use disorders on treatment outcome for anxiety disorders. *Journal of Anxiety Disorders, 22*(6), 1087-1098.

McGregor, C., Machin, A., & White, J. M. (2003). In-patient benzodiazepine withdrawal: comparison of fixed and symptom triggered taper methods. *Drug & Alcohol Review, 22*, 175-180.

McHorney, C. A., Ware, J. E. J., Lu, J. F. R., & Sherbourne, C. D. (1994). The MOS 36 Item Short Form Health Survey (SF 36): 3.Tests of data quality, scaling assumptions, and reliability across diverse patient groups. *Medical Care, 32*, 40-66.

McHorney, C. A., Ware, J. E. J., & Raczek, A. E. (1993). The MOS 36 Item Short Form Health Survey (SF36): 2. Psychometric and clinical tests of validity measuring physical and mental health constructs. *Medical Care, 31*, 247-263.

McKibbin, C., Patterson, T. L., & Jeste, D. V. (2004). Assessing Disability in Older Patients With Schizophrenia Results From the WHODAS-II. *Journal of Nervous and Mental Disease, 192*(6), 405-413.

McLellan, A., Luborsky, L., Woody, G., & O'Brien, C. (1980). An improved diagnostic evaluation instrument for substance abuse patients. The Addiction Severity Index. *Journal of Nervous and Mental Disease, 168*, 26-33.

McLellan, A. T., Cacciola, J. S., & Alterman, A. I. (2004). Commentaries on Mäkelä: The ASI as a still developing instrument: Response to Mäkelä. *Addiction, 99*(4), 411-412.

McLellan, A. T., Kushner, H., Metzger, D., Peters, R., Smith, I., Grissom, G., et al. (1992). The Fifth Edition of the Addiction Severity Index. *Journal of Substance Abuse Treatment, 9*(3), 199-213.

McLellan, A. T., Luborsky, L., Cacciola, J., Griffiths, J., Evans, F., Barr, H. L., et al. (1985). New data from the Addiction Severity Index; Reliability and validity in three centers. *Journal of Nervous and Mental Disorders, 173*, 412-423.

McMurran, M., & Hollins, C. R. (1989). The Short Alcohol Dependence Data (SADD) Questionnaire: Norms and reliability data for male young offenders. *British Journal of Addiction, 84*, 315-318.

McNaught, M., Caputi, P., Oades, L. G., & Deane, F. P. (2007). Testing the validity of the Recovery Assessment Scale using an Australian sample. *Australian & New Zealand Journal of Psychiatry, 41*(5), 450-457.

McShane, G., Bazzano, C., Walter, G., & Barton, G. (2007). Outcome of patients attending a specialist educational and mental health service for social anxiety disorders. *Clinical Child Psychology & Psychiatry, 12*(1), 117-124.

Meehan, J. P., Webb, M. G. T., & Unwin, A. R. (1985). The severity of alcohol dependence questionnaire in a sample of Irish problem drinkers. *British Journal of Addiction, 80*, 57-63.

Melberg, H. O. (2004). Three problems with the ASI composite scores. *Journal of Substance Use, 9*(3-4), 120-126.

Mellor, D. (2004). Furthering the use of the Strengths and Difficulties Questionnaire: Reliability with younger child respondents. *Psychological Assessment, 16*(4), 396-401.

Mellor, D. (2005). Normative data for the Strengths and Difficulties Questionnaire in Australia. *Australian Psychologist, 40*(3), 215-222.

Mellor, D., & Stokes, M. (2007). The factor structure of the Strengths and Difficulties Questionnaire. *European Journal of Psychological Assessment, 23*(2), 105-112.

Miele, G. M., Carpenter, K. M., Cockerham, M. S., Trautman, K. D., Blaine, J., & Hasin, D. (2000a). Substance Dependence Severity Scale (SDSS): reliability and validity of a clinician-administered interview for DSM-IV substance use disorders. *Drug and Alcohol Dependence, 59*(1), 63-75.

Miele, G. M., Carpenter, K. M., Cockerham, M. S., Trautman, K. D., Blaine, J., & Hasin, D. S. (2000b). Concurrent and predictive validity of the Substance Dependence Severity Scale (SDSS). *Drug and Alcohol Dependence, 59*(1), 77-88.

Miele, G. M., Carpenter, K. M., Cockerham, M. S., Trautman, K. D., Blaine, J., & Hasin, D. S. (2001). Substance Dependence Severity Scale: Reliability and validity for ICD-10 substance use disorders. *Addictive Behaviors, 26*(4), 603-612.

Miles, H., Johnson, S., Amponsah-Afuwape, S., Finch, E., Leese, M., & Thornicroft, G. (2003). Characteristics of subgroups of individuals with psychotic illness and a comorbid substance use disorder. *Psychiatry Services, 54*(4), 554-561.

Miles, J. N. V., Marshall, G. N., & Schell, T. L. (2008). Spanish and English versions of the PTSD Checklist-Civilian Version (PCL-C): Testing for differential item functioning. *Journal of Traumatic Stress, 21*(4), 369-376.

Mills, K. L., Teesson, M., Darke, S., Ross, J., & Lynskey, M. (2004). Young people with heroin dependence: Findings from the Australian Treatment Outcome Study (ATOS). *Journal of Substance Abuse Treatment, 27*(1), 67-73.

Milne, D., Reichelt, K., & Wood, E. I. (2001). Implementing HoNOS: An eight stage approach. *Clinical Psychology and Psychotherapy, 8*(2), 106-116.

Mintz, L. B., & O'Halloran, M. S. (2000). The Eating Attitudes Test: Validation with DSM-IV eating disorder criteria. *Journal of Personality Assessment, 74*(3), 489-503.

Mirzamani, S.-M., Mohammadi, M. R., Mahmoudi-Gharaei, J., & Mirzamani, M.-S. (2007). Validity of The PTSD Symptoms Scale Self Report (PSS-SR) in Iran. *Iranian Journal of Psychiatry 2*, 120-123.

Mischke, H. D., & Venneri, R. L. (1987). Reliability and validity of the MAST, Mortimer Filkins questionnaire and CAGE in DWI assessment. *Journal of Studies on Alcohol, 48*, 492-501.

Mohr, S., Simon, A., Favrod, J., Fokianos, C., & Ferrero, F. (2004). Validation de la version française du Profil des compétences de vie (Life Skills Profile) chez des personnes atteintes de schizophrénie. [Validation of the French Version of the Life Skills Profile with people suffering of schizophrenia.]. *L'Encéphale, 30*(4), 343-351.

Moller, T., & Linaker, O. M. (2006). The long-term level of functioning in patients admitted for psychotic disorders with and without substance abuse. *Nordic Journal of Psychiatry, 60*(2), 121-125.

Moraes, C. L., Viellas, E. F., & Reichenheim, M. E. (2005). Assessing alcohol misuse during pregnancy: Evaluating psychometric properties of the CAGE, T-ACE and TWEAK in a Brazilian setting. *Journal of Studies on Alcohol, 66*(2), 165-173.

Morgan, T. J., Morgenstern, J., Blanchard, K. A., Labouvie, E., & Bux, D. A. (2003). Health-related quality of life for adults participating in outpatient substance abuse treatment. *American Journal of Addictions, 12*, 98-210.

Morlan, K. K., & Tan, S.-Y. (1998). Comparison of the Brief Psychiatric Rating Scale and the Brief Symptom Inventory. *Journal of Clinical Psychology, 54*(7), 885-894.

Morley, K. C., Teesson, M., Reid, S. C., Sannibale, C., Thomson, C., Phung, N., et al. (2006). Naltrexone versus acamprosate in the treatment of alcohol dependence: A multi-centre, randomized, double-blind, placebo-controlled trial. *Addiction, 101*(10), 1451-1462.

Morosini, P., Gigantesco, A., Mazzarda, A., & Gibaldi, L. (2003). HoNOS-Rome: An expanded, customized, and longitudinally oriented version of the HoNOS. *Epidemiologia e Psichiatria Sociale, 12*, 53-62.

Morris-Yates, A., Barber, H., Harris, M., & Zapart, S. (1999). *Trainer's Manual for the Health of the Nation Outcome Scales, Version 5, Adult*. Canberra: Commonwealth Department of Health and Aged Care.

Morrow-Howell, N., Proctor, E., Choi, S., Lawrence, L., Brooks, A., Hasche, L., et al. (2008). Depression in public community long-term care: Implications for intervention development. *The Journal of Behavioral Health Services & Research, 35*(1), 37-51.

Morton, J. L., Jones, T. V., & Manganaro, M. A. (1996). Performance of alcoholism screening questionnaires in elderly veterans. *American Journal of Medicine, 101*(2), 153-159.

Mossbarger, B. (2005). A Modified Global Assessment of Functioning (GAF) Scale for Use in Long-Term Care Settings. *Educational Gerontology, 31*(9), 715-725.

Mote, T. A., Natalicio, L. F., & Rivas, F. (1971). Comparability of the Spanish and English editions of the Spielberger State-Trait Anxiety Inventory. *Journal of Cross-Cultural Psychology, 2*(2), 205-206.

Mozley, C. G., Huxley, P., Sutcliffe, C., Bagley, H., Burns, A., Challis, D., et al. (1999). 'Not knowing where I am doesn't mean I don't know what I like': Cognitive impairment and quality of life responses in elderly people. *International Journal of Geriatric Psychiatry, 14*(9), 776-783.

Mubarak, A. R., & Barber, J. G. (2003). Emotional expressiveness and the quality of life of patients with schizophrenia. *Social Psychiatry and Psychiatric Epidemiology, 38*(7), 380-384.

Mueser, K., Curran, P., & McHugo, G. J. (1997). Factor structure of the Brief Psychiatric Rating Scale in schizophrenia. *Psychological Assessment, 9*, 196-204.

Mueser, K. T., Rosenberg, S. D., Fox, L., Salyers, M. P., Ford, J. D., & Carty, P. (2001). Psychometric evaluation of trauma and posttraumatic stress disorder assessments in persons with severe mental illness. *Psychological Assessment, 13*(1), 110-117.

Mugno, D., Ruta, L., D'Arrigo, V. G., & Mazzone, L. (2007). Impairment of quality of life in parents of children and adolescents with pervasive developmental disorder. *Health & Quality of Life Outcomes, 5*, 22.

Mullick, M. S. I., & Goodman, R. (2001). Questionnaire screening for mental health problems in Bangladeshi children: A preliminary study. *Social Psychiatry and Psychiatric Epidemiology, 36*(2), 94-99.

Muris, P., Mayer, B., & Merckelbach, H. (1998). Trait anxiety as a predictor of behaviour therapy outcome in spider phobia. *Behavioural and Cognitive Psychotherapy, 26*(1), 87-91.

Muris, P., Meesters, C., & van den Berg, F. (2003). The Strengths and Difficulties Questionnaire (SDQ): Further evidence for its reliability and validity in a community sample of Dutch children and adolescents. *European Child and Adolescent Psychiatry, 12*(1), 1-8.

Murphy, B., Herrman, H., Hawthorne, G., Pinzone, T., & Evert, H. (2000). *Australian WHOQoL instruments: User's manual and interpretation guide*. Melbourne, Australia: Australian WHOQoL Field Study Centre.

Myrick, H., & Brady, K. (2003). Current review of the comorbidity of affective, anxiety and substance use disorders. *Current Opinion in Psychiatry, 16*, 261-270.

Mystakidou, K., Tsilika, E., Parpa, E., Galanos, A., & Vlahos, L. (2007). Psychometric properties of the Impact of Event Scale in Greek cancer patients. *Journal of Pain and Symptom Management, 33*(4), 454-461.

Nagel, B. J., Schweinsburg, A. D., Phan, V., & Tapert, S. F. (2005). Reduced hippocampal volume among adolescents with alcohol use disorders without psychiatric comorbidity. *Psychiatry Research: Neuroimaging, 139*(3), 181-190.

Najavits, L. M., Weiss, R. D., Reif, S., Gastfriend, D. R., Siqueland, L., Barber, J. P., et al. (1998). The Addiction Severity Index as a screen for trauma and posttraumatic stress disorder. *Journal of Studies on Alcohol, 59*(1), 56-62.

Nakai, Y. (2003). [Validity of Eating Attitudes Test (EAT).]. *Seishin Igaku (Clinical Psychiatry), 45*(2), 161-165.

Navarro, P., Ascaso, C., Garcia-Esteve, L., Aguado, J., Torres, A., & Martín-Santos, R. (2007). Postnatal psychiatric morbidity: A validation study of the GHQ-12 and the EPDS as screening tools. *General Hospital Psychiatry, 29*(1), 1-7.

Nelson, C. B., & Lotfy, M. (1999). *The World health Organization's WHOQOL-Bref quality of life assessment of psychometric qualities. Results of a field trial*: WHO (MNH/MHP/99.7).

Neumärker, K. J., Bettle, N., Bettle, O., Dudeck, U., & Neumärker, U. (1998). The Eating Attitudes Test: Comparative analysis of female and male students at the Public Ballet School of Berlin. *European Child & Adolescent Psychiatry, 7*(1), 19-23.

