The Psychological Therapy Outcome Scale – Intellectual Disabilities (PTOS-ID): the development of a psychological therapy outcome measure for adults with intellectual disabilities

Nikolaos Vlissides

Supervised by:

Laura Golding
Nigel Beail

Date of submission: 9th July 2014

Word count (excluding references): 13,832

Submitted in partial fulfilment of the Doctorate in Clinical Psychology at the Division of Clinical Psychology, University of Liverpool

I would like to thank my supervisors – Nigel Beail and Laura Golding – for their guidance and feedback. I would also like to thank everyone who was involved in this project, as well as my family and friends, for their support.

Contents

Appendices

List of Tables & Boxes

Pg.

Chapter 1: Systematic Literature Review

Chapter 2: Empirical Paper

Introduction: Thesis Overview

People with intellectual disabilities are more likely to experience living circumstances and life events associated with an increased risk of mental health problems, compared to the general population (Hulbert-Williams & Hastings, 2008). These include: lack of meaningful relationships, stigmatization, unemployment, and discrimination (Martorell et al., 2009; Thornicroft, 2006). Prevalence of mental health problems are difficult to determine within this population, with estimates ranging from 10% to 39% (Emerson & Hatton, 2007).

Pharmacological and behavioural approaches have often been used in the treatment of mental health problems for people with intellectual disabilities (Vereenooghe & Langdon, 2013). Bender (1993) argued that the lack of psychotherapeutic interventions available was due to a perceived “therapeutic disdain” towards this population. Difficulties identifying mental health problems in people with intellectual disabilities, perceived lack of training amongst practitioners, and lack of research evidence have also been cited as reasons for lack of provision of psychotherapy services (Emerson, Moss & Kiernan, 1999; Royal College of Psychiatrists, 2004; Taylor & Knapp, 2013).

These views have been challenged by clinicians, and there is a growing evidence base of the effectiveness of psychological therapies with people with intellectual disabilities (Willner, 2005). The research has come from both single case studies (Kellett, Beail, Bush, Dyson & Wilbram, 2009; Salvadori & Jackson, 2009) and controlled clinical trials (Taylor, Novaco, Gillmer, Robertson & Thorne, 2005).

Recent systematic reviews and meta-analyses have also highlighted the effectiveness of Cognitive Behavioural Therapy (CBT) and psychodynamic therapy with this population (James & Stacey, 2014; Nicoll, Beail & Saxon, 2013; Vereenooghe & Langdon, 2013). In one of the most comprehensive reviews to date, Prout and Nowak-Drabik (2003) reviewed the outcome and effectiveness of 92 studies and found that there was a moderate benefit of psychotherapy for people with intellectual disabilities. Recent legislation in the United Kingdom has responded by recommending improved access to psychological therapy for people with intellectual disabilities (Department of Health (DoH), 2007; 2009).

The emergence of practice-based evidence has also meant that there is a growing expectation that services that provide psychological therapies show some evidence for the effectiveness of what they do (DoH, 2010). Coupled with this is the Payment by Results (PbR) initiative, where commissioners will pay healthcare providers dependent on the number of patients seen and outcomes achieved (DoH, 2013). One of the difficulties for providers of psychological therapies for people with intellectual disabilities is the availability of valid and reliable therapy outcome measures that can be easily used in service settings and accurately assess the effectiveness of interventions (Skelly, 2011; Weston, Elsworth & Stacey, 2011).

Thesis

This thesis aims to address the difficulties in measuring outcomes of psychological therapy with people with intellectual disabilities. The first paper will systematically review the quality of outcome measures that have been used in psychological therapies with adults with intellectual disabilities. The second paper will assess the psychometric properties of the Psychological Therapy Outcome Scale – Intellectual Disabilities (PTOS-ID), a newly developed therapy outcome measure designed specifically for use with people with intellectual disabilities.

Timeline of Work

Elements of research reported in this thesis were completed before the commencement of the Doctorate in Clinical Psychology programme. These were the development of the PTOS-ID, which included selection of the dependent variables through the use of focus groups; the selection of the item pool; and the development of the response format (please see the Introduction of the research paper for further details).

The assessment of the psychometric properties of the PTOS-ID (i.e. construct validity, concurrent validity and internal consistency) was completed for the research study reported in this thesis in partial fulfilment of the award of Doctor of Clinical Psychology.

Data were collected through service audits both prior to, and during, the period of study for this doctorate.

References

Bender, M. (1993). The unoffered chair: The history of therapeutic disdain towards people with a learning difficulty. *Clinical Psychology Forum. 54(1),* 7-12.

Department of Health (2007). *Services for people with learning disabilities and challenging behaviour or mental health needs.* London: Department of Health.

Department of Health (2009). *Learning disabilities positive practice guide.* London: Department of Health.

Department of Health (2010). *Liberating the NHS: Transparency in outcomes: A framework for the NHS.* London: Department of Health.

Department of Health (2013). *Payment by Results in the NHS: tariff for 2013 to 2014.* London: Department of Health.

Emerson, E. & Hatton, C. (2007). Mental health of children and adolescents with intellectual disabilities in Britain. *The British Journal of Psychiatry, 191*, 493–499.

Emerson, E., Moss, S. & Kiernan, C. (1999). The relationship between challenging behaviour and psychiatric disorders in people with severe developmental disabilities. In N. Bouras (Ed.). *Psychiatric & behavioural disorders in developmental disabilities & mental retardation* (pp. 38-48). Cambridge, UK: Cambridge University Press.

Hulbert-Williams, L. & Hastings, R. P. (2008). Life events as a risk factor for psychological problems in individuals with intellectual disabilities: A critical review. *Journal of Intellectual Disability Research*, *52(11),* 883–895.

James, C. W. & Stacey, J. M. (2014). The effectiveness of psychodynamic interventions for people with learning disabilities: A systematic review. *Tizard Learning Disability Review, 19(1),* 17-24.

Kellett, S., Beail, N., Bush, A., Dyson, G. & Wilbram, M. (2009). Single case experimental evaluations of psychodynamic and cognitive-behavioural psychotherapy: Examples of methods and outcomes. *Advances in Mental Health & Learning Disabilities, 3(4),* 36-44

Martorell, A., Tsakanikos, E., Pereda, A., Gutierrez-Recacha, P., Bouras, N. & Ayosu-Mateos, J. L. (2009). Mental health in adults with mild and moderate intellectual disabilities. *American Journal of Mental Retardation, 113,* 307-318.

Nicoll, M., Beail, N. & Saxon, D. (2013). Cognitive behavioural treatment for anger in adults with intellectual disabilities: A systematic review and meta-analysis. *Journal of Applied Research in Intellectual Disabilities, 26,* 47-6.

Prout T. H. & Nowak-Drabik K. M. (2003) Psychotherapy with persons who have intellectual disabilities: An evaluation of effectiveness. *American Journal on Intellectual Disabilities, 108,* 82–93.

Royal College of Psychiatrists (2004). *Psychotherapy and learning disability*. London: Council Report.

Salvadori, A. & Jackson, T. (2009). A structured account of short-term psychodynamic psychotherapy with a man with learning disabilities. *Advances in Mental Health & Learning Disabilities, 3(4),* 31-35

Skelly, A. (2011). Outcomes measurement in routine practice for learning disability psychology services. *Clinical Psychology & People with Learning Disabilities, 9(4),* 2-6.

Taylor, J. L. & Knapp, M. (2013). Mental health and emotional problems in people with intellectual disabilities. In J. L. Taylor, W. R. Lindsay, R. P. Hastings & C. Hatton (Eds.), *Psychological therapies for adults with intellectual disabilities* (pp. 1-14). Chichester: Wiley & Sons

Taylor, J.L., Novaco, R.W., Gillmer, B.T., Robertson, A. and Thorne, I. (2005). Individual cognitive-behavioural anger treatment for people with mild-borderline intellectual disabilities and histories of aggression: A controlled trial. *British Journal of Clinical Psychology, 44*, 367-382.

Thornicroft, G. (2006). *Shunned: Discrimination against people with mental illness.* Oxford: Oxford University Press.

Vereenooghe, L. & Langdon, P. E. (2013). Psychological therapies for people with intellectual disabilities: A systematic review and meta-analysis. *Research in Developmental Disabilities, 34,* 4085-4102.

Weston, C., Elsworth, J. & Stacey, J. (2011). The development of a tool for the routine measurement of clinical outcomes in a learning disabilities psychology department: Ridgeway routine outcome measure (RROM). *Clinical Psychology & People with Learning Disabilities, 9(2),* 7-13.

Willner, P. (2005). The effectiveness of psychotherapeutic interventions for people with learning disabilities: A critical overview. *Journal of Intellectual Disability Research, 49(1),* 73-85.

A Systematic Review of the Outcome Measures used in Psychological Therapies with Adults with Intellectual Disabilities[1]

Nik Vlissides

University of Liverpool

Laura Golding

University of Liverpool

Nigel Beail

South West Yorkshire Partnership NHS Foundation Trust

---

[1] To be submitted to: *Journal of Applied Research in Intellectual Disabilities* (word count, 7000)

Abstract

Background. Individual psychological therapy is increasingly being offered to people with intellectual disabilities. The current study aims to review the quality of the therapy outcome measures that are currently in use with this population.

Method. A literature search found eleven studies detailing the development and psychometric properties of ten self-report therapy outcome measures used with people with intellectual disabilities. The quality of these outcome measures was examined using the Fitzpatrick et al. (1998) criteria.

Results. The review revealed a number of single- and multi-trait outcome measures currently used in individual psychological therapy with adults with intellectual disabilities. The psychometric properties of these measures suggests a need for more robust and rigorous measures to be developed

Conclusions. The number of available therapy outcome measures is encouraging. However, further work assessing the construct validity and involving service users is needed.

Keywords: intellectual disabilities, learning disabilities, psychological therapy, outcome measure

Introduction

There is a well-established evidence base for the effectiveness of psychological therapies in the treatment of mental health difficulties in the general population (Green & Latchford, 2012; Grissom, 1996). Historically people with intellectual disabilities have been denied access to psychological therapies (Benson, 2004). Bender (1993) argued that this is because there has been disdain amongst professionals towards providing psychological therapy for people with intellectual disabilities. This is despite people in this population being more likely to experience events associated with an increase risk of mental health difficulties, including stigmatization and discrimination, and lack of meaningful relationships (Martorell et al., 2009; Thornicroft, 2006). The research assessing psychological therapies with people with intellectual disabilities is promising, but there is need for more high quality clinical trials to provide a stronger evidence-base of their effectiveness (Beail, 2003; Veerenooghe & Langdon, 2013). Trials require high quality therapy outcome measures, and currently there are no established benchmark measures in the field of intellectual disabilities (Hatton & Taylor, 2013). For example, Froyd et al. (1996) reviewed psychological therapy outcome studies and found 1430 measures had been used, of which 851 were used just once. The aim of the current paper, therefore, is to systematically identify and review the quality of outcome measures that have been used in published reports of psychological therapies with adults with intellectual disabilities.

Psychological Therapies

Psychological therapies are “an interpersonal process designed to bring about modifications of feelings cognitions, attitudes and behaviour which have proved troublesome to the person seeking help from the trained professional” (Strupp, 1978, p. 3). The key objectives of psychological therapies is to help people gain a better understanding of the issues that are troubling them (e.g. mental health difficulties, relationship difficulties) and generate procedures for relieving distress (NHS Choices, 2013; Roth & Fonagy, 2005).

People with intellectual disabilities who experience mental health difficulties are increasingly being offered psychological therapies (Department of Health (DoH), 2007; 2009a). The range of therapies available includes: psychodynamic therapy (e.g. James & Stacey, 2014), Cognitive Behavioural Therapy (CBT, McCabe et al., 2006), Cognitive Analytic Therapy (CAT, Psaila & Crowley, 2005), and mindfulness-based therapy (Singh et al., 2008). A number of studies have provided promising evidence for the effectiveness of psychological therapies for people with intellectual disabilities (Whitehouse et al., 2006; Willner, 2005). A review of 92 studies by Prout and Nowak-Drabik (2003) found that psychological therapies were of moderate benefit for people with intellectual disabilities.

Psychological Therapy Outcome Measures

There are currently a plethora of psychological therapy outcome measures available to researchers and clinicians (Barkham et al., 1998). They can be used to assess a number of possible outcomes, such as psychological well-being, social well-being and quality of life (Fitzpatrick et al., 1998). They can also be used to assess the outcome of specific problems (e.g. depression, anxiety disorders) or of general mental health (Green & Latchford, 2012). For example, the General Anxiety Disorder Assessment (GAD-7; Spitzer et al., 2006) would be used to assess changes in symptoms associated with anxiety, whereas the Clinical Outcomes in Routine Evaluation Outcome Measure (CORE-OM; Evans et al., 2000) assesses general mental health.

Outcome measures can also be model and/or population specific (Lueger & Barkham, 2010). For example, the Beck Depression Inventory – second edition (BDI-II; Beck et al., 1996) is designed for use in Cognitive Behavioural Therapy (CBT) with adults who are experiencing depression, whereas the Beck Depression Inventory – Youth (BDI-Y; Beck Youth Inventories, 2005) is designed for use in CBT with young people, aged 7-18, who are experiencing depression. Here, there are two measures that are model, problem and population specific.

Finally, therapy outcome measures can be completed by the client, clinician and/or third party (e.g. relative, carer). The Health of the Nation Outcome Scale (HoNOS; Wing et al., 1998), for example, is a 12-item measure assessing the health and social functioning of people with mental illness that is completed by clinicians.

