The University Of Sheffield.

Investigating Psychological Symptoms and Patient Report Outcome Measures (PROMs) for People with Intellectual Disabilities: Validating and Abbreviating the Outcomes and Wellbeing Distress Scale for Adults with ID (OWLS-ID).

Emily Kerry

A thesis submitted in partial fulfilment of the requirements for the award of Doctorate in Clinical Psychology (DClinPsy)

Clinical and Applied Psychology Unit

Department of Psychology

The University of Sheffield

May 2024

Declaration

This thesis has not been submitted to any other degree or institution.

Structure and Word Counts

Abstracts

Lay Summary	288
Literature review	341
Empirical study	362

Part I: Literature Review

Excluding references and tables	8000
Including references and tables	12,757

Part II: Empirical Study

Excluding references and tables	7931
Including references and tables	13,075

Total Word Count

Excluding references and tables	15,931
Including references and tables	25,832

Lay Summary

Individuals with intellectual disabilities (ID) are more likely to experience common mental health problems, including depression. As such, understanding the factors related to depression in this population is important to inform assessment and treatment. Psychological therapies are a recommended approach to treating depression. Patient reported outcome measures (PROMs) are tools which can assess symptoms and evaluate psychological therapies.

Part I of this thesis identified factors in the literature which are associated with depression in populations with ID. Research in depression and ID is significantly less than that of the general population. The literature often debates whether depression presents differently in this clinical group. As such, it would be helpful to understand what factors are associated with depression to support the development of targeted interventions. Part I also considered which PROMs were used to measure depressive symptoms in these studies. This highlighted a wide range of measures, few of which were specifically developed for those with ID.

Part II investigated the reliability and validity of an existing measure of psychological distress in people with ID, the Psychological Therapies Outcome Scale–ID (PTOS-ID-II). The measure was made specifically for this population and has been used in many clinical and research settings. However, its validity and reliability had yet to be examined in its current form. Validation analysis identified two items that were unnecessary and subsequently removed. Overall, 27-items remained which showed evidence of reliability and validity. This new measure was named ‘Outcomes for Wellbeing and Distress Scale’ (OWLS-ID). Part II then considered the feasibility of creating a shorter version of this measure for quick administration in practice. A 10-item version of the measure was created which also showed evidence of reliability and validity. This was named the OWLS-Mini.

Acknowledgements

I would like to extend thanks to my supervisors, Dr Gregg Rawlings, Dr Nik Vlissides and Professor Nigel Beail for allowing me to be involved in this great project. Thank you for the support, and guidance which has been offered throughout. I would also like to thank Thomas Gourley and Georgia Moreland for your contribution to the project. It has been a pleasure working with all in the LD team.

I would like to thank my parents, John, and Alison for shaping me into who I am today. Instilling the confidence in me to walk an unfamiliar path and holding my hand throughout. Your unconditional love, support and friendship has been pivotal in all of my successes since being little. You continue to drive me to be the best person that I can be, modelled upon the hardworking, loving, thoughtful and reliable role models you have consistently been throughout my life.

To my extended family, Grandma, Charlotte, Melissa, Archie, Billy, Nellie, and Tallulah. Thank you for the distractions and the reminder that I am not just a doctorate student but a granddaughter, sister, and an auntie too. I would like to acknowledge my friends who have been incredibly patient and understanding whilst I have been more absent than I would ever wish to be.

Thank you to my fiancé Jake, who has supported me in this journey since the first conception of the idea of becoming a Clinical Psychologist almost 12 years ago. You gave me the belief that I could do it allowing the idea to become viable and encouraged me each and every step of the way. I am truly thankful for the time you have invested in helping my aspirations become a reality.

Finally, to the colleagues and friends I have shared the DClin experience with, thank you for the support. Special thanks to the soon to be awarded Dr L Whitaker, Dr J Isgar, Dr K Neal, and Dr C Wojnarowski. Thank you and well done!

Table of Contents

Declaration ii
Structure and Word Counts iii
Lay Summary iv
Acknowledgements v
Part I: Literature Review
Abstract 2
Introduction 4
Method 8
Literature Search 8
Inclusion and Exclusion Criteria 9
Quality Appraisal 10
Data Extraction and Analysis 11
Results 12
Systematic Search 12
Summary of Studies 13
Participants 13
Measures of Depression 14
Quality Appraisal 16
Narrative Synthesis 21
Discussion 33
Strengths and Limitations 35
Implications 35
Conclusion 37
References 37
Appendices 52
Appendix A: Prospero Registration 52
Appendix B: Narrative Summary of Quality Appraisal 65
Appendix C: Key Findings from Identified Studies 69
Appendix D: Thematic Map 71
Appendix E: Reference List of Inaccessible Articles 72
Appendix F: PRISMA checklist 73

Part II: Empirical Project

Abstract 77
Introduction 79
Method 87
Setting and Approvals 87
Sample 87
Measures 88
Data Analysis 89
Statistical Analysis 90
Patient and Public Involvement 94
Results: Stage 1-Validation 94
Step Six: Principal Components Analysis 96
Step Seven: Confirmatory Factor Analysis 101
Step Eight: Reliability and Validity Analysis 104
Results: Stage 2-Abbreviation 107
Validation (OWLS-Mini) 109
Discussion 113
Strengths and Limitations 115
Implications 117
Future Research 118
Conclusions 119
References 120
Appendices 133
Appendix A: PTOS-ID-II 133
Appendix B: Permissions 138
Appendix C: STROBE Checklist 142
Appendix D: ANOVA Results 144
Appendix E: Correlation Matrix 147
Appendix F: Summary of Factor Loadings 153
Appendix G: Histograms and Q-Q Plots 157
Appendix H: Correlational Analyses 163

This Page is Intentionally Left Blank

Part I: Literature Review

Correlates of Depression in Adults with an Intellectual Disability: A Systematic Review of Quantitative Studies.

Abstract

Objectives

Evidence demonstrates that up to 39% of individuals with intellectual disabilities (ID) may experience symptoms of depression. Examining factors associated with depression in this group may help to identify targets for prevention and management. The current systematic review investigated factors that have been examined in the literature related to self-reported symptoms of depression in adults with ID.

Method

A systematic search of four databases (Web of Science, PubMed, PsychInfo and Cochrane Library) was conducted in October 2023. Search terms relating to correlates (e.g., association, correlation and/or regression) of depression in people with ID formed the search strategy. Quality appraisal was conducted using the Mixed Method Appraisal Tool (MMAT). Results are narratively synthesised.

Results

The search yielded seventeen eligible papers. Factors associated with depression were synthesised into seven categories: psychological, social, environmental (life events), health-related, behavioural, treatment-related, and demographic factors. Ten different patient report outcome measures of depression were used, only two of which were developed specifically for use with populations with ID.

Conclusion

Research has tended to focus more on psychological and social risk factors of depression in this clinical group, when compared to environmental, health-related, behavioural, treatment-related, and demographic factors. The variety of self-report depression measures utilised suggests that there is not a clear consensus on measures of depressive symptoms in populations with ID.

Practitioner Points

- Identifying correlates of depression can support formulation of why adults with ID develop depression, as well as identify possible targets for treatments.
- The review found depression in people with ID to be associated with psychological factors such as cognitions, anxiety, hopelessness, and self-esteem.
- Social factors such as social support and changes in relationships were also found to be related to depression.
- Clinicians should be sensitive to such correlates in assessments to scaffold questions around particular risk areas, for example, paying greater attention to factors such as quality and frequency of social support.
- Clinicians should also use this information to target treatments, for example, supporting positive social connections to increase social support and/or target co-occurring anxiety.

Key words: 'Intellectual Disability' 'Depression', 'Correlates', 'Associates'

Introduction

The American Psychiatric Association (American Psychiatric Association [APA], 2013) defines depression as feelings of extreme sadness or despair which interferes with daily life and can lead to a variety of additional difficulties such as pain, changes to weight, disrupted sleep, and lethargy. The Diagnostic and Statistical Manual (DSM-V) states an individual must experience five or more of the following symptoms during a two week period to receive a diagnosis of depression; (a) depressed mood most of the day, (b) markedly diminished interest in almost all activities of the day, (c) significant change in appetite, (d) slowing down of thought/reduction of physical movement, (e) fatigue/loss of energy, (f) feelings of worthlessness, (g) diminished ability to concentrate and, (h) recurrent thoughts of death (APA, 2013).

A population-based study involving 17,152 people in the United Kingdom (UK) found the prevalence of depressive symptoms to range from 11.3% for mild symptoms and 3.3% for severe symptoms (Arias de la Torre et al., 2021). More widely, the World Health Organisation (WHO, 2023) estimate around 280 million people in the world have depression; this makes depression one of the most experienced mental health disorders.

Depression and Intellectual Disability (ID)

ID is a neurodevelopmental condition characterised by significant difficulties in cognitive and social/adaptive functioning, which are present before the age of 18 (APA, 2013). ID can be categorised based on IQ scores as mild (50-70), moderate (35-50), severe (20-35), or profound (<20; Boat et al., 2015). Individuals with an ID are at a greater risk of experiencing mental health difficulties than the general population, including depression (Cooper et al., 2007). However, the evidence relating to the prevalence rate for depression in people with ID is unclear, with studies citing rates from 2.2% to 39% (Deb et al., 2001; McGillivray & McCabe, 2007). Such variance may be explained by the way in which

prevalence is recorded, and differences in samples e.g., severity of disability, residential status, and age, or methodologies such as how depression has been examined or operationalised in this group (Scheirs et al., 2023). Indeed, it is only within the last 30-40 years that affective disorders in this clinical group have been acknowledged (Nezu et al., 1995). In the 1980's, researchers questioned whether those with an ID experienced affective disorders at all (Sovner & Hurley, 1983). Currently, researchers agree that depression is not only experienced by those with an ID, but similar to the general population, depression is the most common mental health difficulty experienced by this clinical group (Melville et al., 2023), with rates appearing higher than those reported in the general populace (Poindexter, 2006).

While current medical nosology provides a clear criterion on how to diagnose depression (APA, 2013), ambiguity remains around the presentation of depression in those with ID. Although there are symptoms of depression that present universally across populations, some researchers have suggested that depression may manifest in *atypical* ways in people with ID (McBrien, 2003). For instance, it has been proposed that self-injury, physical aggression, and verbal aggression may be 'behavioural equivalent' symptoms of depression in those with severe/profound ID (Jawed et al., 1993; Marston et al., 1997). Though, the evidence is contradictory (Tsouris, 2001; Tsouris et al., 2003) and some argue that a reliance on behavioural symptoms may lead to 'false positive' diagnoses (Sturmey et al., 2010; McBrien, 2003). This is problematic as accurate diagnoses is essential in providing appropriate treatment. As such, the decision to only focus on self-report measures of depression in the current review was made to gain a clearer perspective from the intended sample and reduce the heterogeneity of included studies. Indeed, there are informant-based measures of depression in adults with ID. We have discussed the implications of excluding this data in the discussion.

Treatment of Depression

There are a number of treatment options for adults with depression in the general population outlined in the National Institute for Clinical Excellence guidance (NICE, 2022). NICE guidance identifies both psychotherapeutic and pharmaceutical interventions as effective treatments. Therapeutic recommendations include cognitive behavioural therapy (CBT), behaviour activation, short term psychodynamic psychotherapy, counselling, and interpersonal psychotherapy. Until the late 1990's, psychological therapies were not considered suitable for individuals with ID (Bender, 1993), with Bender terming this phenomenon 'the unoffered chair' providing a prominent metaphor of the attitudes of psychotherapy towards individuals with ID (Greenhill, 2011). Reviews on the efficacy of psychotherapy during this time failed to reference adults with ID at all (Roth et al., 1996; Bergin & Garfield, 1994; Irvine & Beail, 2016). However, throughout the 1990's-2000's, much has been done to understand the mental health needs of those with ID. Resultantly, the evidence of the use of psychotherapy in ID settings is growing (Beail, 2016). A proportion of the evidence-base was included in a recent meta-analysis involving 19 studies and 698 people with ID. Analyses identified a small and significant effect size (g=0.43, CI[0.47, 1.32], p=<0.05) for the efficacy of psychological therapies for those with ID presenting with a range of mental health conditions. However, the authors conclude that methodological limitations are preventing strong conclusions about effectiveness being made (Tapp et al., 2023).

Predictors of Depression

There is a plethora of studies focusing on predictors of depression in the general population, which have identified significant correlates including social isolation, poor health, poor mobility, higher rates of family burden, gender, abuse, lower educational attainment, substance use, age, weight, chronic disease, inactivity, and smoking (Handing et al., 2022; Peters et al., 2015; Tanaka et al., 2011). While some factors are likely to play a

similar role in adults with ID, it is possible that some factors may interact differently given the different profile and life experiences of an individual with ID. Understanding what factors increase an individuals' risk of depression could support prevention and inform treatment options. It may also help to identify what to examine when assessing depression in this clinical group. Indeed, while some authors have individually examined factors associated with symptoms of depression in adults with ID, we are unaware of systematic review on the topic.

Current Study

Thus far there have been substantial delays in the identification and treatment of depression within this population noted, combined with an ambiguity around symptomatic presentation of depression in people with ID (Bakkan, 2021). Clarity around the risk factors of depression in this population is important in supporting prevention and treatment. While the literature has been synthesised involving the general population (Gutierrez-Rojas et al., 2020), there is no recent review to the authors knowledge that explores the association between risk factors and the subsequent development of depression for individuals with ID.

The aim of the current paper is to systematically review the literature to identify and synthesise the factors associated with depression in populations with ID, which will be presented narratively. A narrative approach was taken as the data was too heterogeneous to group for quantitative analyses due to the lack of standardisation in the administration of measures meaning the studies were not sufficiently similar. As part of this aim, we intend to recognise the different methods and tools used for assessing depressive symptoms in people with ID by reviewing the measures utilised throughout the studies that address our initial aim. Finally, the quality of studies will be reviewed.

Method

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Tricco et al., 2018) was used to support the development and write up of this review. This review was pre-registered on PROSPERO (Appendix A; project ID CRD42023423013).

Literature Search

The PICO (Population, Intervention, Control and Outcomes) method was used to support the development of the search criteria (Richardson et al., 1995). A control condition was not required due to the correlational nature of the studies being explored, therefore, only Population, Intervention and Outcomes were used to develop the searches. The finalised search terms can be seen in Table 1.

A systematic literature search of Web of Science, PubMed, PsychInfo and Cochrane Library were conducted on the 27th of October 2023. There were no limits or restrictions on the dates published. Papers identified by the systematic search were exported to the online systematic review tool, Rayyan (Ouzzani et al., 2016). Duplicates were detected and removed, and the remaining papers were screened by the lead author (EK) based on titles and abstracts for relevance to the current review (see inclusion and exclusion criteria section). Next, full text articles that met the criteria were reviewed for inclusion, of which, 20% were randomly allocated to a second reviewer to assess against the eligibility criteria (GHR). Inter-rater agreement was 91.6%, which increased to 100% following discussion between reviewers. Articles that did not meet the criteria were excluded. Forward and backward searches took place of suitable articles to identify any additional relevant studies.

Table 1

Search Terms

PICO Criteria	Search Terms (***Boolean Operators***)
Population	intellectual disabilit* (*OR*) learning disabilit* (*OR*) mental retard* (*OR*) handicap*
	(*AND*)
Intervention	Depress*
	(*AND*)
Outcomes	Correlation (*OR*) correlate* (*OR*) associat* (*OR*) predict* (OR) regression (*OR*) determinant

Inclusion and Exclusion Criteria

The eligibility criteria were also based on the PICO (or in this case PIO) mnemonic. Studies were considered eligible for inclusion if they met the following criteria: (a) the majority of participants were aged 18 or above; (b) participants had a diagnosed ID (any severity of ID borderline-profound); (c) individuals presented with symptoms of depression, scored above the clinical threshold on measures of depression, or were reported as being diagnosed with depression; (d) measures used were self-report and; e) studies explored associations or correlates.

Studies that were not available in English language, included non-human samples, included participants that did not have a diagnosed ID, relied on informant-rated measures and grey literature were excluded. Grey literature was excluded due to the absence of rigorous quality control measures through the peer review process.

The definition of ID can vary across cultures, as such, all full text articles were checked for the approach to which ID was diagnosed. Those based on an IQ below 70 and adaptive behaviour below 70 were accepted as stated in the Diagnostic and Statistical Manual-Fifth Edition (DSM-V; APA, 2013). Studies in which there was no clear evidence

Table 2

Inclusion and Exclusion Criteria

PICO (PIO)	Include	Exclude
Participants	• Participants aged 18+ • Participants had a diagnosis of ID (any severity and/or comorbidity included) • Individuals presented with depression/above clinical threshold of a depression measure	• Non-human samples • Participants did not have a diagnosed ID • Incorrect definition of ID utilised
Intervention	• Self-report measures of depression used	• Informant based measure of depression used
Outcome	• Study exploring associations or correlations	• Grey literature • Not available in English language

that this definition of ID was utilised were excluded. Some individuals with a 'borderline' ID were included despite being slightly above this threshold due to still accessing ID health services. Borderline is indicative of individuals who scored near to the threshold of ID but did not surpass it, therefore, do not have a diagnosable ID yet comprise of a vulnerable group often also struggling with adaptive and social functioning (Wieland & Zitman, 2016).

Quality Appraisal

Quality appraisal was conducted (see results) to understand the quality of research in this area and how this interacts with the study's findings (Delavari et al., 2023). The Mixed Method Appraisal Tool (MMAT; Hong et al., 2018) was used to assess the quality of studies. The MMAT was selected as it is intended for use in systematic reviews regardless of study design; indeed, it was expected that data from studies utilising a range of methodologies may be identified i.e., clinical trials, longitudinal and cross-sectional studies. Authors of the MMAT discourage providing quantitative ratings; instead, it is advised that narrative summaries of each domain are conducted (Hong et al., 2018). The MMAT suggests a response criterion of 'yes', 'no' or 'can't tell' as a basis for each summary which informed

the allocation of a Red, Amber, Green rating of overall quality for that domain (red='no', amber='can't tell', green= 'yes'). Full narrative summaries can be seen in Appendix B. An overview of the quality appraisal can be seen in Table 4.

The primary researcher (EK) appraised 100% of the articles. An independent second reviewer reviewed 20% of the papers at random (KN). From this, an agreement of 88.6% was found which progressed to 100% after discussion of the discrepancies. The primary research then re-reviewed the remaining 80% of articles considering such discussions.

Data Extraction and Analysis

The lead researcher (EK) read each paper thoroughly and extracted information relevant to the review question. A data charting table was created on Microsoft Word and used to input the relevant data from each study. The data extracted included, the aim of the study, study design, location, type of analysis, sample size, level of ID, how ID was diagnosed, self-report measure of depression used, any other measures used, ethnicity, gender, age range, and living situation. The results of the associative study in relation to the review question and the main findings were summarised (Appendix C). Three types of analysis were commonly used throughout the studies: correlation, regression, and odds-ratio. Odds ratio is a measure of association based on likelihood (Tenny & Hoffman, 2023).

Each study was coded based on their key findings to aid the development of categorisation into themes. Themes were developed based on the frequency and saliency of codes as opposed to study characteristics (e.g., sample size) or quality appraisal ratings. This was completed by the lead researcher (EK) and a preliminary thematic map was shared with the research team. Discussion with the research team led to the finalised themes as seen in appendix D. These findings were narratively synthesised into emergent themes and sub-themes. Themes are reported in order of distinction, with those with the greatest salience reported first and more novel findings reported later. Salience was determined by a number of

factors such as the frequency of evidence for the finding and the methodology used. For example, correlational analysis with the addition of regression would be considered more salient making the finding more robust.

Data relating to the MMAT appraisal questions were also extracted and placed in a Microsoft Excel spreadsheet. This included data related to the sample, measures used, missing data, confounding factors, adherence to the intervention, randomisation processes, comparability of groups and the process of blinding.

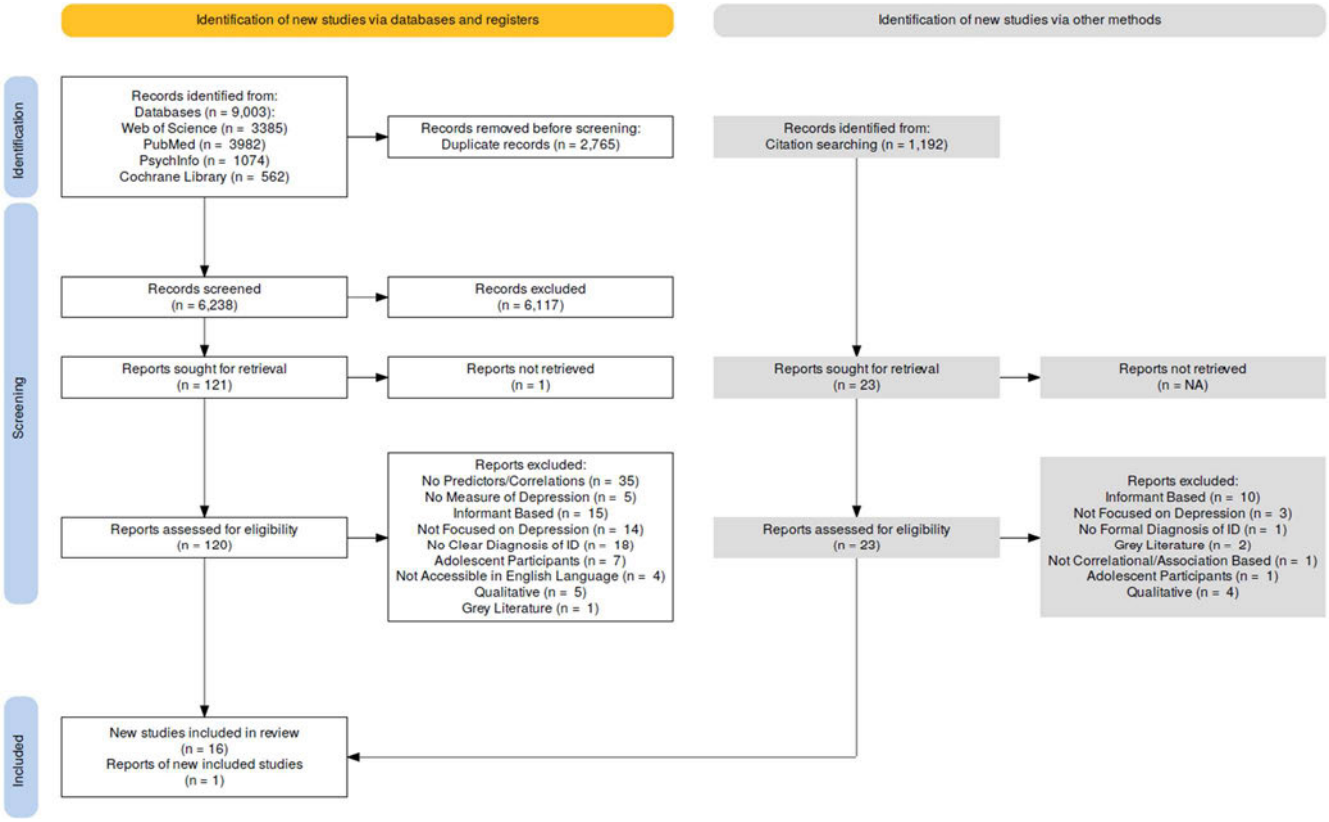
Results

Systematic Search

Overall, 9003 papers were identified by the systematic search. Duplicates were detected and removed (n=2765). The remaining 6238 papers were screened. From this, 121 articles remained which underwent full text review, of which 24 (20%) were randomly allocated to a second reviewer. One paper could not be retrieved despite two attempts to contact the author via email (See Appendix E). In total, 104 texts failed to meet the eligibility criteria. The remaining 16 texts were included in the systematic review. Identification of new studies by other methods included forward and backward searches, this yielded one new study.