Newcombe, D. A. L., Humeniuk, R. E., & Ali, R. (2005). Validation of the World Health Organization Alcohol, Smoking and Substance Involvement Screening Test (ASSIST): Report of results from the Australian site. *Drug and Alcohol Review, 24*(3), 217-226.

Ng, F., Trauer, T., Dodd, S., Callaly, T., Campbell, S., & Berk, M. (2007). The validity of the 21-item version of the Depression Anxiety Stress Scales as a routine clinical outcome measure. *Acta Neuropsychiatrica, 19*(5), 304-310.

Nielsen, L. F., Middelboe, T., Werdelin, G., Petersen, L., Mackeprang, T., Mortensen, E. L., et al. (1999). Need assessment in long-term psychiatric patients. *Nordic Journal of Psychiatry, 53*(5), 371-376.

Nielsen, L. F., Petersen, L., Werdelin, G., Hou, B. J., & Lindhardt, A. (2000). Patients who are difficult to place: A description of the patient characteristics, admission patterns, and need for care. *Nordic Journal of Psychiatry, 54*(2), 135-141.

Nielssen, O., & Misrachi, S. (2005). Prevalence of psychoses on reception to male prisons in New South Wales. *Australian and New Zealand Journal of Psychiatry, 39*(6), 453 - 459.

Nimeus, A., Traskman Bendz, L., & Alsen, M. (1997). Hopelessness and suicidal behavior. *Journal of Affective Disorders, 42*(2-3), 137-144.

Norman, R. M. G., Malla, A. K., McLean, T., Voruganti, L. P., Cortese, L., MacIntosh, E., et al. (2000). The relationship of symptoms and level of functioning in schizophrenia to general wellbeing and the Quality of Life Scale. *Acta Psychiatrica Scandinavica, 102*(4), 303-309.

Norton, J., de Roquefeuil, G., Benjamins, A., Boulenger, J. P., & Mann, A. (2004). Psychiatric morbidity, disability and service use amongst primary care attenders in France. *European Psychiatry, 19*(3), 164-167.

NSW Department of Health. (2004). *Clinician's reference guide to the BTOM-C v1.0: Methadone/Buprenorphine treatment* Canberra: Australian Government Department of Health and Ageing.

NSW Department of Health. (2007). *Mental health reference resource for drug and alcohol workers*. Sydney: NSW Department of Health.

NSW Health. (2008). *Drug and alcohol psychosocial interventions: Professional practice guidelines*. Sydney: NSW Department of Health.

Nunes, M. A., Camey, S., Olinto, M. T. A., & Mari, J. J. (2005). The validity and 4-year test-retest reliability of the Brazilian version of the Eating Attitudes Test-26. *Brazilian Journal of Medical and Biological Research, 38*(11), 1655-1662.

O'Carroll, R. E., Smith, K., Cousten, M., Cossar, J. A., & Hayes, P. C. (2000). A comparison of the WHOQOL-100 and the WHOQOL-Bref in detecting change in quality of life following liver transplantation. *Quality of Life Research, 9*, 121-124.

O'Connor, D. W., & Parslow, R. A. (2008). Different responses to K-10 and CIDI suggest that complex structured psychiatric interviews underestimate rates of mental disorder in old people. *Psychological Medicine*, Published online by Cambridge University Press 02 Dec 2008.

O'Hare, T., Sherrer, M. V., LaButti, A., & Emrick, K. (2004). Validating the Alcohol Use Disorders Identification Test With Persons Who Have a Serious Mental Illness. *Research on Social Work Practice, 14*(1), 36-42.

O'Hare, T., & Tran, T. V. (1997). Predicting problem drinking in college students: Gender differences and the CAGE questionnaire. *Addictive Behaviors, 22*(1), 13-21.

O'Leary, T. A., Rohsenow, D. J., Martin, R., Colby, S. M., Eaton, C. A., & Monti, P. M. (2000). The relationship between anxiety levels and outcome of cocaine abuse treatment. *American Journal of Drug and Alcohol Abuse, 26*(2), 179-194.

Ober, C., & Schlesinger, C. (2005). *Indigenous Risk Impact Screen (IRIS) user manual*. Brisbane: Queensland Health, Queensland Government.

Ocker, L. B., Lam, E. T. C., Jensen, B. E., & Zhang, J. J. (2007). Psychometric properties of the Eating Attitudes Test. *Measurement in Physical Education and Exercise Science, 11*(1), 25-48.

Olde, E., Kleber, R. J., van der Hart, O., & Pop, V. J. M. (2006). Childbirth and posttraumatic stress responses: A validation study of the Dutch Impact of Event Scale-Revised. *European Journal of Psychological Assessment, 22*(4), 259-267.

Orrell, M., Yard, P., Handysides, J., & Schapira, R. (1999). Validity and reliability of the Health of the Nation Outcome Scales in psychiatric patients in the community. *British Journal of Psychiatry, 174*, 409-412.

Oslin, D. W., Slaymaker, V. J., Blow, F. C., Owen, P. L., & Colleran, C. (2005). Treatment outcomes for alcohol dependence among middle-aged and older adults. *Addictive Behaviors, 30*(7), 1431-1436.

Osman, A., Hoffman, J., Barrios, F. X., Kopper, B. A., Breitenstein, J. L., & Hahn, S. K. (2002). Factor structure, reliability, and validity of the Beck Anxiety Inventory in adolescent psychiatric inpatients. *Journal of Clinical Psychology, 58*(4), 443-456.

Osman, A., Kopper, B. A., Barrios, F., Gutierrez, P. M., & Bagge, C. L. (2004). Reliability and validity of the Beck Depression Inventory--II with adolescent psychiatric inpatients. *Psychological Assessment, 16*(2), 120-132.

Osman, A., Kopper, B. A., Barrios, F. X., Osman, J. R., & Wade, T. (1997). The Beck Anxiety Inventory: Reexamination of factor structure and psychometric properties. *Journal of Clinical Psychology, 53*(1), 7-14.

Ostroff, J. S., Woolverton, K. S., Berry, C., & Lesko, L. M. (1996). Use of the Mental Health Inventory with adolescents: A secondary analysis of the Rand Health Insurance Study. *Psychological Assessment, 8*(1), 105-107.

Ouimette, P., Wade, M., Prins, A., & Schohn, M. (2008). Identifying PTSD in primary care: Comparison of the Primary Care-PTSD screen (PC-PTSD) and the General Health Questionnaire-12 (GHQ). *Journal of Anxiety Disorders, 22*(2), 337-343.

Overall, J., & Gorham, D. R. (1962). The Brief Psychiatric Rating Scale. *Psychological Reports, 10*, 799-812.

Ownby, R. L., & Seibel, H. P. (1994). A factor analysis of the Brief Psychiatric Rating Scale in an older psychiatric population: Exploratory and confirmatory analyses. *Multivariate Experimental Clinical Research, 10*(3), 145-156.

Oyefeso, A., Brown, S., Chiang, Y., & Clancy, C. (2008). Self-injurious behaviour, traumatic life events and alexithymia among treatment-seeking opiate addicts: Prevalence, pattern and correlates. *Drug and Alcohol Dependence, 98*(3), 227-234.

Padaiga, Z., Subata, E., & Vanagas, G. (2007). Outpatient methadone maintenance treatment program. Quality of life and health of opioid-dependent persons in Lithuania. *Medicina (Kaunas, Lithuania), 43*(3), 235-241.

Page, A. C., Hooke, G. R., & Rutherford, E. M. (2001). Measuring mental health outcomes in a private psychiatric clinic: Health of the Nation Outcome Scales and Medical Outcomes Short Form SF-36. *Australian and New Zealand Journal Psychiatry, 35*(3), 377-381.

Pan, P. C., & Goldberg, D. P. (1990). A comparison of the validity of GHQ-12 and CHQ-12 in Chinese primary care patients in Manchester. *Psychological Medicine, 20*(4), 931-940.

Panos, P. T. (2004). The Validation of the factor structure of the Brief Psychiatric Rating Scale-Expanded Version (BPRS-E) with geriatric and nongeriatric psychiatric inpatients. *Research on Social Work Practice, 14*(3), 180-190.

Pariente, P. D., Lépine, J. P., Boulenger, J. P., Zarifian, E., & et al. (1989). The Symptom Check-List 90R (SCL-90-R) in a French general psychiatric 708 outpatient sample: Is there a factor structure? *European Psychiatry, 4*(3), 151-157.

Parker, G., & Hadzi-Pavlovic, D. (1995). The capacity of a measure of disability (the LSP) to predict hospital readmission in those with schizophrenia. *Psychological Medicine, 25*(1), 157-163.

Parker, G., O'Donnell, M., Hadzi-Pavlovic, D., & Proberts, M. (2002). Assessing outcome in community mental health patients: A comparative analysis of measures. *International Journal of Social Psychiatry, 48*(1), 11-19.

Parker, G., Rosen, A., Emdur, N., & Hadzi-Pavlov, D. (1991). The Life Skills Profile: psychometric properties of a measure assessing function and disability in schizophrenia. *Acta Psychiatrica Scandinavica, 83*(2), 145-152.

Parslow, R. A., & Jorm, A. F. (2006). Tobacco use after experiencing a major natural disaster: Analysis of a longitudinal study of 2063 young adults. *Addiction, 101*(7), 1044-1050.

Parslow, R. A., & Jorm, A. F. (2007). Pretrauma and posttrauma neurocognitive functioning and PTSD symptoms in a community sample of young adults. *American Journal of Psychiatry, 164*(3), 509-515.

Parslow, R. A., Jorm, A. F., & Christensen, H. (2006). Associations of pre-trauma attributes and trauma exposure with screening positive for PTSD: Analysis of a community-based study of 2085 young adults. *Psychological Medicine, 36*(3), 387-395.

Parsons, K. J., WallBrown, F. H., & Myers, R. W. (1994). Michigan Alcoholism Screening Test: Evidence supporting general as well as specific factors. *Educational and Psychological Measurement, 54*(2), 530-536.

Patterson, D. A., & Lee, M. S. (1995). Field trial of the Global Assessment of Functioning Scale--Modified. *American Journal of Psychiatry, 152*(9), 1386-1388.

Pedersen, E. R., & LaBrie, J. W. (2006). A Within-Subjects Validation of a Group-Administered Timeline Followback for Alcohol Use. *Journal of Studies on Alcohol, 67*(2), 332-335.

Peel, N. M., Bartlett, H. P., & Marshall, A. L. (2007). Measuring quality of life in older people: Reliability and validity of WHOQOL-OLD. *Australasian Journal on Ageing, 26*(4), 162-167.

Percy, A., McCrystal, P., & Higgins, K. (2008). Confirmatory factor analysis of the Adolescent Self-Report Strengths and Difficulties Questionnaire. *European Journal of Psychological Assessment, 24*(1), 43-48.

Perera-Diltz, D. A. (2007). *An investigation of the psychometric properties of the Sinhala version of the Impact of Event Scale-Revised (Sri Lanka).* ProQuest Information & Learning: US.

Perini, S. J., Slade, T., & Andrews, G. (2006). Generic effectiveness measures: Sensitivity to symptom change in anxiety disorders. *Journal of Affective Disorders, 90*(2-3), 123-130.

Pernegerv, T. V., Lepledge, A., & Etter, J. F. (1999). Cross-cultural adaptation of a psychometric instrument: two methods compared. *Journal of Clinical Epidemiology, 52*, 1037-1046.

Perrin, S., Meiser-Stedman, R., & Smith, P. (2005). The children's revised impact of event scale (CRIES): Validity as a screening instrument for PTSD. *Behavioural and Cognitive Psychotherapy, 33*(4), 487-498.

Peters, R. H., Greenbaum, P. E., Steinberg, M. L., Carter, C. R., Ortiz, M. M., Fry, B. C., et al. (2000). Effectiveness of screening instruments in detecting substance use disorders among prisoners. *Journal of Substance Abuse Treatment, 18*(4), 349-358.

Petersen, D. J., Bilenberg, N., Hoerder, K., & Gillberg, C. (2006). The population prevalence of child psychiatric disorders in Danish 8- to 9-year-old children. *European Child & Adolescent Psychiatry, 15*(2), 71-78.

Petrakis, I. L., Poling, J., Levinson, C., Nich, C., Carroll, K., Ralevski, E., et al. (2006). Naltrexone and Disulfiram in Patients with Alcohol Dependence and Comorbid Post-Traumatic Stress Disorder. *Biological Psychiatry, 60*(7), 777-783.

Peveler, R. C., & Fairburn, M. G. (1990). Measurement of neurotic symptoms by self-report questionnaire: Validity of the SCL-90-R. *Psychological Medicine, 20*, 873-879.

Phelan, M., Slade, M., Dunn, G., Holloway, G., Strathdee, G., Thornicroft, G., et al. (1993). *Camberwell Assessment of Need* London: London Institute of Psychiatry.

Phelan, M., Slade, M., Thornicroft, G., Dunn, G., Holloway, F., Wykes, T., et al. (1995). The Camberwell Assessment of Need: the validity and reliability of an instrument to assess the needs of people with severe mental illness. *British Journal of Psychiatry, 167*(5), 589-595.