There is also an increasing pressure for measures to be used to assess the outcomes of psychological therapies (DoH, 1996). The Department of Health (1996) reviewed psychological therapy services in the UK and emphasized the need for evidence from routine clinical settings (effectiveness data) to complement the outcomes of randomised-controlled trials (efficacy data) to inform practice. Coupled with this there is a need to assess service users’ perspectives when evaluating any healthcare intervention (Slevin et al., 1988). Therefore, good quality self-report outcome measures are required to assess therapy outcomes from both randomised-control trials and routine clinical settings, and provide service users with a voice about interventions provided (Barkham et al., 2001).

There are limited reviews of psychological therapy outcome measures used with adults with intellectual disabilities. Hatton & Taylor (2013) reviewed mental health assessment tools for use with adults with intellectual disabilities, and found that measures varied in their design (e.g. for use with the general population or specifically for people with intellectual disabilities), target population (mild, moderate intellectual disabilities etc) and psychometric properties (e.g. reliability and validity).

More recently, McGurk and Skelly (2014) conducted a review of clinical outcome measures used with people with intellectual disabilities. Again, they found that there was a huge variation in the reliability and validity of the measures, and that there were limited measures assessing social outcomes.

Both of these recent reviews are important in highlighting the available outcome measures, and identifying strengths and weaknesses of them. However, neither review systematically identified outcome measures that have been used to assess individual psychological therapy with adults with intellectual disabilities. Nor did they use an identified quality appraisal tool to assess the quality of the outcome measures. Therefore, the current review will appraise the quality of therapy outcome measures that have been used in individual psychological therapy with adults with intellectual disabilities. This will be limited to self-report outcome measures due to increasing recognition of involving service users’ perspectives in healthcare interventions (Slevin et al., 1988) and measures that assess psychological well-being (i.e. positive well-being, psychological illness/distress) and social well-being (i.e. family and intimate relationships; Fitzpatrick et al., 1998)

One of the first steps in assessing whether a measure is appropriate for use is ensuring that it has sound psychometric properties (Kraus & Castonguay, 2010). Psychometric properties of a measure are usually assessed through its reliability and validity (Beck et al., 1988).

Reliability refers to an outcome measure’s ability to produce similar results from the same respondents in consistent conditions (Field, 2013). The reliability of therapy outcome measures are often assessed through test re-test reliability (degree of which scores are consistent from one administration to the next) and internal consistency (consistency of results across items in a test) (Fitzpatrick et al., 1998).

Validity refers to the degree to which an outcome measure measures what it aims to assess (Rose & Sullivan, 1996). For example, its ability to measure psychological distress in people who are experiencing psychological distress. Validity of therapy outcome measures is often assessed through: construct validity (refers to the extent that a measure actually measures the constructs it claims to); concurrent validity (when a measure is administered at the same time as a pre-existing one and they are correlated); discriminant validity (when a measure has low levels of correspondence with another measure that represents another) (Cahill et al., 2008; Fitzpatrick et al., 1998)

Fitzpatrick et al. (1998) identify five additional criteria for assessing the quality of outcome measures. Cahill et al. (2008) developed a quality appraisal tool designed to assess outcome measures based on the Fitzpatrick et al. (1998) criteria. These will be discussed further in the Method section.

Current Paper

The current paper aims to systematically assess the quality of self-report therapy outcome measures that have been used in individual psychological therapy with adults with intellectual disabilities. The review will focus on therapy outcome measures that assess psychological well-being and social well-being as the outcome of psychological therapies. This will be done in a three-stage process:

1. Systematically identify therapy outcome measures that have been used in published peer reviewed studies assessing the outcome of individual psychological therapies with adults with intellectual disabilities

2. Identify the key papers that report the development/psychometric assessment of the identified self-report therapy outcome measure with adults with intellectual disabilities

3. Use the Fitzpatrick et al. (1998) criteria (Box 1) to assess the quality of the self-report therapy outcome measures that have been used in individual psychological therapy with adults with intellectual disabilities.

Methods

Search Strategy

The initial strategy involved searching three major electronic databases (PsycINFO, Scopus and MEDLINE) for studies that had assessed the effectiveness of individual psychological therapy with adults with intellectual disabilities. This first step was used to ensure that any therapy outcome measures found had been used in individual psychological therapy with adults with intellectual disabilities.

Keywords anywhere (title, abstract, journal) for the terms 'psycho* therap*' and 'outcome*' returned 390,164 and 3,428,084 references. To limit the search to the desired population, keywords anywhere for the terms 'intellectual disabilit*', 'learning disabilit*', 'developmental disabilit*' and 'mental retardation' returned 253,966 references. The population terms were then combined with 'psycho* therap*' and 'outcome*', yielding 522 references. References were then limited to those published in English, in peer-reviewed journals, and did not contain populations other than adults (18+) with intellectual disabilities (i.e. children with intellectual disabilities were excluded from the review). This returned 212 studies.

The 212 studies were screened for content by applying relevant inclusion and exclusion criteria. Studies were included if: i) they used a self-report outcome measure to report change; ii) the outcome measure was concerned with psychological and/or social well-being; and iii) service-users presented with a mental health problem of a psychological nature (e.g. depression, anxiety).

Studies were excluded if: i) they reported interventions that did not use individual psychological therapy (i.e. pharmacological interventions, group/family therapy); ii) service users did not present with mental health problems; iii) they did not measure mental health as the primary dependent variable (i.e. offending behaviours, challenging behaviours); iv) they did not use a self-report outcome measure to assess change (i.e. no psychometric measures were used); and v) they were reviews or other non primary research.

Of the 212 studies, 201 did not meet the criteria and were removed from the review. The search strategy, therefore, yielded eleven relevant studies.

The eleven studies that met the inclusion criteria were reviewed to identify the therapy outcome measures used (Table 1). The primary articles concerned with the measure development and validation within the target population were then identified. If these did not exist (i.e. the psychometric properties had not been assessed within the target population), the therapy outcome measure was excluded from the analysis. This was the case for the Posttraumatic Stress Diagnostic Scale (PDS; Foa et al., 1997) identified in the Dilly (2014) paper, and the Beck Anxiety Inventory – Youth (BAI-Y, Beck Youth Inventories, 2005) and the Beck Depression Inventory – Youth (BDI-Y, Beck Youth Inventories, 2005) identified in the Hassiotis et al. (2013) paper. All three of these measures were not included in the quality appraisal.

Table 1 Studies that have used outcome measures in individual psychological therapy with adults with intellectual disabilities that met the inclusion criteria

Author and Year	**Design**	**Presenting Difficulties**	**Setting**	**Sample**	**Outcome Measure(s)**
Alim, (2010)	Case study	Anger	Community	34 year-old male with mild intellectual disabilities	Novaco Anger Scale (NAS) Provocation Inventory (PI) Brief Symptom Inventory (BSI) Inventory of Interpersonal Problems, 32nd ed (IIP-32) Rosenberg Self-Esteem Scale (RSES)
Atnas & Lippold (2013)	Case study	Anxiety, Psychogenic non-epileptic seizures	Community	20 year-old female with mild intellectual disability	Glasgow Anxiety Scale – Intellectual Disabilities (GAS-ID) Glasgow Depression Scale – Learning Disabilities (GDS-LD)

Beail et al. (2005)	Open trial	Aggression, Sexually inappropriate behaviour, Psychosis, Depression, OCD, Bulimia, Self injury	Community	17 men 3 women Mild intellectual disabilities	Symptom Checklist 90-Revised (SCL-90-R) IIP-32 RSES
Brooks et al. (2013)	Not reported	Not reported	Community, High-security NHS, Other NHS	272 clinical (110 men, 162 women) 52 non-clinical (22 men, 30 women) Mild-moderate	Clinical Outcomes in Routine Evaluation – Learning Disability (CORE-LD)

				intellectual disability	
Dilly (2014)	_Case study_	_Trauma, Self-harm_	_Inpatient_	_25 year-old man with severe intellectual disabilities_	_Posttraumatic Stress Diagnostic Scale (PDS)*_
Hassiotis et al. (2013)	_Randomised Control Trial_	_Mood disorder_	_Community_	_16 treatment (5 men, 11 women)_ _16 control (7 men, 9 women)_ _Mild-moderate intellectual disability_	_Beck Depression Inventory-Youth (BDI-Y)*_ _Beck Anxiety Inventory-Youth (BAI-Y)*_
Kellett et al. (2009)	3 single case experimental design	Hypochondriasis, Ambulophobia, Anger	Community	40-year old woman with moderate intellectual disabilities,	BSI IIP-32 RSES

				43-year old man with intellectual disabilities, 27-year old man with mild intellectual disabilities	
Newman & Beail (2002)	Case study	Anger	Community	25-year old male with intellectual disabilities	SCL-90-R
Rose (2013)		Anger	Community	37 treatment (25 men, 12 women) Level of disability not reported	Anger Inventory (AI)
Taylor et al. (2002)	Delayed	Anger	Low and	9 treatment (all male)	PI

	waiting-list control trial		medium secure inpatient hospitals	10 control (all male) Mild-borderline intellectual disability	
Taylor et al. (2005)	Delayed waiting-list control trial	Anger	Low and medium secure hospitals	16 treatment (all male) 20 control (all male) Mild-borderline intellectual disability	NAS PI

* Removed from quality appraisal

This meant that the quality appraisal consisted of eleven primary articles and eleven studies, reporting the development of ten self-report therapy outcome measures that have been used in individual psychological therapy with people with intellectual disabilities (see Table 2). The primary articles were then reviewed for data extraction as outlined below. Additional information was taken from the studies reporting the use of the therapy outcome measures (Table 1) if needed.

Table 2 Psychological therapy outcome measures identified for quality appraisal

Measure	Areas assessed	Number of items	Response Scale	General population measure or specific for people with intellectual disabilities
Anger Inventory (AI) Rose & Gerson (2009)*	Anger: reactivity to a number of anger provoking scenarios	35	4-point Likert scale: higher scores suggest higher anger levels	Specific
Brief Symptom Inventory (BSI) Kellett et al., (2003; 2004)*	Global indices of psychological distress (GSI, PSDI, PST) Somatization Obsessive-compulsive Interpersonal sensitivity Depression Anxiety	53	5-point Likert scale: higher scores suggest higher levels of distress	General (adapted)

	Hostility Phobic anxiety Paranoid ideation Psychoticism			
Clinical Outcome Routine Evaluation – LD (CORE-LD) Brooks, Davies & Twigg (2013)*	Feelings	14	3-point Likert scale: higher scores suggest feeling worse	Specific
Glasgow Anxiety Scale – ID (GAS-ID) Mindham & Espie (2003)*	Overall anxiety including indices of worries, specific fears and physiological symptoms	27	3-point Likert scale: higher scores suggest higher levels of anxiety	Specific
Glasgow Depression Scale – ID (GDS-ID) Cuthill, Espie & Cooper	Depression	20	3-point Likert scale: higher scores suggest higher levels of anxiety	Specific

(2003)*				
Inventory of Interpersonal Problems – 32 (IIP-32) Kellett et al. (2005)*	Indexes difficulties adults experience in their interpersonal relationships	32	5-point Likert scale: higher scores suggest higher levels of interpersonal difficulties	General (adapted)
Novaco Anger Scale (NAS) Novaco & Taylor (2004)*	Indexes cognitive arousal and behavioural substrates of anger	48	3-point Likert scale: higher scores suggest higher levels of anger	General (adapted)
Provocation Inventory (PI) Novaco & Taylor (2004)*	Indexes anger intensity and generality across a range of provocative situations	25	3-point Liker scale: higher scores suggest higher levels of anger intensity	General (adapted)
Rosenberg Self-Esteem Scale (RSES) Davis et al. (2009)*	Self-esteem	10	Dichotomous scoring – ‘yes’ or ‘no’: higher scores suggest higher levels of self-esteem	General

Symptom Checklist – 90R (SCL-90-R) Kellett et al. (1999)*	Global indices of psychological distress (GSI, PSDI, PST) Somatization Obsessive-compulsive Interpersonal sensitivity Depression Anxiety Hostility Phobic anxiety Paranoid ideation Psychoticism	90	5-point Likert scale: higher scores suggest higher levels of distress	General (adapted)

* studies that report the psychometric properties of the outcome measures with adults with intellectual disabilities

Quality Appraisal

Fitzpatrick et al. (1998) identified a series of desirable attributes for patient based outcome measures (Box 1). Cahill et al. (2008) identified a way of assessing the criteria and developed a rating tool to address this (Appendix A). An adapted version of this rating tool was used to assess the quality of the candidate therapy outcome measures for use in individual psychological therapy with adults with intellectual disabilities (Appendix B). Table 3 summarises the criteria that were used to assess the quality of the therapy outcome measures. The Interpretability criterion is incorporated into Precision in the assessment of the quality of the outcome measures. The Responsiveness was assessed from the studies reporting the use of the therapy outcome measure in individual psychological therapy with adults with intellectual disabilities found in the initial systematic literature search.

Box 1 Fitzpatrick criteria taken from Cahill et al. (2008)

Reliability	A reliable measure is one that produces consistent results from the same respondents at different times where there exists no evidence of change
Validity	The extent to which a measure really measures the concept that it purports to measure
Responsiveness	Addresses the question: does the instrument detect changes over time that matter to the patient? It can be discriminative (between individuals) or evaluative (within an individual across time)
Acceptability	Addresses the question: is the measure acceptable to users?