Figure 1

A PRISMA Flow Diagram (Moher et al., 2009) Outlining the Screening Process



Summary of Studies

The studies reviewed mostly used correlational analyses (n=10), with a smaller number applying regression analyses (n=6), odds-ratio (n=3) and chi-square (n=1). Overall, seven studies took place in the United States of America (USA), five in the UK, two in Ireland, two in Australia and, one across both USA & Canada. An overview of the studies can be seen in table 5. Key findings are summarised in appendix C.

Participants

The total number of participants across all the included studies was 2093 of which 26 were considered 'borderline', which is not within the formal ID range. The sample included 992 (47.4%) individuals reported as female and 935 (44.7%) males. One study did not report demographic data on gender; therefore, the gender of 166 people is unknown (Collishaw et al., 2004). Ethnicity was recorded in nine of the 17 studies; 36.6% (n=572) reported to be

Caucasian/white, 1.9% (n=40) African American, 1.5% (n=32) Black, 0.43% (n=9) Hispanic, 0.1% (n=2) Native American, and 0.1% (n=2) Asian. Ethnicity was not available for the majority of participants (68.5%, n=1433).

In total, 16 studies referred to the participant's level of ID: 40.9% (n=856) were reported to have a mild ID, 20.3% (n=424) moderate ID, 8.3% (n=174) severe/profound ID, and 1.3% (n=26) borderline ID. For the remaining participants (29.3%, n=613) level of ID was not specified. This may reflect our criteria stipulating symptoms of depression must be self-reported, which may be challenging for some individuals with more severe ID.

The living situation of participants was reported in 13 of the 17 studies: 33.4% (n=699) lived in a group home or supported accommodation, 16.2% (n=340) lived in a residential/institutional setting, 13.1% (n=274) of individuals lived alone or within the family home, and 0.6% (n=12) lived with a host family. Data was not available for all participants.

Measures of Depression

Ten different self-report measures of depression were used, which ranged from 18–53 items. The most commonly used self-report measure was the Glasgow Depression Scale (GDS-LD; Cuthill et al., 2003) which was administered in five studies (Austin et al., 2018; Bond et al., 2019; Bond et al., 2020; Hartley et al., 2008; Melville et al., 2023) to n=1274 (60.9% of the sample) individuals. This was followed by the Zung Depression Scale (Zung, 1965) which was used in three studies involving n=133 (6.4%) adults (Dagnan & Sandhu, 1999; Laman & Reiss, 1987; Reiss & Benson, 1985).

Other measures administered included the Brief Symptom Inventory (BSI; Derogatis, 1975) to 1.8% of the sample (n=38; Hulbert-Williams et al., 2011), the Birleson Depression Scale (BDS-S; Birleson 1981) to 6.2% (n=129; Hartley et al., 2009; Lunsky & Benson, 2001), The Reynolds Adolescent Depression Scales (RADS; Reynolds, 1987) to 9.4% (n=196; Laman & Reiss, 1985; McGillivaray et al., 2007), The Reynolds Child Depression

Scales (RCDS; Reynolds, 1989), to 2.2% (n=46; Glenn et al., 2003), Beck Depression Inventory (BDI-II; Beck, Steer & Brown, 1996) to 12.3 % (n=258; McGillivray & McCabe, 2007; Nezu et al 1995), The Self-Report Depression Questionnaire (SRDQ; Reynolds & Baker, 1988) to 3.5% (n=73, Esbensen & Benson, 2005), The Malaise Inventory (Grant et al., 1990; Rutter et al., 1970) to 4.1% (n=86; Collishaw et al., 2004), and the Psychopathology Inventory for Mentally Retarded Adults (PIMRA-ID; Sturmey & Ley, 1990) to 5.1% (n=107; Nezu et al., 1995). One study required participants to have a formal diagnosis of depression (Larson et al., 2011). Of these measures, only two were specifically developed for people with ID (PIMRA-ID and GDS-ID). Other studies opted to use child/adolescent measures such as the RADS or shortened questionnaires such as the BDS-S.

Table 3

Summary of Outcome Measures used Including Psychometric Properties Reported

Measure Name	Number of items	Psychometric Properties
Glasgow Depression Scale (GDS-LD)	20	Bond (2019) reported a correlation of .88 with the BDI-II and 'good' test re-test reliability (r=.97) in their sample. The recommended cut-off (>13) yields 96% sensitivity and 90% specificity in discriminating between those with and without depression (Bond et al., 2019; Bond et al., 2020; Cuthill et al., 2003). Internal consistency was shown to be between .90 and .97 (Austin, 2018; Bond et al., 2019; Hartley et al., 2008)
Zung Depression Scale	20	Test re-test reliability between .61 and .75 was reported (Dagnan, 1999; Reiss, 1985). Laman (1987) report concurrent validity in this population but no evidence of discriminant and factor validity.
Birleson Depression Scale (BDS-S)	18	Lunsky (2001) reports significant correlation with informant ratings.
Brief Symptom Inventory (BSI)	53	Hulbert-Williams (2011) Internal consistency between .63-.78 for each subscale when used with individuals with mild-borderline ID.
Reynolds Adolescent Depression Scale (RADS)	30	Laman (1987) reports internal consistency coefficient of .87. McGillivray (2007) reports good reliability and validity but not in current sample.
Reynolds Child Depression Scale (RCDS)	30	Glenn (2003) reports internal consistency .92.
Beck Depression Inventory (BDI-II)	21	Nezu (1995) reports previous evidence of correlation between self-report and informant reports of BDI, no evidence from current sample provided.

The Self Report Depression Questionnaire (SRDQ)	32	Esbensen (2005) reports internal consistency of the measure is high, test re-test reliability is moderate and criterion validity has been established, however, these are not founded in the current sample/population.
The Malaise Inventory	24	Collishaw (2004) showed internal consistency of .75-.87 for those with mild ID.
Psychopathology inventory for Mentally Retarded Adults (PIMRA-ID)	56	Nezu (1995) reports test re-test reliability estimates of .69. In addition to significant correlation with the Hamilton Rating Scale for Depression (Hamilton, 1960). Evidence in the current population is unclear.

Shading used to enhance clarity for reader.

Quality Appraisal

All 17 papers passed the two screening questions 'Are there clear research questions?' and 'Do the collected data allow to address the research questions?' of the MMAT. All studies were assessed using the relevant criteria from the MMAT: most of the studies (16) were cross sectional, therefore, the MMAT quantitative non-randomised appraisal questions were used. One study, Melville et al (2023), used a randomised control trial (RCT) and therefore the MMAT quantitative RCT appraisal questions were used (labelled 1a, 2a... etc).

All studies represented the population (item 1) with the presence of an ID diagnosis, however, some had small samples (<60) and others failed to show a range of ID severity (mild-profound) which may not be representative of the wider ID population, thus limiting generalisability. For item 2, 'Are measurements appropriate regarding both the outcome and intervention?', 12 studies lacked clarity in this area. Although the measures used were mostly standardised, the findings were weakened by the use of general population measures informally adapted for ID or child measures with limited psychometric evidence of being used in adults with ID. Studies which used ID specific measures or showed psychometric evidence for the tools used received higher appraisal ratings. Four studies performed poorly on item 3, 'Are there complete outcome data?' due to presence of missing data, whilst eight studies failed to comment on missing data at all. Studies performed weakest on item 4 'Are the confounders accounted for in the design and analysis?'; eight studies performed well on this by using comparison groups or matched controls, which can reduce extraneous

participant confounds; five studies performed poorly, failing to consider confounding variables in the design of the study. All studies appeared to deliver the research as intended (item 5) which was a common strength of the research.

Table 4

Results of MMAT Quality Appraisal

Primary Author (Year)	Screening questions		Quantitative Design	Q1	Q2	Q3	Q4	Q5
Austin (2018)	✓	✓	Cross sectional	Yes	Can't tell	Yes	Yes	Yes
Bond (2019)	✓	✓	Longitudinal Cross sectional	Yes	Can't tell	No	Can't tell	Yes
Bond (2020)	✓	✓	Longitudinal Cross sectional	Yes	Can't tell	No	Can't tell	Yes
Collishaw (2004)	✓	✓	Longitudinal Cross sectional	Yes	Can't tell	Yes	Yes	Yes
Dagnan (1999)	✓	✓	Cross sectional	No	Can't tell	Can't tell	No	Yes
Esbensen (2005)	✓	✓	Cross sectional	Yes	Can't tell	Can't tell	Yes	Yes
Glenn (2003)	✓	✓	Cross sectional	Can't tell	Can't tell	Can't tell	No	Yes
Hartley (2008)	✓	✓	Cross sectional	Yes	Yes	Can't tell	Yes	Yes
Hartley (2009)	✓	✓	Cross sectional	Yes	Can't tell	Can't tell	Yes	Yes
Hulbert Williams (2011)	✓	✓	Cross sectional	Can't tell	Yes	No	Yes	Yes
Laman (1987)	✓	✓	Cross sectional	Can't tell	Yes	Can't tell	Can't tell	Yes
Larson (2011)	✓	✓	Cross sectional	No	Can't tell	Yes	No	Yes
Lunsky 2001)	✓	✓	Longitudinal T1 vs T2	No	Can't tell	Yes	No	Yes
McGillivray (2007)	✓	✓	Cross sectional	Yes	Can't tell	Can't tell	No	Yes
Nezu (1995)	✓	✓	Cross sectional	Yes	Can't tell	Can't tell	Yes	Yes
Reiss (1985)	✓	✓	Cross sectional	Can't tell	Yes	No	Yes	Yes
Author (year)	**Screening Questions**		**Design**	**1a**	**2a**	**3a**	**4a**	**5a**
Melville (2023)	✓	✓	RCT	Can't tell	Can't tell	Can't tell	No	Yes

MMAT Quantitative nonrandomised questions (1) Are the participants representative of the target population? (2) Are measurements appropriate regarding both the outcome and intervention? (3) Are there complete outcome data? (4) Are the confounders accounted for in the design and analysis? (5) During the study period, is the intervention administered as intended?

MMAT Quantitative randomised control trial questions (1a) Is randomisation appropriately performed? (2a) Are the groups comparable at baseline? (3a) Are there complete outcome data? (4a) Are outcome assessors blinded to the intervention provided? (5a) Did participants adhere to the assigned intervention?

Table 5

Data Charting Table Showing Study Characteristics

Primary author and year	Country	Aim	Type of analysis	N	Gender	Age (mean, std deviation, range)	Ethnicity	Living situation	Level of ID	Self-Report Measure of depression
Austin (2018)	Australia	Investigate factors which may predict anxiety and depression in young adults with ID.	Chi Square, Correlation and ANOVA	137	Female-27 Male-28	18-30	NR	Family/parents- 49 Alone- 3 With friends- 2 With partner- 1	Mild ID-35 Moderate-13 Borderline-7	Glasgow Depression Scale for people with a LD (GDS-LD)
Bond (2019)	Ireland	Examine the relationship of life events and mental ill health in the older ID population.	Correlational/Odds Ratio	598	Female-332 Male-266	59.21 (SD=8.83)	NR	Independent/family-91 Community group home-239 Institutional-268	Mild ID-132 Moderate-255 Severe/profound-163	Glasgow Depression Scale for people with LD (GDS-LD)
Bond (2020)	Ireland	To determine biopsychosocial factors associated with depression and anxiety in older adults with ID.	Correlational/Odds Ratio	291	Female-170 Male-121	<50=32 50-64=187 65+=72	NR	Community group home-141 Independent/family-78 Residential care 72	Mild ID=111 Moderate=138 Severe/profound=11	Glasgow Depression Scale for People with LD (GDS-LD)
Collishaw (2004)	UK	The extent to which adult socio-economic disadvantage and ill health contribute to the risk of affective disorder	Regression and Odds Ratio	86	NR	Cross section at age 43	NR	NR	Mild ID	The Malaise Inventory (Rutter et al., 1970; Grant et al., 1990)
Dagnan (1999)	UK	Explores the relationship between social comparison processes, self-esteem, and depression in people with intellectual disability	Correlational	43	Female-18 Male-25	35.1 (SD=10.2)	NR	living with family-24, Group home- 12, Independently-3, Foster family- 2, Not known-2	Mild-Moderate ID	Zung Depression Scale (Zung 1965)
Esbensen (2005)	USA	To examine cognitive variables related to the cognitive triad and hopelessness theories of depression	Correlational	73	Female-36 Male-37	40.6 (SD=12.2) 20-76	Caucasian-63	Lived at home with support-64	Mild-56 Moderate-9 Borderline-12 Not specified-3	The Self-Report Depression Questionnaire (SRDQ; Reynolds & Baker, 1988)
Glenn (2003)	USA	Explore the relationship between depression and anxiety in persons with mental retardation	Regression model	46	Female-21 Male-25	36.41 (SD=9.19) 21-59	Caucasian-39 African American-7	Supervised apartments-46	Mild ID-30 Moderate-9 Borderline-7 Mean IQ-66.1 (SD=8.53)	The Reynolds Child Depression Scale (RCDS; Reynolds, 1989)

Primary author and year	Country	Aim	Type of analysis	N	Gender	Age (mean, std deviation, range)	Ethnicity	Living situation	Level of ID	Self-Report Measure of depression
Hartley (2008)	USA	Examined the relationship between excessive reassurance seeking, negative and rejecting social interactions and depression in adults with mild ID.	Regression	87	Female-45 Male-42	40.1 (SD=14.1)	Caucasian-76 Hispanic-3 African American-4 Native American-2 Asian-2	Alone/with roommate-17 Group home- 64 Family/host family- 6	Mild ID Mean IQ 62.9 (SD=8.8) Range-55-70	The Glasgow Depression Scale for people with Learning Disability (GDS; Cuthill et al. 2003)
Hartley (2009)	USA	Investigated whether stressful social interactions, negative attributions and maladaptive coping maintain depression in ID populations as seen in the general population.	Correlational	52	23-Female 24-Male	42.62 (SD=12.87)	Caucasian-42 Non-Caucasian-5	Group home-31 Family/host family-4 Alone or with friend/partner-12	Mild IQ Mean IQ 62.94 (SD=5.64)	The Birleson Depressive Short Form Self-Rating Scale (Birleson, 1981)
Hulbert-Williams (2011)	UK	Aimed to replicate the reported relationship between life-events and psychological problems in people with intellectual disabilities, using self-report data.	Correlational	38	Females-12 Males-26	40.7 (SD=9.8) 18-59	NR	Lived alone-6 With family-15 Supported accommodation-14 With partner-1 Refused to answer-2	NR	Brief Symptom Inventory (BSI; Derogatis, 1975)
Laman (1987)	USA	To build upon previous research exploring social support and ID.	Correlational	45	Female-22 Male-23	NR	White-24 Black-18 Hispanic-3	NR	Mild ID	The Zung Self-Rating Depression Inventory (Zung, 1965); The Reynolds Adolescent Depression Scale (Reynolds, 1989)
Larson (2011)	UK	Investigate self-reported attachment styles in ID and whether there is a relationship between attachment, challenging behaviour and mental health difficulties	Chi square	56	Female-27 Male-29	NR	NR	NR	NR	Diagnosis of depression

Primary author and year	Country	Aim	Type of analysis	N	Gender	Age (mean, std deviation, range)	Ethnicity	Living situation	Level of ID	Self-Report Measure of depression
Lunsky (2001)	USA & Canada	To explore whether social support and social strain were associated with quality of life, symptoms of depression and somatic complaints 6 months later	Regression	77	Female-43 Male-41	38 20-65	White-66 African American-11	Lived alone-22 Live with 1 other-23 Lived with 2 others- 31	Mild ID	The Birleson Depressive Short Form Self-Rating Scale (BDS-S; Birleson 1981)
McGillivray (2007)	Australia	To determine the presentation and risk factors for depression in adults with mild/moderate ID	Regression	151	Female-68 Male-83	36.17 (SD=10.59) 19-68	NR	Living with parents- 59 Supported accommodation-40 Independent shared house-19 Living alone-18 Living with partner- 11 Living with siblings-4	Mild-Moderated ID	Beck Depression Inventory (BDI-II; Beck et al., 1996) and The Reynolds Adolescent Depression Scale (RADS; Reynolds 1987)
Melville (2023)	UK	This study takes an exploratory approach to examine whether pre-treatment variables are predictors and/or moderators of outcomes experienced by adults with IDs and depression treated with behavioural activation or guided self-help	Regression	161	Female- 85 Male-76	Group 1-40.3 (SD=11.7) Group 2-40.1 (SD=12)	White-156 Other-3 Unknown-2	Less than daily support- 49 Daily support- 112	Mild – Moderate ID	The Glasgow Depression Scale (GDS-LD; Cuthill et al. 2003)
Nezu (1995)	USA	Aimed to investigate the role of cognitive factors in depression in ID	Correlational	107	Female-43 Male-64	35.8 18-71	Caucasian-76 African American-29 Hispanic- 2	NR	Mild ID	Beck Depression Inventory (BDI-R), the Psychopathology Inventory for Mentally Retarded Adults (PIMRA-D)
Reiss (1985)	USA	To explore psychosocial correlates of depression in ID	Correlational	45	Female-20 Male-25	NR	White-30 Black-14 Hispanic-1	NR	Mild ID	The Zung Self-Rating Depression Inventory (Zung, 1965)

NR=Not reported, SD=Standard deviation, N=Number of participants, UK=United Kingdom, USA= United States of America, IQ=Intelligence Quotient.

Narrative Synthesis

To address the primary review question, the factors associated with depression in adults with ID were synthesised. Based on preliminary stages of the analysis, it was anticipated that correlates could be grouped into a biopsychosocial model and a deductive approach was taken. However, the rigidity of this approach meant that themes were not appropriately captured. For example, correlates relating to challenging behaviour did not neatly fit into any category as the cause of such behaviour could be argued to be biological, psychological or social. Therefore, an inductive approach was used to work more closely with the data. The extracted data was coded by the lead author and grouped into themes based on frequency e.g., the higher the frequency of a correlate and the wider range of methodologies used made for a more salient theme. Discussion with the research team led to the iterative final thematic map which can be seen in Appendix D. Seven themes were identified: social, psychological, life-events, behavioural, health, treatment, and demographic factors. Themes relating to psychological and social factors had the most evidence.

Theme One: Psychological Factors

Cognitions

A number of psychological concepts were researched, with the most cited being cognitive factors (e.g., automatic thoughts). Nezu (1995) considered the correlation between depression scores on the BDI-R with reports of negative automatic thoughts as measured by the Automatic Thoughts Questionnaire (ATQ; Hollon & Kendall, 1980) and found a significant positive correlation (r=.61, *p*=<.001). This suggests that as negative automatic thoughts increase so does depression scores and vice versa.

Glenn et al., (2003) also conducted a correlational study and found similar results utilising a different self-report measure of depression. They found that self-report symptoms of depression in the RCDS significantly correlated with automatic thoughts as rated by the

ATQ (r=.89, *p*=<.05). This was also replicated with a different measure of automatic thoughts, the cognitions checklist (Beck et al., 1987) which was slightly amended for use in this population. They found scores on the cognitions checklist significantly correlated with the score of depression as measured by the RCDS (r=.76, *p*=<.05) suggesting reliability of this finding. This was supported further by Esbensen & Benson (2005) who found a significant correlation between automatic thoughts as measured by the ATQ and depression scores on the SRDQ (r=.75, *p*=<.01). Esbensen was interested in exploring correlations related to Beck's (1967) Cognitive Triad Theory of Depression which suggests that individuals with depression have a negative view of themselves, the world, and the future (Beck, 1967, 1970, 1976). To explore this, the authors used the Cognitive Triad Inventory for Children (Kaslow et al., 1992) to investigate these three areas. A significant correlation was found between scores on the cognitive triad measure and score of depression (r=.48, *p*=<.01).

McGillivray and McCabe (2007) investigated cognitive factors particularly negative automatic thoughts further using regression analysis in a sample of 151 adults with mild-moderate ID. They developed a regression model which accounted for 58.1% of the variance in depression scores on the BDI-II with negative automatic thoughts being found to have the greatest significant impact on the regression model (β=.55). Other significant correlates were identified such as self-esteem and frequency of social support which will be expanded upon in later sections.

Although, none of these studies used an ID specific measure of depression which was considered a weakness, the finding has been replicated numerous times and with a relatively large sample across all four studies (n=337).

Anxiety

Other psychological factors explored included anxiety. Glenn et al., (2003) was interested in the interaction between depression and anxiety in people with ID and found a

significant correlation between self-reported depression on the RCDS and the Beck Anxiety Inventory (Beck & Steer, 1991; r=.74, *p*=<.001) in a sample of 46 adults with ID. However, of these, seven individuals were 'borderline' which is not considered a formal ID diagnosis, thus weakening the quality and generalisability of the findings.

Austin et al., (2018) investigated the relationship between anxiety and depression further in young adults aged 18-30 years with ID. To do this, spearman's correlations between GDS-LD scores and scores on the Glasgow Anxiety Scale for people with ID (GAS-ID; Mindham & Espie, 2003) were conducted. It was found that self-reported depression scores significantly correlated with scores on the GAS-ID (rho=.63, *p*=<.001).

Finally, Melville et al., (2023) created a regression model to explore associates of CBT-based treatments for depression in adults with ID. The model included severity of anxiety symptoms as well as other factors which will be discussed later. Greater severity of anxiety symptoms at baseline was associated with poorer outcomes on the GDS-LD at 12 months.

Two studies (Austin et al., 2018; Glenn et al., 2003) included a small number of individuals with borderline ID in the analysis (n=14), which may impact the findings as these individuals do not have a formal diagnosis of ID. Austin et al., (2018) strengthen this finding through replication with ID specific measures of both anxiety and depression which have more evidence of reliability and validity in this population.

Hopelessness

Esbensen & Benson (2005) examined cognitive variables related to the two most common theories of depression: 1) Beck's Cognitive Triad Theory as previously mentioned and 2) The Learned Hopelessness Theory which suggests that depression stems from negative expectations about outcomes and feelings of helplessness in the ability to change these outcomes (Abramson et al., 1978, 1989). Esbensen and Benson completed correlational

analysis of self-reported depression scores on the SRDQ, and measures of hopelessness as measured by The Hopelessness Scale for Children (Kazdin et al., 1986), observing a significant correlation (r=.26, *p*=<.05). This supported earlier findings by Nezu et al., (1995) who found that self-reported depression in the BDI-R was significantly correlated with scores on the Hopelessness Scale for Children (r=.36, *p*=<.001). Esbensen and Benson's sample of 73 adults included 12 individuals considered 'borderline' thus weakening the quality of the findings. Nezu's research strengthens this finding due to the relatively large sample (n=107) of individuals with mild ID.

Self-Esteem

Dagnan & Sandhu (1999) conducted a correlational study which found scores on the Zung Depression Scale negatively correlated with self-esteem on the Rosenberg Self-Esteem Scale (Rosenberg et al., 1989; r=-.39, *p*=<.01). In other words, as depression scores go up, self-esteem goes down and vice versa. This was further supported by Dagnan & Sandhu (1999), who found a significant positive correlation between negative self-esteem and scores on the ZDS (r=.41, *p*=<.01) showing that the higher score of negative self-esteem, the higher the score of depression. Similarly, Esbensen & Benson (2005) found a significant correlation between SRDQ scores and self-esteem measure by the Children's Self Concept Scale (Piers, 1969; r=.66, *p*=<.01).