Philip, G., Greg, W., & Khelifa, N. A. (2007). Quality Of Life And Disability In Alcohol And Drug Dependent Patients
Undergoing Treatment At Depaul House. *International e-Journal of Science, Medicine & Education, 1*, 35-40.

Phillips, K. A. (1996). *The broken mirror: Understanding and treating body dysmorphic disorder*. New York: Oxford University Press.

Phillips, K. A., Atala, K. D., & Pope, H. G. (1995). *Diagnostic instruments for body dysmorphic disorder. New Research Program and Abstracts.* Paper presented at the American Psychiatric Association 148th Annual Meeting, Miami.

Piersma, H. L., & Boes, J. L. (1997). The GAF and psychiatric outcome: a descriptive report. *Community Mental Health Journal, 33*, 35-41.

Pietrantonio, F., Gennaro, L. D., Di Paolo, M. C., & Solano, L. (2003). The Impact of Event Scale: Validation of an Italian version. *Journal of Psychosomatic Research, 55*(4), 389-393.

Piontek, D., Kraus, L., & Klempova, D. (2008). Short scales to assess cannabis-related problems: A review of psychometric properties. *Substance Abuse Treatment, Prevention and Policy, 3*(25).

Pirkis, J., Burgess, P., Kirk, P., Dodson, S., & Coombs, T. (2005). *An Australian Government funded initiative Review of standardised measures used in the National Outcomes and Casemix Collection (NOCC)*: Australian Mental Health Outcomes and Classification Network.

Plotzker, R. E., Metzger, D. S., & Holmes, W. C. (2007). Childhood sexual and physical abuse histories, PTSD, depression, and HIV risk outcomes in women injection drug users: A potential mediating pathway. *The American Journal on Addictions, 16*(6), 431-438.

Pokorny, A. D., Miller, B. A., & Kaplan, H. B. (1972). The brief MAST: A shortened version of the Michigan Alcoholism screening test. *American Journal of Psychiatry, 129*, 342-345.

Politi, P. L., Piccinelli, M., & Wilkinson, G. (1994). Reliability, validity and factor structure of the 12-item General Health Questionnaire among young males in Italy. *Acta Psychiatrica Scandinavica, 90*(6), 432-437.

Pollack, L. E., & Cramer, R. D. (2000). Perceptions of problems in people hospitalized for bipolar disorder: implications for patient education. *Issues in Mental Health Nursing, 21*(8), 765-778.

Pollack, L. E., Cramer, R. D., & Varner, R. V. (2000). Psychosocial functioning of people with substance abuse and bipolar disorders. *Substance Abuse, 21*(3), 193-203.

Pomeroy, I. M., Clark, C. R., & Philp, I. (2001). The effectiveness of very short scales for depression screening in elderly medical patients. *International Journal of Geriatric Psychiatry, 16*(3), 321-326.

Pösl, M., Cieza, A., & Stucki, G. (2007). Psychometric properties of the WHODASII in rehabilitation patients. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care & Rehabilitation, 16*(9), 1521-1531.

Powell, J. E., & McInness, E. (1994). Alcohol use among older hospital patients. *Drug and Alcohol Review, 13*, 5-12.

Power, M., Quinn, K., & Schmidt, S. (2005). Development of the WHOQOL-Old module. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care & Rehabilitation, 14*(10), 2197-2214.

Prather, R. C., & Williamson, D. A. (1988). Psychopathology associated with bulimia, binge-eating, and obesity. *International Journal of Eating Disorders, 7*, 177-184.

Preston, N. J. (2000a). The Health of the Nation Outcome Scales: Validating factorial structure and invariance across two health services. *Australian and New Zealand Journal Psychiatry, 34*(3), 512-519.

Preston, N. J. (2000b). Predicting community survival in early psychosis and schizophrenia populations after receiving intensive case management. *Australian and New Zealand Journal of Psychiatry, 34*(1), 122-128.

Prieto, L., Alonson, J., Ferrer, M., & Antò, J. M. (1997). Are results of the SF 36 Health Survey and the Nottingham Health Profile similar?: A comparison in COPD patients. *Journal of Clinical Epidemiology, 50*, 463-473.

Prins, A., Ouimette, P., Kimerling, R., Cameron, R. P., Hugelshofer, D. S., & Shaw-Hegwer, J. (2003). The primary care PTSD screen (PC-PTSD): Development and operating characteristics. *Primary Care Psychiatry, 9*(1), 9-14.

Pulkkinen, L., Feldt, T., & Kokko, K. (2006). Adaptive behavior in childhood as an antecedent of psychological functioning in early middle age: Linkage via career orientation. *Social Indicators Research, 77*(1), 171-195.

Quello, S. B., Brady, K. T., & Sonne, S. C. (2005). Mood disorders and substance use disorders: A complex comorbidity. *Science and Practice Perspectives 3*, 13-24.

Raistrick, D., Bradshaw, J., Tober, G., Weiner, J., Allison, J., & Healey, C. (1994). Development of the Leeds Dependence Questionnaire (LDQ): A questionnaire to measure alcohol and opiate dependence in the context of a treatment evaluation package. *Addiction, 89*, 563-572.

Raistrick, D., Dunbar, G., & Davidson, R. (1983). Development of a questionnaire to measure alcohol dependence. *British Journal of Addiction, 78*, 89-95.

Raistrick, D., Tober, G., Heatherward, N., & Clark, J. A. (2007). Validation of the Social Satisfaction Questionnaire for outcome evaluation in substance use disorders. *Psychiatric Bulletin, 31*(9), 333-336.

Rajendran, R., & Cheridan, R. R. (1990). Tamil adaption of the Alcohol Dependence Scale. *Journal of the Indian Academy of Applied Psychology, 16*, 26-28.

Rajpal, V. (2006). *Impact of family risk factors and psychopathology on suicidal behavior in adolescents with borderline personality disorder.* ProQuest Information & Learning: US.

Ralevski, E., Ball, S., Nich, C., Limoncelli, D., & Petrakis, I. (2007). The impact of personality disorders on alcohol-use outcomes in a pharmacotherapy trial for alcohol dependence and comorbid axis I disorders. *The American Journal on Addictions, 16*(6), 443-449.

Rash, C. J., Coffey, S. F., Baschnagel, J. S., Drobes, D. J., & Saladin, M. E. (2008). Psychometric properties of the IES-R in traumatized substance dependent individuals with and without PTSD. *Addictive Behaviors, 33*(8), 1039-1047.

Rauter, U. K., Leonard, C. E., & Swett, C. P. (1996). SCL-90-R factor structure in an acute, involuntary, adult psychiatric inpatient sample. *Journal of Clinical Psychology, 52*(6), 625-629.

Ray, L. A., Hutchison, K. E., & Bryan, A. (2006). Psychosocial predictors of treatment outcome, dropout, and change processes in a Pharmacological clinical trial for alcohol dependence. *Addictive Disorders & Their Treatment, 5*(4), 179-190.

Read, J. P., Bollinger, A. R., & Sharkansky, E. (2003). Assessment of comorbid substance use disorder and posttraumatic stress disorder. In P. Ouimette & P. J. Brown (Eds.), *Trauma and substance abuse: Causes, consequences and treatment of comorbid disorder* (pp. 111-125). Washington, DC: American Psychological Society.

Rees, A., Richards, A., & Shapiro, D. A. (2004). Utility of the HoNOS in measuring change in a community mental health care population. *Journal of Mental Health, 13*(3), 295-304.

Reinert, D. F., & Allen, J. P. (2002). The Alcohol Use Disorders Identification Test (AUDIT): A review of recent research. *Alcoholism: Clinical and Experimental Research, 26*(2), 272-279.

Reinert, D. F., & Allen, J. P. (2007). The Alcohol Use Disorders Identification Test: An update of research findings. *Alcoholism: Clinical and Experimental Research, 31*(2), 185-199.

Remschmidt, H., Martin, M., Fleischhaker, C., Theisen, F. M., Hennighausen, K., Gutenbrunner, C., et al. (2007). Forty-two-years later: the outcome of childhood-onset schizophrenia. *Journal of Neural Transmission, 114*(4), 505-512.

Resnick, B., & Nahm, E. S. (2001). Reliability and validity testing of the revised 12-item Short-Form Health Survey in older adults. *Journal of Nursing Management, 9*, 151-161.

Resnick, B., & Parker, R. (2001). Simplified scoring and psychometrics of the revised 12-item Short-Form Health Survey. *Outcomes Management for Nursing Practice, 5*, 161-166.

Rey, J. M., Starling, J., Wever, C., Dossetor, D. R., & Plapp, J. M. (1995). Inter-rater reliability of Global Assessment of Functioning in a clinical setting. *Journal of Child Psychology and Psychiatry, 36*(5), 787-792.

Reynolds, T., Thronicroft, G., Abas, M., Woods, B., Hoe, J., Leese, M., et al. (2000). Camberwell Assessment of Need for the Elderly (CANE): Development, validity and reliability. *British Journal of Psychiatry, 176*, 444-452.

Richardson, K., Baillie, A., Reid, S., Morley, K., Teesson, M., Sannibale, C., et al. (2008). Do acamprosate or naltrexone have an effect on daily drinking by reducing craving for alcohol? *Addiction, 103*(6), 953-959.

Richter, P., Werner, J., Heerlein, A., Kraus, A., & Sauer, H. (1998). On the validity of the Beck Depression Inventory: A review. *Psychopathology, 31*(3), 160-168.

Ries, R. K. (1995). *Assessment and treatment of patients with coexisting mental illness and alcohol and other drug abuse. (Treatment Improvement Protocol (TIP) Series 9.)*. Rockville, MD: Substance Abuse and Mental Health Services Administration.

Rikoon, S. H., Cacciola, J. S., Carise, D., Alterman, A. I., & McLellan, A. T. (2006). Predicting DSM-IV dependence diagnoses from Addiction Severity Index composite scores. *Journal of Substance Abuse Treatment, 31*(1), 17-24.

Riskind, J. H., Beck, A. T., Brown, G., & Steer, R. A. (1987). Taking the measure of anxiety and depression. Validity of the reconstructed Hamilton scales. *Journal of Nervous and Mental Disease, 175*(8), 474-479.

Robbins, I., & Hunt, N. (1996). Validation of the IES as a measure of the long-term impact of war trauma. *British Journal of Health Psychology, 1*(Part 1), 87-89.

Roberts, R. E., Roberts, C. R., & Xing, Y. (2006). Prevalence of youth-reported DSM-IV psychiatric disorders among African, European, and Mexican American adolescents. *Journal of the American Academy of Child & Adolescent Psychiatry, 45*(11), 1329-1337.

Roche, A. M., & Pollard, Y. (2006). *Improved services for people with drug and alcohol problems and mental illness*. Adelaide: National Centre for Education and Training on Addiction (NCETA).

Roesch, S. C., & Vaughn, A. A. (2006). Evidence for the Factorial Validity of the Dispositional Hope Scale Cross-Ethnic and Cross-Gender Measurement Equivalence. *European Journal of Psychological Assessment, 22*(2), 78-84.

Rogalski, C. J. (1987). Factor Structure of Addiction Severity Index in an inpatient detoxification sample. *Archives of General Psychiatry, 45*, 1069-1077.

Ronan, G. F., Dreer, L. E., & Dollard, K. M. (2000). Measuring patient symptom change on rural psychiatry units: Utility of the Symptom Checklist-90 Revised. *Journal of Clinical Psychiatry, 61*(7), 493-497.

Ronning, J. A., Handegaard, B. H., Sourander, A., & Morch, W.-T. (2004). The Strengths and Difficulties Self-Report Questionnaire as a screening instrument in Norwegian community samples. *European Child Adolescent Psychiatry, 13*(2), 73-82.

Rosen, A., Hadzi-Pavlovic, D., & Parker, G. (1989). The Life Skills Profile: a measure assessing function and disability in schizophrenia. *Schizophrenia Bulletin, 15*, 325-337.

Rosen, A., Hadzi-Pavlovic, D., Parker, G., & Trauer, T. (2006). The Life Skills Profile: Background, items and scoring for the LSP–39, LSP–20 and the LSP–16.   Retrieved 19th February, 2009, from http://www.blackdoginstitute.org.au/docs/LifeSkillsProfile.pdf

Rosen, A., & Teesson, M. (2001). Does case management work? The evidence and the abuse of evidence-based medicine. *Australian & New Zealand Journal of Psychiatry, 35*(6), 731-746.

Rosen, A., Trauer, T., Hadzi-Pavlovic, D., & Parker, G. (2001). Development of a brief form of the Life Skills Profile: the LSP-20. *Australian and New Zealand Journal of Psychiatry, 35*, 677-683.

Rosen, C. S., Henson, B. R., Finney, J. W., & Moos, R. H. (2000). Consistency of self-administered and interview-based Addiction Severity Index composite scores. *Addiction, 95*(3), 149-425.