Feasibility	Is the measure easy to administer and process?
Precision	How precise is the measure?
Interpretability	How interpretable are the scores of the measure?

Table 3 Criteria to assess quality of the measures (based on Cahill et al., 2008)

Criterion	**Definition**
Reliability	
Internal consistency	As measured by Cronbach's alpha, split-half reliability estimates
Test-retest reliability	Determines the consistency across time. Measured using correlational analysis
Validity	
Construct validity	Hypotheses are generated and a measure tested to determine whether it actually reflects these prior hypotheses
Concurrent validity	Where a new measure is administered at the same time as a pre-existing one and the two are correlated
Convergent validity	A measure converges with other indications of the same concept
Discriminant validity	A measure demonstrates low levels of correspondence with a measure that represents another concept
Responsiveness	Addresses the question: does the instrument detect changes over time that matter to the patient? It can be discriminative (between individuals) or evaluative (within individual across time)

Acceptability	Addresses the question: is the measure acceptable to users? Practicality of administration Time taken to complete Length of instrument Translations Access by ethnic minorities Reading age
Feasibility	Is the measure easy to administer and process? Cost and burden to administrative staff Electronic scanning options Scoring systems Training package Training manual Support from measure developers
Precision	Interpretability Normative data

Analysis

Each measure was critically evaluated using data from the studies and the primary articles (Tables 1 and 2). The relevant information from the studies and the primary articles were entered into a summary sheet (Appendix C). The first author then used the adapted rating tool (Appendix B) to score up the measure against each of the six criteria. Finally, the coding instructions were applied to assess the quality of each measure (see Table 4). The codings provided an overall estimate for each of the six criteria. So, the overall reliability score for each measure was used when possible. When there were multiple reliability estimates and/or only estimates for each subscale, the range of the reliability scores was reported.

The number of validity tests that had been used to analyse the measure were used to assess the quality of that criteria. The types of validity analysis that were considered are outlined in Table 3. Each type of validity analysis had to meet significant levels and have an appropriate sample size to be considered acceptable for inclusion. For example, Floyd and Widaman (1995) recommend a minimum of 1:5 items to cases ratio for factor analytic techniques. So assessment of the construct validity of any measure would require this ratio to be considered as an acceptable validity test.

Table 4 Coding instructions for the quality assessment of the outcome measures

Fitzpatrick Criteria	Coding	Explanation
Reliability	Adequate	> 0.7
	Partial	> 0.5 < 0.7
	Inadequate	< 0.5
	Unknown	Reliability not supplied
Validity	Adequate	Reports >3 types of validity tests
	Partial	Reports 2 types of validity tests
	Inadequate	Reports 1 validity test
	Unknown	Validity estimates not supplied
Responsiveness	Adequate	Significant differences found between groups or within individuals
	Partial	Non-significant trends found between groups or within individuals

	Inadequate	Not addressed
Acceptability	Adequate	All of the components described
	Partially addressed	At least one of the components described
	Not addressed	None of the components described
Feasibility	Adequate	All of the components described
	Partially addressed	At least one of the components described
	Not addressed	None of the components described
Precision	Adequate	All of the components described
	Partially addressed	At least one of the components described
	Not addressed	None of the components described

Results

After the removal of measures that had not been used in individual psychological therapy with people with intellectual disabilities, ten measures were included in the quality appraisal. The psychometric properties of all the measures had been assessed on populations within the UK. Table 2 summarises each of the measures and the primary articles that assessed the psychometric properties with adults with intellectual disabilities. There were six single-trait measures: three assessed anger, one assessed depression, one assessed anxiety, and one assessed self-esteem. There were also four multi-trait measures: two assessed psychological distress and nine indices of symptoms, one assessed interpersonal difficulties, and one assessed feelings. The response scales varied between a 2-, 3-, 4- and 5-point Likert scales, with 50% using a 3-point Likert scale. Five of the outcome measures were adapted from use with the general population, four were designed specifically for use with people with intellectual disabilities, and one was designed for use with the general population.

Table 5 summarises the key psychometric properties and the quality of each measure in relation to the Fitzpatrick et al. (1998) criteria. The review will discuss the findings of the quality appraisal in relation to each of the criteria, below.

Reliability

All measures showed adequate levels of internal consistency ($\alpha > 0.7$) apart from the Rosenberg Self-Esteem Scale (RSES; Rosenberg, 1965), which reported partial levels ($\alpha = 0.64$). The Clinical Outcome in Routine Evaluation - Learning Disabilities (CORE-LD; Brooks & Davies, 2007; Brooks et al., 2013), the Glasgow Anxiety Scale – Intellectual Disabilities (GAS-ID; Mindham & Espie, 2003), the Glasgow Depression Scale – Learning Disabilities (GDS-LD; Cuthill et al., 2003), Inventory of Interpersonal Problems – 32 (IIP-32; Barkham et al., 1996), the Novaco Anger Scale (NAS; Novaco & Taylor, 2004), the Provocation Inventory (PI; Novaco & Taylor, 2004) and the RSES showed very good test re-test reliability ($r = 0.52 – 0.97$). The GAS-ID and the GDS-LD displayed the best overall reliability, with a combination of excellent internal consistency and test re-test reliability scores.

However, sample sizes used for a number of these tests were very small. For example, the internal consistency of the GAS-ID was analysed on 35 participants with mild to moderate intellectual disabilities (‘anxious’ group = 19, non-anxious = 16). The test re-test reliability was only assessed on a subsample of 17 participants from the overall group. Such small sample sizes create difficulties in the ability to extrapolate the findings to the wider intellectual disability population.

Validity

There was a large discrepancy in the number and quality of the various validity assessments. For example, no validity assessments have been run on the CORE-LD and two assessments of validity (concurrent and convergent) were done on the GAS-ID. The Anger Inventory (AI; Rose & Gerson, 2009) was found to have no relationship with the staff measure (concurrent validity) at assessment or pre-treatment. Overall, no measure met the ‘adequate’ level of quality in line with the quality appraisal criteria, and only the GAS-ID, IIP-32 and RSES met the partially evidenced criteria.

Construct validity analysis (factor analysis) was only conducted on the Brief Symptom Inventory (BSI, Derogatis, 1993), IIP-32 and RSES. Interestingly, the analysis of these measures found that they did not factor in the same way as they did in their analysis with the general population. The IIP-32, for example, has been found to consist of eight 4-item subscales with general adult population (Barkham et al., 1996). Kellett et al. (2005) found only four interpretable factors in their analysis with 255 adults with intellectual disabilities; three factors which mapped onto the same as with the general population (Hard to be Assertive, Hard to be Supportive, and Too Aggressive) and one which was an amalgamation of two factors (Hard to be Involved and Hard to be Sociable) from Barkham et al.’s (1996) analysis.

Table 5 Quality appraisal of the outcome measures

	Reliability		**Validity**				**Responsiveness**	**Acceptability**	**Feasibility**	**Precision**
Measure	Internal Consistency (Cronbach's a/split- half)	Test re-test	Discriminant	Concurrent	Convergent	Construct				
Anger Inventory Rose (2013); Rose & Gerson (2009)*	0.923 *study 1* 0.93 *study 2*	No details	No details	No relationship with staff measure	No details	No details	Adequate *evaluative*	Not addressed	Partially addressed *describes how administered*	Partially addressed *Describes how data is scored and mean scores*
Brief Symptom Inventory Alim, (2010); Kellett et al., (2003; 2004;	0.63 – 0.78 / r = 0.66 – 0.79	No details	No details	No details	No details	8 factor structure	Adequate *evaluative*	Not addressed	Partially addressed *Describes assisted completion format and training*	Adequate *Differences found between different population, how data is scored and cites*

2009)*									*required to administer*	*benchmarks*
Clinical Outcome in Routine Evaluation – LD Brooks et al. (2013)•	0.80	r = 0.64	No details	No details	No details	No details	Adequate *evaluative*	Partially addressed *Adapted questions for population, developed with service users*	Not addressed	Not addressed
Glasgow Anxiety Scale – ID Atnas & Lippold (2013); Mindham & Espie (2003)*	0.96 / r = 0.93	r = 0.95	No details	Significant correlation with BAI	Significant correlation of physiological subscale and heart rate	No details	Adequate *evaluative*	Partially Addressed *Developed with service user and time taken to complete described*	Partially addressed *Describes how administered and how long it takes to complete*	Adequate *Differences found between different populations, how data is scored and cites benchmarks*

Glasgow Depression Scale – LD Atnas & Lippold (2013); Cuthill, Espie & Cooper (2003)*	0.90	r = 0.97	No details	Significant correlation with BDI (r = 0.94) and carer form (r = 0.93)	No details	No details	Partial *Evaluative but non-significant changes*	Partially addressed *Developed with service user and time taken to complete described*	Partially addressed *Describes how administered and how long it takes to complete*	Adequate *Differences found between different populations, how data is scored and cites benchmarks*
Inventory of Interpersonal Problems – 32 Alim, (2010); Beail et al. (2005); Kellett et al.	0.89	r = 0.84	No details	Significant correlation with Interpersonal Sensitivity subscale on BSI (r = 0.35)	No details	4 factor structure	Adequate *Evaluative and discriminative*	Not addressed	Partially addressed *Describes assisted completion format and training required to administer*	Partially addressed *Differences found between different populations*

(2005, 2009)*										
Novaco Anger Scale Alim, (2010); Novaco & Taylor (2004); Taylor et al. (2005)*	0.92	r = 0.52	No details	Significant correlations with subscales of STAXI	No details	No details	Adequate *Evaluative (but no scores given) and discriminative*	Partially addressed *Adapted questions for population and reported why people did not complete measure*	Partially addressed *Described how administered*	Not addressed
Provocation Inventory Alim, (2010); Novaco & Taylor (2004);	0.92	r = 0.57	No details	No details	No details	No details	Adequate *Evaluative (no scores given) and discriminative*	Partially addressed *Adapted questions for population and reported why people*	Partially addressed *Described how administered*	Not addressed

Taylor et al. (2002, 2005)*								*did not complete measure*		
Rosenberg Self-Esteem Scale Alim, (2010); Beail et al. (2005); Davis et al. (2009); Kellett et al. (2009)*	0.64	r = 0.63	Negative correlation with IIP-32 (r = -0.32)	No details	No details	2 factor structure	Adequate *Evaluative (no scores given) and discriminative*	Not addressed	Partially addressed *Describes assisted completion format and training required to administer*	Partially addressed *Differences found between different populations*
Symptom Checklist – 90R Beail et al. (2005); Kellett et al.	0.75 – 0.86 / r = 0.71 – 0.86	No details	No details	No details	No details	No details	Adequate *Evaluative and discriminative*	Not addressed	Partially addressed *Describes assisted completion format and*	Partially addressed *Differences found between different populations*

(1999); Newman & Beail (2002)*	*training required to administer*

* primary articles and studies used to appraise the quality of the outcome measures

One possible reason for this is the small sample sizes often used to assess the psychometric properties of the outcome measures. Floyd and Widaman (1995) recommend a minimum of 1:5, item: case ratio for factor analytic techniques. This would mean that the 48-item NAS would have to be completed by at least 240 participants to be able assess the construct validity of the measure.

Responsiveness

All measures showed either adequate or partial responsiveness due to the initial search criteria limiting the quality appraisal to measures that have been used in individual therapy with people with intellectual disabilities. Many of the studies reported changes within individuals over the course of therapy (evaluative) rather than between groups (descriptive). Only the NAS and PI have been used in controlled research designs (Taylor et al., 2002; 2005). In these studies, changes in mean NAS and PI scores over an 18-session intervention were compared to a waiting list control. This is consistent with literature regarding assessment of psychological therapies with people with ID, where there are still very few controlled trials (Beail, 2003; Willner 2005).

Acceptability

No measures reached adequate levels of acceptability for service users. This was often due to lack of information rather than lack of Acceptability. For example, only the GAS-ID, GDS-LD and CORE-LD were developed specifically for use with people with intellectual disabilities, with the help of service users to generate the item pool. The acceptability (e.g. reading age, length of instrument etc) of these measures for adults with intellectual disabilities may be implicit through their design, but no data on this were reported.

In terms of populations that the measures may not be suitable for, the psychometric properties of all of the measures were assessed on people with ‘mild’ to ‘moderate’ intellectual disabilities. Moreover, only the NAS and PI reported why people did not complete the measure. Taken together, this means that there is still some uncertainty for whom or under what circumstances the measures are not suitable for use. For example, could the measures be used with people with ‘severe’ intellectual disabilities?

Feasibility

No measures reached adequate levels of Feasibility. Again, this was due to lack of information rather than poor Feasibility. All of the measures except the CORE-LD described administration instructions. These were an assisted completed method, whereby the administrator would read each question to the individual and then ask them to rate themselves using a pictorial version of the Likert scale. However, no administration manuals or training packages for the measures were described, and only the GAS-ID and GDS-LD reported time taken to administer the measures.

Other areas that could not be determined included: level of training required to administer the measures, cultural or language translations or adaptations and details on scoring instructions or availability of scanning options.