As previously mentioned, McGillivray and McCabe (2007) conducted a regression analysis of cognitive associates of depression scores on the BDI-II which was found to account for 58.1% of the variance overall. This significant regression model also included self-esteem, as measured by the Rosenberg Self-Esteem scale which were found to have a significant impact on depression scores (β=-.18). Albeit to a lesser extent then quality of social support (β =.22) and automatic negative thoughts (β =.55) which had the greatest impact.

Replicability of self-esteem as a correlate in these studies spans the UK, USA and Australia which is promising, however, the use of child and general population measures limits the validity of the results.

Other Psychological Correlates

Other psychological correlates with less available evidence included coping, insight, stigmatisation, attributional styles, expectations, and self-reinforcement.

Self-reinforcement, which the authors describe as the act of giving or withholding self-rewards based on own self-evaluation (Nezu et al, 1995), was measured using the Frequency of Self Reinforcement Questionnaire (FSRQ; Heiby, 1983). They found a significant positive correlation between FSRQ scores and depression scores on the BDI-R (r=.58, *p*=<.01).

Esbensen and Benson (2005) were interested in the Learned Hopelessness Theory of Depression and measured attributional styles which they describe as a moderator between non-contingent stimuli and future expectations. They found a significant correlation between attributional styles as measured by the Children's Attributional Styles questionnaire (Seligman et al., 1984) and depression scores on the SRDQ (r=.33, p<.01).

Austin et al., (2018) found that unmet expectations of adult milestones and achievements was significantly correlated with depression scores on the GDS-LD (rho=.21, *p*=<.05) based on a measure of educational, vocational, familial, social communication and decision-making domains (Olin & Jansson, 2009; Soenens et al., 2007).They also explored both maladaptive coping (rho=.45) and reduced insight (rho=.-.5, *p*=<.01) finding significant correlations when compared to scores on the GDS-LD. Finally, Reiss and Benson (1985) found no significant correlation between depression scores on the ZDS and stigmatisation. Although, the authors note it would be premature to draw firm conclusions from one correlational study.

In summary, there are a number of psychological correlates of depression in people with ID such as anxiety, self-esteem, hopelessness, and prior expectations. Some of these findings were strengthened by inclusion of both correlational and regression analysis (anxiety, self-esteem, expectations) while others were strengthened by replication of correlational studies (negative automatic thoughts). Stigmatisation was the only psychological factor which was not found to significantly correlate with depression scores (Reiss & Benson, 1985). Common weaknesses of the studies include the use of child/adapted general population measures and inclusion of 'borderline' participants which may contaminate the findings.

Theme Two: Social Factors

Social Support

Different methods of assessing levels of social support in relation to depression was explored in six of the seventeen studies (Hartley & Maclean, 2009; Hulbert-Williams et al., 2011; Laman & Reiss, 1987; McGillivray et al., 2007; Nezu 1995; Reiss & Benson, 1985). Reiss and Benson (1985) aimed to explore the psychosocial correlates of depression in ID reporting scores on the ZDS were negatively correlated with levels of social support (r=-.41, *p*=<.001). This suggests as depression scores increase, social support reduces and vice versa. This was replicated by Laman & Reiss (1987) who found lower levels of social support correlated with ZDS scores (r=.41, *p*=<.01). Both Reiss and Laman's research was limited by small samples (n=45). That being said, Nezu (1995) replicated this finding in a sample of 107 individuals with mild ID, demonstrating significant positive correlation between negative social support and depression scores on the BDI-R (r=.29, *p*=<.01).

However, a correlational study by Hartley & Maclean (2009) did not replicate these findings in a sample of 52 adults with mild ID (r=.29, *p*=>.05). There are a number of differences between these studies which could have led to the contradictions. Firstly, research

by Laman (1987) and Reiss (1985) were conducted over 30 years ago suggesting systems of social support could have changed over the years contributing to contradictory results. Secondly, different self-report measures of depression where used, none of which were developed for individuals with ID calling into question the validity of the findings.

More recently, Hulbert-Williams et al., (2011) explored social factors as correlates of depression in ID in the UK and found high social support criticism (rho.5, p=<.001) and low social support closeness (rho=.56, p=<.05) significantly correlated with depression scores on the BSI. This provides further evidence of social factors playing a role in depression in ID.

McGillivray et al., (2007) expanded on these findings using regression analysis to explore the risk factors for depression in people with ID in an Australian sample. They found that perceived quality of social support (β=.22) and frequency of social support (β=.21) were significant associates of depression scores in the BDI-II contributing to the final model explaining 58.1% of the variance in BDI scores.

Weighted together, there is sufficient evidence to suggest social support is a factor relating to depression scores across nationalities (UK, USA, and Australia) despite contradictory findings by Hartley & Maclean (2009), with such findings being understood in the context of their different methodologies and limited sample size.

Relational

Bond et al., (2019) investigated a variety of scenarios and depression using odds-ratio. They found that changes in staff (OR=3.24, p=<.05) and changes to frequency of visits from loved ones (OR=4.32, p<.05) showed significant associations with depression on the GDS-LD. This suggests that changes to relational contact may play a role in depression scores. However, new residents joining the home, changing key workers, illness of a loved one, and changes in the service were not found to be significantly (>.05) associated with scores on the GDS-LD.

Other Social Factors

Social comparison, social disadvantage, and social strain were also identified as social correlates of depression in people with ID, albeit with lesser supporting evidence. Dagnan and Sandhu (1999) completed a correlational analysis of scores on the ZDS and scores on the Social Comparison Scale (Gilbert & Allen, 1994) and found a significant positive correlation (r=.50, p=<.001). However, McGillivray and McCabe (2007) included social comparison in their regression model and found social comparison to insufficiently contribute to the model. This discrepancy may be explained by the different methodologies. For example, the correlation between social comparison and depression scores found by Dagnan & Sandhu (1999) could be mediated by many factors not accounted for in this analysis.

Collishaw et al., (2004) aimed to explore the extent to which socio-economic disadvantage impacted affective disorders using a self-created 5-point tool. To do this, they used odds-ratio and regression methods. The authors found social disadvantage was significantly associated with higher levels of depression (OR=11.6, *p*=<.05).

Lunsky and Benson, (2001) completed regression analysis to assess whether social factors were associated with symptoms of depression in 77 adults with ID residing in supported living accommodation. They found social strain, defined as interpersonal interactions which lead to distress rather than enhance wellbeing (Rook, 1984), accounted for a significant proportion of the variance in depressive symptoms as rated on the BDS-S.

Finally, Larson et al., (2011) investigated self-rated attachment styles and depression using chi-square analysis. They found a non-significant association with formal diagnosis of depression and insecure-avoidant attachment (x^2[2]=5.59, *p*=.06).

In summary, social support showed the most evidence, with relational changes also featuring significantly. The strength of novel findings such as social comparison, social

strain, socio-economic disadvantage, and attachment styles were limited by the lack of replication.

Theme Three: Environmental

Life Events

Hulbert-Williams et al., (2011) explored the relationships between life events and psychological difficulties in people with ID using correlational analysis. They found total unique life events (rho=.52, *p*=<.001), negative life events (rho=.52, *p*=<.001), and weighted life events (which accounts for the frequency of events occurring; rho=.58, *p*=<.001) were significantly correlated with depression scores on the BSI.

Bond et al., (2019) used odds-ratio to assess associations between specific life events (such as death of a parent, major illness, experience of crime, moving home, or breakup from a relationship) and mental health in older adults with ID (mean age-59.2 years). They found the death of a relative (OR=11.6, *p*=<.01) and death of a friend (OR=2.86, *p*=<.05) were significantly associated with self-reported depression symptoms on the GDS-LD.

To conclude, there is evidence of life events relating to self-report depression scores. Although, this is limited to two studies and with a lack of exploration of other possible life events for example, trauma.

Theme Four: Health-Related Factors

Health

Three studies reported on health-related correlates of depression (Bond et al., 2019; Collishaw et al., 2004; Melville et al., 2023). Collishaw, Maughan & Pickles, (2004) explored whether ill-health influenced affective disorders using data from the National Child Development Study (Bynner et al., 2000) whereby participants were assessed regularly from birth to adulthood. Using odds-ratio, a cross-section of 86 adults with ID at aged 43 found

that poor health (OR=3.5, *p*=<.0001) was significantly associated with higher levels of depression on the Malaise Inventory.

On the contrary, Bond et al., (2019) conducted a similar analysis using odds-ratio and found that major illness/injury was not significantly associated with depression scores on the GDS-LD. Discrepancies could be present due to differences between the studies methodologies. For example. Bond et al., (2019) used a much larger sample (n=598) as well as an ID specific measure of depression symptoms (GDS-LD) thus strengthening the research. In contrast, Collishaw's findings were from a sample of 43-year-olds and data on living status, gender and ethnicity were not stated thus limiting the interpretations of external validity.

As mentioned above, Melville et al., (2023) created a regression model to consider the variance in BDI scores. The final significant model included hearing impairment with the overall model explaining 35.5% of the variance in GDS-LD scores at 12-months post-therapy.

In sum, there is mixed and limited evidence of the role of health in depression. Significant regression results were found for a specific health condition only (hearing impairment; Melville et al., 2023) whilst more general investigation of health and illness produced conflicting results (Bond et al., 2019; Collishaw et al., 2004).

Theme Five: Behavioural Factors

Behaviour

Of the seventeen studies, two identified behavioural correlates of depression (Bond et al., 2020; Hartley et al., 2008). Bond et al., (2020) explored biopsychosocial factors such as sleep, obesity, challenging behaviour, and unemployment and their relationship with depression and anxiety in older adults with ID (>40 years) using correlational analysis and odds-ratio. They found that the presence of challenging behaviour as measured by the

Behaviour Problems Inventory (Rojahn et al., 2001; OR=4.35, *p*=.04) was associated with self-reported depression scores on the GDS-LD. Due to the nature of associative analysis, it is not clear whether presence of challenging behaviour is the cause of higher depression scores or the effect of having higher depression scores.

Hartley, Lickel and Maclean, (2008) examined social behaviours and found reassurance seeking (β=.40, *p*=<.01) was associated with depression scores on the GDS-LD when inputted into a regression model. They also completed a correlational analysis on adaptive behaviour scores and the GDS-LD, which was not significant (*p*=.09). Together, there is novel evidence for relationships between challenging behaviour and reassurance seeking on depression scores.

Theme Six: Treatment Factors

Treatment-Related

Two studies explored treatment-related correlates including attendance to therapy and using pharmaceutical medication (Bond et al., 2020; Melville et al, 2023). Bond et al., (2020) found that taking mood stabilisers (OR=6.11, *p*=.02) was associated with depression scores on the GDS-LD, whilst, taking anxiolytics was not. In other words, those taking mood stabilisers were more likely to score higher on the depression measure. Based on the current analysis, it is not clear whether this is the cause or effect of the presence of depression.

Melville et al., (2023) examined pre-therapy variables to consider the factors related to outcome for adults with ID treated for depression. They found the greater percentage of therapy sessions attended were significantly related to more positive outcomes in depression treatments, demonstrating how those who attended more therapy had better outcomes. They also found that higher expectations of therapy as rated by a two-item Therapy Expectation Measure (Kilbane & Jahoda, 2011), were related to more positive outcomes for depression treatments. The final model included baseline depression symptoms, IQ, baseline anxiety

symptoms, hearing impairment, expectation of change and percentage of therapy sessions attended which explained 35.3% of the overall variance in GDS-LD scores at 12-months post therapy.

Theme Seven: Demographic Factors

Demographics

The relationship between age, gender, IQ, and ID status and depression has been explored by four studies (Bond et al., 2020; Collishaw et al., 2004; Hartley et al., 2008; Reiss & Benson, 1985). Reiss and Benson (1985) reported no significant difference in scores on the ZDS between men (r=.45, *p*=>.05) and women (r=.39, *p*=>.05). This was supported by Bond et al., (2020) who also found no significant correlation between age or gender and depression scores on the GDS-LD. Hartley et al., (2008) replicated this and found no significant association with age (r=-.08) or IQ (r=-.08). Whilst, Collishaw, Maughan & Pickles, (2004) found that the presence of a mild ID increased the risk of depression in both males (OR=3.92) and females (OR=2.38) compared to those without ID.

In sum, demographic factors, namely age, gender, and IQ were not found to be significant correlates of depression in people with ID. There is some evidence that the presence of mild ID may be a risk factor when compared to the general population (Collishaw et al., 2004); however, differences in IQ scores within an ID sample was not found to play a role in depression scores (Hartley et al., 2008).

Table 6

Crosstabulation Showing Each Theme and the Studies which Contributed to its Development.

✓-indicates significant evidence of association found
X-indicates no significant association was found

Study (Primary Author and Year)	Psychological Factors	Social Factors	Environmental (Life Events)	Health-Related	Behavioural Factors	Treatment-Related Factors	Demographic Factors
Austin (2018)	✓						
Bond (2019)		✓	✓	✓			
Bond (2020)					✓	✓	X
Collishaw (2004)		✓		✓			✓
Dagnan (1999)	✓	✓					
Esbensen (2005)	✓						
Glenn (2003)	✓						
Hartley (2008)					✓		X
Hartley (2009)		✓					
Hulbert-Williams (2011)		✓	✓				
Laman (1987)		✓					
Larson (2011)		✓					
Lunsky (2001)		✓					
McGillivray (2007)	✓	✓					
Melville (2023)	✓			✓		✓	
Nezu (1995)	✓	✓					
Reiss (1985)	X	✓					X

Discussion

The current review is the first systematic review investigating the factors related to self-reported depression in people with ID. The review summarises a range of correlates which were categorised as psychological, social, environmental, health-related, demographic, behavioural, and treatment-related. Psychological and social factors were most salient. The secondary aim was to report on the type of self-report measures of depression that have been used in this population.

Research has progressed from querying whether depression exists in this population (Sovner & Hurley, 1983) to considering the correlates of depression in this clinical group. Correlates of depression included self-concept, attributional styles, and unmet expectations.

Some correlates were strengthened by multiple investigations e.g. automatic-thoughts, life-events, and hopelessness. Though others were single event findings e.g. insight, social disadvantage and reassurance seeking. Demographic factors such as age, gender, and IQ were not found to relate to depression in those with ID. Social and psychological factors were found to be the most studied.

Research has found social isolation and poor health to be the biggest associates of depression in the general population (Handing et al., 2022). In the present review, social factors including frequency and quality of social support were found to be associated with depression scores. However, health-related factors were lesser explored despite higher prevalence of physical health difficulties in those with ID (Liao et al., 2021). Quality of life and trauma are two further examples of correlates which feature heavily in the general population literature (Cho et al., 2019; Mandelli et al., 2015) but did not feature in the current review suggesting more research is needed.

The current review focused on self-report measures of depression as research has shown that informant-rated questionnaires fail to capture the internal psychological experiences of the individual (Mileviciute & Hartley, 2015). They explored the correlation between self-report and informant-rated depressive symptoms in adults with ID and found that self-reported scores tended to show higher frequency of affective and cognitive symptoms, whereas staff reported more physical symptoms. This is problematic considering psychological factors were most apparent in this review. While we recognise informant measures can offer valuable insight, particularly when individuals are unable to complete the measures independently, self-reported measures provide unique psychological information which may otherwise be missed and an opportunity for the service-users experiences to be considered (Barkham et al., 2001).

The secondary aim of the review was to explore the patient-report outcome measures used. This highlighted the use of ten different self-report depression inventories; however, the GDS-LD and PIMRA-ID were the only two measures developed for use in this clinical population. Instead, authors opted to adapt general population measures (e.g., by removing questions and shortening the Likert responses) or utilise child measures. Furthermore, authors regularly failed to report on the reliability and validity of these measures in this specific population (see Table 3). This is problematic as research has found that outcome measures cannot be easily inducted into populations they were not developed for, with both the BSI and Rosenberg Self-Esteem (Rosenberg, 1965) measures found to factor differently with people who have ID to that of the general population (Davis et al., 2009; Kellet et al., 2004).

The GDS-LD showed the most evidence of psychometric properties in the intended sample. However, its psychometric properties were determined from an initial analysis of only 38 people with ID, highlighting issues of generalisability (Vlissides et al., 2016). Future research should consider the development of a validated measure made specifically for those with ID to provide more accurate reporting of psychological symptoms in this population.

Strengths and Limitations

Correlational research is limited by its inability to evidence cause and effect. For example, it is not clear whether the incidence of challenging behaviour is the cause of higher depression scores or an effect having depression. Similarly, its cross-sectional nature means that assessment of association is at a single point in time. Regression is advantageous as it provides more information on the relationship between factors. However, significant factors may not be inputted into the regression model and therefore missed. Therefore, the focus on associations is a limitation of this review.

Three studies reported a small amount of the sample to be 'borderline', which equates to 26 individuals in the overall sample. Borderline individuals may on occasion still access

ID services, hence, there inclusion in some studies. Nevertheless, caution should be taken when interpreting these results.

The data extraction and quality appraisal processes were improved by having a second reviewer thus reducing researcher bias. The quality appraisal stages encouraged the findings to be considered holistically in light of their strengths and weaknesses. The review was also strengthened by the systematic and comprehensive nature of the searches. However, the exclusion of non-English language studies may have reduced the scope of the findings.

Implications

The current review highlights a number of reliably significant associates of depression in ID which could support clinical practice as target areas for assessment and intervention. For example, the higher percentage of therapy sessions attended were predictive of more positive outcomes in depression treatments suggesting potential barriers to attending therapy could be explored in assessment and targeted for intervention. Furthermore, supporting positive social connections to increase social support, offering education/support to reduce co-occurring anxiety, encouraging joint work with other professionals (e.g., social workers and nurses) to holistically address health and social factors, and linking in with third sector/charity organisations to reduce socio-economic disadvantage. Furthermore, clinicians should be sensitive to such correlates in assessments and use this information to scaffold questions around particular risk areas, for example, paying greater attention to social factors such as quality and frequency of social support and psychological factors such as hopelessness, and self-esteem. Systemically, services could manage possible distress caused by systemic changes such as staff changes/changes to visitation and encourage a culture receptive to the risk factors of depression in those with ID.

Conclusion

There are a number of psychological and social correlates to depression in people with ID. Risk factors relating to the environment, health, behaviour, and treatment, were found with lesser support. Demographic factors were not found to be significant correlates. Clinical implications include use in practice to target systemic and individual interventions and to support assessment of depression in ID. Few of the measures used were intended for those with ID, which may reduce the validity of the results. Research exploring the causality of these factors would be beneficial. Plus, the development of a validated measure of distress for use with this clinical group.

References

***-identifies papers included in the analysis*

Abramson, L. Y., Metalsky, G. I., & Alloy, L. B. (1989). Hopelessness depression: A theory-based subtype of depression. *Psychological Review, 96*, 358–372. https://doi.org/10.1037/0033-295X.96.2.358

Abramson, L. Y., Seligman, M. E., & Teasdale, J. D. (1978). Learned helplessness in humans: critique and reformulation. *Journal of Abnormal Psychology*, *87*(1), 49–74. https://doi.org/10.1037/0021-843X.87.1.49

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

Arias de la Torre, J., Vilagut, G., Ronaldson, A., Dregan, A., Ricci-Cabello, I., Hatch, S. L., Serrano-Blanco, A., Valderas, J. M., Hotopf, M., & Alonso, J. (2021). Prevalence and age patterns of depression in the United Kingdom. A population-based study. *Journal of Affective Disorders*, *279*, 164–172. https://doi.org/10.1016/j.jad.2020.09.129

**Austin, K. L., Hunter, M., Gallagher, E., & Campbell, L. E. (2018). Depression and anxiety symptoms during the transition to early adulthood for people with intellectual disabilities. *Journal of Intellectual Disability Research*, *62*(5), 407–421. https://doi.org/10.1111/jir.12478

Bakken, T. L. (2021). Editorial for the special issue 'Intellectual Disabilities and Mental health: Assessment and Treatment.' *International Journal of Developmental Disabilities*, *67*(5), 307–309. https://doi.org/10.1080/20473869.2021.1971887

Barkham, M., Margison, F., Leach, C., Lucock, M., Mellor-Clark, J., Evans, C., Benson, L., Connell, J., Audin, K., & McGrath, G. (2001). Service profiling and outcomes benchmarking using the CORE-OM: Toward practice-based evidence in the

psychological therapies. *Journal of Consulting and Clinical Psychology*, *69*(2), 184–196. https://doi.org/10.1037//0022-006x.69.2.184

Beail, N. (2016). Psychological therapies and people who have intellectual disabilities: A new report from the Faculties for People who have Intellectual Disabilities of the British Psychological Society and the Royal College of Psychiatrists. *FPID Bulletin: The Bulletin of the Faculty for People with Intellectual Disabilities*, *14*(2), 5–7. https://doi.org/10.53841/bpsfpid.2016.14.2.5

Beck, A. T. (1967). *The Diagnosis and Management of the Emotional Disorders*. University of Pennsylvania Press, Philadelphia, PA.

Beck, A. T. (1970). *Depression: Causes and Treatment*. University of Pennsylvania Press, Philadelphia, PA.

Beck, A. T. (1976). *Cognitive Therapy and the Emotional Disorders*. International Universities Press, New York.

Beck, A. T., Brown, G., Steer, R. A., Eidelson, J. I., & Riskind, J. H. (1987). Differentiating anxiety and depression: a test of the cognitive content-specificity hypothesis. *Journal of Abnormal Psychology*, *96*(3), 179. https://doi.org/10.1037/0021-843X.96.3.179

Beck, A. T., & Steer, R. A. (1991). Relationship between the Beck anxiety inventory and the Hamilton anxiety rating scale with anxious outpatients. *Journal of Anxiety Disorders*, *5*(3), 213-223. https://doi.org/10.1016/0887-6185(91)90002-B

Beck, A. T., Steer, R. A., & Brown, G. (1996). *Beck Depression Inventory–II (BDI-II)* [Database record]. APA PsycTests. https://doi.org/10.1037/t00742-000

Bender, M. (1993). The unoffered chair: the history of therapeutic distain towards people with a learning difficulty. *Clinical Psychology Forum, 54*, 7 – 12. DOI:10.53841/bpscpf.1993.1.54.7

Bergin, A. E., & Garfield, S. L. (1994). *Handbook of psychotherapy and behavior change*. J. Wiley.

Birleson, P. (1981). The validity of depressive disorder in childhood and the development of a self-rating scale: a research report. *Journal of Child Psychology and Psychiatry*, *22*(1), 73-88. https://doi.org/10.1111/j.1469-7610.1981.tb00533.x

Boat, T. F., & Wu, J. T. (2015). *Mental disorders and disabilities among low-income children*. National Academies Press.