Rosen, D., Smith, M. L., & Reynolds III, C. F. (2008). The prevalence of mental and physical health disorders among older methadone patients. *American Journal of Geriatric Psychiatry, 16*(6), 488-497.

Rosenberg, S. D., Drake, R. E., Wolford, G. L., Mueser, K. T., Oxman, T. E., Vidaver, R. M., et al. (1998). Dartmouth Assessment of Lifestyle Instrument (DALI): a substance use disorder screen for people with severe mental illness. *American Journal of Psychiatry, 155*(2), 232-238.

Rosenthal, T. L., Downs, J. M., Arheart, K. L., Deal, N., & Rosenthal, R. H. (1991). Similarities and differences on five inventories among mood and anxiety disorder patients. *Behaviour Research and Therapy, 29*(3), 239-247.

Ross, H. E., Gavin, D. R., & Skinner, H. (1990). Diagnostic validity of the MAST and the Alcohol Dependence Scale in the assessment of DSM-III alcohol disorders. *Journal of Studies on Alcohol, 51*, 506-513.

Ross, H. E., & Glaser, F. B. (1989). Psychiatric screening of alcohol and drug patients: The validity of the GHQ-60. *American Journal of Drug and Alcohol Abuse, 15*, 429-442.

Rost, K., Smith, G., Burnam, M. A., & Burns, B. J. (1992). Measuring outcomes of care for mental health problems: the case of depressive disorders. *Medical Care, 30(Suppl)*, MS266-MS273.

Rothenberger, A., Becker, A., Erhart, M., Wille, N., & Ravens-Sieberer, U. (2008). Psychometric properties of the Parent Strengths and Difficulties Questionnaire in the general population of German children and adolescents: Results of the BELLA study. *European Child & Adolescent Psychiatry, 17*(Suppl 1), 99-105.

Roy-Byrne, P., Dagadakis, C., Unutzer, J., & Ries, R. (1996). Evidence for limited validity of the revised global assessment of functioning scale. *Psychiatric Services, 47*, 864-866.

Roy, B. V., Veenstra, M., & Clench-Aas, J. (2008). Construct validity of the five-factor Strengths and Difficulties Questionnaire (SDQ) in pre-, early, and late adolescence. *Journal of Child Psychology and Psychiatry, 49*(12), 1304-1312.

Ruggeri, M., Lasalvia, A., Nicolaou, S., & Tansella, M. (1999). La versione italiana della Camberwell Assessment of Need (CAN), un'intervista per l'identificazione dei bisogni di cura. [The Italian version of the Camberwell Assessment of Need (CAN), an interview for the identification of needs of care.]. *Epidemiologia e Psichiatria Sociale, 8*(2), 135-167.

Ruggiero, K. J., Del Ben, K., Scotti, J. R., & Rabalais, A. E. (2003). Psychometric properties of the PTSD Checklist--Civilian Version. *Journal of Traumatic Stress, 16*(5), 495-502.

Russell, M., & Bigler, L. (1979). Screening for alcohol-related problems in an outpatient obstetric-gynecologic clinic. *American Journal of Obstetrics and Gynaecology, 134*, 4-12.

Russell, M., Martier, S. S., Sokol, R. J., & Mudar, P. (1994). Screening for pregnancy risk-drinking. *Alcoholism Clinical and Experimental Research, 18*, 1156-1161.

Russell, M., Martier, S. S., Sokol, R. J., Mudar, P., Jacobson, S., & Jacobson, J. (1996). Detecting risk drinking during pregnancy: a comparison of four screening questionnaires. *American Journal of Public Health, 86*(10), 1435-1439.

Russo, J., Roy-Byrne, P., Jaffe, C., Ries, R., Dagadakis, C., Dwyer-O'Connor, E., et al. (1997). The relationship of patient-administered outcome assessments to quality of life and physician ratings: Validity of the BASIS-32. *Journal of Mental Health Administration, Spr 24*(2), 200-214.

Ruz, F. I., Gonzalez, S. F., & Ruiz, A. F. (1998). Drug dependent patients in the methadone maintenance program: evaluation in primary care of psychosocial and organic severity. *Aten Primaria, 21*(6), 384-388.

Ryan, C. F., & White, J. M. (1996). Health status at entry to methadone maintenance treatment using the SF-36 health survey questionnaire. *Addiction, 91*(1), 39-45.

Ryan, J. M., Corry, J. R., Attewell, R., & Smithson, M. J. (2002). A comparison of an electronic version of the SF-36 General Health Questionnaire to the standard paper version. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care & Rehabilitation, 11*(1), 19-26.

Ryff, C. (1989a). Beyond Ponce de Leon and life satisfaction: New directions in quest of successful aging. *International Journal of Behavioural Development, 12*, 35-55.

Ryff, C. (1989b). Happiness is everything, or is it?. Explorations on the meaning of psychological well-being. *Journal of Personality and Social Psychology, 57*, 1069-1081.

Ryff, C. D., & Keyes, C. L. (1995). The structure of psychological well-being revisited. *Journal of Personality and Social Psychology, 69*, 719-727.

Sack, W. H., Seeley, J. R., Him, C., & Clarke, G. N. (1998). Psychometric properties of the Impact of Events Scale in traumatized Cambodian refugee youth. *Personality and Individual Differences, 25*(1), 57-67.

Sacks, J. A. Y., Drake, R. E., Williams, V. F., Banks, S. M., & Herrell, J. M. (2003). Utility of the Time-Line Follow-Back to assess substance use among homeless adults. *Journal of Nervous and Mental Disease, 191*(3), 145-153.

Sahebi, A., Asghari, M. J., & Salari, R. S. (2004). [Validation of depression anxiety and stress scale (DASS-21) for an Iranian population.]. *Journal of Iranian Psychologists, 1*(1), No Pagination Specified.

Saltstone, R., Halliwell, S., & Hayslip, M. A. (1994). A multivariate evaluation of the Michigan Alcoholism Screening Test and the Drug Abuse Screening Test in a female offender population. *Addictive Behaviors, 19*(5), 455-462.

Salvi, G., Leese, M., & Slade, M. (2005). Routine use of mental health outcome assessments: Choosing the measure. *British Journal of Psychiatry, 186*(2), 146-152.

Salyers, M. P., Bosworth, H. B., Swanson, J. W., Lamb-Pagone, J., & Osher, F. C. (2000). Reliability and validity of the SF-12 health survey among people with severe mental illness. *Medical Care, 38*, 1141-1150.

Sanderson, K., & Andrews, G. (2002a). Prevalence and severity of mental health-related disability and relationship to diagnosis. *Psychiatric Services, 53*, 80-86.

Sanderson, K., & Andrews, G. (2002b). The SF-12 in the Australian population: cross-validation of item selection. *Australian and New Zealand Journal of Public Health, 26*(4), 343-345.

Sanderson, K., Andrews, G., & Jelsma, W. (2001). Disability measurement in the anxiety disorders: comparison of three brief measures. *Journal of Anxiety Disorders, 15*, 333-344.

Sannibale, C., Fucito, L., O'Connor, D., & Curry, K. (2005). Process evaluation of an out-patient detoxification service. *Drug and Alcohol Review, 24*(6), 475-481.

Sanson-Fisher, R. W., & Perkins, J. J. (1998). Adaptation and validation of the SF-36 Health Survey for use in Australia. *Journal of Clinical Epidemiology, 51*, 961-967.

Saunders, J. B., Aasland, O. G., Amundsen, A., & Grant, M. (1993). Alcohol consumption and related problems among primary health care patients : WHO Collaborative Project on Early Detection of Persons with Harmful Alcohol Consumption - I. *Addiction, 88*, 349-362.

Saunders, J. B., Aasland, O. G., Babor, T. F., Fuente, J. R. d. l., & Grant, M. (1993). Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption - II. *Addiction, 88*, 791-804.

Savage, C. L., Wray, J. N., Ritchey, P. N., & Fulmer, M. (2002). Measurement of maternal alcohol consumption in a pregnant population. *Substance Abuse, 23*(4), 211-214.

Saxena, S., Carlson, D., Billington, R., & Orley, J. (2001). The WHO quality of life assessment instrument (WHOQOL-Bref): The importance of its items for cross-cultural research. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care & Rehabilitation, 10*(8), 711-721.

Saxon, A. J., Kivlahan, D. R., Doyle, S., & Donovan, D. M. (2007). Further validation of the Alcohol Dependence Scale as an index of severity. *Journal of Studies on Alcohol and Drugs, 68*(1), 149-156.

Schaffer, D., Gould, M. S., Brasic, J., Ambrosini, P., Fisher, P., Bird, H., et al. (1983). A children's global assessment scale (CGAS). *Archives of General Psychiatry, 40*, 1228-1231.

Schanowitz, J. Y., & Nicassio, P. M. (2006). Predictors of positive psychosocial functioning of older adults in residential care facilities. *Journal of Behavioral Medicine, 29*(2), 191-201.

Scher, C. D., McCreary, D. R., Asmundson, G. J. G., & Resick, P. A. (2008). The structure of post-traumatic stress disorder symptoms in three female trauma samples: A comparison of interview and self-report measures. *Journal of Anxiety Disorders, 22*(7), 1137-1145.

Schinka, J. A., Brown, L. M., Borenstein, A. R., & Mortimer, J. A. (2007). Confirmatory factor analysis of the PTSD checklist in the elderly. *Journal of Traumatic Stress, 20*(3), 281-289.

Schlesinger, C. M., Ober, C., McCarthy, M. M., Watson, J. D., & Seinen, A. (2007). The development and validation of the Indigenous Risk Impact Screen (IRIS): a 13-item screening instrument for alcohol and drug and mental risk. *Drug and Alcohol Review, 26*, 109-117.

Schmidt, P., Küfner, H., Hasemann, S., Löhnert, B., Kolb, W., Zemlin, U., et al. (2007). Ist der European Addiction Severity Index ein sinnvolles disgnoseinstrument bei alkhollabhängigkeit? [Is the European Addiction Severity Index a useful tool in the diagnostic routines of alcohol dependence?]. *Fortschritte der Neurologie, Psychiatrie, 75*(9), 541-548.

Schmitz, N., Kruse, J., Heckrath, C., Alberti, L., & Trees, W. (1999). Diagnosing mental disorders in primary care: the General Health Questionnaire (GHQ) and the Symptom Check List (SCL-90-R) as screening instruments. *Social Psychiatry and Psychiatric Epidemiology, 34*(7), 360-366.

Schmitz, N., Kruse, J., & Tress, W. (1999). Psychometric properties of the General Health Questionnaire (GHQ-12) in a German primary care sample. *Acta Psychiatrica Scandinavica, 100*(6), 462-468.

Schneider, J., Wooff, D., Carpenter, J., Brandon, T., & McNiven, F. (2002). Service organisation, service use and costs of community mental health care. *Journal of Mental Health Policy and Economics, 5*(2), 79-87.

Schofield, M. J., & Mishra, G. (1998). Validity of the SF-12 Compared with the SF-36 Health Survey in Pilot Studies of the Australian Longitudinal Study on Women's Health. *Journal of Health Psychology, 3*(2), 259-271.

Schorre, B. E. H., & Vandvik, I. H. (2004). Global assessment of psychosocial functioning in child and adolescent psychiatry. A review of three unidimensional scales (CGAS, GAF, GAPD). *European Child & Adolescent Psychiatry, 13*(5), 273-286.

Schwartz, R. C. (2007). Concurrent validity of the Global Assessment of Functioning Scale for clients with schizophrenia. *Psychological Reports, 100*(2), 571-574.

Schwartz, R. C., Hilscher, R. L., & Hayhow, P. (2007). Substance abuse and psychosocial impairments among clients with schizophrenia. *American Journal of Orthopsychiatry, 77*(4), 610-615.

Schwarzwald, J., Solomon, Z., Weisenberg, M., & Mikulincer, M. (1987). Validation of the impact of event scale for psychological sequelae of combat. *Journal of Consulting and Clinical Psychology, 55*(2), 251-256.

Sciolla, A., Patterson, T. L., Wetherell, J. L., McAdams, L. A., & Jeste, D. V. (2003). Functioning and well-being of middle-aged and older patients with schizophrenia: measurement with the 36-item short-form (SF-36) health survey. *American Journal of Geriatric Psychiatry, 11*, 629-637.

Scogin, F. R. (2003). Introduction: The status of self-administered treatments. *Journal of Clinical Psychology, 59*(3), 247-249.

Scogin, F. R., Hanson, A., & Welsh, D. (2003). Self-administered treatment in stepped-care models of depression treatment. *Journal of Clinical Psychology, 59*(3), 341-349.

Scott, J., Chant, D., Andrews, G., & McGrath, J. (2006). Psychotic-like experiences in the general community: the correlates of CIDI psychosis screen items in an Australian sample. *Psychological Medicine, 36*, 231-238.

Scott, K. M., Sarfali, D., Tobias, M. I., Martin, I., & Haslett, S. J. (2000). A challenge to the cross-cultural validity of the SF-36 health survey: Factor structure in Maori, Pacific and New Zealand ethnic groups. *Social Science & Medicine, 51*, 1655-1664.