Precision

The analysis of the Precision - the ability to detect differences between different populations, information on how the measure should be scored and ability to cite benchmarks to facilitate interpretation of the scores – was mixed. Analysis to assess the ability to detect differences between different populations was conducted on the BSI, GAS-ID, GDS-LD, IIP-32, RSES and Symptom Checklist – 90R (SCL-90R, Derogatis, 1983). This often consisted of comparing clinical (meeting diagnostic criteria, referred for mental health difficulties) and non-clinical (referred for eligibility assessments, not in receipt of mental health services) samples. The GAS-ID and the GDS-LD also analysed data from the general population. The measures that were assessed in this way showed good psychometric properties within each of the specific populations and were able to detect differences between them. For example, Kellett et al. (2003) compared data on the BSI across non-clinical, clinical and forensic sample, and found that "The reliability results…illustrate that the nine symptom dimensions remain broadly reliable according to context" (p. 130). They also found that there were significant differences in reported symptoms and overall psychopathology across the three groups. The non-clinical group were the least symptomatic, followed by the forensic group, and the clinical group reported the highest levels of symptoms and overall psychopathology. The CORE-LD, NAS and PI did not compare scores across different populations.

Research analysing the GDS-LD, GAS-ID and BSI reported how data can be scored into an overall score and dimension scores. For example, Kellett et al. (2004) reported that the BSI can be scored up into nine symptom dimensions and three global indices of psychopathology: the Global Severity Index (GSI) as an indicator of psychological distress; the Positive Symptom Distress Index (PSDI) as an indicator of symptom intensity corrected for number of present symptoms; and the Positive Symptom Total (PST) as a count for the number of positive symptoms for an individual.

They also reported various meaningful benchmarks in terms of normative or comparative data to facilitate interpretation. For example, Cuthill et al. (2003) identified clinical cut-off score of 13-15 on the GDS-LD to identify people with “possible depression” (p. 350).

All of the other measures reviewed did not report scoring instructions or meaningful benchmarks. They alluded to higher scores suggesting higher levels of distress in the domains they were assessing (e.g. higher scores on the NAS suggest higher levels of anger), but did not report how the administrator would score. For example, whether one should record total scores, mean scores or scores from only the positively rated items. For this reason they were not rated as having met these criteria in the quality appraisal.

Discussion

The main aim of the review was to systematically evaluate the quality of outcome measures that have been used in individual psychological therapy with adults with intellectual disabilities. The number of measures initially identified was small (n = 13). The psychometric properties of three of these measures had not been assessed on the target population and were removed. Therefore, ten outcome measures were identified for the quality appraisal.

Quality Appraisal

The review highlighted that assessment of the construct validity of the available measures is lacking. This is important when one considers the measures that had their construct validity assessed - BSI, IIP-32, RSES – were found to factor differently with adults with intellectual disabilities compared to the general population. For example, Davis et al. (2009) found that the RSES consisted of two factors (Self-Worth and Self-Criticism) when assessed with 219 adults with intellectual disabilities, compared to the uni-dimensional structure found through it's initial development with adolescents without intellectual disabilities (Rosenberg, 1965). Findings like this suggest that adults with intellectual disabilities may experience mental health difficulties in a different way to the general population. This means that outcome measures designed to assess specific mental health difficulties (e.g. depression, anxiety disorders) in adults with intellectual disabilities may not be assessing the constructs we believe they are.

In assessing the responsiveness of measures the review revealed that a lot of the research assessing the outcomes of psychological therapy with adults with intellectual disabilities consists of case reports and non-controlled studies. This is an often-reported problem in assessing the effectiveness of psychological therapies with adults with intellectual disabilities (Nicoll et al., 2013; James & Stacey, 2014). It also limits the quality of the available therapy outcome measures, as their discriminative responsiveness cannot be assessed. Only the NAS and PI have been used in a waiting list controlled study (Taylor et al., 2002; 2005). More controlled trials are needed to both assess the effectiveness of psychological therapies with adults with intellectual disabilities and the quality of the outcome measures used.

The acceptability of outcome measures with the target population was another poorly addressed area identified by the current review. The AI, GAS-ID, GDS-LD and CORE-LD had been developed specifically for use with adults with intellectual disabilities. The remaining measures had been adapted for use with this population. There were no reports on service user involvement in the adaptation of these measures and/or how they found the experience of completing the measures. Only the CORE-LD reported any service user feedback on what it was like to complete (Brooks et al., 2013). There is a current drive for greater service user involvement in intellectual disability healthcare services (DoH, 2000; 2009b). Co-production of services, resources and assessment tools is seen as essential in delivering effective services (Greenhill & Whitehead, 2010; Roberts et al., 2011; 2013). Therefore, service user inclusion in the development and/or feedback of is essential to ensure acceptability of an outcome measure.

Government legislation has argued that psychological therapies need to be available for all people with intellectual disabilities (DoH, 2007; 2009a). The review has identified that the psychometric properties of all of the currently available measures have been assessed on people with ‘mild’ to ‘moderate’ intellectual disabilities. This raises issues as to whether the measures are acceptable for use with people with ‘severe’ or ‘specific’ (e.g. autism, Down’s syndrome) intellectual disabilities. Further work assessing the psychometric properties in different populations needs to be done to improve the Acceptability of the measures.

Limited information was provided on time taken to administer outcome measures and administration instructions. This may be less problematic for well-established measures such as the BSI, IIP-32, RSES and SCL-90R that have administration manuals and detailed ‘assisted completion format’ administration instructions (see Kellett et al., 1999). The combination of these sources of information may be enough to aid administration of these measures. However, without clear instructions clinicians may adapt items for service users. This may affect the meaning of the items and the overall validity of the outcome measure.

A number of the outcome measures assessed individual mental health difficulties. For example, the NAS and the RSES assess anger and self-esteem, respectively. This can create difficulties in routine care, where co-morbidity of mental health difficulties is common. Trying to use a number of single-trait measures to capture the difficulties of a service user experiencing multiple mental health difficulties would be time consuming and burdensome even if each measure was relatively short to complete on its own (e.g. GDS-LD, GAS-ID).

Finally, there was limited reporting of benchmarks and/or cut-offs in the reviewed outcome measures. This is important to be able to assess the severity of a difficulty in relation to normative data. Similarly there is the need to be able to identify clinically significant change to assess the effectiveness of an intervention (Evans et al., 2002; Jacobson & Traux, 1991).

Limitations

One of the major limitations of the current review was the quality of the research assessing the effectiveness of psychological therapies with adults with intellectual disabilities. A number of studies identified in the systematic search did not report the use of any outcome measures to assess change (despite this being one of the search terms), and out of the 212 identified, only two had used a comparative waiting-list control. This is a major issue as ability to detect change is an essential criterion for a therapy outcome measure (Fitzpatrick et al., 1998). Historically, researchers have been interested in developing diagnostic and screening tools for use with adults with intellectual disabilities (Kellett et al., 2003). However, it is now imperative that they develop measures that can be used to assess therapy outcomes, and design studies where change can be quantified.

Another limitation was the lack of information about the outcome measures in the research papers. Quality assessing the Acceptability and Feasibility was particularly difficult because of this. For example, only studies assessing the NAS and PI reported why people did not complete the measures. Without information like this it is difficult to know for whom or when the measures are suitable for use.

There was also limited information on how each measure was scored. All of the papers implied that higher scores equated to higher levels of distress, but did not report how to score them (i.e. mean score or total score), or whether to take scores from specific items or all of them. Equally, there was no information on whether items were reversed, or if clinicians needed to reverse them when they were scored up. A good example of this is from the development of the CORE-LD (Brooks et al., 2013), where they report “the mean pre-therapy scores are 0.9 and the mean post therapy score is 0.5” (p. 327), but do not report what these mean scores are taken from.

Conclusions

Results from the current review reveal a number of single- and multi-trait outcome measures that are currently used in individual psychological therapies with adults with intellectual disabilities.

The quality appraisal of these measures highlighted that the outcome measures currently used have good psychometric properties, but there is still work to be done to improve their overall quality. The three key areas that need to be addressed are:

1. Assessment of the construct validity of the available outcome measures

2. Use more rigorous designs to assess the effectiveness of psychological therapies with adults with intellectual disabilities and the ability of therapy outcome measures to detect change through intervention (e.g. controlled trials)

3. Further involvement of service users in all areas of measure development (item pool generation, wording of questions, feedback on length, administration procedure etc).

References

Alim, N. (2010). Therapeutic progressions of client and therapist throughout a course of psychodynamic therapy with a man with mild learning disabilities and anger problems. *Advances in Mental Health & Learning Disabilities, 4(1),* 42-49

Atnas, C. I. & Lippold, T. (2013). Cognitive behaviour therapy with a woman in her twenties with a mild intellectual disability presenting with psychogenic non-epileptic seizures. *Advances in Mental Health & Intellectual Disabilities, 7(4),* 245-250

Barkham, M., Evans, C., Margison, F., McGrath, G., Mellor-Clark, J., Milne, D. & Connell, J. (1998). The rationale for developing and implementing core outcome batteries for routine use in service settings and psychotherapy outcome research. *Journal of Mental Health, 7(1),* 35-47.

Barkham, M., Hardy, G. E., & Startup, M. (1996). The Inventory of Problems-32: A short version of the Inventory of Interpersonal Problems. *British Journal of Clinical Psychology, 35,* 21–35

Barkham, M., Margison, F., Leach, C., Lucock, M., Mellor-Clark, J., Evans, C., McGrath, G., Benson, L., Connell, J. & Audin, K. (2001). Service profiling and outcomes benchmarking using the CORE-OM: Toward practice-based evidence in the psychological therapies. *Journal of Consulting & Clinical Psychology, 69(2),* 184-196

Beail, N. (2003). What works for people with mental retardation? Critical commentary on cognitive-behavioural and psychodynamic psychotherapy research. *Mental Retardation, 41,* 468-472

Beail, N., Warden, S., Morsley, K. and Newman, D. W. (2005). Naturalistic evaluation of the effectiveness of psychodynamic psychotherapy with adults with intellectual disabilities. *Journal of Applied Research in Intellectual Disabilities, 18,* 245-251.

Beck A. T., Steer R. A. & Brown G. K. (1996). *Beck Depression Inventory – 2nd Edition Manual*. San Antonio: The 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 Youth Inventories, 2005. 2nd ed. Retrieved From: http://www.pearsonassessments.com/HAIWEB/Cultures/enus/Productdetail.htm?Pid=015-8014-19.

Bender, M. (1993). The unoffered chair: The history of therapeutic disdain towards people with a learning difficulty. *Clinical Psychology Forum. 54(1),* 7-12

Benson, B.A. (2004). Psychological interventions for people with intellectual disability and mental health problems. *Current Opinion in Psychiatry, 17(5),* 353-67

Brooks, M. & Davies, S. (2007). Pathways to participatory research in developing a tool to measure feelings. *British Journal of Learning Disabilities, 36,* 128-133.

Brooks, M., Davies, S. & Twigg, E. (2013). A measure for feelings – Using inclusive research to develop a tool for evaluating psychological therapy (clinical outcomes in routine evaluation – learning disability). *British Journal of Learning Disabilities, 41,* 320-329.

Cahill, J., Barkham, M., Hardy, G., Gilbody, S., Richards, D., Bower, P., Audin, K. & Connell, J. (2008). A review and critical appraisal of measures of therapist-patient interactions in mental health settings. *Health Technology Assessment, 12(24),* 1-68.

Cuthill, F. M., Espie, C. A. & Cooper, S-A. (2003). Development and psychometric properties of the Glasgow depression scale for people with a learning disability: Individual and carer supplement versions. *British Journal of Psychiatry, 182,* 347-353

Davis C., Kellett S. & Beail, N. (2009). Utility of the Rosenberg self-esteem scale. *American Journal on Intellectual & Development Disabilities, 114(3),* 172-178

Department of Health (1996). *NHS psychotherapy services in England: Review of strategic policy*. London: Department of Health

Department of Health (2000). *The NHS plan*. . London: Department of Health.

Department of Health (2007). *Services for people with learning disabilities and challenging behaviour or mental health needs.* London: Department of Health.

Department of Health (2009a). *Learning disabilities positive practice guide.* London: Department of Health

Department of Health (2009b). *Valuing people now: From progress to transformation – a consultation on the next three years of learning disability polic*y. London: Department of Health

Derogatis, L.R. (1983) *SCL-9OR; Administration, scoring and procedures: Manual 11.* Towson, MD: Clinical Psychometric Research

Derogatis, L.R. (1993). *Brief symptom inventory: Administration scoring and procedures manual* (3rd ed.). Minneapolis: National Computer Systems.

Dilly, R. (2014). Eye movement desensitisation and reprocessing in the treatment of trauma with mild intellectual disabilities: A case study. *Advances in Mental Health & Intellectual Disabilities, 8(1),* 63-71

Evans, C., Connell, J., Barkham, M., Margison, F., McGrath, G., Mellor-Clark, J. & Audin, K. (2002). Towards a standardised brief outcome measure: Psychometric properties and utility of the CORE-OM. *British Journal of Psychiatry, 180,* 51-60.

Evans, C., Mellor-Clark, J., Margison, F., Barkham, M., McGrath, G., Connell, J. & Audin, K. (2000). Clinical outcomes in routine evaluation: The CORE-OM. *Journal of Mental Health, 9,* 247-255.

Field, A. P. (2013). *Discovering statistics using IBM SPSS statistics: (and sex and drugs and rock 'n' roll).* London: Sage.

Fitzpatrick, R., Davey, C., Buxton, M. J. & Jones, D. R. (1998). Evaluating patient-based outcome measures for use in clinical trials. *Health Technology Assessment, 2(14),* 1-74

Floyd, F. & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment, 7,* 286-299

Foa, E., Cashman, B., Jaycox, L. & Perr, K. (1997). The validation of a self-report measure of posttraumatic stress disorder: The posttraumatic stress diagnosis scale. *Psychological Assessment, 9(4),* 445-451.