**Bond, L., Carroll, R., Mulryan, N., O'Dwyer, M., O'Connell, J., Monaghan, R., Sheerin, F., McCallion, P., & McCarron, M. (2020). Biopsychosocial factors associated with depression and anxiety in older adults with intellectual disability: results of the wave 3 Intellectual Disability Supplement to The Irish Longitudinal Study on Ageing. *Journal of Intellectual Disability Research*, *64*(5), 368–380. https://doi.org/10.1111/jir.12724

**Bond, L., Carroll, R., Mulryan, N., O'Dwyer, M., O'Connell, J., Monaghan, R., Sheerin, F., McCallion, P., & McCarron, M. (2019). The association of life events and mental ill health in older adults with intellectual disability: results of the wave 3 Intellectual Disability Supplement to The Irish Longitudinal Study on Ageing. *Journal of Intellectual Disability Research*, *63*(5), 454–465. https://doi.org/10.1111/jir.12595

Bynner, J., Butler, N., Ferri E., Shepherd, P., Smith, K. (2000). The design and conduct of the 1999 2000 surveys of the National Child Development Study and the 1970 British Cohort Study. Centre for Longitudinal Studies: Working Paper 1. user_guide_1999-2000_vol_1_design_and_conduct_of_the_1999-2000_surveys.pdf (ukdataservice.ac.uk)

Cho, Y., Lee, J. K., Kim, D. H., Park, J. H., Choi, M., Kim, H. J., Nam, M. J., Lee, K. U., Han, K., & Park, Y. G. (2019). Factors associated with quality of life in patients with depression: A nationwide population-based study. *PloS one*, *14*(7), e0219455. https://doi.org/10.1371/journal.pone.0219455

**Collishaw, S., Maughan, B., & Pickles, A. (2004). Affective problems in adults with mild learning disability: the roles of social disadvantage and ill health. *British Journal of Psychiatry*, *185*(4), 350–351. https://doi.org/10.1192/bjp.185.4.350

Cooper, S. A., Smiley, E., Morrison, J., Williamson, A., & Allan, L. (2007). An epidemiological investigation of affective disorders with a population-based cohort of 1023 adults with intellectual disabilities. *Psychological Medicine*, *37*(6), 873–882. https://doi.org/10.1017/s0033291707009968

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. *The British Journal of Psychiatry: The Journal of Mental Science*, *182*, 347–353. https://doi.org/10.1192/bjp.182.4.347

**Dagnan, D., & Sandhu, S. (1999). Social comparison, self-esteem and depression in people with intellectual disability. *Journal of Intellectual Disability Research*, *43*(5), 372–379. https://doi.org/10.1046/j.1365-2788.1999.043005372.x

Davis, C., Kellett, S., & Beail, N. (2009). Utility of the Rosenberg self-esteem scale. *American Journal on Intellectual and Developmental Disabilities*, *114*(3), 172–178. https://doi.org/10.1352/1944-7558-114.3.172

Deb, S., Thomas, M., & Bright, C. (2001). Mental disorder in adults with intellectual disability. 1: Prevalence of functional psychiatric illness among a community-based

population aged between 16 and 64 years. *Journal of Intellectual Disability Research*, *45*(6), 495-505. https://doi.org/10.1046/j.1365-2788.2001.00374.x

Delavari, S., Pourahmadi, M., & Barzkar, F. (2023). What Quality Assessment Tool Should I Use? A Practical Guide for Systematic Reviews Authors. *Iranian Journal of Medical Sciences*, *48*(3), 229–231. https://doi.org/10.30476/IJMS.2023.98401.3038

Derogatis, L. R. (1975). Brief Symptom Inventory. Baltimore, MD: Clinical Psychometric Research.

**Esbensen, A. J., & Benson, B. A. (2005). Cognitive variables and depressed mood in adults with intellectual disability. *Journal of Intellectual Disability Research*, *49*(7), 481–489. https://doi.org/10.1111/j.1365-2788.2005.00710.x

Gilbert, P. & Allen, S. (1994). Assertiveness, submissive behaviour and social comparison. *British Journal of Clinical Psychology, 33,* 295-306. https://doi.org/10.1111/j.2044-8260.1994.tb01125.x

**Glenn, E., Bihm, E. M., & Lammers, W. J. (2003). Depression, anxiety, and relevant cognitions in persons with mental retardation. *Journal of Autism and Developmental Disorders*, *33*(1), 69–76. https://doi.org/10.1023/a:1022282521625

Grant, G., Nolan, M., & Ellis, N. (1990). A reappraisal of the Malaise Inventory. *Social Psychiatry and Psychiatric Epidemiology: The International Journal for Research in Social and Genetic Epidemiology and Mental Health Services, 25*(4), 170–178. https://doi.org/10.1007/BF00782957

Greenhill, B. (2011). "They have behaviour, we have relationships?". *Reformulation*, Winter, pp.10-15. "They have behaviour, we have relationships?" (acat.me.uk)

Gutiérrez-Rojas, L., Porras-Segovia, A., Dunne, H., Andrade-González, N., & Cervilla, J. A. (2020). Prevalence and correlates of major depressive disorder: a systematic

review. *Revista brasileira de psiquiatria (Sao Paulo, Brazil : 1999)*, *42*(6), 657–672. https://doi.org/10.1590/1516-4446-2020-0650

Handing, E. P., Strobl, C., Jiao, Y., Feliciano, L., & Aichele, S. (2022). Predictors of depression among middle-aged and older men and women in Europe: A machine learning approach. *The Lancet Regional Health - Europe*, *18*, 100391. https://doi.org/10.1016/j.lanepe.2022.100391

**Hartley, S. L., Lickel, A. H., & MacLean Jr, W. E. (2008). Reassurance seeking and depression in adults with mild intellectual disability. *Journal of Intellectual Disability Research*, *52*(11), 917–929. https://doi.org/10.1111/j.1365-2788.2008.01126.x

**Hartley, S.L. and MacLean, W.E. (2009). Depression in Adults With Mild Intellectual Disability: Role of Stress, Attributions, and Coping. *American Journal on Intellectual and Developmental Disabilities*, [online] 114(3), pp.147–160. https://doi.org/10.1352/1944-7588-114.3.147.

Heiby, E. M. (1983). Assessment of frequency of self-reinforcement. *Journal of Personality and Social Psychology, 44,* 1304-1307. https://doi.org/10.1037/0022-3514.44.6.1304

Hollon, S. D., & Kendall, P. C. (1980). Cognitive self-statements in depression: Development of an automatic thoughts questionnaire. *Cognitive Therapy and Research*, *4*, 383-395. https://doi.org/10.1007/BF01178214

Hong, Q. N., Pluye, P., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M. P., Griffiths, F., Nicolau, B., O'Cathain, A., Rousseau, M. C., Vedel, I. Mixed Methods Appraisal Tool (MMAT), version 2018. Registration of Copyright (#1148552), Canadian Intellectual Property Office, Industry Canada.

**Hulbert-Williams, L., Hastings, R.P., Crowe, R. and Pemberton, J. (2011). Self-Reported Life Events, Social Support and Psychological Problems in Adults with Intellectual

Disabilities. *Journal of Applied Research in Intellectual Disabilities*, *24*(5), pp.427–436. https://doi.org/10.1111/j.1468-3148.2011.00624.x.

Irvine, M., & Beail, N. (2016). *Identifying and meeting the emotional and mental health needs of people who have intellectual disabilities through psychological therapie* (pp. 11–21). Faculties for Intellectual Disabilities of the Royal College of Psychiatrists and the Division of Clinical Psychology of the British Psychological Society.

Jawed S. H., Krishnan V. H., Prasher V. P. & Corbett J. A. (1993). Worsening of pica as a symptom of depressive illness in a person with severe mental handicap. *British Journal of Psychiatry 162*, 835–7. https://doi.org/10.1192/bjp.162.6.835

Kaslow, N. J., Stark, K. D., Printz, B., Livingston, R. & Tsai, S. L. (1992). Cognitive Triad Inventory for Children: development and relation to depression and anxiety. *Journal of Clinical Child Psychology 21*, 339–47. https://doi.org/10.1207/s15374424jccp2104_3

Kazdin, A. E., Rodgers, A., & Colbus, D. (1986). The hopelessness scale for children: psychometric characteristics and concurrent validity. *Journal of Consulting and Clinical Psychology*, *54*(2), 241. https://doi.org/10.1037//0022-006x.54.2.241

Kellett, S., Beail, N., Newman, D. W., & Hawes, A. (2004). The Factor Structure of the brief symptom inventory: Intellectual disability evidence. *Clinical Psychology & Psychotherapy*, *11*(4), 275–281. https://doi.org/10.1002/cpp.410

Kilbane, A. L., & Jahoda, A. (2011). Therapy expectations: preliminary exploration and measurement in adults with intellectual disabilities. *Journal of Applied Research in Intellectual Disabilities*, *24*(6), 528-542. https://doi.org/10.1111/j.1468-3148.2011.00630.x

**Laman, D.S. and Reiss, S. (1987). Social skill deficiencies associated with depressed mood of mentally retarded adults. *PubMed, 92*(2), pp.224–9. Social skill deficiencies associated with depressed mood of mentally retarded adults - PubMed (nih.gov)

**Larson, F. V., Alim, N., & Tsakanikos, E. (2011). Attachment style and mental health in adults with intellectual disability: self-reports and reports by carers. *Advances in Mental Health and Intellectual Disabilities*, *5*(3), 15–23. https://doi.org/10.1108/20441281111142585

Liao, P., Vajdic, C., Trollor, J., & Reppermund, S. (2021). Prevalence and incidence of physical health conditions in people with intellectual disability - a systematic review. *PloS one*, *16*(8), e0256294. https://doi.org/10.1371/journal.pone.0256294

**Lunsky, Y., & Benson, B. A. (2001). Association between perceived social support and strain, and positive and negative outcome for adults with mild intellectual disability. *Journal of Intellectual Disability Research*, *45*(2), 106–114. https://doi.org/10.1046/j.1365-2788.2001.00334.x

Mandelli, L., Petrelli, C., & Serretti, A. (2015). The role of specific early trauma in adult depression: A meta-analysis of published literature. Childhood trauma and adult depression. *European Psychiatry*, *30*(6), 665-680. https://doi.org/10.1016/j.eurpsy.2015.04.007

Marston, G. M., Perry, D. W., & Roy, A. (1997) Manifestations of depression in people with intellectual disabilities. *Journal of Intellectual Disability Research, 41*, 476-80. https://doi.org/10.1111/j.1365-2788.1997.tb00739.x

McBrien, J. A. (2003). Assessment and diagnosis of depression in people with intellectual disability. *Journal of Intellectual Disability Research*, *47*(1), 1–13. https://doi.org/10.1046/j.1365-2788.2003.00455.x

**McGillivray, J. A., & McCabe, M. P. (2007). Early detection of depression and associated risk factors in adults with mild/moderate intellectual disability. *Research in Developmental Disabilities*, *28*(1), 59–70. https://doi.org/10.1016/j.ridd.2005.11.001

**Melville, C. A., Hatton, C., Beer, E., Hastings, R. P., Cooper, S.-A., McMeekin, N., Dagnan, D., Appleton, K., Scott, K., Fulton, L., Jones, R. S. P., McConnachie, A., Zhang, R., Knight, R., Knowles, D., Williams, C., Briggs, A., and Jahoda, A. (2023). Predictors and moderators of the response of adults with intellectual disabilities and depression to behavioural activation and guided self-help therapies. *Journal of Intellectual Disability Research*, *67*, 986–1002. https://doi-org.sheffield.idm.oclc.org/10.1111/jir.13063.

Mileviciute, I., & Hartley, S. L. (2015). Self-reported versus informant-reported depressive symptoms in adults with mild intellectual disability. *Journal of Intellectual Disability research: JIDR*, *59*(2), 158–169. https://doi.org/10.1111/jir.12075

Mao, Y., Zhang, N., Liu, J., Zhu, B., He, R., & Wang, X. (2019). A systematic review of depression and anxiety in medical students in China. *BMC medical education*, *19*(1), 327. https://doi.org/10.1186/s12909-019-1744-2

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS medicine*, *6*(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097

**Nezu, C. M., Nezu, A. M., Rothenberg, J. L., DelliCarpini, L., & Groag, I. (1995). Depression in adults with mild mental retardation: Are cognitive variables involved? *Cognitive Therapy and Research*, *19*(2), 227–239. https://doi.org/10.1007/bf02229696

NICE. (2022, June 29). *Overview | Depression in adults: treatment and management | Guidance | NICE*. Www.nice.org.uk. https://www.nice.org.uk/guidance/ng222Odessa, FL: Psychological Assessment Resources, Inc

Olin, E., & Jansson, B. R. (2009). On the outskirts of normality: Young adults with disabilities, their belonging and strategies. *International Journal of Qualitative studies on Health and Well-being*, *4*(4), 256-266.

Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan—a Web and Mobile App for Systematic Reviews. *Systematic Reviews*, *5*(1). https://doi.org/10.1186/s13643-016-0384-4

Peters, A. T., Shankman, S. A., Deckersbach, T., & West, A. E. (2015). Predictors of first-episode unipolar major depression in individuals with and without sub-threshold depressive symptoms: A prospective, population-based study. *Psychiatry Research*, *230*(2), 150–156. https://doi.org/10.1016/j.psychres.2015.08.030

Piers, E. V. (1969). *The Piers-Harris children's self-concept scale:(The way I feel about myself)*. Nashville: Counselor Recordings and Tests.

Poindexter, A. R. (2006). Diagnosis of depression in people with developmental disabilities: Progress and problems. *International Review of Research in Mental Retardation*, *32*, 261-281. https://doi.org/10.1016/S0074-7750(06)32009-5

**Reiss, S. and Benson, B. A. (1985). Psychosocial correlates of depression in mentally retarded adults: I. Minimal social support and stigmatization. *PubMed*, 89(4), pp.331–7. Psychosocial correlates of depression in mentally retarded adults: I. Minimal social support and stigmatization - PubMed (nih.gov)

Reynolds, W. M. (1987). Reynolds Adolescent Depression Scale: Professional Manual. Odessa, FL:

Reynolds, W. M. (1989). *Reynolds child depression scale: Professional manual*. Lutz, FL: Psychological Assessment Resources.

Reynolds, W. M., & Baker, J. A. (1988). Assessment of depression in persons with mental retardation. *American journal of mental retardation: AJMR*, *93*(1), 93–103. Assessment of depression in persons with mental retardation - PubMed (nih.gov)

Reynolds, W.M. (1988). Suicidal Ideation Questionnaire: Professional manual. Odessa, FL.

Richardson, W. S., Wilson, M. C., Nishikawa, J., & Hayward, R. S. (1995). The well-built clinical question: a key to evidence-based decisions. *ACP Journal Club*, *123*(3), A12–A13. The well-built clinical question: a key to evidence-based decisions - PubMed (nih.gov)

Rojahn J., Matson J. L., Lott D., Esbensen A. J. & Smalls Y. (2001). The Behaviour Problems Inventory: an instrument for the assessment of self-injury, stereotyped behaviour, and aggression/destruction in individuals with developmental disabilities. *Journal of Autism and Developmental Disorders 31*, 577–88. https://doi.org/10.1023/a:1013299028321

Rosenberg, M. (1965). *Society and the Adolescent Self-image*. Princeton University Press, Princeton, NJ.

Roth, A., Fonagy, P., Parry, G., Target, M., et al. (1996). *What works for whom? A critical review of psychotherapy research.* Guilford Press.

Rutter, M., Tizard, J., & Kingsley Whitmore. (1970). *Education, health and behaviour*. Robert E. Krieger.

Scheirs, J. G. M., Muller, A., Manders, N. C. P., & van der Zanden, C. D. (2023). The prevalence and diagnosis of depression in people with mild or borderline intellectual

disability: Multiple instrument testing tells us more. *Journal of Mental Health Research in Intellectual Disabilities*, *16*(1), 54-66. https://doi.org/10.1080/19315864.2022.2029642

Seligman, M. E., Kaslow, N. J., Alloy, L. B., Peterson, C., Tanenbaum, R. L., & Abramson, L. Y. (1984). Attributional style and depressive symptoms among children. *Journal of Abnormal Psychology*, *93*(2), 235. https://doi.org/10.1037//0021-843x.93.2.235

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. https://doi.org/10.1038/bjc.1988.20

Soenens, B., Vansteenkiste, M., Lens, W., Luyckx, K., Goossens, L., Beyers, W., & Ryan, R. M. (2007). Conceptualizing parental autonomy support: adolescent perceptions of promotion of independence versus promotion of volitional functioning. *Developmental Psychology*, *43*(3), 633. https://doi.org/10.1037/0012-1649.43.3.633

Sovner, R., & Hurley, A. (1983). Do the mentally retarded suffer from affective illness? *Psychiatric Aspects of Mental Retardation Reviews, 5,* 16-21. doi:10.1001/archpsyc.1983.01790010063008

Sturmey, P., & Ley, T. (1990). The Psychopathology Instrument for Mentally Retarded Adults: Internal consistencies and relationship to behaviour problems. *The British Journal of Psychiatry, 156,* 428–430. https://doi.org/10.1192/bjp.156.3.428

Sturmey, P., Laud, R. B., Cooper, C. L., Matson, J. L., & Fodstad, J. C. (2010). Challenging behaviors should not be considered depressive equivalents in individuals with intellectual disabilities. II. A replication study. *Research in Developmental Disabilities*, *31*(5), 1002–1007. https://doi.org/10.1016/j.ridd.2010.04.018

Tanaka, H., Sasazawa, Y., Suzuki, S., Nakazawa, M., & Koyama, H. (2011). Health status and lifestyle factors as predictors of depression in middle-aged and elderly Japanese adults: A seven-year follow-up of the KOMO-ise cohort study. *BMC Psychiatry*, *11*(1). https://doi.org/10.1186/1471-244x-11-20

Tapp, K., Vereenooghe, L., Hewitt, O., Scripps, E., Gray, K. M., & Langdon, P. E. (2023). Psychological therapies for people with intellectual disabilities: An updated systematic review and meta-analysis. *Comprehensive Psychiatry*, *122*, 152372.

Tenny, S., & Hoffman, M. R. (2023). Odds Ratio. In *StatPearls*. StatPearls Publishing.

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., & Lewin, S. (2018). PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Annals of Internal Medicine*, *169*(7), 467–473. https://doi.org/10.7326/M18-0850

Tsiouris J. A. (2001) Diagnosis of depression in people with severe/profound intellectual disability. *Journal of Intellectual Disability Research 45*, 115-20. https://doi.org/10.1046/j.1365-2788.2001.00333.x

Tsiouris, J. A., Mann, R., Patti, P. J., & Sturmey, P. (2003). Challenging behaviours should not be considered as depressive equivalents in individuals with intellectual disability. *Journal of Intellectual Disability Research*, *47*(1), 14–21. https://doi.org/10.1046/j.1365-2788.2003.00456.x

Vlissides, N., Goulding, L., and Beail, N. (2016). *A systematic review of the outcome measures used in psychological therapies with adults with ID*, in Beail, N. (Ed),

Psychological Therapies and People Who Have Intellectual Disabilities, British Psychological Society, Leicester.

Wieland, J., & Zitman, F. G. (2016). It is time to bring borderline intellectual functioning back into the main fold of classification systems. *BJPsych Bulletin*, *40*(4), 204–206. https://doi.org/10.1192/pb.bp.115.051490

World Health Organization. (2023, March 31). *Depressive Disorder (depression)*. World Health Organization; World Health Organization. https://www.who.int/news-room/fact-sheets/detail/depression

Zung, W. W. K. (1965). A Self-Rating Depression Scale. *Archives of General Psychiatry*, *12*(1), 63. https://doi.org/10.1001/archpsyc.1965.01720310065008

Appendices

Appendix A

PROSPERO Registration

UNIVERSITY of York
Centre for Reviews and Dissemination

Systematic review

A list of fields that can be edited in an update can be found here

1. * Review title.

Give the title of the review in English

Correlates of Depression in Adults with Intellectual Disability: a Systematic Review

2. Original language title.

For reviews in languages other than English, give the title in the original language. This will be displayed with the English language title.

3. * Anticipated or actual start date.

Give the date the systematic review started or is expected to start.

31/05/2023

4. * Anticipated completion date.

Give the date by which the review is expected to be completed.
30/04/2024

5. * Stage of review at time of this submission.

This field uses answers to initial screening questions. It cannot be edited until after registration.

Tick the boxes to show which review tasks have been started and which have been completed.

Update this field each time any amendments are made to a published record.

The review has not yet started: Yes

Review stage	Started	Completed

Preliminary searches	No	No
Piloting of the study selection process	No	No
Formal screening of search results against eligibility criteria	No	No
Data extraction	No	No
Risk of bias (quality) assessment	No	No
Data analysis	No	No

Provide any other relevant information about the stage of the review here.

6. * Named contact.

The named contact is the guarantor for the accuracy of the information in the register record. This may be any member of the review team.

Emily Kerry

Email salutation (e.g. "Dr Smith" or "Joanne") for correspondence:

Emily

7. * Named contact email.

Give the electronic email address of the named contact.

ekerry1@sheffield.ac.uk

8. Named contact address

Give the full institutional/organisational postal address for the named contact.

1 Vicar Lane, Sheffield City Centre, Sheffield, S1 2LT

9. Named contact phone number.

Give the telephone number for the named contact, including international dialling code.

10. * Organisational affiliation of the review.

Full title of the organisational affiliations for this review and website address if available. This field may be completed as 'None' if the review is not affiliated to any organisation.

The University of Sheffield and South West Yorkshire Partnership NHS Trust

Organisation web address:

11. * Review team members and their organisational affiliations.

Give the personal details and the organisational affiliations of each member of the review team. Affiliation refers to groups or organisations to which review team members belong. **NOTE: email and country now MUST be entered for each person, unless you are amending a published record.**

Miss Emily Kerry. The University of Sheffield
Dr Nik Vlissides. South West Yorkshire Partnership Trust

Dr Gregg Rawlings. South West Yorkshire Partnership Trust
Professor Nigel Beail. The University of Sheffield

12. * Funding sources/sponsors.

Details of the individuals, organizations, groups, companies or other legal entities who have funded or sponsored the review. Not applicable

Grant number(s)

State the funder, grant or award number and the date of award

13. * Conflicts of interest.

List actual or perceived conflicts of interest (financial or academic).

None

14. Collaborators.

Give the name and affiliation of any individuals or organisations who are working on the review but who are not listed as review team members. **NOTE: email and country must be completed for each person, unless you are amending a published record.**

15. * Review question.

State the review question(s) clearly and precisely. It may be appropriate to break very broad questions down into a series of related more specific questions. Questions may be framed or refined using PI(E)COS or similar where relevant.

This review aims to answer the question, 'What are the common correlates of depression in adults with Intellectual Disability?'

Problem/Population (P)- Depression in people with Intellectual Disabilities

Intervention (I)- Not applicable

Comparison/Control (C)- Not applicable
Outcome (O)- Factors/correlates leading to the development of Depression

16. * Searches.

State the sources that will be searched (e.g. Medline). Give the search dates, and any restrictions (e.g. language or publication date). Do NOT enter the full search strategy (it may be provided as a link or attachment below.)

Electronic database searches on:

PsycINFO Web of Science

PubMed

Cochrane Library

No starting limit will be placed on the date range. All searches will be complete by April 2024. Concept one

intellectual disabilit* OR learning disabilit* OR mental retard* OR handicap*

AND

Concept two

Depress*

AND

Concept four
Correlation OR correlate* OR associat* OR predict* OR regression OR determinant

17. URL to search strategy.

Upload a file with your search strategy, or an example of a search strategy for a specific database, (including the keywords) in pdf or word format. In doing so you are consenting to the file being made publicly accessible. Or provide a URL or link to the strategy. Do NOT provide links to your search **results**.
Alternatively, upload your search strategy to CRD in pdf format. Please note that by doing so you are consenting to the file being made publicly accessible.
Do not make this file publicly available until the review is complete

18. * Condition or domain being studied.

Give a short description of the disease, condition or healthcare domain being studied in your systematic review.

Depression in adults with intellectual disabilities.

Intellectual Disability will be defined by those with an official diagnosis of Intellectual Disability based on the criteria in the Diagnostic and Statistical Manual (DSM-5). The DSM-5 highlights three criteria for the diagnosis of a learning disability. 1) significant impairment in intellectual functioning (as measured by an individually administered fully standardised IQ test. For example, a score less than 70 on the Wechsler Adult Intelligence Scale – Fourth Edition). 2) Significant impairment in adaptive behaviour/social functioning (via assessment of conceptual/communication abilities, activities of daily living, and social skills. 3) The disability originates before age 18.

Depression will be defined as those scoring above the clinical threshold for Depression or low mood on validated outcome measures of depression such as the Patient Health Questionnaire (PHQ-9) or accessing services or support for difficulties with Depression.