Searles, J. S., Alterman, A. I., & Purtill, J. (1990). The detection of alcoholism in hospitalised schizophrenics: A comparison of the MAST and the MAC. *Alcoholism Clinical and Experimental Research, 14*, 557-560.

Searles, J. S., Helzer, J. E., Rose, G. L., & Badger, G. J. (2002). Concurrent and retrospective reports of alcohol consumption across 30, 90 and 366 days: Interactive voice response compared with the timeline follow back. *Journal of Studies on Alcohol, 63*(3), 352-362.

Sederer, L. I. (1992). Case-based reimbursement for psychiatric hospital care. *Hospital and Community Psychiatry, 43*(11), 1120-1126.

Self-Brown, S. R., LeBlanc, M., Kelley, M. L., Hanson, R., Laslie, K., & Wingate, A. (2006). Effects of Community Violence Exposure and Parental Mental Health on the Internalizing Problems of Urban Adolescents. *Violence and Victims, 21*(2), 183-198.

Selzer, M. L. (1971). The Michigan alcoholism screening test: The quest for a new diagnostic instrument. *American Journal of Psychiatry, 127*, 85-94.

Selzer, M. L., Vinokur, A., & van Rooijen, L. (1975). A self-administered Short Michigan Alcoholism Screening Test (SMAST). *Journal of Studies on Alcohol, 36*(1), 117-126.

Shane, P., Jasiukaitis, P., & Green, R. S. (2003). Treatment outcomes among adolescents with substance abuse problems: The relationship between comorbidities and post-treatment substance involvement. *Evaluation and Program Planning, 26*, 393-402.

Shanks, J. (1994). How are things? Developing outcome measures for mental health services. *Health Trends (in press)*.

Sharma, V. K., Wilkinson, G., & Fear, S. (1999). Health of the Nation Outcome Scales: a case study in general psychiatry. *British Journal of Psychiatry, 174*, 395-398.

Sharples, L. D., Todd, C. J., Caine, N., & Tait, S. (2000). Measurement properties of the Nottingham Health Profile and Short Form 36 Health status measures in a population sample of elderly people living at home:Results from ELPHS. *British Journal of Health Psychology, 5*, 217-233.

Shearer, J., Wodak, A., Mattick, R. P., Van Beek, I., Lewis, J., Hall, W., et al. (2001). Pilot randomized controlled study of dexamphetamine substitution for amphetamine dependence. *Addiction, 96*(9), 1289-1296.

Sheeran, T., & Zimmerman, M. (2002). Screening for posttraumatic stress disorder in a general psychiatric outpatient setting. *Journal of Consulting and Clinical Psychology, 70*, 961-966.

Sherbourne, C. D., Wells, K. B., & Ludd, L. L. (1996). Functioning and well being of patients with panic disorder. *American Journal of Psychiatry, 153*, 213-218.

Shergill, S. S., Shankar, K. K., Seneviratna, K., & Orrell, M. W. (1999). The validity and reliability of the Health of the Nation Outcome Scales (HoNOS) in the elderly. *Journal of Mental Health (UK), 8*(5), 511-521.

Shields, A. L., & Caruso, J. C. (2003). Reliability generalization of the Alcohol Use Disorders Identification Test. *Educational and Psychological Measurement, 63*(3), 404-413.

Shields, A. L., & Caruso, J. C. (2004). A reliability induction and reliability generalization study of the CAGE questionnaire. *Educational and Psychological Measurement, 64*(2), 254-270.

Shields, A. L., Howell, R. T., Potter, J. S., & Weiss, R. D. (2007). The Michigan Alcoholism Screening Test and its shortened form: A meta-analytic inquiry into score reliability. *Substance Use & Misuse, 42*(11), 1783-1800.

Sica, C., Ghisi, M., & Lange, M. A. (2007). The Italian versions of the Beck Anxiety Inventory and the Beck Depression Inventory-II: Psychometric properties and discriminant power. In *Leading-edge psychological tests and testing research.* (pp. 27-50): Nova Science Publishers: Hauppauge.

Siegel, K., Gluhoski, V. L., & Karus, D. (1997). Coping and mood in HIV-positive women. *Psychological Reports, 81*(2), 435-442.

Sim, K., Gwee, K. P., & Bateman, A. (2005). Case Formulation in Psychotherapy: Revitalizing Its Usefulness as a Clinical Tool. *Academic Psychiatry, 29*, 289-292.

Simith, S. D., Schwartz, R. C., George, R. G., & Panke, D. (2004). Convergent validity of the Beck depression inventory for youth. *Psychological Reports, 94*(3,Pt2), 1444-1446.

Simpson, M., Lawrinson, P., Copeland, J., & Gates, P. (2007). *The Australian Alcohol Treatment Outcome Measure (AATOM-C): Psychometric properties*. Sydney: National Drug and Alcohol Reserach Centre, Technical Report Number 288.

Simpson, M., Lawrinson, P., Copeland, J., & Gates, P. (2009). The Alcohol Treatment Outcome Measure (ATOM): A new clinical tool for standardising outcome measurement for alcohol treatment. *Addictive Behaviors, 34*(1), 121-124.

Sims, M., Abramson, M., Forbes, A., Glass, D., Ikin, J., Ittak, P., et al. (2003). *Australian Gulf War Veterans' Health Study*. Canberra: Department of Veterans Affairs.

Singh, J., Mattoo, S. K., Sharan, P., & Basu, D. (2005). Quality of life and its correlates in patients with dual diagnosis of bipolar affective disorder and substance dependence. *Bipolar Disorders, 7*(2), 187-191.

Singleton, E. G., Trotman, A. J. M., Zavahir, M., Taylor, R. C., & Heishman, S. J. (2002). Determination of the reliability and validity of the Marijuana Craving Questionnaire using imagery scripts. *Experimental and Clinical Psychopharmacology, 10*(1), 47-53.

Skevington, S. M., Sartorius, N., & Amir, M. (2004). Developing methods for assessing quality of life in different cultural settings: The history of the WHOQOL instruments. *Social Psychiatry and Psychiatric Epidemiology, 39*(1), 1-8.

Skinner, H. (1982). The Drug Abuse Screening Test. *Addictive Behaviours, 7*, 363-371.

Skinner, H., & Goldberg, A. E. (1986). Evidence for a drug dependence syndrome among narcotic users. *British Journal of Addiction, 81*, 479-484.

Skinner, H. A., & Horn, J. L. (1984). *Alcohol Dependence Scale (ADS): Users Guide*. Toronto: Addiction Research Foundation.

Skipsey, K., Burleson, J. A., & Kranzler, H. R. (1997). Utility of the AUDIT for identification of hazardous or harmful drinking in drug-dependent patients. *Drug and Alcohol Dependence, 45*(3), 157-163.

Skowrow, S., & Smith, D. I. (1986). Survey of homelessness, alcohol consumption and related problems amongst Aboriginals in the Hedland area. Western Australian Alcohol and Drug Authority.

Slade, M., Beck, A., Bindman, J., Thornicroft, G., & Wright, S. (1999). Routine clinical outcome measures for patients with severe mental illness: CANSAS and HoNOS. *British Journal of Psychiatry, 174*, 404-408.

Slade, M., Phelan, M., Thornicroft, G., & Parkman, S. (1996). The Camberwell Assessment of Need (CAN): comparison of assessments by staff and patients of the needs of the severely mentally ill. *Social Psychiatry and Psychiatric Epidemiology, 31*(3), 109-113.

Smedje, H., Broman, J. E., Hetta, J., & von Knorring, A. L. (1999). Psychometric properties of a Swedish version of the 'Strengths and Difficulties Questionnaire'. *European Child and Adolescent Psychiatry, 8*(2), 63-70.

Smelson, D. A., Losonczy, M. F., Davis, C. W., Kaune, M., Williams, J., & Ziedonis, D. (2002). Risperidone decreases craving and relapses in individuals with schizophrenia and cocaine dependence. *The Canadian Journal of Psychiatry / La Revue canadienne de psychiatrie, 47*(7), 671-675.

Smelson, D. A., Losonczy, M. F., Ziedonis, D., Sussner, B. D., Castles-Fonseca, K., Rodrigues, S., et al. (2007). A brief community linkage intervention for veterans with a persistent mental illness and a co-occurring substance abuse disorder. *European Journal of Psychiatry, 21*(2), 143-152.

Smith, A. R. (1986). Alcoholism and gender: Patterns of diagnosis and response. *The Journal of Drug Issues, 16*, 407-420.

Smith, M. Y., Egert, J., Winkel, G., & Jacobson, J. (2002). The impact of PTSD on pain experience in persons with HIV/AIDS. *Pain, 98*(1-2), 9-17.

Snow, M., Thurber, S., & Hodgson, J. M. (2002). An adolescent version of the Michigan Alcoholism Screening Test. *Adolescence, 37*(148), 835-840.

Snyder, C. R., Harris, C., Anderson, J. R., Holleran, S. A., Irving, L. M., Sigmon, S. T., et al. (1991). The will and the ways: Development and validation of an individual-differences measure of hope. *Journal of Personality and Social Psychology, 60*, 570-588.

Snyder, C. R., Hoza, B., Pelham, W. E., Rapoff, M., Ware, L., Danovsky, M., et al. (1997). The development and validation of the children's hope scale. *Journal of Pediatric Psychology, 22*, 399-421.

Snyder, C. R., Sympson, S. C., Ybasco, F. C., Tyrone F. Borders, Babyak, M. A., & Higgins, R. L. (1996). Development and Validation of the State Hope Scale. *Journal of Personality and Social Psychology, 70*(2), 321-335.

Sobell, L. C., Agrawal, S., Annis, H., Ayala-Velazquez, H., Echeverria, L., Leo, G. I., et al. (2001). Cross-cultural evaluation of two drinking assessment instruments: Alcohol Timeline Followback and Inventory of Drinking Situations. *Substance Use and Misuse, 36*, 313-331.

Sobell, L. C., Brown, J., Leo, G. I., & Sobell, M. B. (1996). The reliability of the Alcohol Timeline Followback when administered by telephone and by computer. *Drug and Alcohol Dependence, 42*(1), 49-54.

Sobell, L. C., & Sobell, M. B. (1995). *Alcohol timeline followback users' manual*. Toronto: Addiction Research Foundation.

Sobell, L. C., & Sobell, M. B. (1996). Timeline Followback (TLFB) for alcohol (Version 4.0b). Toronto, Canada: Addiction Research Foundation.

Sobell, L. C., & Sobell, M. B. (2000). Alcohol Timeline Followback (TLFB). In American Psychiatric Association (Ed.), *Handbook of psychiatric measures* (pp. 477-479). Washington, DC: American Psychiatric Association.

Sokol, R. J., Martier, S. S., & Ager, J. W. (1989). The T-ACE questions: Practical prenatal detection of risk-drinking. *American Journal of Obstetrics and Gynaecology, 160*, 863-870.

Solís, L., Cordero, M., Cordero, R., & Martínez, M. (2007). Caracterización del nivel de dependencia al alcohol entre habitantes de la Ciudad de México. [Characterization of level of alcohol dependence in Mexico City inhabitants.]. *Salud Mental, 30*(6), 62-68.

Sourander, A., Helenius, H., & Piha, J. (1996). Child psychiatric short-term inpatient treatment: CGAS as follow-up measure. *Child Psychiatry and Human Development, 27*(2), 93-104.

Sourander, A., Leijala, H., Lehtila, A., Kanerva, A., Helenius, H., & J., P. (1996). Short-term child psychiatric inpatient treatment. Place of residence as one-year outcome measure. *European Child and Adolescent Psychiatry, 5*(1), 38-43.

Sourander, A., & Piha, J. (1997). Parent, teacher and clinical ratings on admission to child psychiatric inpatient treatment: A study of cross-informant correlations using the CBCL, Rutter scales and the CGAS. *Nordic Journal of Psychiatry, 51*(5), 365-370.

Sousa, S. A., & Frazier, R. (2004). A nursing tool for adherence and recovery in psychosis: A pilot study. *Journal of Psychosocial Nursing and Mental Health Services, 42*(3), 28-36.

Spasojevic, J., Heffer, R. W., & Snyder, D. K. (2000). Effects of posttraumatic stress and acculturation on marital functioning in Bosnian refugee couples. *Journal of Traumatic Stress, 13*(2), 205-217.

Spear, J., Chawla, S., O'Reilly, M., & Rock, D. (2002). Does the HoNOS 65+ meet the criteria for a clinical outcome indicator for mental health services for older people? *International Journal of Geriatric Psychiatry, 17*(3), 226-230.

Spielberger, C. D. (1983). *Manual for the State-Trait Anxiety Inventory*. Paolo Alto: Consulting Psychologists Press.

Spies, G., Stein, D. J., Roos, A., Faure, S. C., Mostert, J., Seedat, S., et al. (2009). Validity of the Kessler 10 (K-10) in detecting DSM-IV defined mood and anxiety disorders among pregnant women. *Archives of Women's Mental Health, Epub ahead of print*.

Spooner, C., Mattick, R. P., & Noffs, W. (2001). Outcomes of a comprehensive treatment program for adolescents with a substance-use disorder. *Journal of Substance Abuse Treatment, 20*(3), 205-213.