Froyd, J. E., Lambert, M. J. & Froyd, J. D. (1996). A review of practicies of psychotherapy outcome measurement. *Journal of Mental Health, 5,* 11-15

Green, D. & Latchford, G. (2012). *Maximising the benefits of psychotherapy: A practice-based evidence approach.* London: Wiley-Blackwell

Greenhill B. & Whitehead R. (2010). Promoting service user inclusion in risk assessment and management: A pilot project developing a human rights based approach. *British Journal of Learning Disabilities, 39,* 277-283

Grissom, R. J. (1996). The magical number .7 +/- .2: Meta-meta-analysis of the probability of superior outcome in comparisons involving therapy, placebo, and control. *Journal of Consulting & Clinical Psychology, 64,* 973-982.

Hassiotis, A., Serfaty, M., Azam, K., Strydom, A., Blizard, R., Romeo, R., Martin, S. & King, M. (2013). Manualised individual cognitive behavioural therapy for mood disorders in people with mild to moderate intellectual disability: A feasibility randomised controlled trial. *Journal of Affective Disorders, 151,* 186-195.

Hatton, C. & Taylor, J. L. (2013). The assessment of mental health problems in adults with intellectual disabilities. In J. L. Taylor, W. R. Lindsay, R. P. Hastings & C. Hatton (Eds.), *Psychological therapies for adults with intellectual disabilities* (pp. 31-54). Chichester: Wiley & Sons.

Jacobson, N. S. & Traux, P. (1991). Clinical significance: A statistical approach to defining change in psychotherapy research. *Journal of Consulting & Clinical Psychology, 59(1),* 12-19.

James, C. W. & Stacey, J. M. (2014). The effectiveness of psychodynamic interventions for people with learning disabilities: A systematic review. *Tizard Learning Disability Review, 19(1),* 17-24

Kellett S., Beail, N., Bush, A., Dyson, G. & Wilbram, M. (2009). Single case experimental evaluations of psychodynamic and cognitive-behavioural psychotherapy: Examples of methods and outcomes. *Advances in Mental Health & Learning Disabilities, 3(4),* 36-44

Kellett, S., Beail, N., & Newman, D. W. (2005). Measuring interpersonal problems in people with mental retardation. *American Journal of Mental Retardation, 110(2)*, 136-144.

Kellett, S. C., Beail, N., Newman, D. W & Frankish, P. (2003). Utility of the brief symptom inventory in the assessment of psychological distress. *Journal of Applied Research in Intellectual Disabilities, 16*, 127-134.

Kellett, S. C., Beail, N., Newman, D. W & Hawes, A. (2004). The factor structure of the brief symptom inventory: Intellectual disability evidence. *Clinical Psychology & Psychotherapy, 11*, 275-281.

Kellett, S. C., Beail, N., Newman, D. W. & Mosley, E. (1999). Indexing psychological distress in people with an intellectual disability: Use of the symptom checklist-90-r. *Journal of Applied Research in Intellectual Disabilities, 12(4),* 323-334.

Kraus, D. & Castonguay, L. G. (2010). Treatment outcome package (TOP) – Development and use in naturalistic settings. In M. Barkham, G. E. Hardy & J. Mellor-Clark (Eds.), *Developing and delivering practice-based evidence: A guide for the psychological therapies* (pp. 155-175). Chichester: Wiley & Sons

Lueger, R. J. & Barkham, M. (2010). Using benchmarks and benchmarking to improve quality of practice and services. In M. Barkham, G. E. Hardy & J. Mellor-Clark (Eds.), *Developing and delivering practice-based evidence: A guide for the psychological therapies* (pp. 223-256). Chichester: Wiley & Sons

Martorell, A., Tsakanikos, E., Pereda, A., Gutierrez-Recacha, P., Bouras, N. & Ayosu-Mateos, J. L. (2009). Mental health in adults with mild and moderate intellectual disabilities. *American Journal of Mental Retardation, 113,* 307-318.

McCabe, M. P., McGillivray, J. A. & Newton, D. C. (2006). Effectiveness of treatment programmes for depression among adults with mild/moderate intellectual disability. *Journal of Intellectual Disability Research, 50,* 239-247.

McGurk, K. A. & Skelly, A. (2014). A quick guide for psychologists working in learning disability on the available clinical outcome measures currently in use. *The Bulletin of the Faculty for People with Intellectual Disabilities, 12(1),* 28-45

Mindham, J. & Espie, C. A. (2003). Glasgow anxiety scale for people with an intellectual disability (GAS-ID): Development and psychometric properties of a new measure for use with people with mild intellectual disability. *Journal of Intellectual Disability Research, 47(1),* 22-30

Newman, D. W. & Beail, N. (2002). Monitoring change in psychotherapy with people with intellectual disabilities: The application of the assimilation of problematic experiences scale. *Journal of Applied Research in Intellectual Disabilities, 15,* 48-6

NHS Choices (2013). *Psychotherapy – How it works.* Retrieved from: http://www.nhs.uk/Conditions/Psychotherapy/Pages/Whatitisusedfor.aspx

Nicoll, M., Beail, N. & Saxon, D. (2013). Cognitive behavioural treatment for anger in adults with intellectual disabilities: A systematic review and meta-analysis. *Journal of Applied Research in Intellectual Disabilities, 26,* 47-62

Novaco, R. W. & Taylor, J. L. (2004). Assessment of anger and aggression in male offenders with developmental disabilities. *Psychological Assessment, 16,* 42-50.

Prout T. H. & Nowak-Drabik K. M. (2003) Psychotherapy with persons who have intellectual disabilities: An evaluation of effectiveness. *American Journal on Intellectual Disabilities, 108,* 82–93.

Psaila, C. L. & Crowley, V. (2005). Cognitive analytic therapy in people with learning disabilities: An investigation into the common reciprocal roles found within this client group. *Mental Health and Learning Disabilities: Research and Practice, 2(2),* 96-108.

Roberts, A., Greenhill, B., Talbot, A. & Cuzak, M. (2011). 'Standing up for my human rights': A group's journey beyond consultation towards co-production. *British Journal of Learning Disabilities, 40,* 292-301

Roberts, A., Townsend, S., Morris, J., Rushbrooke, E., Greenhill, B., Whitehead, R., Matthews, T. & Golding, L. (2013). Treat me right, treat me equal: Using national policy and legislation to create positive changes in local health services for people with intellectual disabilities. *Journal of Applied Research in Intellectual Disabilities, 26,* 14-25

Rose, J. L. (2013). A preliminary investigation into the influence of therapist experience on the outcome of individual anger interventions for people with intellectual disabilities. *Behavioural & Cognitive Therapy, 41,* 470-478

Rose, J. L. & Gerson, D. F. (2009). Assessing anger in people with intellectual disability. *Journal of Intellectual & Developmental Disability, 34(2),* 116-122

Rose, D. & Sullivan, O. (1996). *Introducing data analysis for social scientists* (2nd Ed.). London: Open University Press.

Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.

Roth, A. & Fonagy, P. (2005). *What works for whom? A critical review of psychotherapy research (2nd Ed.).* New York: Guildford Press

Singh, N. N., Lancioni, G. E., Winton, A. S. W., Singh, A. N., Adkins, A. D. & Singh, J. (2008). Clinical and benefit-cost outcomes of teaching a mindfulness-based procedure to adult offenders with intellectual disabilities. *Behaviour Modification, 32,* 622-637.

Slevin, M.L., Plant, H., Lynch, D., Drinkwater, J. & Gregory, W. M. (1988). Who should measure quality of life, the doctor or the patient? *British Journal of Cancer*, *57***,** 109–112.

Spitzer, R. L., Kroenke, K., Williams, J. B. & Lowe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archives of Internal Medicine, 166(10),* 1092-1097.

Strupp, H. H. (1978). Psychotherapy research and practice: an overview. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behaviour change* (2nd Ed., pp. 3-22). New York: Wiley

Taylor, J. L., Novaco, R. W., Gillmer, B., Thorne, I. (2002). Cognitive-behavioural treatment of anger intensity among offenders with intellectual disabilities. *Journal of Applied Research in Intellectual Disabilities, 15,* 151-165

Taylor, J. L., Novaco, R. W., Gillmer, B. T., Robertson, A. & Thorne, I. (2005). Individual cognitive-behavioural anger treatment for people with mild-borderline intellectual disabilities and histories of aggression: A controlled trial. *British Journal of Clinical Psychology, 44,* 367-382.

Thornicroft, G. (2006). *Shunned: Discrimination against people with mental illness.* Oxford: Oxford University Press

Vereenooghe, L. & Langdon, P. E. (2013). Psychological therapies for people with intellectual disabilities: A systematic review and meta-analysis. *Research in Developmental Disabilities, 34,* 4085-4102

Whitehouse, R. M., Tudway, J. A., Look, R. & Kroese, S. (2006). Adapting individual psychotherapy for adults with intellectual disabilities: A comparative review of the cognitive behavioural and psychodynamic literature. *Journal of Applied Research in Intellectual Disabilities, 19,* 55-65.

Willner, P. (2005). The effectiveness of psychotherapeutic interventions for people with learning disabilities: A critical overview. *Journal of Intellectual Disability Research, 49(1),* 73-85.

Wing, J. K., Beevor, A. S., Curtis, R.H. ET AL. (1998). Health of the nation outcome scales (HoNOS): Research and development. *British Journal of Psychiatry*, *172*, 11-18

Development and Psychometric Properties of the Psychological Therapy Outcome Scale – Intellectual Disabilities (PTOS-ID)[2]

Nik Vlissides

University of Liverpool

Laura Golding

University of Liverpool

Nigel Beail

South West Yorkshire Partnership NHS Foundation Trust

---

[2] To be submitted to: *Journal of Applied Research in Intellectual Disabilities* (word count, 7000)

Abstract

Background. There are few valid and reliable therapy outcome measures available for use with people with intellectual disabilities. The current study aims to examine the validity and internal consistency of a new scale, the Psychological Therapy Outcome Scale – Intellectual Disabilities (PTOS-ID).

Method. The PTOS-ID was administered to 175 service users accessing specialist intellectual disabilities services.

Results. Three factors emerged from the principal components analysis with high levels of internal consistency: a) emotional and behavioural discomfort ($\alpha = 0.82$), b) positive well-being ($\alpha = 0.81$), and c) anxiety ($\alpha = 0.76$). Factors a) and b) were combined to measure Psychological Distress ($\alpha = 0.85$), which correlated strongly with the Global Severity Index on the Brief Symptom Inventory ($r = 0.85$).

Conclusions. This preliminary study suggests that the PTOS-ID is a psychometrically robust measure that can be used with people with intellectual disabilities. Further research is required to assess its ability to detect change over therapy.

Keywords: intellectual disabilities, learning disabilities, PTOS-ID, outcome measure, psychological therapy

Introduction

People with intellectual disabilities are more likely to experience life events and living circumstances associated with an increased risk of mental health problems, including stigmatization, lack of meaningful friendships, unemployment and discrimination (Martorell et al., 2009; Marshall & Willoughby-Booth, 2007; Thornicroft, 2006). Clinicians have responded to this by increasing the availability of psychological therapies for people with intellectual disabilities (Willner, 2005). There is also a growing evidence base for the effectiveness of psychological therapies for people with intellectual disabilities (Beail, 2003; Hatton, 2002; Whitehouse et al., 2006). For example, recent reviews have found that psychological therapies have moderate benefit for people with intellectual disabilities (Nicoll et al., 2013; Prout & Nowak-Drabik, 2003; Taylor et al., 2008) and that the changes are maintained over time (Beail et al., 2005).

Within the psychotherapeutic community, there is a shift towards practice based evidence approaches as well as using evidence-based approaches (Barkham, Hardy & Mellor-Clark, 2010). There is an expectation that service providers must show some evidence for the effectiveness of what they do (Department of Health, 2010). One of the difficulties for the providers of psychological therapies for people with intellectual disabilities is the availability of valid and reliable outcome measures that are appropriate for use with this population, are easy to use in service settings, and are inexpensive (Skelly, 2011; Weston et al., 2011).

To date, most of the measures for assessing psychological therapies with people who have intellectual disabilities fall into three categories: measures developed for a general population and used with people with intellectual disabilities; general population measures adapted for use with people who have intellectual disabilities; and measures developed specifically for people with intellectual disabilities

There have been a number of single- and multi-trait measures that have been developed for use with the general population, but have been used with people with intellectual disabilities in an assisted completion format. For example, Kellett et al. (1999) detailed an assisted completion protocol of the Symptom Checklist-90R (SCL-90R, Derogatis, 1983) with people with intellectual disabilities. They also found that the measure had very good levels of internal consistency (α >0.7 for all subscales) and discriminant validity between clinical and non-clinical populations.

Other measures that have been developed for use with the general population include: the Brief Symptom Inventory (BSI, Derogatis, 1993), the Inventory of Interpersonal Problems (IIP-32, Barkham et al., 1996), the Rosenberg Self-Esteem Scale (RSES, Rosenberg, 1965), the Beck Depression Inventory – 2nd edition (BDI-II, Beck et al., 1996), and the Beck Anxiety Inventory (BAI, Beck et al., 1988). The IIP-32 was found to have acceptable reliability and validity with people with mild intellectual disabilities (Kellett et al., 2005). The RSES was found to have less than satisfactory reliability and validity from a sample of 219 people with intellectual disabilities (Davis et al., 2009). The BDI-II and BAI were found to have good levels of internal consistency, but no information on the validity was gathered (Lindsay, 2007).