19 [2 changes]. * Participants/population.

Specify the participants or populations being studied in the review. The preferred format includes details of both inclusion and exclusion criteria.

Inclusion:

Any severity of ID. dults (Aged 18+)
Diagnosis of Intellectual Disability (May include comorbidity with other conditions)

Accessing services for treatment/support with symptoms of
Depression Score above the clinical threshold on measures of
Depression.
Self-report measures

Exclusion: Peer-reviewed articles

People without a diagnosis of ID.

People under the age of 18.

Individuals who do not meet the clinical threshold on measures of Depression.

Non-human samples (e.g. animal studies), Informant responses on measures, Grey literature/non-peer reviewed articles

20. * Intervention(s), exposure(s).

Give full and clear descriptions or definitions of the interventions or the exposures to be reviewed. The preferred format includes details of both inclusion and exclusion criteria.

Not applicable.

21. * Comparator(s)/control.

Where relevant, give details of the alternatives against which the intervention/exposure will be compared (e.g. another intervention or a non-exposed control group). The preferred format includes details of both inclusion and exclusion criteria.

Not applicable.

22 [1 change]. * Types of study to be included.

Give details of the study designs (e.g. RCT) that are eligible for inclusion in the review. The preferred format includes both inclusion and exclusion criteria. If there are no restrictions on the types of study, this should be stated.

Inclusion:
Cross Sectional Published in English.

Longitudinal

Trials

Any study using quantitative design
Correlational or Regression Analysis
Exclusion:

Qualitative Design

23. Context.

Give summary details of the setting or other relevant characteristics, which help define the inclusion or exclusion criteria.

Inclusion:

Any research considering the correlates of depression in adults with ID. For the purpose of this review, 'correlates' refers to any internal or external factor which is associated with depression in individuals with ID.

Exclusion: not accessible in
English Language.

24. * Main outcome(s).

Give the pre-specified main (most important) outcomes of the review, including details of how the outcome is defined and measured and when these measurement are made, if these are part of the review inclusion criteria.

Commonly cited predictors of Depression in Intellectual Disability Populations.

This review is interested in identifying the common predictors/factors/correlates which increase the likelihood of depression in ID populations. This includes any psychosocial factors (internal/external) which influence depression in this population.

Measures of effect

Please specify the effect measure(s) for you main outcome(s) e.g. relative risks, odds ratios, risk difference, and/or 'number needed to treat.

25. * Additional outcome(s).

List the pre-specified additional outcomes of the review, with a similar level of detail to that required for main outcomes. Where there are no additional outcomes please state 'None' or 'Not applicable' as appropriate to the review

In addition to the main outcome, this review is also interested in identifying which measures are currently being used to assess/monitor depression in ID populations at present.

Measures of effect

Please specify the effect measure(s) for you additional outcome(s) e.g. relative risks, odds ratios, risk difference, and/or 'number needed to treat.

26. * Data extraction (selection and coding).

Describe how studies will be selected for inclusion. State what data will be extracted or obtained. State how this will be done and recorded.

Study Selection:

One reviewer will apply eligibility criteria selecting studies for inclusion of the systematic review. The remaining research team will check these decisions. Any disagreements will be resolved through discussion as a team until a solution is found.

Data extraction:

Author names, sample characteristics, study design, measures used, scores on measures of depression will be extracted. The lead researcher will extract this data and work will be checked="checked" value="1" by the remaining research team.

The lead researcher will aim to contact authors for any unreported or missing data. If this is unsuccessful, a discussion with the wider research team to agree on how to manage missing data will take place if necessary.
Data extraction will be recorded in an excel spreadsheet on a password protected computer. The spreadsheet will be stored on a trust laptop in a shared file where the wider research team can access this.
This is protected by the trust VPN system and only those with relevant permissions can access this.

27. * Risk of bias (quality) assessment.

State which characteristics of the studies will be assessed and/or any formal risk of bias/quality assessment tools that will be used.

The Mixed Method Appraisal Tool (MMAT) will be used to formally assess risk of bias.

The lead researcher will review 100% of the studies. A second member of the research team will review a proportion of the studies (10-30%). Any disagreements will be discussed as a research time until a shared agreement is achieved.

28 [1 change]. * Strategy for data synthesis.

Describe the methods you plan to use to synthesise data. This **must not be generic text** but should be **specific to your review** and describe how the proposed approach will be applied to your data. If

metaanalysis is planned, describe the models to be used, methods to explore statistical heterogeneity, and software package to be used.

Demographic data, data on the measures used, the outcome (correlates of depression in LD) and risk of bias/quality appraisal will be extracted and a narrative approach will be used. Outcomes will be grouped based on the nature of the correlate. For example, demographics, psychological correlates, environmental correlates.

Because of the assumed heterogeneity in the studies identified (e.g. population, sample, variables used, outcome measures), a narrative approach would most likely be suitable. No minimum number of studies required given it is a narrative analysis.

As this is not a meta analytic approach caution will be taken around the conclusions drawn. However, we aim to report on the common correlates of depression in ID populations, prevalence of measures used and from this make recommendations for future research and consider the clinical implications. Clinical implications could include identifying correlates which could be more susceptible to change which in turn could direct clinical intervention or identify methods of 'upstream' psychosocial support.

29. * Analysis of subgroups or subsets.

State any planned investigation of ‘subgroups’. Be clear and specific about which type of study or participant will be included in each group or covariate investigated. State the planned analytic approach.

Not applicable.

30. * Type and method of review.

Select the type of review, review method and health area from the lists below.

Type of review

Cost effectiveness
No

Diagnostic
No

Epidemiologic
No

Individual patient data (IPD) meta-analysis
No

Intervention
No

Living systematic review
No

Meta-analysis
No

Methodology
No

Narrative synthesis
No

Network meta-analysis
No

Pre-clinical

No
Prevention
No

Prognostic
No

Prospective meta-analysis (PMA)
No

Review of reviews
No

Service delivery

Synthesis of qualitative studies
No

Systematic review
Yes

Other
No

Health area of the review

Alcohol/substance misuse/abuse
No

Blood and immune system
No

Cancer
No

Cardiovascular
No

Care of the elderly
No

Child health
No

Complementary therapies
No

COVID-19
No

Crime and justice
No

Dental
No

Digestive system
No

Ear, nose and throat
No

Education
No

Endocrine and metabolic disorders Eye disorders
No

General interest
No

Genetics
No

Health inequalities/health equity
No

Infections and infestations
No
International development
No
Mental health and behavioural conditions
Yes
Musculoskeletal
No
Neurological
No
Nursing
No
Obstetrics and gynaecology
No
Oral health
No
Palliative care
No
Perioperative care
No
Physiotherapy
No
Pregnancy and childbirth
No
Public health (including social determinants of health)
No
Rehabilitation

Respiratory disorders
No
Service delivery
No
Skin disorders
No
Social care
No
Surgery
No
Tropical Medicine
No
Urological
No
Wounds, injuries and accidents
No
Violence and abuse
No

31. Language.

Select each language individually to add it to the list below, use the bin icon to remove any added in error.

There is not an English language summary

32. * Country.

Select the country in which the review is being carried out. For multi-national collaborations select all the countries involved.

England

33. Other registration details.

Name any other organisation where the systematic review title or protocol is registered (e.g. Campbell, or The Joanna Briggs Institute) together with any unique identification number assigned by them. If extracted data will be stored and made available through a repository such as the Systematic Review Data Repository (SRDR), details and a link should be included here. If none, leave blank.

34. Reference and/or URL for published protocol.

If the protocol for this review is published provide details (authors, title and journal details, preferably in Vancouver format)

Add web link to the published protocol.
Or, upload your published protocol here in pdf format. Note that the upload will be publicly accessible.

No I do not make this file publicly available until the review is complete
Please note that the information required in the PROSPERO registration form must be completed in full even if access to a protocol is given.

35. Dissemination plans.

Do you intend to publish the review on completion?

Yes
Give brief details of plans for communicating review findings.?

36. Keywords.

Give words or phrases that best describe the review. Separate keywords with a semicolon or new line. Keywords help PROSPERO users find your review (keywords do not appear in the public record but are included in searches). Be as specific and precise as possible. Avoid acronyms and abbreviations unless these are in wide use.

37. Details of any existing review of the same topic by the same authors.

If you are registering an update of an existing review give details of the earlier versions and include a full bibliographic reference, if available.

38. * Current review status.

Update review status when the review is completed and when it is published.New registrations must be ongoing so this field is not editable for initial submission.

Please provide anticipated publication date

Review_Ongoing

39. Any additional information.

Provide any other information relevant to the registration of this review.

40. Details of final report/publication(s) or preprints if available.

Leave empty until publication details are available OR you have a link to a preprint (NOTE: this field is not editable for initial submission). List authors, title and journal details preferably in Vancouver format.

Give the link to the published review or preprint.

Appendix B

Narrative Summary of Quality Appraisal

		Screening Questions		Quantitative non-randomised?	Quality Appraisal				
Primary author and year	**Title**	Are there clear research questions?	Do the collected data allow to address the research questions?	Cohort/Cross sectional/case study?	Are the participants representative of the target population?	Are measurements appropriate regarding both the outcome and intervention (or exposure)?	Are there complete outcome data?	Are the confounders accounted for in the design and analysis?	During the study period, is the intervention administered (or exposure occurred) as intended?
Austin (2018)	Depression and anxiety symptoms during the transition to early adulthood for people with intellectual disabilities.	Yes	Yes	Cross Sectional	Large sample with ID ranging from Borderline-Moderate ID. Severe/profound Id not represented	some appropriate LD specific measures used at times. Some measures based on children and not ID.	Missing data clearly reported and replaced with means. Process clearly outlined.	age matched controls	Yes
Bond (2019)	The association of life events and mental ill health in older adults with intellectual disability: results of the wave 3 Intellectual Disability Supplement to The Irish Longitudinal Study on Ageing.	Yes	Yes	Longitudinal Cross Sectional	Yes, large sample taken from the Irish Longitudinal study of ID and ageing	Some appropriate ID specific measures used. One adapted measure of life events, identified by author as a limitation	Some missing data but clearly reported. Minimal data missing.	Wide range of factors included. Confounding factors not explicitly addressed	Yes
Bond (2020)	Biopsychosocial factors associated with depression and anxiety in older adults with intellectual disability: Results of the wave 3	Yes	Yes	Longitudinal Cross Sectional	Yes, large sample taken from the Irish Longitudinal study of ID and ageing. Range of ID.	Appropriate ID specific measures used where possible. Some standardised measures and some non-standardised questionnaire items used.	Some missing data but clearly reported.	Wide range of biopsychosocial correlates included. Confounds not clearly mentioned.	Yes

	Intellectual Disability Supplement to the Irish Longitudinal Study on Ageing.								
Collishaw (2004)	Affective problems in adults with mild learning disability: The roles of social disadvantage and ill health.	Yes	Yes	Longitudinal Cross Sectional	Large sample of mild ID only. This was the targeted population.	Measures are appropriate although not ID specific measures, reliability and validity for the population is cited.	Yes	Comparison group used	Yes
Dagnan (1999)	Social comparison, self-esteem and depression in people with intellectual disability.	Yes	Yes	Cross Sectional	Small sample. Mod-Mild ID only	Measures are appropriate but no ID specific measures are used	Presumed but not explicitly stated.	not discussed	Yes
Esbensen (2005)	Cognitive variables and depressed mood in adults with intellectual disability.	Yes	Yes	Cross Sectional	Large sample with range from borderline to moderate ID. Severe/profound not included	Some ID specific measures, some child specific measures. Outcomes relevant to question.	Presumed but not explicitly stated	Matched controls	Yes
Glenn (2003)	Depression, anxiety, and relevant cognitions in persons with mental retardation.	Yes	Yes	Cross Sectional	relatively small sample of Borderline-Moderate ID	Measures are appropriate but no ID specific measures are used	Presumed but not explicitly stated.	Confounds mentioned in discussion but not in design e.g. inclusion of borderline ID	Yes
Hartley (2008)	Reassurance seeking and depression in adults with mild intellectual disability.	Yes	Yes	Cross Sectional	Large sample. Mild ID only, target population.	Some ID specific measures, some adapted meaures. Reliability and validity on intended population cited.	Presumed but not explicity stated	Mediation analysis included	Yes
Hartley (2009)	Depression in adults with mild intellectual disability: role of stress, attributions, and coping.	Yes	Yes	Cross Sectional	Yes, DSM-5 criteria of ID, recruited from wide range of disability service (n=10). 52 participant's relatively small sample. IQ ranged from 55-70	Some standardised measures used. Some open-ended questions used.	Five individuals failed to pass the pre-test and were excluded. Remaining outcome data is assumed complete but not clearly reported.	Yes, ensure subject characteristics other than age, gender, and IQ could not account for findings. Controlled for adaptive behaviour and living status. Contribution of comorbid psychiatric diagnosis is unknown. Participants were matched on subject characteristics.	Yes
Hulbert-Williams (2011)	Self-Reported Life Events, Social Support and	Yes	Yes	Cross Sectional	Formal IQ testing is beyond scope of study. Diagnosis of ID is based	BLESID-SR (ID based measure) BSI (general population measure). Social	high level of missing data on SNM (over 50%).	accounts for total life events, negative life events and weighted life events.	Yes

	Psychological Problems in Adults with Intellectual Disabilities				on access to relevant services and carer reports. Appropriate but not excellent. 38 participants, small sample.	Network Map (general population). Collects relevant data, some general population and some ID specific measures. SNM was difficult to use in this population.	38 out of 41 completed the necessary measures.	Considers moderating effects of social support on life events.	
Laman (1987)	Social skill deficiencies associated with depressed mood of mentally retarded adults.	Yes	Yes	Cross Sectional	Yes,all participants met the American Assosciation on Mental Deficiency criteria for mental retardation. 45 participants, relatively small sample.	Yes, relevant for measuring outcomes and reference evidence for reliability and validity in this population	Some missing data. Those who stated 'don't know' on the Zung Depression Scale. Frequencies of this are unknown.	Unclear, little mention of confounding variables.	Yes
Larson (2011)	Attachment style and mental health in adults with intellectual disability: Self-reports and reports by carers.	Yes	Yes	Cross Sectional	IQ based on limits of funding authority, relatively small sample, level of ID unknown across sample	Measures appropriate for outcomes were used however reliability and validity of attachment questionnaire is unknown	4 participants excluded from analysis due to missing data. Clearly reported.	Confounds discussed in limitations e.g., ability to self-report but not considered in design/analysis	Yes
Lunsky (2001)	Association between perceived social support and strain, and positive and negative outcome for adults with mild intellectual disability.	Yes	Yes	Longitudina l T1 vs T2	Relatively small sample. Mild ID only but this was the target population. Supported living only.	One measure made specifically for ID. Some questions developed for this study. Other adapted or child standardised assessments used.	Yes	Mentioned in discussion only, not included in analysis.	Yes
McGillivr ay (2007)	Early detection of depression and associated risk factors in adults with mild/moderate intellectual disability.	Yes	Yes	Cross Sectional	Large sample with good age range. Severe/profound ID not included.	Appropriate standardised outcome measures but few are ID specific	Presumed but not explicitly stated.	not discussed	Yes
Nezu (1995)	Depression in adults with mild mental retardation: Are cognitive variables involved?	Yes	Yes	Cross Sectional	Large sample Mild ID only. This was target population	Measures appropriate. Some ID specific measures, some adapted measures and child measures	Presumed but not explicitly stated.	Matched controls	Yes
Reiss (1985)	Psychosocial correlates of depression in mentally retarded adults: I. Minimal	Yes	Yes	Cross Sectional	Participants had mild ID but diagnosis details not explicit. 45 participants, relatively small sample.	Yes, relevant measures for outcomes, evidence provided for use in the population.	Some missing data but clearly reported.	Confounding factors are considered, decided time frames of measures to reduce effects of any ongoing interventions.	Yes

	social support and stigmatization.								
Primary Author and Year	**Title**	**Are there clear research questions ?**	**Do the collected data allow to address the research questions ?**	**Quantitative randomised control trial**	**(1a) Is randomisation appropriately performed?**	**(2a) Are the groups comparable at baseline?**	**(3a) Are there complete outcome data?**	**(4a) Are outcome assessors blinded to the intervention provided?**	**(5a) Did participants adhere to the assigned intervention?**
Melville (2023)	Predictors and moderators of the response of adults with intellectual disabilities and depression to behavioural activation and guided self-help therapies	Yes	Yes	RCT	Randomisation complete but details of randomisation not explicit in the report. Details can be found in supplementary materials.	unknown if randomisation is appropriately performed as supplementary material not aviable.	Demographic descriptive data available. Statistical analyses of two groups not complete.	Single blind design used. Researcher aware of allocations.	Participants adhered to treatment. Supported by carer support offer.

Appendix C

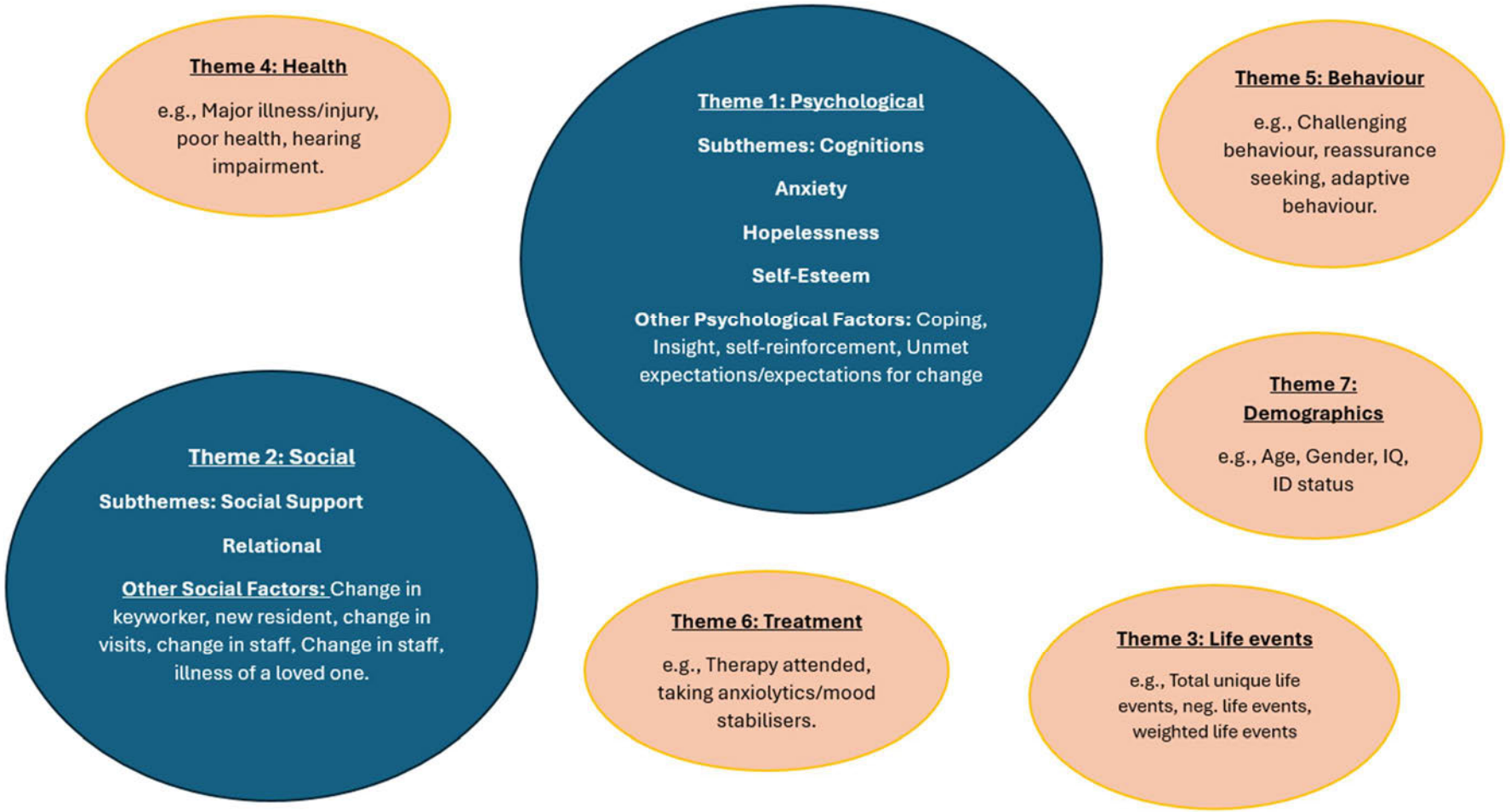
Key Findings from Identified Studies

Primary author and year	Key findings relating to review question
Austin (2018)	Self-reported depression scores were found to significantly correlate with anxiety (rho=.63, p=<.001), Maladaptive coping (rho=.45, p<.01), hopelessness (rho=.12, p=<.01), reduced insight (rho=-.5, p=<.01) and, unmet expectations of adulthood milestones/achievement (rho=-.21, p=<.05). Regression analysis identified maladaptive coping had the most predictive power accounting for 17% of the variance of depression.
Bond (2019)	Odds ratio showed significant associations with self-report depressive symptoms and a change of staff in home/service (OR=3.24, p=<.05), Death of a relative (OR=11.56, p=<.01), Change in frequency of visits from family/friends (OR=4.32, p=<.05) and death of a friend (OR=2.86, p=<.05). New resident, change of key worker, major illness of loved one, change in service and major illness/injury were not significant
Bond (2020)	Age, gender, taking anxiolytics, trouble sleeping and maintaining enthusiasm were not found to significantly correlate with self-reported depression scores. Challenging behaviour (OR=4.35, p=.04) and taking mood stabilisers (OR=6.11, p=.02) were found to be significantly associated with self-reported depression scores.
Collishaw (2004)	Mild learning disability significantly increased the risk of depression in both males (OR=3.92) and females (OR=2.38). Social disadvantage (OR=11.6, p=<.05) and poor health (OR=3.5, p=<.0001) were significantly associated with higher levels of depression.
Dagnan (1999)	Self-reported depression scores correlated significantly with self-esteem (r=-.39, p=<.01), social comparison (r=.50, p<.001) and negative self-esteem (r=.41, p=<.01).
Esbensen (2005)	Self-reported depression scores significantly correlated with automatic thoughts (r=.75, -<.01), hopelessness (r=.26, p=<.05), attributional styles (r=.33, p=<.01), the cognitive triad inventory (r=-.48, p=<.01) and self-concept (r=-.66, p=<.01).
Glenn (2003)	Self-reported depression was significantly correlated with anxiety (r=.74, p=<.001), automatic thoughts (r=.89, p=<.05) and depression cognition checklist (r=.76, p=<.05).
Hartley (2008)	Significant positive correlations were found between informant based and self-report (r=.45, p=<.01). Age (r=-.08), IQ (r=-.08), adaptive behaviour (r=-.09) where not found to be significantly correlated with depressive symptoms (>.05). Reassurance seeking was found to significantly predict depression scores (β=.40, p=<.01).