Springer, K. W., & Hauser, R. M. (2006). An assessment of the construct validity of Ryff's Scales of Psychological Well-Being: Method, mode, and measurement effects. *Social Science Research, 35*(4), 1080-1102.

Sprinkle, S. D., Lurie, D., Insko, S. L., Atkinson, G., Jones, G. L., Logan, A. R., et al. (2002). Criterion validity, severity cut scores, and test-retest reliability of the Beck Depression Inventory-II in a university counseling center sample. *Journal of Counseling Psychology, 49*(3), 381-385.

Sriram, T. G., Chandrashekar, C. R., Isaac, M. K., & Shanmugham, V. (1989). The General Health Questionnaire (GHQ): Comparison of the English version and a translated Indian version. *Social Psychiatry and Psychiatric Epidemiology, 24*(6), 317-320.

Stafrace, S. (2002). Doubts about HoNOS. *Australian and New Zealand Journal Psychiatry, 36*(2), 270.

Staley, D., & El-Guebaly, N. (1990). Psychometric properties of the Drug Abuse Screening Test in a psychiatric patient population. *Addictive Behaviors, 15*(3), 257-264.

Stalker, C. A., Gebotys, R., & Harper, K. (2005). Insecure attachment as a predictor of outcome following inpatient trauma treatment for women survivors of childhood abuse. *Bulletin of the Menninger Clinic, 69*(2), 137-156.

Stanley, M. A., Beck, J. G., & Zebb, B. J. (1996). Psychometric properties of four anxiety measures in older adults. *Behaviour Research and Therapy, 34*(10), 827-838.

Starcevic, V., Bogojevic, G., & Marinkovic, J. (2000). The SCL-90-R as a screening instrument for severe personality disturbance among outpatients with mood and anxiety disorders. *Journal of Personality Disorders, 14*(3), 199-207.

Startup, M., Jackson, M. C., & Bendix, S. (2002). The concurrent validity of the Global Assessment of Functioning (GAF). *British Journal of Clinical Psychology, 41*, 417-422.

Stasiewicz, P. R., Vincent, P. C., Bradizza, C. M., Connors, G. J., Maisto, S. A., & Mercer, N. D. (2008). Factors affecting agreement between severely mentally ill alcohol abusers' and collaterals' reports of alcohol and other substance abuse. *Psychology of Addictive Behaviors, 22*(1), 78-87.

Stastny, D., & Potter, M. (1991). Alcohol abuse by patients undergoing methadone treatment programs. *British Journal of Addictions, 86*, 307-310.

Stedman, T., Yellowlees, P., Drake, S., Chant, D., Clarke, R., & Chapple, B. (2000). The perceived utility of six selected measures of consumer outcomes proposed for routine use in Australian mental health services. *Australian and New Zealand Journal of Psychiatry, 34*(5), 842-849.

Stedman, T., Yellowlees, P., Mellsop, G., Clarke, R., & Drake, S. (1997). *Measuring consumer outcomes in Mental Health: Field testing of selected measures of consumer outcome in mental health*. Canberra: Department of Health and Family Services.

Steer, R. A., & Clark, D. A. (1997). Psychometric characteristics of the Beck Depression Inventory-II with college students. *Measurement and Evaluation in Counseling and Development, 30*(3), 128-136.

Steer, R. A., Ranieri, W. F., Beck, A. T., & Clark, D. A. (1993). Further evidence for the validity of the Beck Anxiety Inventory with psychiatric outpatients. *Journal of Anxiety Disorders, 7*(3), 195-205.

Steer, R. A., Rissmiller, D. J., Ranieri, W. F., & Beck, A. T. (1993). Structure of the computer-assisted Beck Anxiety Inventory with psychiatric inpatients. *Journal of Personality Assessment, 60*(3), 532-542.

Steer, R. A., Rissmiller, D. J., Ranieri, W. F., & Beck, A. T. (1995). Use of the computer-administered Beck Depression Inventory and Hopelessness Scale with psychiatric patients. *Computers in Human Behavior, 10*, 223-229.

Steer, R. A., & Schut, J. (1979). Types of psychopathology displayed by heroin addicts. *American Journal of Psychiatry, 136*, 1463-1465.

Stein, G. S. (1999). Usefulness of the Health of the Nation Outcome Scales. *British Journal of Psychiatry, 174*, 375-377.

Stein, M. D., Herman, D. S., & Anderson, B. J. (2009). A motivational intervention trial to reduce cocaine use. *Journal of Substance Abuse Treatment, 36*(1), 118-125.

Steiner, S., Baumeister, S. E., & Kraus, L. (2008). Severity of Dependence Scale: Establishing a cut-off point for cannabis dependence in the German adult population. *Sucht: Zeitschrift für Wissenschaft und Praxis, 54*(Suppl1), S57-S63.

Steinhausen, H.-C. (1987). Global assessment of child psychopathology. *Journal of the American Academy of Child and Adolescent Psychiatry, 26*(2), 203-206.

Steinhausen, H.-C., & Metzke, C. W. (2001). Global measures of impairment in children and adolescents: Results from a Swiss community survey. *Australian and New Zealand Journal of Psychiatry, 35*(3), 282-286.

Steinweg, D. L., & Worth, H. (1993). Alcoholism: the keys to the CAGE. *American Journal of Medicine, 94*, 520-523.

Stengler-Wenzke, K., Kroll, M., Riedel-Heller, S., Matschinger, H., & Angermeyer, M. C. (2007). Quality of life in obsessive-compulsive disorder: the different impact of obsessions and compulsions. *Psychopathology, 40*(5), 282-289.

Stevenson, J. S., & Masters, J. A. (2005). Predictors of Alcohol Misuse and Abuse in Older Women. *Journal of Nursing Scholarship, 37*(4), 329-335.

Stewart, K. J., Turner, K. L., Bacher, A. C., DeRegis, J. R., Sung, J., Tayack, M., et al. (2003). Are fitness, activity, and fatness associated with health-related quality of life and mood in older persons? *Journal of Cardiopulmonary Rehabilitation, 23*, 115-121.

Stewart, S. E., Stack, D. E., & Wilhelm, S. (2008). Severe obsessive-compulsive disorder with and without body dysmorphic disorder: Clinical correlates and implications. *Annals of Clinical Psychiatry, 20*(1), 33-38.

Stieglitz, R.-D., Frommberger, U., Foa, E. B., & Berger, M. (2001). Evaluation of the German version of the PTSD Symptom Scale (PSS). *Psychopathology, 34*, 128-133.

Stockwell, T., Hodgson, R., Edwards, G., Taylor, C., & Rankin, H. (1979). The development of a questionnaire to measure severity of alcohol dependence. *British Journal of Addiction to Alcohol & Other Drugs., 74*(1), 79-87.

Stockwell, T., Murphy, D., & Hodgson, R. (1983). The severity of alcohol dependence questionnaire: its use, reliability and validity. *British Journal of Addiction, 78*, 145-155.

Stockwell, T., Sitharthan, T., McGrath, D., & Lang, E. (1994). The measurement of alcohol dependence and impaired control in community samples. *Addiction, 89*, 167-174.

Stoffelmayr, B. E., Mavis, B. E., & Kasim, R. M. (1994). The longitudinal stability of the Addiction Severity Index. *Journal of Substance Abuse Treatment, 11*(4), 373-378.

Strain, E. C. (2002). Assessment and treatment of comorbid psychiatric disorders in opioid-dependent patients. *Clinical Journal of Pain, 18*(4 Suppl), S14-27.

Sullivan, T. P., & Holt, L. J. (2008). PTSD symptom clusters are differentially related to substance use among community women exposed to intimate partner violence. *Journal of Traumatic Stress, 21*(2), 173-180.

Sumnall, H. R., Wagstaff, G. F., & Cole, J. C. (2004). Self-reported psychopathology in polydrug users. *Journal of Psychopharmacology, 18*(1), 75-82.

Sundin, E. C., & Horowitz, M. J. (2002). Impact of Event Scale: Psychometric properties. *British Journal of Psychiatry, 180*(3), 205-209.

Sussner, B. D., Smelson, D. A., Rodrigues, S., Kline, A., Losonczy, M., & Ziedonis, D. (2006). The validity and reliability of a brief measure of cocaine craving. *Drug and Alcohol Dependence, 83*(3), 233-237.

Svikis, D. S., McCaul, M. E., Haug, N. A., & Boney, T. Y. (1996). Detecting alcohol problems in drug-dependent women of childbearing age. *American Journal of Drug & Alcohol Abuse, 22*(4), 563-575.

Swallow, B. L., Lindow, S. W., Masson, E. A., & Hay, D. M. (2003). The use of the General Health Questionnaire (GHQ-28) to estimate prevalence of psychiatric disorder in early pregnancy. *Psychology, Health & Medicine, 8*(2), 213-218.

Swanson, J. W., Swartz, M. S., Essock, S. M., Osher, F. C., Wagner, H. R., Goodman, L. A., et al. (2002). The Social-Environmental Context of Violent Behavior in Persons Treated for Severe Mental Illness. *American Journal of Public Health, 92*(9), 1523-1531.

Swett, C., Cohen, C., Surrey, J., & Compaine, A. (1991). High rates of alcohol use and history of physical and sexual abuse among women outpatients. *American Journal of Drug and Alcohol Abuse, 17*, 49-60.

Swift, W., Hall, W., Didcott, P., & Reilly, D. (1998). Patterns and correlates of cannabis dependence among long-term users in an Australian rural area. *Addiction, 93*(8), 1149-1160.

Swift, W., Maher, L., & Sunjic, S. (1999). Transitions between routes of heroin administration: A study of Caucasian and Indochinese heroin users in south-western Sydney, Australia. *Addiction, 94*(1), 71-82.

Szobot, C. M., Ketzer, C., Parente, M. A., Biederman, J., & Rohde, L. A. (2004). The acute effect of methylphenidate in Brazilian male children and adolescents with ADHD: a randomized clinical trial. *Journal of Attention Disorders, 8*(2), 37-43.

Tait, R. J., French, D. J., & Hulse, G. K. (2003). Validity and psychometric properties of the General Health Questionnaire-12 in young Australian adolescents. *Australian and New Zealand Journal of Psychiatry, 37*(3), 374-381.

Tanaka, E., Sakamoto, S., Ono, Y., & Fujihara, S. (1996). Hopelessness in a community population in Japan. *Journal of Clinical Psychology, 52*(6), 609-615.

Taylor, A. W., Wilson, D. H., Dal Grande, E., Ben-Tovim, D., Elzinga, R. H., Goldner, R. D., et al. (2000). Mental health status of the South Australian population. *Australian & New Zealand Journal of Public Health, 24*, 29-34.

Taylor, J. R., & Wilkinson, G. (1997). HoNOS v. GP opinion in a shifted out-patient clinic. *Psychiatric Bulletin, 21*(8), 483-485.

Teesson, M., Clement, N., Copeland, J., Conroy, A., & Reid, A. (2000). *The measurement of outcome in alcohol and other drug treatment: A review of available instruments*. Sydney: NDARC Technical Report No 92. NDARC.

Teesson, M., & Gallagher, J. (1999). Evaluation of a treatment programme for serious mental illness and substance use in an inner city area. *Journal of Mental Health, 8*(1), 19-28.

Teesson, M., Hodder, T., & Buhrich, N. (2004). Psychiatric disorders in homeless men and women in inner Sydney. *Australian and New Zealand Journal of Psychiatry, 38*(3), 162-168.

Teitelbaum, L., & Mullen, B. (2000). The validity of the MAST in psychiatric settings: A metaanalytic integration. *Journal of Studies on Alcohol, 61*(2), 254-261.

Teitelbaum, L. M., & Carey, K. B. (1996). Alcohol assessment in psychiatric patients. *Clinical Psychology: Science and Practice, 3*(4), 323-338.

Teitelbaum, L. M., & Carey, K. B. (2000). Temporal stability of alcohol screening measures in a psychiatric setting. *Psychology of Addictive Behaviors, 14*(4), 401-404.

Thabet, A. A., Stretch, D., & Vostanis, P. (2000). Child mental health problems in Arab children: Application of the Strengths and Difficulties Questionnaire. *International Journal of Social Psychiatry, 46*(4), 266-280.

Thomas, B. A., & McCambridge, J. (2008). Comparative psychometric study of a range of hazardous drinking measures administered online in a youth population. *Drug and Alcohol Dependence, 96*(1-2), 121-127.

Thomas, S. D. M., Slade, M., McCrone, P., Harty, M.-A., Parrott, J., Thornicroft, G., et al. (2008). The reliability and validity of the forensic Camberwell Assessment of Need (CANFOR): A needs assessment for forensic mental health service users. *International Journal of Methods in Psychiatric Research, 17*(2), 111-120.

Thompson, K. M., Crosby, R. D., Wonderlich, S. A., Mitchell, J. E., Redlin, J., Demuth, G., et al. (2003). Psychopathology and sexual trauma in childhood and adulthood. *Journal of Traumatic Stress, 16*(1), 35-38.