The BSI (Derogatis, 1993) has been found to have good reliability in assessing psychological distress in people with intellectual disabilities (Kellett et al., 2003). Kellett et al. (2004) conducted an exploratory factor analysis on 335 completed BSIs with people with intellectual disabilities. They found acceptable validity for five out of the original nine primary symptom dimensions. However, there was some evidence of items unexpectedly loading onto scales. For example, the depression scale consisted of items from the Depression, Interpersonal sensitivity, Somatization and Psychoticism scales from the original BSI analysis (Kellett et al., 2004). This suggests that although the BSI may provide an appropriate measure of psychological distress, specific symptoms may be experienced differently by people with intellectual disabilities compared to the general population.

An alternative approach to the measurement of the outcome of psychological therapy is to adapt scales for use with the general population for people with intellectual disabilities. Examples of this approach include the Clinical Outcome in Routine Evaluation – Learning Disabilities (CORE-LD, Marshall & Willoughby-Booth, 2007), the Health of the Nation Outcome Scale – Learning Disabilities (HoNOS-LD, Skelly & D'Antonio, 2008), and the Anxiety, Depression and Mood scales (Hermans et al., 2012). The Anxiety, Depression and Mood scale was found to have very good levels of internal consistency ($\alpha > 0.80$) and excellent test-retest reliability ($r > 0.75$). The HoNOS-LD and the CORE-LD were found to have good test-retest reliability with people with intellectual disabilities.

The CORE-LD (Brooks & Davies, 2007), the Glasgow Anxiety Scale – Intellectual Disabilities (GAS-ID, Mindham & Espie, 2003), and the Glasgow Depression Scale – Learning Disabilities (GDS-LD, Cuthill, et al., 2003) have been identified as a therapy outcome measure specifically designed for people with intellectual disabilities. The GAS-ID and the GDS-LD have very good internal consistency ($\alpha > 0.80$) and test-retest reliability ($r > 0.75$). The CORE-LD was developed through combining adapted questions from the Clinical Outcome in Routine Evaluation – Outcome Measure (Evans et al., 2002) with questions developed from focus groups with people with intellectual disabilities on their experience of having an intellectual disability and how this impacts on feelings (Brooks & Davies, 2007). Brooks et al. (2013) analysed data from 486 completed measures and found that the CORE-LD had good levels of internal consistency ($\alpha = 0.80$), test-retest reliability ($p = 0.85$) and ability to detect change between the start and end of therapy.

The measures described above offer some positive advances in assessing outcomes of psychological therapy with people with intellectual disabilities. However, there are some limitations to the currently available measures. First, a number of the measures are single problem or diagnosis (e.g. GAS-ID, BDI-II etc) and can, therefore, only evaluate change in one problem area. This is problematic when co-morbidity and dual diagnoses are common. Although service users can be asked to complete a number of measures this can often be burdensome and time consuming.

Second, the psychometric properties of some of the current measures have often been standardised with small and unrepresentative samples. For example, the CORE-LD (Marshall & Willoughby-Booth, 2007) was standardised with twenty-two people with “mild to moderate” (p. 35) intellectual disabilities. Equally the psychometric properties of the GDS-LD were determined from analysis with 38 people with intellectual disabilities. This raises the issue of the generalisability of some of the measures and their use in services with people with markedly different levels of cognitive ability.

Third, measures like the HoNOS-LD are completed by clinicians. This raises the question of how accurately the scores capture service users’ emotional distress and psychological well-being. The language used in some of the items in many of the measures needs some rewording to enable a person with intellectual disabilities to understand it. Measures such as the BSI and IIP-32 are the subject of copyright and cannot be modified or reproduced in a different way. Thus the administrator has to use the instrument in a creative way to aid client understanding (Kellet et al, 1999). Kellett et al. (2005) and Davis et al. (2009) also showed that people with intellectual disabilities still had difficulties with some of the questions. This may mean that the items are not measuring the intended construct accurately and may affect the reliability and validity of the measure.

Fourth, some of the measures assess factors that psychological therapy may not have a direct impact on. For example, the HoNOS-LD includes some areas such as memory and orientation, which may not be applicable to therapeutic change. Therefore, any outcome measure needs to be sensitive to the factors that direct therapy can have an impact on and the change that may take place through the course of treatment.

Finally, some of the measures described above are copyrighted and are not freely available to services. In the current economic climate, the cost of these measures may be too much and may limit their use in public sector clinical services.

Although there have been some positive developments in the assessment of psychological therapies with adults with intellectual disabilities the above review suggests there are still some areas for improvement. Specifically, a therapy outcome measure needs to: a) be short and easy enough to complete in routine clinical practice, b) accurately measure the difficulties that people with intellectual disabilities present with and that psychological therapies can have a direct impact on, c) be completed in an assisted format so is accessible to most people with an intellectual disability, d) be trans-theoretical so it is appropriate for use with a number of therapeutic models, and e) be freely available for services to use.

Development of the Psychological Therapy Outcome Scale – Intellectual Disabilities (PTOS-ID)

The development of a new measure, the PTOS-ID (Appendix D), has followed a number of discreet stages. First, focus groups identified service systems and key dependent variables for the measure. Second, an item pool was developed from existing diagnostic manuals (e.g. DSM-IV, American Psychiatric Association, 1994) and available outcome measures. Finally, an appropriate response format was developed.

Focus Groups

One-hundred and ten clinical psychologists, trainee clinical psychologists and assistant psychologists participated in the focus group stage. Their primary aim was to focus on what dependent variables would demonstrate that their psychological intervention had been effective. A large pool of dependent variables focussing on both psychological symptoms (e.g. anxiety, depression, challenging behaviour and psychosis) and other areas such as quality of life and psychological well-being were generated. Service providers required the measure to cover a wide range of areas, but also wanted something that was quick and easy to administer and within their available resources. It was felt that there was a need to reduce the number of dependent variables to meet this need. Quality of life items were removed as it was deemed that it had a less direct impact from psychological therapy. Challenging behaviour is a frequently reported problem for care staff and services. However, interventions for such behaviours tend to be at a systems level and were, therefore, not considered here. Of the remaining areas, those that were mentioned most by participants were retained. These were: anxiety, depression, anger, interpersonal well-being, psychological well-being and self-esteem.

Development of Item Pool

Available diagnostic manuals (i.e. DSM-IV) and research on measures validated for use with adults with intellectual disabilities were reviewed to identify a descriptive pool of items for the depression, anxiety and anger subscales. Items that overlapped across the measures and contributed the most to their respective scales in terms of their psychometric properties were considered for inclusion.

The interpersonal functioning item pool was derived from research on the IIP-32 and research by Kellett et al. (2005). This helped identify the items that worked best for people with intellectual disabilities.

Items for self-esteem were taken from secondary analysis of the data reported in research by Davis et al. (2009) on the Rosenberg Self-Esteem Scale. This study showed that the scale was not uni-dimensional when used with adults who have intellectual disabilities and some items proved difficult for them. Secondary analysis of the self-worth items did not produce a scale with satisfactory internal consistency so better-worded items were developed from the three with the best psychometric properties.

Developing an item pool for psychological well-being proved difficult as there were no measures currently in use with people with intellectual disabilities that assessed this. The Warwick-Edinburgh Mental Well-Being Scale (WEMWBS, Tennant et al., 2007) was piloted in an outpatient clinic. It was found that the wording of some of the questions was too difficult for service users. There was also considerable overlap of some of the questions. Items that appeared to have the best face validity and measure different aspects of well-being were included in the item pool.

The item pool was then reduced through evaluation of their face validity and clinical importance in diagnostic manuals. This resulted in six items for anxiety and anger, seven items for depression, five for interpersonal well-being and six for psychological well-being. This provided a short (30 item) measure which could produce individual scores for depression, anxiety, anger, interpersonal well-being and psychological well-being.

Response Format

Finally, the style and length of the response format was developed. Kellett et al. (1999; 2003; 2005) demonstrated that a 0 to 4, 5-point Likert scale can be used in assisted completion format with adults with intellectual disabilities. Self-report scales developed for use with adults with intellectual disabilities tend to use shorter response formats such as 3- or 4-point Likert scales (Cuthill, et al., 2003; Novaco & Taylor, 2004). For the current measure, the response format needed to have sufficient sensitivity to detect change, but was also easy enough for recipients to understand. Analysis of 493 administrations of the BSI (Derogatis, 1993) with people with intellectual disabilities found that the “quite a lot” rating was checked the least often for 50 of the 53 items, less than 10% of the time for 50 items, and at most was checked only 12.6% of the time. Based on this the “quite a lot” point was removed from the response scale, leaving a 4-point Likert scale allowing for responses “not at all”, “a little”, “sometimes” and “a lot”.

Aim

The overall aim of the current study was to examine the psychometric properties of the PTOS-ID when completed by adults with intellectual disabilities. This includes assessment of: a) its factor structure, b) internal consistency, and c) concurrent validity through comparisons with a measure validated for use with adults with intellectual disabilities.

Methods

The psychometric properties of the PTOS-ID were assessed through analysis of service audit databases where the measure had been used to assess outcomes of psychological therapies as part of routine clinical practice. Initially, researchers contacted the services to ask for access to their anonymised audit databases. When services agreed to this NHS ethics approval was sought. Approval was granted to access the service audit databases in September 2013 (Appendix E). Data had been collected by the services from June 2009, and researchers accessed the audit databases in March 2014.

Sample

The sample consisted of 175 service users who completed the PTOS-ID as part of their routine care. Service user data were included if they were either registered or in the process of being registered for intellectual disability services. The sample had a mean age of 29.43 years (SD = 11.31) with a range of 17-62 years. The sample consisted of 91 males with an average age of 30.15 years (SD = 12.43) and 84 females with an average age of 28.66 years (SD = 10.01). The overall sample was achieved through aggregation of three (two NHS and one third sector) service audit data sets from intellectual disability services in the United

Kingdom. Data from 167 service users were taken from NHS services and eight from a third sector service. IQ scores were available for 127 service users, and ranged from 45 to 72 (mean = 59.85, SD = 6.39). There were a number of incomplete forms, so specific sample numbers are quoted for each part of the analysis.

Measures

Psychological Therapy Outcome Scale – Intellectual Disabilities

This is a 30-item scale designed to measure anxiety (six items), anger (six items), depression (seven items), interpersonal well-being (five items), and psychological well-being (six items). Each item is rated on a 4-point Likert scale anchored by “not at all” to “a lot”. The measure assesses indices of psychological distress (anxiety, anger, depression) and positive well-being (interpersonal well-being and psychological well-being). See above for information about the development of this new measure.

Brief Symptom Inventory

The BSI (Derogatis, 1993) is a 53-item self-report inventory designed to reflect the psychological symptom patterns of people experiencing psychiatric problems. Each item is scored on a 5-point Likert scale ranging from ‘not at all’ to ‘extremely’. The scores are interpreted in terms of nine dimensions, including, Somatization, Obsessive-compulsive, Interpersonal sensitivity, Depression, Anxiety, and Hostility. It also provides three global indices of psychological distress. The Global Severity Index (GSI) combines data on number of symptoms and level of distress to provide a single summary of psychopathology. The Positive Symptom Distress Index (PSDI) is a measure of symptom intensity accounting for number of symptoms. The Positive Symptom Total (PST) provides a count of the number of symptoms selected by the individual. The BSI has been found to reliably discriminate between community and clinical populations (Kellett et al., 2003), used to evaluate mental health interventions with people with an intellectual disability (Newman et al., 2003) and retains the majority of its factor structure with people with mild intellectual disabilities (Kellett et al., 2004).

Procedure

The PTOS-ID was completed in a one-to-one format, in a private setting (e.g. consulting room). Before completing the measure service users were told that they would be asked questions about how they had been feeling over the past week, including the day of assessment. They were informed that each question could be responded to according to one of four answers. Service users were provided with a copy of the response scale depicting the four possible answers in numerical, written and pictorial form to aid their completion.

The questions were read verbatim and in chronological order from the PTOS-ID. If service users had difficulty understanding the question they were read a simplified version of the question pertaining to the same difficulty. For example, *"are you feeling anxious?"* could be replaced with *"are you feeling scared/nervous?"*. If service users still had difficulty understanding the question, it was left blank. Administration instructions are reproduced in Appendix F.

In some services the BSI was administered alongside the PTOS-ID. Here, the BSI was administered in accordance to the 'assisted completion format' used by Kellett et al. (1999; 2003; 2004). In instances where the PTOS-ID and BSI were completed together service users were informed that although some of the questions may be similar, they were not being asked them again because they had got them wrong and that they should answer each question in accordance to how they were feeling. As the PTOS-ID and the BSI were administered in routine clinical practice the order in which they were administered was not recorded.

The measures were administered at a number of time points in each of the services. These were: assessment of eligibility for service, pre-therapy and post-therapy. The first time an individual had completed the PTOS-ID was used in the analysis. This meant that the analysis predominantly consisted of data from eligibility assessments and pre-therapy data.

Data Analysis

A number of statistical tests were used to assess both the reliability and the validity of the measure. Construct validity was assessed through exploratory factor analysis. Floyd and Widaman (1995) recommend a minimum of 1:5, item: case ratio for factor analytic techniques. The current study achieved 1:5.83 ratio. Exploratory factor analysis techniques were selected because there were no hypotheses based on theory about the composition of the subscales and were therefore used to help identify the latent constructs that underpin the measure. The internal reliability was assessed through internal consistency analysis. The concurrent validity was assessed through comparison with the BSI.