Hartley (2009)	Social support was not found to be a significant predictor of self-reported depressive symptoms (r=-.29, p=>.05), however, social support was found to be a significant predictor on informant measures.
Hulbert-Williams (2011)	High Social Support Criticism correlated with depression scores on the BSI (rho=.5, p=<.001) along with Low Social Support closeness (rho=.56, p=<.05), Total unique life events (rho=.52, p=<.001), Negative life events (rho=.54, p=<.001) and weighted life events (rho=.58, p=<.001).
Laman (1987)	Low levels of social support is associated with self-reported depression scores (r=.41, p=<.01).
Larson (2011)	There was a strong association with depression and I-AV attachment but this was not significant (x^2=[2]= 5.59, p=.06).
Lunsky (2001)	Social strain accounted for a significant proportion of additional variance in depressive symptoms beyond gender, age and social support [AR2 = 0.05, AF (1, 71) = 4.19, P < 0.05]. Only when social strain was added to the equation (step 3) did the model become significant [F (4, 71) = 2.45, P < 0.05]. Once social strain was added to the equation, the model accounting for somatic complaints became significant [F (5, 71) = 3.58).
McGillivray (2007)	The multiple regression identified that social comparison, self-esteem (β=.18), automatic negative thoughts (β=.55), social readjustment (β=.14) and perceived quality (β=.22) and frequency of social support (β=.21) were significant predictors of depression scores in the BDI accounting for 58.1% of the variance. Negative automatic thoughts had the greatest impact whilst social comparison did not significantly contribute thot the regression model.
Melville (2023)	Greater severity Of depressive symptoms at baseline (-0.160, p=0.001), higher expectation of (-1.013, p=0.005) and greater percentage of therapy sessions attended (-0.058, p= 0.007) were predictors of more positive outcomes for treatment after adjusting for randomised group allocation. The final model included severity of depressive and anxiety symptoms, lower FIASI performance, hearing impairment, higher expectation of change and percent of therapy sessions attended and explained 35.5% of the variance in the total GDS-LD score at 12 months (r2 = 0.353, F4, 128 = 17.24, p=0.001)
Nezu (1995)	Self-reported depression symptoms significantly correlated with negative automatic thoughts (r=.61, p=<.001), hopelessness (r=.36, p=<.001), frequency of self-reinforcement (r=.58, p=<.001) and negative social support (r=.28, p=<.01). self-reported depression was not significantly correlated with emotional or practical social support received.
Reiss (1985)	Social support and self-reported depressive symptoms were negatively correlated (r=-.41, p=<.001). There was no significant difference between men (r=-.45) and women (r=-.39). Stigmatisation was not found to correlate with self-reported depression.

Appendix D

Thematic Map Illustrating the Themes Summarised in the Narrative Synthesis



Appendix E

Reference List of Inaccessible Papers

MacMahon, P., & Jahoda, A. (2008). Social comparison and depression: People with mild and moderate intellectual disabilities. *American Journal on Mental Retardation*, *113*(4), 307. https://doi.org/10.1352/0895-8017(2008)113[307:scadpw]2.0.co;2

Email correspondence with the author in an attempt to obtain a copy of the article was unsuccessful

Appendix F

PRISMA Checklist

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	Page. 1
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	Page. 2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	Page. 4
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	Page. 7
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	Page. 9
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	Page. 8
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	Page. 8
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	Page. 8
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	Page. 11
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	Page. 11
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	Page. 12
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	Page. 10
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	n/a
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	n/a

Section and Topic	Item #	Checklist item	Location where item is reported
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	n/a
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	Page. 11
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	Page. 11
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	n/a
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	n/a
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	Page. 10
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	Page. 10
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	Page. 12
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	n/a
Study characteristics	17	Cite each included study and present its characteristics.	Page. 18
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	Page. 16
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	n/a
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	Page. 15
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	n/a
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	n/a
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	n/a
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	Page. 15
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	Page. 15
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	Page. 34

Section and Topic	Item #	Checklist item	Location where item is reported
	23b	Discuss any limitations of the evidence included in the review.	Page. 36
	23c	Discuss any limitations of the review processes used.	Page. 36
	23d	Discuss implications of the results for practice, policy, and future research.	Page. 37
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	Page. 7
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	Page. 7
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	Appendix. A
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	n/a
Competing interests	26	Declare any competing interests of review authors.	n/a
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	n/a

From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71

For more information, visit: http://www.prisma-statement.org/

Part II: Empirical Project

Validation and Abbreviation of a Patient Reported Outcome Measure of Psychological Distress and Positive Wellbeing for Adults with Intellectual Disabilities.

Abstract

Objectives

The Psychological Therapies Outcome Scale (PTOS) was developed to assess psychological distress and positive wellbeing in adults with an intellectual disability (ID). Following initial validation of the PTOS by Vlissides et al., (2017), the measure was revised taking into consideration the removal and rewording of specific items. This was subsequently called the PTOS-ID-II consisting of 29-items. The current project seeks to (i) validate the PTOS-ID-II and (ii) abbreviate the measure to allow for quick administration in clinical settings.

Design

A quantitative cross-sectional design was utilised to test the psychometric properties of the PTOS-ID-II and abbreviate the measure.

Method

Data gathered from a community ID health service (n=879) was used to explore the psychometric properties of the PTOS-ID-II. An existing primer on the development of a health outcome measure was used to guide the analysis. This included the use of principal component analysis, confirmatory factor analysis and reliability and validity testing.

Results

Three factors emerged as seen in the original PTOS, however, two items were deemed redundant and removed. The new 27-item measure is re-branded 'the Outcomes for Wellbeing and Distress Scale' (OWLS-ID). The measure was then abbreviated to a ten-item version for quick administration, named the 'OWLS-Mini'. Both the OWLS-ID and OWLS-Mini showed evidence of internal consistency, concurrent validity, sensitivity, and specificity.

Conclusion

The OWLS-ID and OWLS-Mini are validated patient report outcome measures developed specifically for individuals with ID to assess psychological distress and positive wellbeing. Future research involves further development of the measures, namely their responsiveness to changes over time.

Practitioner Points

- Patient reported outcome measures (PROMs), such as the Psychological Therapies Outcome Scale (PTOS-ID-II), can be used in practice to quantify the progress of therapy, screen for symptoms and assess service user's subjective accounts.
- The OWLS-ID is the most validated ID specific PROM measuring wellbeing and distress.
- Additional analyses were run to abbreviate the OWLS-ID creating a 10-item measure named the OWLS-Mini. This should be used when quicker administration in practice would be beneficial e.g., ongoing outcome monitoring or lower tolerance levels.
- Practitioners should use the OWLS measures (OWLS-ID and OWLS-mini) in practice in place of the PTOS-II.

Key words: 'Patient Report Outcome Measures', 'Wellbeing', 'Distress', 'Intellectual Disability'

Introduction

Intellectual disability (ID) is a neurodevelopmental condition characterised by significant difficulties in cognitive and social/adaptive functioning, present before the age of 18 (American Psychiatric Association, 2013). People with ID are at a greater risk of experiencing mental health problems than the general population (Marshall & Willoughby-Booth, 2007; Cooper et al., 2007; Emerson & Hatton, 2007). Rates in people with ID have been estimated as 22.4%, which is much higher than the 8% prevalence rate reported in the general population (Hughes-McCormack et al., 2017). Resultantly, there has been an increase in the need for effective psychological therapies suitable for people with ID (Willner, 2005; Beail 2016). Assessing the effectiveness of treatments in this group has been a difficulty (Makrinioti et al., 2019), especially given they are more likely to access barriers to successful treatment such as diagnostic overshadowing and overreliance on pharmaceutical interventions (Tapp et al., 2023).

Patient reported outcome measures (PROMs) can be used to help quantify the progress of therapies (Churruca et al., 2021). Such measures may be disease-specific; for example, the Generalised Anxiety Disorder scale (GAD-7; Spitzer et al., 2006), or more generic measures assessing constructs such as psychological distress (CORE-OM; Evans et al., 2000). PROMs are essential in clinical practice and can be used to screen for symptoms of mental health difficulties, assess need, inform treatment, and evaluate treatment effectiveness (Skelly, 2011). PROMS are also used to evaluate service user's experience of healthcare interventions (Slevin et al., 1988; Vlissides et al., 2017). As such, PROMs are necessary not only to assess therapy outcomes, but to encourage the service users' experiences to be considered (Barkham et al, 2001).

There is limited research investigating PROMs in adults with ID, particularly when compared to the general population. Challenges of creating an outcome measure suitable for

this group may arise due to additional barriers such as identifying, labelling, and communicating their distress, amongst the ability to understand others. Vlissides et al. (2016), completed a systematic review of outcome measures for people with ID and evaluated these against Fitzpatrick's (1998) measure of psychometric properties. Fitzpatrick's criterion was developed based on a series of desirable attributes for self-report outcome measures. Cahill et al., (2008) expanded on this by developing a rating tool to assess the quality of each domain proposed by Fitzpatrick (see Table 1). Vlissides et al., (2016) concluded that the status of PROMs used for people with ID were generally poor in terms of all seven of the above criteria. This is problematic as it impacts our ability to reliably assess the effectiveness of care.

Other forms of outcome measures for this clinical group are available, including informant, observation, or clinician-based measures (Moss, 2019; Cuthill et al., 2003). While we recognise the importance of these tools, the current study focuses on measures completed by patients.

Table 1

Fitzpatrick's (1998) Criteria of Psychometric Properties with Cahill et al's., (2008) Rating Tool.

	Fitzpatrick's criteria	Cahill's (2008) rating tool
Reliability	Produces consistent results from the same respondents over time. • Internal consistency • Test-retest reliability	• Adequate= >0.7 • Partial= 0.5-0.7 • Inadequate= <0.5 • Unknown= Reliability not supplied
Validity	The extent to which a measure really measures the concept that it purports to measure. • Discriminant validity • Convergent validity • Concurrent validity • Construct 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	Does the instrument detect changes over time that matter to the patient? • Discriminative (between individuals) or evaluative (within individuals).	• Adequate= Significant differences found between groups or within individuals • Partial= non-significant trends found between groups or within individuals • Inadequate= not addressed
Acceptability	Is the measure acceptable to users? • Practicality of administration • Time taken to complete • Length of instrument • Translations • Access by ethnic minorities • Reading age	• Adequate= all of the components described • Partially addressed=at least one of the components described • Not addressed= None of the components described.
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	• Adequate= All of the components described • Partially addressed= At least one of the components described • Not addressed= None of the components described
Precision	How precise is the measure? • Interpretability • Normative data	• Adequate= All of the components described • Partially addressed= At least one of the components described • Not addressed= None of the components described

Shading used to enhance clarity for reader.

Adapted Measures

Adapted measures are PROMs developed for use with the general population that have been administered to people with ID (Novaco & Taylor., 2004; Vlissides et al., 2017). Such measures are often psychometrically sound in the general population, and with assisted delivery can be administered to those with ID. For example, Kellett et al., (1999, 2003) used the Symptom Checklist-90R (Derogatis., 1983), the Brief Symptom Inventory (Derogatis, 1993) and the Rosenberg Self-Esteem Questionnaire (Rosenberg., 1965) via assisted completion in an outpatient service for adults with ID. However, this mode of administration can be problematic as it requires skilful adaption by the administrator (Beail et al., 2012). Furthermore, these scales are developed and normed upon the general population. As such, the symptoms being measured may be experienced differently by those living with ID (Vlissides et al., 2017). This was evidenced by Kellett et al., (2004) and Davis et al., (2009) who found that the BSI and the Rosenberg Self-Esteem Questionnaire factored differently with people who have ID to that of the general population, meaning the factor structure underpinning the measure differed between the two groups. This indicates that general population measures may not accurately measure what they intend to in individuals with ID.

Specific Measures

It appears that a logical advancement in the field of PROMs for people with ID would be to develop measures specifically for the population (Schwarts et al., 2021). Subsequently, a growing number of specific measures for people with ID have been developed assessing a variety of clinical outcomes. These measures are typically developed via consultation with professionals specialising in caring for this clinical group. Before testing the measures' reliability and validity on data collected from those accessing ID healthcare services. A recent review looking exclusively at ID specific measures found that the evidence of robust psychometric testing was sparse. Gourley and Yates (2022) used Fitzpatrick's (1998) criteria

and found none of the ID specific measures reviewed showed evidence of psychometric quality across all seven domains.

PTOS-ID

The PTOS-ID is an ID specific measure developed in collaboration with over 100 clinical psychologists in the field of ID. Items intend to measure anxiety, anger, depression, interpersonal wellbeing, and psychological wellbeing. The 29-item PTOS-ID was administered in a psychological outpatient service with 175 adults with ID as part of the initial validation (Vlissides et al., 2017). A three-factor model with adequate construct validity was found with each factor showing evidence of internal consistency (Positive Wellbeing, α =0.81; Emotional and Behavioural Discomfort (EBD), α=0.82; Anxiety, α=0.76). The latter two factors were combined to create the psychological distress index. Concurrent validity was evident from correlation with the BSI scales of distress (r=.85, p<.001) and acceptability was identified through consultation with service users and professionals (Beail et al., 2023). As the PTOS-ID was trialled in practice, amendments were made including changes to the wording and order of questions, addition of a time frame (over the past two weeks) and a reduction in items. This led to the second edition of the measure (PTOS-ID-II; Appendix A). However, this new measure has not undergone psychometric testing. As such, the current research will focus on the psychometric development of the PTOS-ID-II.

Abbreviated measures are seldom found in ID services despite the characteristics of a shorter interview being of particular benefit to this population. As such, a secondary aim is to abbreviate the measure for efficient administration. Both the validation and abbreviation stages are recommended to ensure that the resultant abbreviated measure is based on an appropriately validated measure. Whilst the different measures are both useful in different contexts (See Discussion). Members of the current research team (Dr. Nik Vlissides and Prof.

Nigel Beail) developed the PTOS measures and have given the necessary authorisation to complete this research.

Current Study

The present study proposes to investigate the psychometric properties of the PTOS-ID-II and abbreviate this for quick administration in practice.

Aims

1. **Validate the PTOS-ID-II.**

Boateng et al., (2018) published a primer outlining the stages required to develop an outcome measure for health, social, and behavioural research. While this primer was not developed on findings from a systematic review, it proposes a comprehensive framework consisting of nine steps which were used to guide the current research:

Steps one–four involve identifying the proposed need of the measure, generating relevant items, and administering the draft measure to the target sample. These steps have already been completed in the development of the PTOS-ID.

Step five focuses on item reduction. This typically involves reviewing the data to identify any items that may be problematic, say, if an item is associated with a high rate of being missed.

Step six involves the extraction of factors. Factor analysis is used to understand the internal structure of a set of items and the extent to which the relationships between items are internally consistent. This is done by extracting factors which represent the shared variance in responses across items (Boateng et al., 2018). Principal components analysis (PCA) is a factor analytic technique aimed to reduce a dataset to its essential features otherwise known as its ‘principal components’ (Greenacre et al., 2022). Research has shown that factor solutions derived from EFA (exploratory factor analysis) or PCA show very little difference

in outcome (Guadagnoli & Velicer, 1988) unless testing fewer than 20 variables with low levels of communality (Stevens, 2002). As such, PCA was adopted.

Step seven includes tests of dimensionality. Specifically, confirmatory factor analysis (CFA) will be used to further validate the factor structure hypothesised in step six. CFA is a form of psychometric assessment that allows systematic comparison of an *a priori* factor structure and estimates the latent constructs (Gerbing & Hamilton, 1996). Boateng et al., (2018) highlight the most common tests for dimensionality include Root Mean Square Error of Approximation (RMSEA), Tucker Lewis Index (TLI), Comparative Fit Index (CFI) and Standardized Root Mean Square Residual (SRMR). These fit statistics will be used to consider whether the model is a good fit for the data.

Steps eight to nine, focus on assessing reliability and validity of the proposed measure. Cronbach's alpha will assess the internal consistency of the items i.e. how closely related are the items per each domain. Concurrent (criterion) validity will be tested comparing scores on the newly validated PTOS-ID-II with scores on a measure assessing a similar construct, specifically, the BSI (Derogatis 1993). Predictive criterion validity is the degree to which scores on one measure correlate and can predict the behaviour of a related measure (Boateng et al, 2018). This will be assessed by Receiver Operating Characteristic (ROC) analysis to identify which scores on the psychological distress domain of the PTOS-ID-II fall above the clinical cut-off on the BSI. Clinical cut-offs are scores at which a hypothetical line is drawn between those scoring within a 'typical/healthy' range and those that could be considered distressed (Pashak et al., 2022). This will provide statistics on the optimal clinical cut-off for the PTOS-ID-II.

2. Abbreviate the PTOS-ID-II.

To the authors knowledge, there is a not a similar primer to abbreviate validated PROMs. Therefore, the literature was viewed for examples to help inform the current approach

(Barkham et al., 2013; McCracken & Dhingra, 2002). The proposed framework includes three steps modelled on the approach reported by Barkham et al (2013): Firstly, results from the validation stages of the PTOS-ID-II will be used to identify any items which failed screening and suitability tests. Secondly, analysis of the PCA factor loadings conducted in step 6 above, will be used to identify key items. Finally, correlation matrices will be used to remove any items which are too highly or poorly correlated with other items in that domain. While regression analysis will be run to test whether the remaining items predict the domain they intend to measure. The abbreviated measure will then undergo reliability and validity checks as seen in steps eight-nine, which will be checked against the Fitzpatrick (1998) and Cahill (2008) criterion. Given the aims of this project, objectives have been stated rather than specific hypotheses.

Table 2

Overview of the Nine Step Framework for Scale Development/Validation (From Boateng et al., 2018)

Stage	Step
Item Development	1- Identification of domain and item generation
	2- Content validity
Scale Development	3- Pre-testing of questions
	4- Sampling and survey administration
	5- Item reduction
	6- Extraction of factors
Scale Evaluation	7- Tests of dimensionality
	8- Tests of reliability
	9- Tests of validity

Method

Setting and Approvals

Analyses of the psychometric properties of the PTOS-ID-II and subsequent abbreviated measure were conducted using a database collected from adults with ID accessing a health service as part of routine practice. The service is situated in the north of England serving a population of approximately 245,000 people. Data from the PTOS-ID-II was collected on entry to the service (e.g., at triage or initial assessment), during psychological therapy, and prior to a diagnostic assessment (ID or autism spectrum disorder). This data was stored on the local database kept securely within the NHS.

NHS and University ethical approval was sought. Approval was granted by Health and Care Research Wales (IRAS ID 324291) and the Department of Psychology Research Ethics Committee within the University granted ethical approval (058517). NHS ethics was not required (Appendix B). The STROBE checklist was used (Appendix C) to ensure key features of cross-sectional research were presented (von Elm et al., 2007).

Sample

Participants were aged 17.5 years or above and had a confirmed diagnosis of ID. To access the service, patients are required to have a formal diagnosis of ID or registration on the local ID register. In order to describe the severity of a person's condition, ID is often categorised based on IQ scores; mild (50-70), moderate (35-50), severe (20-35), or profound (<20; Boat et al., 2015; Lee, Cascella & Marwha, 2023). In the current dataset, this will have been tested using a standardised assessment of intelligence such as the Weschler Adult Intelligence scale (WAIS; Weschler, 2008). Information on gender, age, severity of ID, and ethnicity was also gathered.

Measures

Psychological Therapies Outcome Scale-ID 2nd Edition

The PTOS-ID-II is a measure of positive wellbeing and psychological distress in adults with ID. Each item begins with 'Over the past week...' proceeded by a question such as 'Have you felt annoyed?'. Each item is rated on a 4-point Likert scale of 'not at all,' 'a little bit', 'sometimes' and 'a lot'.

Brief Symptom Inventory

The BSI (Derogatis, 1993) is a 53-item self-report measure aimed to quantitatively assess psychological distress and symptoms of psychiatric disorders (Adawi et al., 2019). The BSI measures nine dimensions which are reflected in three global indices. These include the Global Severity Index (GSI); a summary of symptoms and distress; The Positive Symptom Total (PST), which provides a total frequency of the number of symptoms reported, and the Positive Symptom Distress Index (PSDI), which is a measure of symptom intensity. The BSI uses a 5-point Likert scale ranging from "not at all to" to "extremely". The BSI manual suggests a clinical cut-off >0.62 on t-scores of the Global Severity Index (GSI), the indices indicative of total distress. This was derived from adult non-patient sample norms (Derogatis, 1993; Pashak et al., 2022). The BSI has been found to maintain the majority of its factor structure with people with ID, can discriminate between community and clinical populations, and demonstrated internal consistency between .71 and .85 (Adawi et al., 2019; Kellet et al., 2003; Kellet et al., 2004; Vlissides et al., 2016).

Procedure

The PTOS-ID-II were typically completed one-to-one in a consulting room at the service. The questions would typically be read aloud verbatim, in chronological order by a clinician. It is likely that a small number were completed by the client independently or remotely. The BSI was administered to participants by qualified psychologists.

Data Analysis

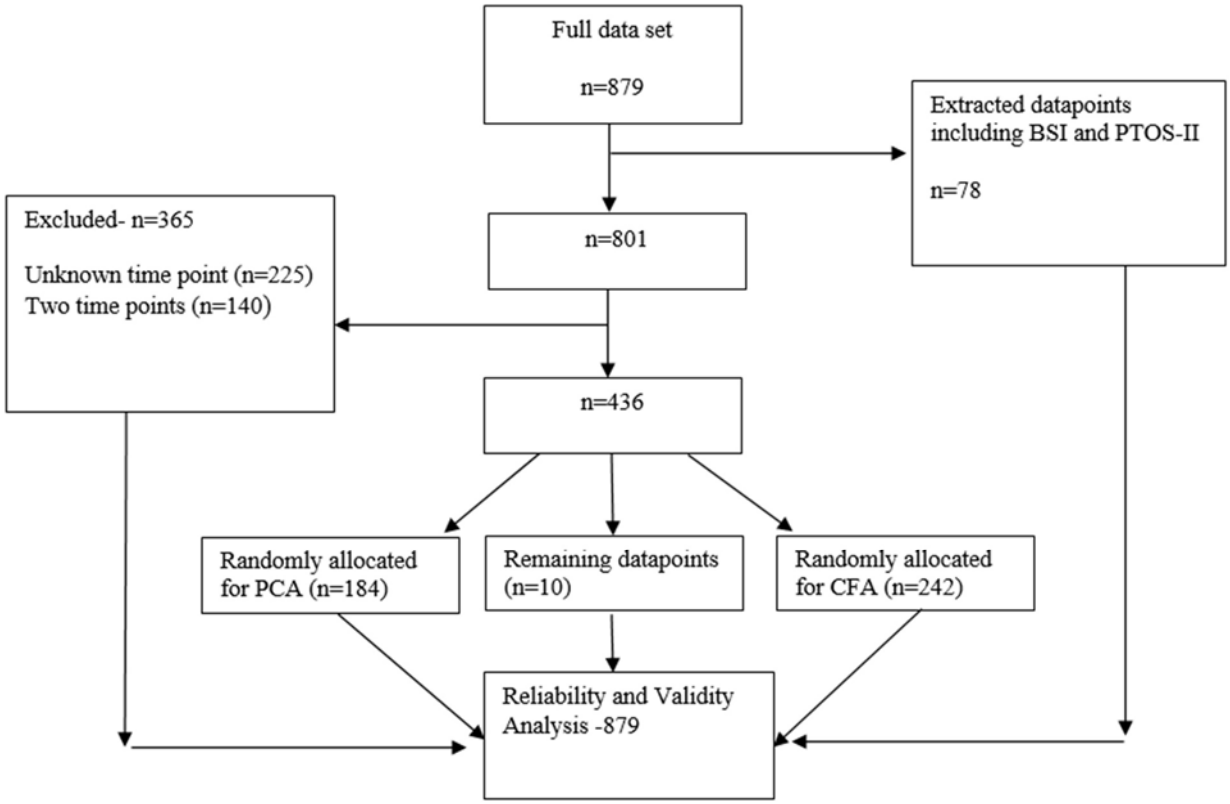
The data from 879 adults with an ID was analysed. The data was split to allow independence in analyses and improve methodological rigour. A flowchart demonstrating the data split can be seen in figure 1. All of the data captured on participants first entry to the service was selected (e.g., initial assessment). Data from 225 participants did not state the stage in therapy that the measure was completed, therefore, were not included in this analysis. Individuals who had also completed a BSI or had completed measures at multiple time points (e.g. during and after therapy) were not included at this stage (n=140). This left 436 data points which was then then split roughly in half. For the validation stage of the PTOS-ID-II (stage one), roughly half (n=184) of the first entry data was used for the PCA, and another half (some BSI data included in this split) used for the CFA (n=242). Samples were uneven due to the exclusions explained above. Samples were allocated randomly via SPSS (IBM Corp, 2023). An ANOVA was run between the PCA and CFA groups to test for significant differences between the two samples.