Tiffany, S. T., Singleton, E., Haertzen, C. A., & Henningfield, J. E. (1993). The development of a cocaine craving questionnaire. *Drug and Alcohol Dependence, 34*(1), 19-28.

Titus, J. C., Dennis, M. L., Lennox, R., & Scott, C. K. (2008). Development and validation of short versions of the internal mental distress and behavior complexity scales in the Global Appraisal of Individual Needs (GAIN). *The Journal of Behavioral Health Services & Research, 35*(2), 195-214.

Tolin, D. F., & Foa, E. B. (1999). Treatment of a police officer with PTSD using prolonged exposure. *Behavior Therapy, 30*(3), 527-538.

Trauer, T. (1998). The Health of the Nation Outcome Scales in outcome measurement: A critical review. *Australasian Psychiatry, 6*(1), 11-14.

Trauer, T. (1999). The subscale structure of the Health of the Nation Outcome Scales (HoNOS). *Journal of Mental Health (UK), 8*(5), 499-509.

Trauer, T. (2003). *Analysis of outcome measurement data from the four Victorian 'round one' agencies*. Victoria: Mental Health Branch, Department of Human Services.

Trauer, T., Callaly, T., Hantz, P., Little, J., Shields, R., & Smith, J. (1999). Health of the Nation Outcome Scales: Results of the Victorian field trial. *The British Journal of Psychiatry, 174*(5), 380-388.

Trauer, T., Duckmanton, R. A., & Chiu, E. (1995). The Life Skills Profile: a study of its psychometric properties. *Australian and New Zealand Journal of Psychiatry, 29*, 492-499.

Trauer, T., Duckmanton, R. A., & Chiu, E. (1997). The assessment of clinically significant change using the Life Skills Profile. *Australian and New Zealand Journal of Psychiatry, 31*(2), 257-263.

Trauer, T., Duckmanton, R. A., & Chiu, E. (1998). Estimation of costs of public psychiatric treatment. *Psychiatric Services, 49*(4), 440-442.

Trauer, T., Eagar, K., Gaines, P., & Bower, A. (2004). *New Zealand Mental Health Consumers and their Outcomes*. Auckland: Health Research Council of New Zealand.

Trauer, T., & Tobias, G. (2004). The Camberwell Assessment of Need and Behaviour and Symptom Identification Scale as routine outcome measures in a psychiatric disability rehabilitation and support service. *Community Mental Health Journal, 40*(3), 211-221.

Trauer, T., Tobias, G., & Slade, M. (2008). Development and evaluation of a patient-rated version of the Camberwell Assessment of Need Short Appraisal Schedule (CANSAS-P). *Community Mental Health Journal, 44*(2), 113-124.

Treatment Protocol Project (TPP). (2000). *Management of Mental Disorders, third edition*. Sydney: World Health Organization Collaborating Centre for Mental Health and Substance Abuse.

Trompenaars, F. J., Masthoff, E. D., Van Heck, G. L., Hodiamont, P. P., & De Vries, J. (2005). Content validity, construct validity, and reliability of the WHOQOL-Bref in a population of Dutch adult psychiatric outpatients. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care & Rehabilitation, 14*(1), 151-160.

Tsai, M.-C., Tsai, Y.-F., Chen, C.-Y., & Liu, C.-Y. (2005). Alcohol Use Disorders Identification Test (AUDIT): Establishment of Cut-off Scores in a Hospitalized Chinese Population. *Alcoholism: Clinical and Experimental Research, 29*(1), 53-57.

Tsey, K., & Every, A. (2000). *Evaluation of an Aboriginal Empowerment Program*. Alice Springs: Cooperative Research Centre for Aboriginal and Tropical Health.

Tsey, K., Travers, H., Gibson, T., Whiteside, M., Cadet-James, Y., Haswell-Elkins, M., et al. (2005). The role of empowerment through life skills development in building comprehensive primary health *Australian Journal of Primary Health, 11*(2), 16-25.

Tsey, K., Wilson, A., Haswell-Elkins, M., Whiteside, M., McCalman, J., Cadet-James, Y., et al. (2007). Empowerment-based research methods: a 10-year approach to enhancing Indigenous social and emotional wellbeing. *Australasian Psychiatry, 15 Supplement*, S34-S38.

Ullman, S. E., & Long, S. M. (2008). Factor structure of PTSD in a community sample of sexual assault survivors. *Journal of Trauma & Dissociation, 9*(4), 507-524.

Ulusoy, M., Sahin, N., & Erkmen, H. (1998). Turkish version of the Beck Anxiety Inventory: psychometric properties. *Journal of Cognitive Psychotherapy, 12*, 163-172.

Uslu, R. I., Kapci, E. G., Oncu, B., Ugurlu, M., & Turkcapar, H. (2008). Psychometric Properties and Cut-off Scores of the Beck Depression Inventory-II in Turkish Adolescents. *Journal of Clinical Psychology in Medical Settings, 15*, 225-233.

Uttaro, T., & Gonzalez, A. (2002). Psychometric properties of the Behavior and Symptom Identification Scale administered in a crisis residential mental health treatment setting. *Psychological Reports, 91*(2), 439-443.

Uys, L. R., & Zulu, R. N. (1996). An evaluation of the implementation and the effectiveness of case management in the rehabilitation of psychiatric outpatients in South Africa. *South African Journal of Psychology, 26*(4), 226-230.

Vakili, S., Sobell, L. C., Sobell, M. B., Simco, E. R., & Agrawal, S. (2008). Using the Timeline Followback to determine time windows representative of annual alcohol consumption with problem drinkers. *Addictive Behaviors, 33*(9), 1123-1130.

Vallejo, M. A., Mañanes, G., Comeche, M. I., & Díaz, M. I. (2008). Comparison between administration via internet and paper and pencil administration of two clinical instruments: SCL-90-R and GHQ-28. *Journal of Behavior Therapy and Experimental Psychiatry, 39*(3), 201-208.

van Balkom, A. J. L. M., van Boeijen, C. A., Boeke, A. J. P., van Oppen, P., Kempe, P. T., & van Dyck, R. (2008). Comorbid depression, but not comorbid anxiety disorders, predicts poor outcome in anxiety disorders. *Depression and Anxiety, 25*(5), 408-415.

van der Ploeg, E., Mooren, T. T. M., Kleber, R. J., van der Velden, P. G., & Brom, D. (2004). Construct validation of the Dutch version of the Impact of Event Scale. *Psychological Assessment, 16*(1), 16-26.

Van der Ploeg, H. M. (1980). [Validity of the Zelf-Beoordelings-Vragenlijst (A Dutch version of the Spielberger State-Trait Anxiety Inventory).]. *Nederlands Tijdschrift voor de Psychologie en haar Grensgebieden, 35*(4), 243-249.

van Dierendonck, D. (2004). The construct validity of Ryff's Scales of Psychological Well-being and its extension with spiritual well-being. *Personality and Individual Differences, 36*(3), 629-643.

van Minnen, A., & Foa, E. B. (2006). The Effect of Imaginal Exposure Length on Outcome of Treatment for PTSD. *Journal of Traumatic Stress, 19*(4), 427-438.

van Widenfelt, B. M., Goedhart, A. W., Treffers, P. D. A., & Goodman, R. (2003). Dutch version of the Strengths and Difficulties Questionnaire (SDQ). *European Child and Adolescent Psychiatry, 12*(6), 281-289.

Varescon, I., Gaugue, J., & Wendland, J. (2007). Alcool et grossesse: Première utilisation du questionnaire de dépistage T-ACE dans une population française. [Alcohol and pregnancy. First use of the T-ACE screening test in a French population.]. *Alcoologie et Addictologie, 29*(3), 221-226.

Varner, R. V., Chen, Y. R., Swann, A. C., & Moeller, F. G. (2000). The Brief Psychiatric Rating Scale as an acute inpatient outcome measurement tool: A pilot study. *Journal of Clinical Psychiatry, 61*(6), 418-421.

Veit, C. T., & Ware, J. E. (1983). The structure of psychological distress and well-being in general populations. *Journal of Consulting and Clinical Psychology, 51*(5), 730-742.

Verthein, U., Bonorden-Kleij, K., Degkwitz, P., Dilg, C., Kohler, W. K., Passie, T., et al. (2008). Long-term effects of heroin-assisted treatment in Germany. *Addiction, 103*(6), 960-966; discussion 967-968.

Victorian Department of Human Services. (2003). *Measuring Consumer Outcomes in Clinical Mental Health Services. A Training Manual for Services in Victoria*. Melbourne: Victorian Government.

Vinson, D. C., Reidinger, C., & Wilcosky, T. (2003). Factors affecting the validity of a Timeline Follow-Back interview. *Journal of Studies on Alcohol, 64*(5), 733-740.

Vleioras, G., & Bosma, H. A. (2005). Are identity styles important for psychological well-being? *Journal of Adolescence, 28*(3), 397-409.

Vostanis, P. (2006). Strengths and Difficulties Questionnaire: Research and clinical applications. *Current Opinion in Psychiatry, 19*(4), 367-372.

Wagner, A., Gandek, B., Aaronson, N. K., Acquadro, C., Alonso, J., Apolone, G., et al. (1998). Cross-cultural comparisons of the content of SF-36 translations across 10 countries: results from the IQOLA Project. International Quality of Life Assessment. *Journal of Clinical Epidemiology, 51*, 925-932.

Walker, E. A., Newman, E., Dobie, D. J., Ciechanowski, P., & Katon, W. (2002). Validation of the PTSD Checklist in an HMO sample of women. *General Hospital Psychiatry, 24*(6), 375-380.

Wallace, C. J., Lecomte, T., Wilde, J., & Liberman, R. P. (2001). CASIG: A consumer-centered assessment for planning individualized treatment and evaluating program outcomes. *Schizophrenia Research, 50*(1-2), 105-119.

Wallerstein, N. (1992). Powerlessness, empowerment, and health: implications for health promotion programs. *American Journal of Health Promotion, 6*(3), 197-205.

Wallerstein, N. (2006). *What is the evidence on effectiveness of empowerment to improve health? Health Evidence Network report*. Copenhagen: WHO Regional Office for Europe.

Walters, J. T. R., Bisson, J. I., & Shepherd, J. P. (2007). Predicting post-traumatic stress disorder: Validation of the Trauma Screening Questionnaire in victims of assault. *Psychological Medicine, 37*(1), 143-150.

Ware, J., Jr., Kosinski, M., & Keller, S. D. (1996). A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Medical Care, 34*(3), 220-233.

Ware, J., Manning, W., Duan, W., Wells, K., & Newhouse, J. (1984). Health status and use of outpatient mental health services. *American Psychologist, 39*, 1090-1100.

Ware, J. E., & Sherbourne, C. D. (1992). The MOS 36 Item Short Form Health Survey (SF-36). *Medical Care, 30*(6), 473-481.

Ware, J. E. J., Kosinski, M., Gandek, B., Aaronson, N. K., Apolone, G., Bech, P., et al. (1998). The factor structure of the SF-36 Health Survey in 10 countries: results from the IQOLA Project. International Quality of Life Assessment. *Journal of Clinical Epidemiology, 51*, 1159-1165.

Waterson, E. J., & Murray-Lyon, I. M. (1988). Are the CAGE questions outdated? (letter). *British Journal of Addiction, 83*, 1113-1115.

Watkins, J. P., Eisele, G. R., & Matthews, K. O. (2000). Occupational medical program alcohol screening. Utility of the CAGE and BMAST. *Journal of Substance Abuse Treatment, 19*(1), 51-57.

Watkins, R. E., Plant, A. J., Sang, D., O´Rourke, T., & Gushulak, B. (2000). Development of a Vietnamese version of the Short form-36 Health Survey. *Asia-Pacific Journal of Public Health, 12*, 118-123.

Watson, C. G., Detra, E., Fox, K. L., Ewing, J. W., Gearhart, L. P., & DeMotts, J. R. (1995). Comparative concurrent validities of five alcoholism measures in a psychiatric hospital. *Journal of Clinical Psychology, 51*(5), 676-684.

Weathers, F. W., Litz, B. T., Herman, D. S., Huska, J. A., & Keane, T. M. (1993). *The PTSD Checklist (PCL): Reliability, validity and diagnostic utility*. Paper presented at the Annual meeting of the International Society for Traumatic Stress Studies.

Weidmann, A., Fehm, L., & Fydrich, T. (2008). Covering the tsunami disaster: Subsequent post-traumatic and depressive symptoms and associated social factors. *Stress and Health: Journal of the International Society for the Investigation of Stress, 24*(2), 129-135.

Weinstein, M. C., Berwick, D. M., Goldman, P. A., Murphy, J. M., & Barsky, A. J. (1989). A comparison of three psychiatric screening tests using receiver operating characteristic (ROC) analysis. *Medical Care, 27*(6), 593-607.

Weiss, D. S., & Marmar, C. R. (1997). The Impact of Event Scale-Revised. In J. P. Wilson & T. M. Keane (Eds.), *Assessing psychological trauma and PTSD* (pp. 399-411). New York: Guilford Press.