Results

Screening Analysis

Prior to conducting the principal components analysis, screening tests were conducted to identify any items for removal. These consisted of: a) assessing the percentage endorsement of each point on each item to ensure sufficient dispersion of items (Tinsley & Tinsley, 1987),

b) ensuring the Kaiser-Meyer-Olkin (KMO) value of each item was above the acceptable limit of 0.5 (Field, 2013), and c) analysis of the correlations between all items to ensure they correlate moderately with each other and the overall scale score (Field, 2013). If items failed to meet more than one of these criteria they were considered for removal from the final analysis.

All items showed adequate endorsement (i.e. each point was selected greater than or equal to 5% of the population) except item 8 – *‘do you care about people and their problems?’* – where ‘*not at all’* was only selected 4% of the time. This item was considered for exclusion from the analysis.

The KMO test indicated that the sample size was appropriate for factor analysis (KMO = .810). All KMO values for individual items were above the acceptable limit of .5 (Field, 2013), except item 8 (KMO = .44). This item was, therefore, removed from the principal components analysis.

Pilot examination of the items via correlational methods indicated that most PTOS-ID items correlated moderately with each other and with the overall scale score (Bartlett’s, $p < .001$). There were also no issues with multicollinearity (¦R¦ > .00002), suggesting that the data was suitable for factor analysis. However, five items did not correlate significantly with over 50% of the other items. This suggested that they may not be measuring the same overall construct as the other items, and were considered for removal, but as all other requirements were met these were retained for analysis.

Principal Components Analysis

Principal components analysis was run on the remaining 29 items (n = 165), using varimax orthogonal rotation. Orthogonal rotation was used because it was hypothesized that the well-being and distress scales would not be related. The analysis yielded a total of eight factors, which accounted for 62.59% of the variance (Table 1).

Table 1 Initial results from principal components analysis

Factor	**Eigenvalue**	**% of Variance**	**Cumulative %**
1	6.77	23.34	23.34
2	2.67	9.20	32.55
3	2.19	7.56	40.10
4	1.65	5.71	45.81
5	1.43	4.92	50.73
6	1.24	4.27	54.99
7	1.15	3.96	58.95
8	1.05	3.63	62.59

Tinsley and Tinsley (1987) suggest using four criteria for calculating how many factors should be extracted from a data set: a) Kaiser’s criterion, b) analysis of the scree plot, c) analysis of the percentage of variance explained by each factor, and d) percentage of variance explained by the overall model. Three interpretable factors were derived using a combination of these criteria and the face validity. The factors accounted for 40.10% of the variance in the model and all had eigenvalues greater than 1 (see Table 2). A detailed summary of the factor loadings >0.35 is outlined in Table 3. Two items failed to factor (items 13 and 14). The three factors that emerged were interpreted and labelled as: 1) Emotional and Behavioural Discomfort, 2) Positive Well-being, and 3) Anxiety.

Table 2 Results of the PTOS-ID principal components analysis

Factor	**Rotated Eigenvalue**	**% of Variance**	**Cumulative %**
1	4.46	15.39	15.39
2	4.01	13.83	29.21
3	3.16	10.89	40.10

The Emotional and Behavioural Discomfort scale consisted of a combination of the questions designed to assess symptoms associated with anger and depression. This included both emotional and behavioural aspects of the respective difficulties. Positive Well-being consisted of the questions aimed at assessing both the psychological and interpersonal well-being, as well negative loadings from two questions designed to assess depressive symptoms (*Do you look forward to things?* and *Are you interested in doing things and meeting people?*). These questions were reversed on the PTOS-ID; with high occurrence (i.e. ‘a lot’) scoring ‘0’ and low occurrence (i.e. ‘not at all’) scoring ‘3’. This may explain why they negatively load onto this factor. The final factor consists exclusively of items aimed at assessing anxiety and one aimed at assessing depressive symptoms (*Do you feel worthless?*). However, this item also loaded onto the positive well-being factor.

Principal Components Analysis with Depression Items Removed

The principal components analysis was rerun with all of the depression items removed. Again, a three-factor model emerged. These factors were identified as: a) anger, b) positive well-being, and c) anxiety. See Appendix H.

Reliability Analysis

Internal consistency analysis was run on the three factors identified in the principal components analysis. Emotional and Behavioural Discomfort (n = 173) and Anxiety (n = 174) had high reliability (α = 0.82 and α = 0.76, respectively). Two of the items (item 5 and item 15) in Positive Well-being were reversed for the analysis because they were negatively scored in the PTOS-ID (see Field, 2013). Positive Well-being (n = 167) was also found to have high reliability with α = 0.81.

Concurrent Validity

Concurrent validity was assessed via correlational analysis of Psychological Distress in the PTOS-ID (n = 172) and the Global Severity Index (GSI) of the BSI (n = 131). Psychological Distress in the PTOS-ID was defined as the mean score of all the items that contributed to the Emotional and Behavioural Discomfort and Anxiety factors. It was found to have high levels of internal consistency (α = 0.85). The GSI refers to the mean score of all the subscale scores from the BSI, and provides a global index of distress (Derogatis, 1993). A significant positive relationship was found between the Psychological Distress index of the PTOS-ID and the GSI of the BSI, $r = .85$, $p < .001$.

Table 3 Item-factor loading matrix for the PTOS-ID items

PTOS-ID Items	**1**	**2**	**3**
Do you have a bad temper?	0.750		
Do you feel like smashing things?	0.678		
Are you feeling annoyed?	0.677		
Are you feeling angry?	0.668		
Do you feel like hitting someone?	0.592		
Do you feel wound up or tense?	0.573		
Are you feeling sad?	0.558		
Do you think about death or dying?	0.490		
Are you sleeping more/less than usual?	0.458		
Are you eating more/less than usual?	0.384		
Can you show love for other people?		0.675	
Do you feel lovable?		0.659	
Do you feel like you can make friends?		0.643	
Can you tell people how you feel?		0.613	
Are you feeling happy?	-0.436	0.597	

Are you satisfied with life?	0.581	
Do you look forward to things?	-0.546	
Do you feel like you are a good person?	0.534	
Are you interested in doing things and meeting people?	-0.512	
Can you do things as well as other people?	0.447	-0.353
Can you stand up for yourself?	0.372	
Do you suddenly feel scared (about things)?		0.787
Do you feel frightened of things, places or activities?		0.773
Do you avoid some places or activities because you are frightened of them?		0.677
Do you feel worthless?	-0.379	0.472
Do you feel faint or dizzy?		0.433
Are you feeling anxious?		0.426

Discussion

The main purpose of this study was to validate a newly developed psychological therapy outcome measure designed for use with people with intellectual disabilities. Analysis of the PTOS-ID identified a 29-item, self-report measure which assesses: a) Emotional and Behavioural Discomfort, b) Positive Well-being, and c) Anxiety. The factors that were identified were unexpected in relation to the item pool and the five proposed scales of depression, anxiety, anger, interpersonal well-being and psychological well-being. However, all identified factors had a Cronbach's alpha of greater than 0.7 suggesting good levels of reliability. The sample size also satisfied statistical power for the exploratory analysis and the overall psychological distress score correlated significantly with the GSI of the BSI. Taken together, this suggests that the PTOS-ID has good levels of construct and concurrent validity.

The Emotional and Behavioural Discomfort factor contains a combination of items from the originally proposed anger and depression scales. It appears to measure mood disorders, including both affect and behavioural components. The Anxiety factor represents a general index of anxious symptomatology, including items assessing phobic anxiety, panic and general anxiety. The item assessing obsessive-compulsive anxiety failed to factor in the final model. Positive Well-being contains items assessing both interpersonal well-being and psychological well-being from the original item pool. This factor is an important addition provided by the PTOS-ID as there are currently no validated therapy outcome measures assessing positive well-being in people with intellectual disabilities. The PTOS-ID also provides a measure of psychological distress, combining the Emotional and Behavioural Discomfort, and Anxiety factors.

The findings from this study are consistent with the existing literature in this area. Kellett et al. (2004) assessed the factor structure of the BSI with people with ID. They found that items from the Depression, Interpersonal sensitivity, Somatization and Psychoticism scales on the original BSI contributed to the Depression scale identified in their analysis with people with intellectual disabilities. Interestingly, they also found that their Hostility factor contained items from the original Anxiety and Depression scales. Equally, the Anxiety factor contained an item from the Interpersonal sensitivity scale from the original analysis. These findings are similar to the present study, where items aimed at assessing symptoms of depression and anger factored together.

Sturmey et al. (1996) conducted a factor analysis of the Reiss Screen for Maladaptive Behaviours (Reiss, 1988) with 180 people with intellectual disabilities. They found that there was “no particular pattern of item loadings related to extra-personal and intra-personal maladaptive behaviours” (p. 289); this included items related to depression and anger. Equally Aman et al. (1986) found that the depression subscale of the Psychopathology Instrument for Mentally Retarded Adults (PIMRA, Matson et al., 1984) was distributed across several factors across the scale. The problem may not be with the scales but the belief that these groupings would work with people who have intellectual disabilities when they do not appear to do so. Thus it may be that the conceptualisation that mental health concerns in people with intellectual disabilities and the general population are the same is wrong.

One explanation for anger and depression items factoring together could be that feelings of anger and sadness are very closely linked for people with intellectual disabilities. Hollins and Sinason (2000) have identified that loss (both actuarial and of sense of self) is often experienced by people with intellectual disabilities and has to be addressed in therapy. Our responses to loss often include both feelings of sadness and anger (Kubler-Ross, 1969). People with intellectual disabilities are often prevented from working through their losses. For example, Hollins & Esterhuyzen (1997) report that roughly 50% do not attend the funeral of a deceased parent. Experiences like this may mean that emotions of anger and sadness are much more intertwined for people with intellectual disabilities because they are denied the opportunity to work through the losses they may experience.

Interestingly, there are emotional and behavioural components that contribute to the first factor. People often use unconscious strategies – defences – to help protect them from intolerable emotions (Shedler, 2006; 2010; Sinason, 1992). The defences used by people with intellectual disabilities are often found to be more externalising and active, such as 'acting out', 'denial', 'hypochondriasis' and 'dissociation' (Newman & Beail, 2010). This may explain why behavioural items such as '*Are you eating more/less than usual?*' and '*Do you feel like hitting someone?'* factor together; they are some of the active defensive processes used to protect against intolerable feeling of loss.

This makes further sense when one considers that the data used for analysis were often collected either during eligibility assessments or pre-therapy. These may be times when loss is experienced more strongly through a formal diagnosis of an intellectual disability or through an actual loss that has resulted in someone seeking out psychological therapies. It has been found that people with intellectual disabilities enter therapy with poorly assimilated problems (thoughts, feelings, memories etc.), meaning that they have less insight into their difficulties and use less helpful strategies to minimize the negative affect (Newman & Beail, 2002; 2005; Stiles et al., 1990). Again this may explain why feelings of anger and sadness factor with distressing behavioural strategies.

Another explanation for the factor structure of the PTOS-ID is that the items developed to measure depression were not measuring it at all. Interestingly, the depression items appear in all three factors. It may be that instead of measuring depression, these items are measuring negative affect (Watson & Clark, 1984). It has been argued that negative affect subsumes a number of negative mood states including anger and fear, and that low negative affect is experienced as calmness (Koch et al., 2013). This may explain why the items developed to measure depression positively load onto both anger and anxiety factors, and negative on to the positive well-being factor.

Analysis of the factor structure of the PTOSS-ID with the depression items removed revealed a three-factor structure, assessing: a) anger, b) positive well-being, and c) anxiety. This further supports the idea that the depression items may not have been measuring a discreet construct, but may have been assessing the negative affect associated with feeling angry or anxious.

Limitations

There are a number of limitations to the study which may have affected the results. Firstly, the use of the present tense in questions may have confused service users. For example, service users may have responded to the first question – *are you feeling anxious?* – in relation to their feelings about completing the PTOS-ID rather than how they were feeling in general. The administration instructions tried to prevent this by asking service users to respond with how they had “been feeling over the past week, including today.”. However, these may have been forgotten during completion of the PTOS-ID. This may have been prevented by including the timeframe before each question, but this would have made questions exceptionally long and may have confused service users further. For example, question one could have read *"Have you been feeling anxious over the past week, including today?"*.

Secondly, the use of supplementary questions may have affected the original meaning of the questions. For example, the question "*Do you think about death or dying?"* had the supplementary question of *"Do you think about suicide?"*. It could be argued that these questions measure slightly different things. So when a service user was asked the supplementary question, their response may have been in relation to something different to a service user who answered the original question. It may be that only allowing the use of one question for each item affects the responses given and the structure of the PTOS-ID.

Finally, it may have been helpful to involve service users in the development of the PTOS-ID. Although service users found the PTOS-ID to be easy to complete (Beail et al., 2012), it may have been helpful to involve them in the item development stage. This may have allowed for areas not identified by the research team to be assessed (e.g. quality of life) and improve the wording of the questions to ensure that they could be understood clearly by the client group.

Future Research

Fitzpatrick et al., (1998) have identified seven criteria for patient rated outcome measures. These are: a) reliability, b) validity, c) responsiveness (ability to detect change over time), d) acceptability (is the measure acceptable to service users?), e) feasibility (is the measure easy to administer and process?), f) precision (can the measure distinguish different populations?), and g) interpretability (are the scores interpretable? Are there benchmarks/cut-offs?).