The whole data set was used for the reliability and validity analysis (n=879) which included some previously unused data e.g., those who did not report the stage of therapy at completion (n=225) and the data gathered at a second time point in the service (n=140).

A fourth dataset was used for validity checks with the BSI. This included only those who completed both a BSI and PTOS-ID-II (n=78). The demographic characteristics of each sample can be found in Table 3.

Figure 1

Flowchart Showing Data Splits Through Analyses



*PCA dataset and Reliability and Validity Dataset used for the abbreviation analyses.

Statistical Analysis

Validation.

Screening tests were conducted to check the suitability of the data (n=184, PCA dataset) for factor analytic techniques. Screening tests included (a) assessing the percentage endorsement on each item to consider dispersion of responses (Tinsley & Tinsley, 1987), items with low levels of endorsement (<5%) were considered for removal, (b) analysing the correlation matrix to ensure all items correlated moderately with one and other. Correlations above 0.3 and below 0.7 were considered ideal (Akoglu, 2018; Field, 2013), (c) computing Variance Inflation Factor (VIF) and tolerance statistics to test for multicollinearity (when two or more variables are highly correlated). VIF below 10 and tolerance statistics above 0.2 were

considered suitable (Menard, 1995; Myers, 1990), and (d) conducting a review of missing cases. Items with more than 5% missing data were considered to have poor acceptability, as this is indicative of a regularly missed item. Any items which failed two or more of these tests were removed from the analysis in replication of the original authors of the PTOS (Vlissides et al., 2017). As the questionnaire responses were in a Likert format, the data was considered ordinal hence parametric assumptions were violated. Spearman's correlations were conducted.

The PCA was run on IBM SPSS version 29 (IBM Corp, 2023). PCA was used to assess construct validity and the reliability of the factor structure proposed by Vlissides et al., (2017). PCA assumptions were tested as follows: 1) the KMO (Kaiser-Meyer-Olkin) was calculated. KMO is the measure of the proportion of variance among the variables derived from common variance (Hair et al., 2018). KMO values range from 0-1, with lower values suggesting partial correlations which could be problematic for PCA. Hair et al., (2018) suggests KMO values below 0.5 are insufficient. 2) the communality of rotated variables was examined. Communality indicates the common variance shared by factors with certain variables. Communalities above 0.4 are generally considered suitable (Costello & Osborne, 2019).

Multiple criteria are required to support extraction of factors from a dataset (Hair et al., 2018 Tinsley & Tinsley, 1987). These included, (a) Kaisers Criterion of eigenvalue of >1, (b) analysis of the scree plot, and (c) analysis of the cumulative percentage of each factor (Williams, Brown & Onsman, 2012). There is no fixed rule for exploring cumulative percentages; however, factors explaining variance between 50-60% are common (Hair et al., 2012). Furthermore, Floyd & Widamann, (1995) suggest at least five data points per one questionnaire item. The current PCA consisted of a sample of 184 participant responses across a 29-item structure. Therefore, the dataset is above this threshold (1:6.34).

The resultant model from the PCA was then analysed using CFA to test whether this model was the best fit for the data. Assumptions of the CFA were tested, which included: (a) data was a random sample, (b) the dataset was larger than n=200, and (c) the model specification was *a priori* based on theoretical deduction and not observed (Statistics Solutions, 2013). Analysis of minimum and maximum values were conducted to assess for extreme outliers and suspicious data as part of the screening process. Similar to the preparations for PCA, percentage of endorsement across all responses and tests of normality were conducted to understand the distribution of the data. Mardia's test of skewness and kurtosis was run using an online web calculator (Cain et al., 2021). Skewness values of more than 3 and -3 are not considered acceptable for factor analytic techniques; Kurtosis values more than 8 and -8 are considered extreme (Byrne, 2010; Griffin & Steinbrecher, 2013; Kline, 2011). Similar to the pre-PCA screening tests, KMO's test of sampling adequacy, Bartletts test of sphericity, VIF and tolerance statistics were run. Correlational analysis was complete to ensure variables were suitably correlated (>.3).

The CFA was run on the statistical computing programme 'R Studio' (R Studio Team, 2020). Fit statistics were used to quantify the best fit for the data. The RMSEA is a measure of the error of the approximate fit (Goretzko et al., 2024; Steiger, 1998) and was assessed based on the following criteria: <0.05-good, 0.05-0.08-acceptable, 0.08-0.1-marginal, >0.1-poor (Fabrigar et al., 1999). TLI and CFI are measures of goodness-of-fit with values above 0.9 and close to 0.9 respectively, considered a 'good' fit (Bentler, 1990). Finally, SRMR is a 'badness of fit' measure, therefore, lower values demonstrate a better fit; with zero demonstrating 'perfect' fit and 1 demonstrating 'poor' fit. According to Hu & Bentler, (1999), SRMR values below 0.08 are considered good. A path diagram was created to visually demonstrate the correlations between items and latent factors.

Internal reliability was assessed via internal consistency analysis (>.8=good; George and Mallary, 2003). Acceptability was assessed via Littles Missing Completely at Random (MCAR) analysis, and concurrent validity was assessed by comparing the PTOS-ID II with the BSI. ROC analysis was also used to assess sensitivity and specificity of the measure (>.80=good; Hosmer & Lemeshow, 2000), based on clinical cut-offs suggested by the BSI. The alpha level (0.05) was accepted.

Abbreviation.

A three-step formula was utilised: 1) any items that fail two or more screening tests in the validation stage (step five) were removed, 2) analysis of factor loadings from the PCA identified key questions with loadings of above .6 retained for further analysis whilst those below this threshold were considered for removal, and 3) a series of correlational analyses took place. Inter-item correlations were run to identify questions that were too highly or too poorly related (correlations below .3 or above .7 were considered for removal). Total score correlations were also run with scores above .7 suggesting items are strongly associated with the domain as intended (Mukaka, 2012), while those equal to or below .7 considered insufficient. Regression analysis was conducted to evaluate the extent to which the retained items could predict overall indices scores. This concluded the abbreviation stage, and the remaining items formed the abbreviated measure.

Preliminary examinations of validity and reliability were run on the abbreviated measure. This included a PCA to test the factor structure, ROC curve analysis to identify sensitivity and specificity of the measure, internal consistency (via Cronbach's alpha), and concurrent validity analysis (via correlation with the BSI) as described above.

To end, psychometric properties of the newly validated measures were assessed against Fitzpatrick (1998) and Cahill's (2008) criteria. Cahill (2008) provides a three-tier

system of ‘inadequate’, ‘partially adequate’ and ‘adequate’. Each domain was colour coded to visually represent the score (red=inadequate, amber=partial, green=adequate).

Patient and Public Involvement (PPI)

The PTOS-ID-II questions were developed from clinicians working in specialist services for adults with ID. Patients or carers were not involved in the item development. Service evaluation was completed to test its acceptability and gain service user feedback (N. Beail, personal communication, July 26th, 2022).

Results: Stage 1-Validation

Participants

The sample (n=879) had a mean age of 28.49 (SD=13.43) years and ranged from 17-75. This consisted of 452 males (51.4%), 385 females (43.8%), and 2 non-binary (0.2%). Data on ethnicity was available for 136 people; 134 individuals (15.2% of total sample) identified as white and two mixed/multiple ethnicities (0.2%). Ethnicity was not reported for the majority of participants (n=743, 84.5%). The sample consisted of 404 individuals with a mild ID (46%) and 44 individuals with a moderate ID (5%). Data on severity of ID was unavailable for 431 individuals (49%). IQ scores were available for 425 individuals (48.4%) and ranged from 41-75 (mean=59.40, SD=7.98), suggesting the majority of participants fell within the mild to moderate range.

An ANOVA and chi-square was run between the PCA (n=184) and CFA (N=242) datasets to test for any differences between the two samples. No significant differences were found in terms of age, gender, or LD severity (Appendix D), suggesting randomisation was successful.

Results have been separated based on the nine-steps proposed by Boateng et al., (2018), steps one-three were completed by authors of the PTOS-ID:

Step Four

Item Dispersion/Missing Data

The following analysis took place on the PCA data set (n=184). Demographic data can be seen in Table 3. Assessment of response dispersion found that of the 29 items on the PTOS-ID-II, no items were endorsed by less than 5% of responders suggesting adequate discrimination across item responses.

Table 3

Demographic Data of Each Dataset

	Full Dataset	**PCA**	**CFA**	**BSI & PTOS**
Characteristics		**N**		
N	879	184	242	78
Demographics				
Age in year (Mean, sd and range)	31.40(*sd=12.55*) 17-73	30.98 (*sd=12.93*) 17-70	29.45 (*sd=11.96*) 17-60	28.26 (*sd=11.76*) 17-60
Gender				
Male	452 (51.4%)	104 (56.5%)	128 (52.9%)	33 (42.3%)
Female	385 (43.8%)	74 (40.2%)	92 (38%)	28 (35.9%)
Non-binary	2 (0.2%)	0 (0%)	0 (0%)	0 (0%)
Not reported	40 (4.6%)	6 (3.3%)	22 (9.1%)	17 (21.8%)
Ethnicity				
White	134 (15.2%)	41 (22.3%)	36 (14.9%)	0 (0%)
Black	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Mixed/Multiple ethnicities	2 (0.2%)	1 (0.5%)	0 (0%)	0 (0%)
Asian	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Not reported	743 (84.5%)	142 (77.2%)	206 (85.1%)	78 (100%)
ID Severity				
Mild	404 (46%)	89 (48.4%)	156 (64.4%)	63 (80.8%)
Moderate	44 (5%)	10 (5.4%)	20 (8.3%)	2 (2.6%)
Severe	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Not reported	431 (49%)	85 (46.2%)	66 (27.3%)	13 (16.6%)
IQ (Mean, sd and range)	59.68 (*sd=7.18*) 36-75	58.97 (sd=7.68) 36-74	60.16 (*sd=7.5*) 41-75	61.74 (*sd=6.82*) 47-75

IQ=Intelligence Quotient, ID=Intellectual Disability, sd=Standard Deviation, PCA=Principal Components Analysis, CFA=Confirmatory Factor Analysis, BSI=Brief Symptom Inventory, PTOS=Psychological Therapies Outcome Scale.

Sampling Adequacy

The Kaiser-Meyer-Olkin (KMO) test was used to measure sampling adequacy (KMO = 0.9) and was considered suitable (>0.5). Bartlett's test of sphericity was found to be significant (p= <0.001) identifying suitability. No items were found to have communality below .4 thus deemed acceptable to progress with factor analytic techniques.

Step Five

Multicollinearity

Correlational analysis was run between all 29 items (Appendix E). Items 11 ("Have you been sleeping less than usual? If no, have you been sleeping more than usual?" and 22 ("Have you been eating more than usual? If no, have you been eating less than usual?") showed poor correlations with almost all other items (>.3). This was this was the first screening tests these items had failed, therefore, were retained for analysis at this stage. VIF was found to be below 10 for each item and tolerance statistics were all above 0.2 suggesting no cause for concern regarding multicollinearity (Menard, 1995; Myers, 1990).

Step Six

Principal Components Analysis

PCA was run on the 29 items (n=184). The initial PCA suggested six factors accounting for 62.7% of the variance based on Eigenvalues above 1. Scree plot analysis took place by identifying the 'elbow point' where the graph levels off and identifying all factors to the left of this point as significant. Based on this, three-factors were retained which accounted for 50.3% of the variance (see Table 4).

Table 4

Initial Results from PCA

Factor	Eigenvalue	Variance (%)	Cumulative Variance (%)
1*	**9.82**	**33.88**	**33.88**
2*	**3.02**	**10.40**	**44.27**
3*	**1.76**	**6.07**	**50.34**
4	1.36	4.70	55.04
5	1.21	4.16	59.20
6	1.03	3.54	62.74

Bold*-Extracted Factors

A summary of the factor loadings is presented in Appendix F. This shows that item 22 ("Have you been eating more than usual? If no, have you been eating less than usual?") did not sufficiently load (<0.3) onto any of the three factors. Due to a high proportion of items loading onto multiple factors, PCA was run again with rotation applied (Table 5). Varimax Orthogonal rotation was used to replicate Vlissides' (2017) work who hypothesised that the items measuring distress and wellbeing would not correlate.

The rotated PCA demonstrated fourteen fewer double factor loadings. Item 12 ("Have you been able to cope with problems?") and 22 ("Have you been eating more than usual? If no, have you been eating less than usual?") failed to factor onto any of the three proposed factors (<0.3). This was the second time item 22 had failed to factor above the recommended threshold, and therefore was removed. A PCA was calculated with item 22 excluded. This time, item 12, 'Have you been able to cope with problems?' and item 15 'Have you been able to stand up for yourself?' failed to load. This was the first test item 15 failed, therefore was retained at this stage. Item 12 had failed to load a second time, therefore, was removed.

Table 5

PCA with item 12 and 22 removed, rotation applied.

Factor	Eigenvalue	Variance (%)	Cumulative Variance (%)
1*	**9.59**	**21.12**	**21.12**
2*	**2.96**	**15.02**	**36.14**
3*	**1.76**	**13.65**	**49.79**
4	1.11	6.60	56.39
5	1.05	4.56	60.95

Table 5b

Factor Loadings from Rotated PCA, two items removed.

Item	Factor 1-Positive Wellbeing	Factor 2-EBD	Factor 3-Anxiety
1-Have you been interested in doing things or meeting people?	***.680***		
5- Have you felt like you can make friends?	***.615***		
7- Have you felt you are a good person?	***.442***		-.391
14- Have you looked forward to things?	***.707***		
15- Have you been able to stand up for yourself?	***.275***		
16- Have you felt you can do things as well as other people?	***.410***		
20- Have you been able to tell people how you feel?	***.307***		
24- Have you felt happy with your life?	***.765***		
26- Have you felt people love or care about you?	**.800**		
27- Have you been able to show people you love or care about them?	**.638**		
29- Have you felt happy?	**.817**		
2- Have you felt sad?	*-.534*	**.425**	.326
3- Have you felt angry?		**.737**	.361
6- Have you felt annoyed?		**.656**	.346
9- Have you felt like smashing things?		**.725**	
11- Have you been sleeping less than usual? If no, have you been sleeping more than usual?		**.343**	
18- Have you felt like you are no good?	*-.566*	**.363**	.346
19- Have you felt like hitting someone?	-.318	**.540**	
23- Have you had a bad temper?		**.736**	

25- Have you thought about death or dying?	*-.517*	**.337**	
28- Have you felt wound up?		**.683**	.360
4-Have you felt frightened of things or places?			**.779**
8-Have you suddenly felt scared?		.307	**.705**
10-Have you felt anxious?		.520	**.563**
13- Have you checked things over and over again?			**.517**
17- Have you felt faint or dizzy?	*-.331*		**.503**
21- Have you stayed away from some places or things because you are frightened of them?			**.776**

This factor structure was tested by running an additional PCA requesting a three-factor model with items 12 and 22 removed. The three factors accounted for 52.96% of the variance confirming a three-factor structure as identified by Vlissides et al., (2017). Factor descriptions coined by the PTOS authors appear fitting; "Positive Wellbeing", "Emotional and Behavioural Discomfort (EBD)" and "Anxiety". Item 25 loaded higher onto factor one (-517), however, based on the inherited scoring system (no negatively scored items) question 25 is grouped onto factor two as this is the highest positively loaded factor (>.3), this allows the previous scoring system to continue but means that some double factor loadings remain. This can also be seen for item 18. The final results of the PCA can be seen in Table 6.

Table 6a

Final Results of the PCA

Factor	Rotated Eigenvalue	% of Variance	Cumulative %
1*	5.92	21.92	21.92
2*	4.56	16.90	38.82
3*	3.82	14.13	52.96

Bold*- Extracted factors.

Table 6b

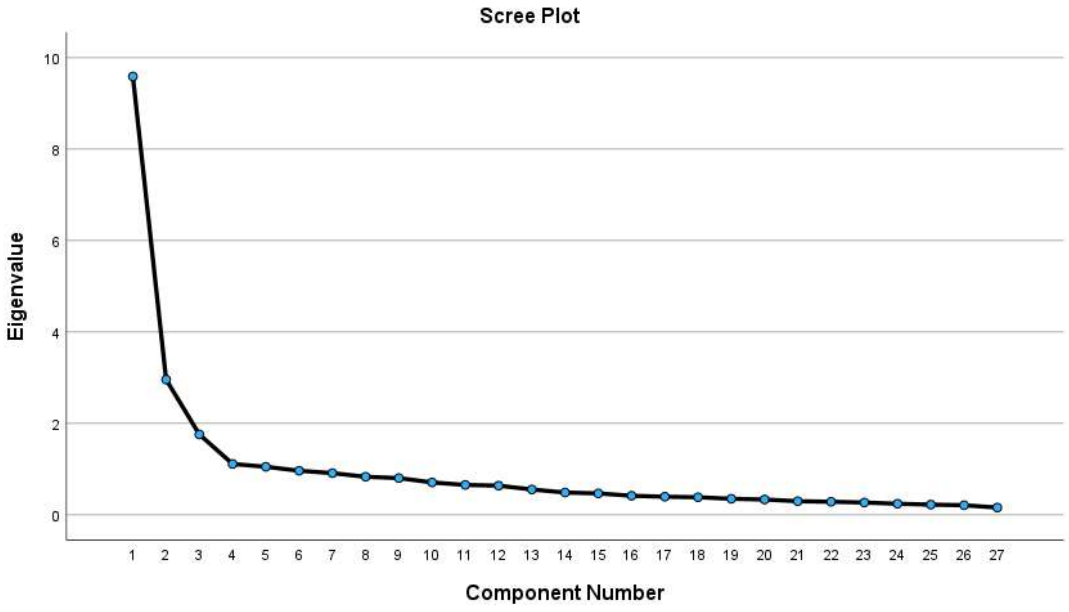
Factor Loadings from Final PCA

Item	Factor 1-Positive Wellbeing	Factor 2-EBD	Factor 3-Anxiety
1-Have you been interested in doing things or meeting people?	***.695***		
5- Have you felt like you can make friends?	***.724***		
7- Have you felt you are a good person?	***.518***		-.438
14- Have you looked forward to things?	***.745***		
15- Have you been able to stand up for yourself?	***.539***		
16- Have you felt you can do things as well as other people?	***.539***		
20- Have you been able to tell people how you feel?	***.567***		
24- Have you felt happy with your life?	***.741***		
26- Have you felt people love or care about you?	**.768**		
27- Have you been able to show people you love or care about them?	**.673**		
29- Have you felt happy?	**.781**		
2- Have you felt sad?	*-.338*	**.534**	.337
3- Have you felt angry?		**.753**	.357
6- Have you felt annoyed?		**.661**	.348
9- Have you felt like smashing things?		**.664**	
11- Have you been sleeping less than usual? If no, have you been sleeping more than usual?		**.516**	
18- Have you felt like you are no good?	*-.434*	**.420**	.372
19- Have you felt like hitting someone?		**.582**	
23- Have you had a bad temper?		**.743**	
25- Have you thought about death or dying?	*-.517*	**.337**	
28- Have you felt wound up?		**.683**	.360
4-Have you felt frightened of things or places?			**.779**
8-Have you suddenly felt scared?		.307	**.705**
10-Have you felt anxious?		.520	**.563**
13- Have you checked things over and over again?			**.517**
17- Have you felt faint or dizzy?	*-.331*		**.503**
21- Have you stayed away from some places or things because you are frightened of them?			**.776**

Bold- highlights the highest positive factor loading.*(Highest positive factor loadings used as a positive scoring system is used for the measure (no negatively scored questions)).*

Figure 2

Scree Plot Analysis



Step Seven

CFA Assumption Testing.

CFA was run on a second dataset (n= 242) to test the proposed three-factor model. Analysis of minimum and maximum values was conducted to assess for extreme outliers and suspicious data, of which none were identified. No items showed endorsement of <5% suggesting acceptability. KMO values of .907 were identified suggesting 'marvellous' adequacy (>.9; Kaiser, 1974). The assumptions of CFA were met suggesting suitability of the dataset.

A correlation matrix revealed that the majority of variables were above 0.3, suggesting the data was a good fit for CFA. Bartletts test showed no concerns regarding homogeneity of variance ($\chi_2(406)= 2795.35$, p=<.001). The determinant value of the CFA (2.62, n=242) dataset was considered non-problematic (>.00001).

Table 7

Results from Mardia's Test of Skewness and Kurtosis

	B	Z	p-value
Skewness	183.10	6988.64	>0.05
Kurtosis	995.17	17.16	>0.05

B= Beta coefficient
Z= Z-score

Tests of Normality.

No item showed kurtosis scores >3 and all variables were within the acceptable range (-8 to 8, Kline, 2011). Items 9, 17, 19, 25, and 26 showed 'moderate skew' (<-1 or >1, Ghaleb & Yaslioglu, 2024). Mardia's coefficient was above the critical value of 3.29 (Wulundari et al., 2021) suggesting the data may not be normally distributed (see Table 7). Histograms and Q-Q plots were visually examined to confirm non-normality (Appendix G). To account for non-normality and heteroscedasticity (unequal variance of residuals across variables), a Satorra-Bentler adjustment was included in the CFA (Brosseau-Liard & Savalei, 2014).

Confirmatory Factor Analysis

The three-factor model was specified with items 1, 5, 7, 14, 15, 16, 20, 24, 26, 27, and 29 relating to factor 1 (Positive Wellbeing), items 2, 3, 6, 9, 11, 18, 19, 23, 25, and 28 relating to factor 2 (EBD) and factor 3 (Anxiety) relating to items 4, 8, 10, 13, 17, and 21. Items with double factor loadings (>.3) were placed into the domain that had the highest positive factor loading. Items 12 and 22 were removed from the model based on the PCA results. A one-factor model (27 items) and a two-factor model (factor 1-positive wellbeing and factor 2- EBD & anxiety) were also run for comparison with the proposed three-factor model to identify the best fit for the data.

The three-factor, 27-item model was found to be the optimal fit (see Table 8). The results showed a CFI of .89 suggesting a 'relatively good' fit (Bentler, 1990). The TLI (.88)

was found to be marginally below the threshold of good fit (>.9). RMSEA (0.058) indicated an acceptable fit based on Fabrigar et al., (1999) criteria (see methods). This was supported in the three-factor model (SRMR=.075) where values below .08 are considered appropriate (Hu & Bentler, 1999). Taken together, this sample had an acceptable fit to the proposed three factor model.

A path diagram was developed (Figure 3) to illustrate the associations modelled by the three factors (27-item) CFA. It can be seen that factor one correlated negatively with factors two (r=0.61) and three (r=-0.49). Whilst factors two and three correlated positively (r=.74).