Weiss, R. D., Griffin, M. L., & Hufford, C. (1995). Craving in hospitalized cocaine abusers as a predictor of outcome. *American Journal of Drug and Alcohol Abuse, 21*(3), 289-301.

Weiss, R. D., Griffin, M. L., Hufford, C., Muenz, L. R., Najavits, L. M., Jansson, S. B., et al. (1997). Early prediction of initiation of abstinence from cocaine: Use of a craving questionnaire. *The American Journal on Addictions, 6*(3), 224-231.

Weiss, R. D., Griffin, M. L., Mazurick, C., Berkman, B., Gastfriend, D. R., Frank, A., et al. (2003). The Relationship Between Cocaine Craving, Psychosocial Treatment, and Subsequent Cocaine Use. *American Journal of Psychiatry, 160*(7), 1320-1325.

Weiss, R. D., Griffin, M. L., & Mirin, S. M. (1989). Diagnosing major depression in cocaine abusers: The use of depression rating scales. *Psychiatry Research, 28*, 335– 343.

Weissman, M. M., Warner, V., & Fendrich, M. (1990). Applying impairment criteria to children's psychiatric diagnosis. *Journal of the American Academy of Child and Adolescent Psychiatry, 29*(5), 789-795.

Wennström, E., & Wiesel, F.-A. (2006). The Camberwell assessment of need as an outcome measure in routine mental health care. *Social Psychiatry and Psychiatric Epidemiology, 41*(9), 728-733.

Wertz, J. S., Cleaveland, B. L., & Stephens, R. S. (1995). Problems in the application of the Addiction Severity Index (ASI) in rural substance abuse services. *Journal of Substance Abuse, 7*(2), 175-188.

Westermeyer, J., Tucker, P., & Nugent, S. (1995). Comorbid anxiety disorder among patients with substance abuse disorders. *American Journal on Addictions, 4*, 97-106.

Wetherell, J. L., & Gatz, M. (2005). The Beck Anxiety Inventory in Older Adults With Generalized Anxiety Disorder. *Journal of Psychopathology and Behavioral Assessment, 27*(1), 17-24.

Whisman, M. A., Perez, J. E., & Ramel, W. (2000). Factor structure of the Beck Depression Inventory-Second Edition (BDI-II) in a student sample. *Journal of Clinical Psychology, 56*(4), 545-551.

White, P., & Chant, D. (2006). The psychometric properties of a psychosis screen in a correctional setting. *International Journal of Law & Psychiatry, 29*(2), 137-144.

Whitney, J. A., Kunik, M. E., Molinari, V., Lopez, F. G., & Karner, T. (2004). Psychological predictors of admission and discharge global assessment of functioning scale scores for geropsychiatric inpatients. *Aging & Mental Health, 8*(6), 505-513.

WHO. (1997). *Composite International Diagnostic Interview (2.0)*. Geneva: World Health Organization.

Wichstrom, L., Anderson, A. M. C., Holte, A., & Wynne, L. C. (1996). Disqualifying family communication and childhood social competence as predictors of offspring's mental health and hospitalization: A 10- to 14-year longitudinal study of children at risk of psychopathology. *Journal of Nervous and Mental Disease, 184*(10), 581-588.

Wiebe, J. S., & Penley, J. A. (2005). A psychometric comparison of the Beck Depression Inventory-II in English and Spanish. *Psychological Assessment, 17*(4), 481-485.

Wiersma, D., Nienhuis, F. J., Giel, R., & Slooff, C. J. (1998). Stability and change in needs of patients with schizophrenic disorders: A 15- and 17-year follow-up from first onset of psychosis, and a comparison between "objective" and "subjective" assessments of needs for care. *Social Psychiatry and Psychiatric Epidemiology, 33*(2), 49-56.

Wijndaele, K., Matton, L., Duvigneaud, N., Lefevre, J., Duquet, W., Thomis, M., et al. (2007). Reliability, equivalence and respondent preference of computerized versus paper-and-pencil mental health questionnaires. *Computers in Human Behavior, 23*(4), 1958-1970.

Wilkins, L. P., & White, M. B. (2001). Interrater reliability and concurrent validity of the Global Assessment of Relational Functioning (GARF) Scale using a card sort method: A pilot study. *Family Therapy, 28*(3), 157-170.

Williams, R. J., Connor, J. P., & Ricciardelli, L. A. (1998). Self-efficacy for refusal mediated by outcome expectancies in the prediction of alcohol-dependence amongst young adults. *Journal of Drug Education, 28*(4), 347-359.

Williamson, D. A., Anderson, D. A., Jackman, L. P., & Jackson, S. R. (1995). Assessment of eating disordered thoughts, feelings, and behaviours. In D. B. Allison (Ed.), *Handbook of assessment*

methods for eating behaviours and weight-related problems: Methods, theory and research (pp. 347-386). Thousand Oaks, CA: Sage.

Williamson, D. A., Prather, R. C., Bennett, S. M., Davis, C. J., Watkins, P. C., & Grenier, C. E. (1989). An uncontrolled evaluation of inpatient and outpatient cognitive behavior therapy for bulimia nervosa. *Behavior Modification, 13*, 340-360.

Willmott, S., Boardman, J., Henshaw, C., & Jones, P. (2008). The predictive power and psychometric properties of the General Health Questionnaire (GHQ-28). *Journal of Mental Health, 17*(4), 435-442.

Winefield, H. R., Goldney, R. D., Winefield, A. H., & Tiggemann, M. (1989). The General Health Questionnaire: Reliability and validity for Australian youth. *Australian and New Zealand Journal of Psychiatry, 23*(1), 53-58.

Wing, J. K., Beevor, A. S., Curtis, R. H., Park, S. B., Hadden, S., & Burns, A. (1998). Health of the Nation Outcome Scales (HoNOS): Research and development. *British Journal of Psychiatry, 172*, 11-18.

Wing, J. K., Lelliott, P., & Beevor, A. S. (2000). Progress on HoNOS. *British Journal of Psychiatry, 176*, 392-393.

Winters, K. (1999). *Screening and assessing adolescents for substance use disorders (Treatment Improvement Protocol (TIP) Series 31)*. Rockville, MD: Substance Abuse and Mental Health Services Administration.

Witteveen, A. B., Bramsen, I., Hovens, J. E., & van der Ploeg, H. M. (2005). Utility of the Impact of Event Scale in screening for posttraumatic stress disorder. *Psychological Reports, 97*(1), 297-308.

Wobrock, T., Sittinger, H., Behrendt, B., D'Amelio, R., Falkai, P., & Caspari, D. (2007). Comorbid substance abuse and neurocognitive function in recent-onset schizophrenia. *European Archives of Psychiatry and Clinical Neuroscience, 257*(4), 203-210.

Wohlfarth, T. D., van den Brink, W., Winkel, F. W., & ter Smitten, M. (2003). Screening for posttraumatic stress disorder: An evaluation of two self-report scales among crime victims. *Psychological Assessment, 15*(1), 101-109.

Wolford, G. L., Rosenberg, S. D., Drake, R. E., Mueser, K. T., Oxman, T. E., Hoffman, D., et al. (1999). Evaluation of methods for detecting substance use disorder in persons with severe mental illness. *Psychology of Addictive Behaviors, 13*(4), 313-326.

Wong, Y.-L. I., Nath, S. B., & Solomon, P. L. (2007). Group and organizational involvement among persons with psychiatric disabilities in supported housing. *The Journal of Behavioral Health Services & Research, 34*(2), 151-167.

Wooff, D., Schneider, J., Carpenter, J., & Brandon, T. (2003). Correlates of stress in carers. *Journal of Mental Health, 12*(1), 29-40.

World Health Organization. (1994). Development of the WHOQOL: Rationale and current status. *International Journal of Mental Health, 23*(3), 24-56.

World Health Organization. (1995). The World Health Organization Quality of Life Assessment (WHOQOL): Position paper from the World Health Organization. *Social Science & Medicine, 41*(10), 1403-1409.

World Health Organization. (1997). *WHOQOL - Measuring Quality of Life*: WHO (WHO/MSA/MNH/PSF/97.4).

World Health Organization. (1998). Group development of the World Health Organization WHOQOL-BREF quality of life assessment. *Psychological Medicine, 28*(3), 551-558.

World Health Organization. (2009). WHO Disability Assessment Schedule II (WHO DAS II).   Retrieved 6th March, 2009, from http://www.who.int/classifications/icf/whodasii/en/index.html

World Health Organization ASSIST Working Group. (2002). The Alcohol, Smoking and Substance Involvement Screening Test (ASSIST): development, reliability and feasibility. *Addiction, 97*(7), 1183-1194.

Wu, A. Y. K., & Chan, W. F. (2007). Clinical determinants and short-term prognosis of suicidal behaviours in Chinese older persons in Hong Kong. *Hong Kong Journal of Psychiatry, 17*(3), 81-86.

Wu, K. K., & Chan, K. S. (2003). The development of the Chinese version of Impact of Event Scale-Revised (CIES-R). *Social Psychiatry and Psychiatric Epidemiology, 38*(2), 94-98.

Wu, K. K., & Chan, S. K. (2004). Psychometric properties of the Chinese version of the Impact of Event Scale-Revised. *Hong Kong Journal of Psychiatry, 14*(4), 2-8.

Wu, Z., Xu, S., & Li, J. (2002). [The construction of the Mental Health Inventory for the Chinese elderly.]. *Chinese Journal of Clinical Psychology, 10*(1), 1-3.

Wyshak, G. (2001). Women's college physical activity and self-reports of physician-diagnosed depression and of current symptoms of psychiatric distress. *Journal of Women's Health and Gender-Based Medicine, 10*(4), 363-370.

Xenitidis, K., Thornicroft, G., Lesse, M., Slade, M., Fotiadou, M., Philp, H., et al. (2000). Reliability and validity of the CANDID--A needs assessment instrument for adults with learning disabilities and mental health problems. *British Journal of Psychiatry, 176*, 473-478.

Xu, J. (1991). Clinical applications and evaluation of the Beck Depression Inventory. *Psychological Science China, 4*, 32-34.

Yamauchi, K., Ono, Y., Baba, K., & Ikegami, N. (2001). The actual process of rating the global assessment of functioning scale. *Comprehensive Psychiatry, 42*(5), 403-409.

Yates, P., Garralda, M. E., & Higginson, I. (1999). Paddington Complexity Scale and Health of the Nation Outcome Scales for Children and Adolescents. *British Journal of Psychiatry, 174*, 417-423.

Yeh, H.-S., Luh, R.-L., Liu, H.-J., Lee, Y.-C., & Slade, M. (2006). Reliability of the Camberwell assessment of need (Chinese version) for patients with schizophrenia at a daycare center of Taiwan. *Social Psychiatry and Psychiatric Epidemiology, 41*(1), 75-80.

Yen, C.-F., Cheng, C.-P., Huang, C.-F., Yen, J.-Y., Ko, C.-H., & Chen, C.-S. (2008). Quality of life and its association with insight, adverse effects of medication and use of atypical antipsychotics in patients with bipolar disorder and schizophrenia in remission. *Bipolar Disorders, 10*(5), 617-624.

Yoon, G., Kim, S. W., Thuras, P., Grant, J. E., & Westermeyer, J. (2006). Alcohol Craving in Outpatients with Alcohol Dependence: Rate and Clinical Correlates. *Journal of Studies on Alcohol, 67*(5), 770-777.

Yudko, E., Lozhkina, O., & Fouts, A. (2007). A comprehensive review of the psychometric properties of the Drug Abuse Screening Test. *Journal of Substance Abuse Treatment, 32*(2), 189-198.

Yule, W., Ten Bruggencate, S., & Joseph, S. A. (1994). Principal components analysis of the Impact of Events Scale in adolescents who survived a shipping disaster. *Personality and Individual Differences, 16*(5), 685-691.

Zanis, D. A., McLellan, A. T., Cnaan, R. A., & Randall, M. (1994). Reliability and validity of the Addiction Severity Index with a homeless sample. *Journal of Substance Abuse Treatment, 11*(6), 541-548.

Zanis, D. A., McLellan, A. T., & Corse, S. (1997). Is the Addiction Severity Index a reliable and valid assessment instrument among clients with severe and persistent mental illness and substance abuse disorders? *Community Mental Health Journal, 33*(3), 213-227.

Zika, S., & Chamberlain, K. (1992). On the relation between meaning in life and psychological wellbeing. *British Journal of Psychology, 83*(1), 133-145.

Zilberg, N. J., Weiss, D. S., & Horowitz, M. J. (1982). Impact of Life Event Scale: A cross-validation study and some empirical evidence supporting a conceptual model of stress response syndromes. *Journal of Consulting and Clinical Psychology, 50*(3), 407-414.

Zizolfi, S. (1997). La versione italiana del Life Skills Profile (LSP), uno strumento per la valutazione del funzionamento e delle disabilità dei pazienti schizofrenici. [The Italian version of Life Skills Profile (LSP), an instrument for evaluating performance and disabilities of schizophrenic patients.]. *Epidemiologia e Psichiatria Sociale, 6*(3), 196-211.