Further psychometric assessments (reliability and validity analysis) are needed to assess the quality of the PTOS-ID. Data for test re-test reliability analyses are currently being collected. Re-assessing the face validity of the items, such as tense and use of supplementary questions may need to be done. Also, reframing the questions in the past tense - *“in the past week have you felt…?”* – may improve comprehension and aid completion.

Assessment of the responsiveness of the PTOS-ID is also needed. The present study has developed a valid and reliable measure for use with people with intellectual disabilities. Future research needs to explore the PTOS-ID’s ability to detect change over the course of therapy. However, it is recommended that the PTOS-ID be reformatted and the redundant items removed before further research assessing this is carried out.

The PTOS-ID has been used with people with ‘mild’ and ‘moderate’ intellectual disabilities. Data from greater numbers within each of these client groups may help identify if there are any differences in the psychometric properties between these groups. Involving service users in the redesign of any items (face validity) may improve the acceptability. Also, development of a form that can be completed by carers/relatives may extend the use of the PTOS-ID to people who do not have the verbal ability to complete it.

The PTOS-ID takes 10-15 minutes to administer, depending on the ability and cooperation of respondent. The available administration instructions will be revised to match the reformatting of the items.

Analysis to assess the ability of the PTOS-ID to discriminate between different populations (i.e. clinical and non-clinical) is needed. The current study could not assess this due to incomplete referral data and lack of clarity between service users who had been referred for an eligibility assessment and for psychological therapies. Finally, identification of benchmarks and scores that represent clinically significant change are needed.

Clinical Implications & Conclusions

The present study has developed a relatively short psychological therapy outcome measure for use with people with intellectual disabilities that can be easily administered in routine clinical practice. The measure assesses both indices of Psychological Distress (Emotional and Behavioural Discomfort and Anxiety) and Positive Well-being (interpersonal and psychological well-being). Although there are some limitations with the present study, there is now a platform from which future research aimed at assessing the impact of psychological therapies for people with intellectual disabilities can be conducted.

References

Aman, M. G., Watson, J. E., Singh, N. N., Turbott, S. H. & Wilsher, C. P. (1986). Psychometric and demographic characteristics of the psychopathology instrument for mentally retarded adults. *Pharmacology Bulletin, 22*, 1972-1976

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders (4th edn) (DSM-IV).*Washington, DC: American Psychiatric Association

Barkham, M. Hardy, G. E. & Mellor-Clark, J. (2010). *Developing and delivering practice-based evidence: A guide for the psychological therapies*. Chichester: Wiley-Blackwell.

Barkham, M., Hardy, G. E., & Startup, M. (1996). The inventory of interpersonal problems-32: A short version of the inventory of interpersonal problems. *British Journal of Clinical Psychology, 35,* 21–35

Beail, N. (2003). What works for people with mental retardation? Critical commentary on cognitive-behavioural and psychodynamic psychotherapy research. *Mental Retardation, 41*, 468-472

Beail, N., Jackson, T., Vlissides, N., Fay, V. & Williams, K. (2012). *Development of psychological therapy outcome scale for people who have intellectual disabilities (PTOS-ID).* Presented at the Seattle Club Conference, Kendal, UK

Beail, N., Warden, S., Morsley, K. and Newman, D. W. (2005). Naturalistic evaluation of the effectiveness of psychodynamic psychotherapy with adults with intellectual disabilities. *Journal of Applied Research in Intellectual Disabilities, 18,* 245-251.

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,* 893–897.

Beck A. T., Steer R. A. & Brown G. K. (1996). *Beck depression inventory – 2nd edition Manual*. San Antonio: The Psychological Corporation

Brooks, M. & Davies, S. (2007). Pathways to participatory research in developing a tool to measure feelings. *British Journal of Learning Disabilities, 36,* 128-133.

Brooks, M., Davies, S. & Twigg, E. (2013). A measure for feelings – Using inclusive research to develop a tool for evaluating psychological therapy (clinical outcomes in routine evaluation – learning disability). *British Journal of Learning Disabilities, 41,* 320-329.

Cuthill, F. M., Espie, C. A. & Cooper, S-A. (2003). Development and psychometric properties of the Glasgow depression scale for people with a learning disability: Individual and carer supplement versions. *British Journal of Psychiatry, 182,* 347-353

Davis C., Kellett S. & Beail, N. (2009). Utility of the rosenberg self-esteem scale. *American Journal on Intellectual and Development Disabilities, 114(3),* 172-178

Department of Health (2010). *Liberating the NHS: Transparency in outcomes: A framework for the NHS.* London: Department of Health.

Derogatis, L.R. (1983) *SCL-9OR; Administration, scoring and procedures: Manual 11*. Towson, MD: Clinical Psychometric Research

Derogatis, L.R. (1993). *Brief symptom inventory: Administration scoring and procedures manual* (3rd ed.). Minneapolis: National Computer Systems.

Evans, C., Connell, C., Barkham, M., Margison, F., McGrath, G., Mellor-Clark, J. & Audin, K. (2002). Towards a standardized brief outcome measure: Psychometric properties and utility of the CORE-OM. *British Journal of Psychiatry, 180,* 51-60

Field, A. P. (2013). *Discovering statistics using IBM SPSS statistics: (and sex and drugs and rock ‘n’ roll).* London: Sage.

Fitzpatrick, R., Davey, C., Buxton, M. J. & Jones, D. R. (1998). Evaluating patient-based outcome measures for use in clinical trials. *Health Technology Assessment, 2(14),* 1-74

Floyd, F. & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment, 7,* 286-299

Green D. & Latchford, G. (2012). *Maximising the benefits of psychotherapy: A practice-based evidence approach.* Chichester: Wiley.

Hatton, C. (2002). Psychosocial interventions for adults with intellectual disabilities and mental health problems. *Journal of Mental Health, 11,* 357-373

Hermans, H., Femke, N. F., Pas, H., & Evenhuis, H. M. (2012). Feasibility, reliability and validity of the Dutch translation of the anxiety, depression and mood scale in older adults with intellectual disabilities. *Research in Developmental Disabilities, 33(2),* 315-323.

Hollins, S. & Esterhuyzen, A. (1997). Bereavement and grief in adults with learning disabilities. *British Journal of Psychiatry, 170,* 497-501.

Hollins, S. & Sinason, V. (2000). Psychotherapy, learning disabilities and trauma: New perspectives. *British Journal of Psychiatry, 176,* 32-36.

Kellet, S., Beail, N., & Newman, D. W. (2005). Measuring interpersonal problems in people with mental retardation. *American Journal on Mental Retardation, 110(2)*, 136-144.

Kellett, S. C., Beail, N., Newman, D. W & Frankish, P. (2003). Utility of the brief symptom inventory in the assessment of psychological distress. *Journal of Applied Research in Intellectual Disabilities, 16*, 127-134.

Kellett, S. C., Beail, N., Newman, D. W & Hawes, A. (2004). The factor structure of the brief symptom inventory: Intellectual disability evidence. *Clinical Psychology & Psychotherapy, 11*, 275-281.

Kellett, S. C., Beail, N., Newman, D. W. & Mosley, E. (1999). Indexing psychological distress in people with an intellectual disability: Use of the symptom checklist-90-r. *Journal of Applied Research in Intellectual Disabilities, 12(4)*, 323-334.

Koch, A. S. Forgas, J. & P. Matovic, D. (2013). Can negative mood improve your conversation? Affective influences on conforming to Grice's communication norms. *European Journal of Social Psychology, 43 (5),* 326–334

Kubler-Ross, E. (1969). *On death and dying.* London: Tavistock Publications Limited

Lindsay W. R. (2007). The Beck depression inventory II and the Beck anxiety inventory in people with intellectual disabilities: Factor analyses and group data. *Journal of Applied Research in Intellectual Disabilities, 20,* 401-408.

Lindsay, W. R. & Mitchie, A. M. (1988). Adaptation of the Zung self-rating scale for anxiety for people with a mental handicap. *Journal of Mental Deficiency Research, 32,* 485-490

Marshall, K. & Willoughby-Booth, S. (2007). Modifying the clinical outcomes in routine evaluation measure for use with people who have a learning disability. *British Journal of Learning Disabilities, 35*, 107-112.

Martorell, A., Tsakanikos, E., Pereda, A., Gutierrez-Recacha, P., Bouras, N. & Ayosu-Mateos, J. L. (2009). Mental health in adults with mild and moderate intellectual disabilities. *American Journal of Mental Retardation, 113,* 307-318.

Matson, J. L., Kazdin, A. E. & Senatore, V. (1984). Psychometric properties of the psychopathology instrument for mentally retarded adults. *Applied Research in Mental Retardation, 5*, 881 889.

Mindham, J. & Espie, C. A. (2003). Glasgow anxiety scale for people with an intellectual disability (GAS-ID): Development and psychometric properties of a new measure for use with people with mild intellectual disability. *Journal of Intellectual Disability Research, 47(1),* 22-30

Newman, D. W. & Beail, N. (2002). Monitoring change in psychotherapy with people with intellectual disabilities: The application of the assimilation of problematic experiences scale. *Journal of Applied Research in Intellectual Disabilities, 15,* 48-60

Newman, D. W. & Beail, N. (2005). Analysis of assimilation during psychotherapy with people who have mental retardation. *American Journal of Mental Retardation, 110(5),* 359-365.

Newman, D. W. & Beail, N. (2010). An exploratory study of the defence mechanisms used in psychotherapy by adults who have intellectual disabilities. *Journal of Intellectual Disability Research, 54(7),* 579-583.

Newman, D. W., Kellett, S. C. & Beail, N. (2003). From research and development to practice based evidence: Clinical governance initiatives in a service for adults with mild intellectual disabilities. *Journal of Intellectual Disability Research, 47,* 68-74

Nicoll, M., Beail, N. & Saxon, D. (2013). Cognitive behavioural treatment for anger in adults with intellectual disabilities: A systematic review and meta-analysis. *Journal of applied Research in Intellectual Disabilities, 26,* 47-62

Novaco, R. W. & Taylor, J. L. (2004). Assessment of anger and aggression in male offenders with developmental disabilities. *Psychological Assessment, 16,* 42-50

Prout T. H. & Nowak-Drabik K. M. (2003) Psychotherapy with persons who have intellectual disabilities: an evaluation of effectiveness. *American Journal on Intellectual Disabilities, 108,* 82–93

Reiss S. (1988). *Reiss screen for maladaptive behaviour*. Worthington, OH: IDS.

Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.

Shedler, J. (2006). *That was then, this is now: Psychoanalytic psychotherapy for the rest of us*. Retrieved from http://psychsystems.net/shedler.html

Shedler, J. (2010). The efficacy of psychodynamic psychotherapy. *American Psychologist, 65(2),* 98-109

Sinason, V. (1992). *Mental handicap and the human condition: New approaches from the Tavistock.* London: Free Association Books

Skelly, A. (2011). Outcomes measurement in routine practice for learning disability psychology services. *Clinical Psychology & People with Learning Disabilities, 9(4),* 2-6.

Skelly, A. & D'Antonio, M. L. (2008). Factor structure of the HoNOS-LD: Further evidence of its validity and use as a generic outcome measure. *Clinical Psychology & People with Learning Disabilities, 6,* 3, 3-7.

Stiles, W. B., Elliott, R., Llewelyn, S. P., Firth-Cozens, J. A., Margison, F. R., Shapiro, D. A. & Hardy, G. (1990). Assimilation of problematic experiences by clients in psychotherapy. *Psychotherapy, 27(3),* 411-420.

Sturmey, P., Jamieson, J., Burcham, J. M., Shaw, B. & Bertman, L. (1996). The factor structure of the Reiss screen for maladaptive behaviours in institutional and community populations. *Research in Developmental Disabilities, 17,* 285-291.

Taylor, J. L., Lindsay, W. & Willner, P. (2008). CBT for people with intellectual disabilities: emerging evidence, cognitive ability and IQ effects. *Behavioural & Cognitive Psychotherapy, 36,* 723-733

Tennant, R., Hiller, L., Fishwick, R., Platt, S., Joseph, S., Weich, S., Parkinson, J., Secker, J. & Stewart-Brown, S. (2007). The Warwick-Edinburgh mental well-being scale (WEMWBS): Development and UK validation. *Health and Quality of Life Outcomes, 5,* 1-13

Thornicroft, G. (2006). *Shunned: Discrimination against people with mental illness.* Oxford: Oxford University Press

Tinsley, H. E . A. & Tinsley, D. J. (1987). Uses of factor analysis in counselling psychology research. *Journal of Counselling Psychology, 34(4),* 414-424

Watson, D. & Clark, L. A. (1984). Negative affectivity: The disposition to experience negative aversive emotional states. *Psychological Bulletin*, *96*, 465–490.

Weston, C., Elsworth, J. & Stacey, J. (2011). The development of a tool for the routine measurement of clinical outcomes in a learning disabilities psychology department: ridgeway routine outcome measure (RROM). *Clinical Psychology & People with Learning Disabilities, 9(2),* 7-13.

Willner, P. (2005). The effectiveness of psychotherapeutic interventions for people with learning disabilities: A critical overview. *Journal of Intellectual Disability Research, 49,* 75-85

Whitehouse, R. M., Tudway, J. A., Look, R. & Kroese, S. (2006). Adapting individual psychotherapy for adults with intellectual disabilities: A comparative review of the cognitive behavioural and psychodynamic literature. *Journal of Applied Research in Intellectual Disabilities, 19,* 55-65