Table 8

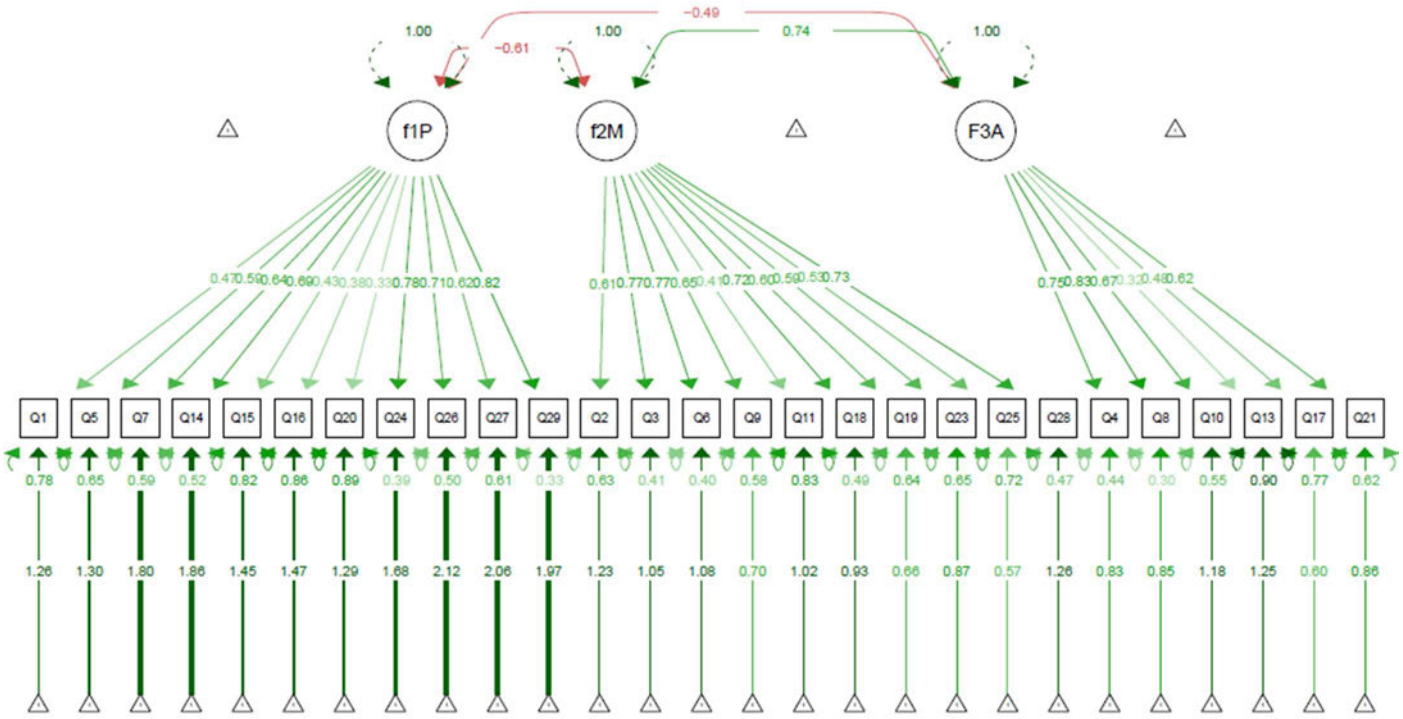
Fit Statistics of the CFA Comparing Four Proposed Models

Model	χ^2	**Df**	***P***	**CFI**	**TLI**	**RMSEA**	**SRMR**	**BIC**
CFA – 3 Factor Model, 27-Items	*569.01*	321	< .001	.89	.88	.058	.075	*17653.15*
CFA- 1 Factor Model, 27-items	*942.68*	324	<.001	.73	.68	.091	.096	*18093.24*
CFA- 2 Factor Model, 27-Items	*654.10*	323	<.001	.85	.84	.067	.075	*17742.45*
CFA- 3 Factor Model, 25 Items**	*502.18*	272	<.001	.89	.88	.061	.077	*16153.96*

***indicates an additional model tested based on findings from the reliability analysis.*
CFA=Confirmatory Factor Analysis.

Figure 3

Path Diagram of Three-Factor, 27-item CFA.



F1P = Positive Wellbeing
F2M= Emotional and Behavioural Discomfort
F3A= Anxiety

Step Eight

Internal Consistency

Reliability testing was completed on a third data set (n=879). Demographic data can be seen in Table 3. Based on George & Mallary's criteria, factor one (positive wellbeing) was found to show good internal consistency (α=.882). Factor two (EBD) showed good internal consistency (α=.853), which would rise to .859 if question 11, "Have you been sleeping less than usual? If no, have you been sleeping more than usual?", was removed. Factor three (anxiety) showed acceptable internal consistency (α=.756) which would rise to .776 if item 13, "Have you checked things over and over again?", was deleted. To test the possibility of removing items 11 and 13, a fourth model was inputted into the CFA to consider how a 25-

item, three factor structure would compare to the other models (see Table 8). The three-factor model with 27-items remained the best fit, therefore, items 11 and 13 were retained.

Step Nine

Concurrent Validity

Concurrent validity was assessed using a fourth data set (n=78). This included all participants who had completed both the PTOS-ID-II and BSI. Correlational analysis of the psychological distress scores of the newly proposed measure (27 items) with the Global Severity Index (GSI) of the Brief Symptom Inventory (BSI) was run. A significant positive relationship was found (r= .903, p=<0.001) suggesting 'strong' concurrent validity (Cohen, 1988).

Predictive Validity

Receiver Operating Characteristic (ROC) analysis was used to consider sensitivity and specificity of the proposed 27-item measure (n=78). The GSI threshold (>0.62) of the BSI was used as a comparison threshold. Area under the curve analysis (AUC=.945, p=<.0001) demonstrated 'outstanding' discrimination. Three proposed cut-offs had sensitivity and specificity scores above the recommended threshold of .80 (see Table 9). The ROC analysis identified an optimum clinical cut-off of 11.5 (sensitivity=.851; specificity=.903); 11.5 was chosen as it had the highest sensitivity and specificity values. Of the 78 participant responses, 44 would score above the clinical threshold of psychological distress (56.4%) on the PTOS-ID-II based on a cut-off of 11.5.

The authors of the PTOS-ID-II differentiated clinical thresholds based on gender (N. Beail, personal communication, February 6th, 2024). As such, further ROC analysis was completed on females (n= 28) and males (n=33) separately. ROC analysis of female participants identified a cut-off of 10.5 or above when compared against both the BSI threshold (Sensitivity=.889, Specificity=1) and the PTOS-ID-II threshold (Sensitivity=1,

Specificity=1), showing outstanding discrimination. For males, the ROC analysis identified scores above 13.5 to be most appropriate compared against the BSI (Sensitivity=.850, Specificity=.846) and the PTOS-ID-II threshold (Sensitivity=1, Specificity=1), again showing 'outstanding' discrimination (Figure 5).

Table 9

ROC Coordinates for Sensitivity and Specificity of Proposed PTOS (27-item) against BSI Threshold

Positive if Greater Than or Equal to	Sensitivity	Specificity
-1	1	0
.50	1	.097
1.50	.979	.161
2.50	.979	.355
3.50	.979	.419
4.50	.979	.516
5.50	.979	.581
6.50	.957	.677
8.00	.957	.774
9.50*	.872	.871
11.50**	.851	.903
13.50*	.830	.935
15.00	.745	.935
16.50	.660	.968
17.50	.660	1
18.50	.617	1
19.50	.532	1
20.50	.426	1
21.50	.404	1
22.50	.384	1
24.50	.340	1
27.00	.277	1
29.00	.191	1
31.50	.149	1
37.50	.043	1
43.00	.000	1

**Highlights scores with sensitivity and specificity above .80.*

*****Highlights** highest sensitivity and specificity values*

Acceptability

Little's Missing Completely at Random (MCAR) test was completed on the BSI data to provide a baseline comparison (n=78). This was found to be non-significant (x^2= 82.2, df=102, *p*=.925), suggesting there are no patterns in the missing data. The maximum missing data for any one question was 2.6%.

Littles MCAR test was then run on the PTOS-ID-II dataset (n=879). This was also found to be non-significant (x^2= = 640.550, df=592, *p*=.082). The maximum percentage of missing data for any item was 1.2%. This is below the comparison of the BSI, suggesting acceptability.

Overall, the 27-item model showed evidence of reliability and validity. To distinguish this newly validated measure from the PTOS-ID-II, it was renamed 'Outcomes of Wellbeing and Distress Scale' (OWLS-ID) and will be referred to as such from this point on.

Results: Stage 2-Abbreviation

1) Screening Tests

Any items that failed two or more tests in the validation stage were removed. This led to the instant removal of item 11 '"Have you been sleeping less than usual? If no, have you been sleeping more than usual?' and 13 'Have you checked things over and over again?'.

2) Identifying Key Questions

Using the final PCA seen in Table 6b (n=184), key questions were identified by high factor loadings (>.6). Factor loadings below this threshold were removed, this led to the removal of 12 items.

3) Correlational and Regression Analyses

Intercorrelations (n=184) identified items which correlated too poorly (<.3) or too highly (<.7). Item 29 'Have you felt happy?' highly correlated (rho=.70, *p*=<0.001) with item 24 'Have you felt happy with your life?' suggesting they were likely measuring a similar

phenomenon and only one item should remain. Item 24 had the lowest factor loading, therefore, was removed.

Correlations between items and the construct they were intending to measure (positive wellbeing or psychological distress; Appendix H) were run. Items 1, 7, 15, 16, 20 & 27 insufficiently correlated (≤.7) with wellbeing; items 4, 8, 9, 17, 19, 21 & 25 insufficiently correlated with distress, and thus removed.

Regression analysis assessed the extent to which the remaining distress related items could predict overall psychological distress scores (n=879). A significant regression was found ($F(6, 868) = 651.16$, $p = <.001$). The R^2 was .818, indicating that distress related items explained approximately 81.8% of the variance in psychological distress scores. Item 2 'Have you felt sad?' was found to be the highest predictor of overall psychological distress scores (β=2.629).

Table 10a

Results of the Regression on Distress Items

	R	R Square	Adjusted R Square	Std Error of the Estimate	R Square Change	F Change	df 1	df2	Sig F Change	Durbin Watson
Model 1	.905	.818	.817	4.770	.818	651.158	6	868	<.001	1.800

df= degrees of freedom

Table 10b

Results of the Regression on Wellbeing Items

	R	R Square	Adjusted R Square	Std Error of the Estimate	R Square Change	F Change	df 1	df2	Sig F Change	Durbin Watson
Model 1	.930	.865	.864	3.057	.865	1390.25	4	869	<.001	1.898

df= degrees of freedom

Regression was run on remaining wellbeing items (see Table 10b). A significant regression was found ($F(4,869) = 1390.25$, $p = <.001$). The R^2 was .865, indicating that wellbeing related items explained approximately 86.5% of the variance in positive wellbeing scores. Item 29 'Have you felt happy?' was found to be the highest predictor of overall wellbeing scores (β=2.630).

Items 5, 14, 26 and 29 (positive wellbeing items) and items 2, 3, 6, 23, 28 & 10 (psychological distress items across the EBD and anxiety domains) remained. This formed the abbreviated ten-item measure coined the OWLS-Mini.

Validation (OWLS-Mini)

Concurrent validity was assessed via correlational analysis of the psychological distress scores on the OWLS-Mini with the GSI score of the BSI (n=78). A significant positive relationship was found (r= .806, p=<0.001). Correlation analysis was also run between the OWLS-ID and OWLS-Mini showing a significant positive relationship (r= .933, p=<0.001) suggesting strong concurrent validity (Cohen, 1988).

PCA with varimax rotation was run to test the factor structure of the OWLS-Mini (n=879). Assumptions were tested prior to the PCA showing suitability of the data (KMO=.88; Bartletts test of sphericity, p=<.001). This identified two clear factors with Eigenvalues above 1 which accounted for 59.36% of the variance (see Table 11). The two factors were labelled "Positive Wellbeing" and "Psychological Distress".

Table 11a

PCA Results of the OWLS-Mini

Factor	Eigenvalue	% of Variance	Cumulative %
1*	3.33	33.32	33.32
2*	2.60	26.03	59.36

Bold*-Extracted Factors

Table 11b

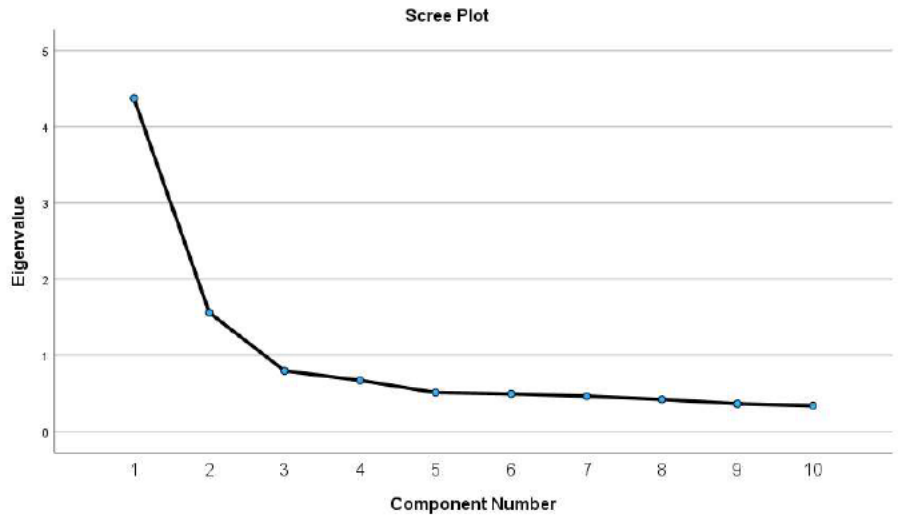
Factor Loadings of PCA on OWLS-Mini

Item	Factor 1- Distress	Factor 2- Wellbeing
2- Have you felt sad?	**.587**	-.314
3- Have you felt angry?	**.800**	-.134
6- Have you felt annoyed?	**.759**	-.196
10-Have you felt anxious?	**.648**	-.210
23- Have you had a bad temper?	**.721**	-.102
28- Have you felt wound up?	**.790**	-.131
5- Have you felt like you can make friends?	-.083	***.752***
14- Have you looked forward to things?	-.108	***.836***
26- Have you felt people love or care about you?	-.249	**.728**
29- Have you felt happy?	-.357	**.763**

Bold- shows highest positive factor loading

Figure 4

Scree Plot Analysis



ROC analysis (Figure 5) was run on the available BSI data (n=78). This identified scores above 7.5 on the OWLS-Mini to be indicative of scores above the clinical threshold of the BSI (Sensitivity=.809, Specificity=.839), with 'excellent' discrimination (AUC=.883). This was supported by ROC analysis of the PTOS-ID-II and OWLS-Mini which also suggested a cut-off of 7.5 or above (Sensitivity= .943, Specificity= .923) demonstrating

'outstanding' discrimination (AUC=.975). Based on the sample of 78 individuals, this would suggest that 43 (55.12%) would score above the clinical threshold of psychological distress.

ROC analysis was run based on gender. No statistically appropriate cut-offs for males were identified. Therefore, it is suggested that gender differences are not used to categorise distress in the abbreviated measure.

Reliability analysis (N=879) found factor one (psychological distress) and factor two (positive wellbeing) were shown to have good internal consistency (α=.836 and α=.807 respectively). The OWLS-ID and OWLS-Mini were assessed against Cahill's scoring of psychometric quality (see Table 12).

Figure 5

ROC curve of OWLS-ID and OWLS-Mini against GSI threshold (indices of BSI).

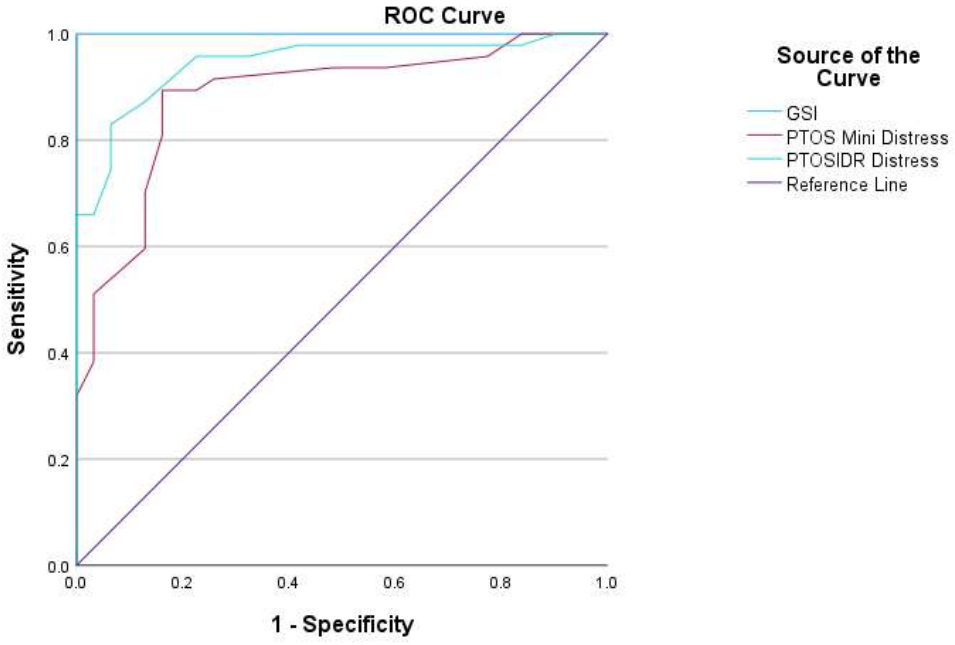


Table 12

OWLS Measures against Fitzpatrick's (1998) Criteria using Cahill's (2008) Scoring

	Cahill's Scoring	
Fitzpatrick's Criteria	**OWLS-ID**	**OWLS-Mini**
Reliability Produces consistent results from the same respondents over time.	• Internal Consistency was considered adequate (>0.7) for all three factors.	• Internal consistency was considered adequate (>0.7) for both factors
Validity The extent to which a measure really measures the concept that it purports to measure.	• Validity considered adequate based on evidence of 3 or more reports validity tests. • Construct validity via PCA and CFA. • 'Excellent' Concurrent validity (r=.903, *p*=<0.001) • Predictive criterion validity= 'outstanding' discrimination (AUC=.945, *p*=<0.001)	• Validity was considered adequate based on evidence of 3 or more reports of validity tests. • Construct validity tested by PCA identifying a 2-factor model. • Concurrent validity tested against BSI was 'good' (r=.806, *p*=<0.001) • Predictive criterion validity= 'excellent' discrimination (AUC=.883, *p*=<0.5) • Regression analysis showed all items predicted the domain they intended to measure.
Responsiveness Does the instrument detect changes over time that matter to the patient? Discriminative (between individuals) or evaluative (within individuals).	• Excellent discrimination identified within and between individuals through ROC analysis with BSI data. • Changes over time to be investigated further, with fixed time points for comparison.	• Excellent discrimination identified through ROC analysis with BSI and OWLS-ID data. • Within individual changes to be investigated further to consider changes over time.
Acceptability Is the measure acceptable to users?	Adequate acceptability of the measure based on components below: • Practicality of administration addressed via littles MCAR test. This identified lower cases of missing data in OWLS-ID compared to BSI suggesting acceptability of the measure. • Instrument has been reduced to 27-items based on feedback from professional consultation. • Previously been translated to different languages for research and clinical purposes. • Some evidence of access by ethnic minorities • Reading age considered (simplified sentences) Administration instructions include being read aloud verbatim by the administrator.	• Adequate acceptability of the measure based on all components below being considered. • Questions are taken from an already acceptable measure therefore has face validity. • It is also based on data from this population therefore acceptability can be assumed. • Shorter measure for quick administration. • Shortened administration time. • Can be easily translated into other languages. • Simplified language used, should be administered by a clinician/researcher.

	Cahill's Scoring	
Fitzpatrick's Criteria	**OWLS-ID**	**OWLS-Mini**
Feasibility Is the measure easy to administer and process?	• Considered adequate based on Cahill's criteria. • Free and easily accessible for clinicians. • Can be completed online/via paper. • Scoring system is simple and has been improved for OWLS-ID. • Item dispersion and missing data analysis identified no issues with feasibility. • Anecdotal evidence from clinic suggested the measure was long which led to the development of the OWLS-Mini. • Training packages can be provided. Clear scoring instructions on the measure.	• Remains free to use and easy to access. • Questionnaire is shorter than then the already feasible OWLS-ID measure therefore feasibility can be assumed. • Can be completed online/via paper. • Scoring system simplified. • Clear scoring instructions provided.
Precision and Interpretability How precise is the measure? How interpretable are the scores of the measure?	• Considered as 'Partially Adequate' as most components considered. Further consideration of population norms for the wellbeing scale are needed. • Overall distress score which can be compared to clinical thresholds. • Optimal clinical cut-offs with sensitivity and specificity above .80. • These were also found to differ based on gender. • Wellbeing scores have no threshold or population norms which is an area for future research.	• Considered as 'Partially Adequate' as most components considered. Further consideration of population norms for the wellbeing scale are needed. • Optimal clinical cut-offs with sensitivity and specificity above .80. • Overall distress score which can be compared to clinical thresholds. • Wellbeing scores have no threshold or population norms which is an area for future research.

Discussion

The current study aimed to (i) validate the psychometric properties of the PTOS-ID-II and (ii) abbreviate the resultant measure to provide a quick to administer PROM. The validation of the PTOS-ID-II identified two redundant items suggesting a more robust 27-item measure. The new 27-item measure was coined the 'Outcomes for Wellbeing and Distress Scale' (OWLS-ID). A new name was proposed to differentiate between historical scores of the PTOS-ID-II recorded in clinical notes. The abbreviation stage led to a ten-item measure coined the OWLS-Mini.

Fitzpatrick's (1998) and Cahill's (2008) criteria were used to guide the psychometric domains to be assessed and rate the quality of each domain (see Table 12). Comparing the current findings to those of previous reviews (Gourley & Yates, 2022; Vlissides et al., 2016), the OWLS-ID appears to be the most robustly tested measure showing evidence of psychometric properties across six of the seven domains. Furthermore, the OWLS-ID has been developed directly for populations with ID, thus should better reflect the populations' experiences of wellbeing and distress. That said, individuals with ID have not been directly consulted to inform the development of these measures, therefore, gaining patient feedback on the measure is advised as an important next step.

Optimal distress scores on the OWLS-ID for females and males were identified. This is in line with research suggesting gender differences in the self-reporting of distress in both the general population and groups with ID (Chester et al., 2013; Dekker et al., 2007; Poutanen et al., 2009;). Cooper et al., (2007) reported more significant life events and lower ability in females with ID suggesting females maybe at higher risk of distress. Whilst prevalence estimates of affective disorders are higher in females (Cooper et al., 2007; Esbensen et al.,2005; Lunsky, 2003), which may explain higher levels of symptom reporting. Conversely, others reported gender was not significantly associated with depression or have had mixed results (Bond et al., 2020; Chester et al., 2013; Reiss & Benson, 1985). As such, when using the OWLS-ID, clinicians are encouraged to exercise clinical judgment in deciding whether to use the generic cut-off (11.5) or distinguish by gender (male-13.5; female-10.5). Specificity and sensitivity are acceptable for both. The ability to retain a generic cut-off is beneficial as this will sometimes be necessary e.g., information on gender is unknown or an individual does not identify with binary gender assignment.

Strengths and Limitations

A key advantage to the current study was the large sample (n=879) considering the previous validation of the PTOS took place with 175 adults (Vlissides et al., 2017). The large sample allowed additional rigour to be implemented into the analysis. The data was divided into three parts allowing the model to be tested on separate participants thus inviting independence into the analysis. However, it is recognised that the decision to promote independence was at a cost to power, appreciating there is a trade-off between the two.

Clear limitations regarding the sample come from the naturalistic dataset, which involved missing data. For example, test-retest reliability was not completed as it was unclear how long participants had waited between first and second administration of the PTOS-ID-II. Marx et al., (2003) recommends an interval of two-weeks for test-retest reliability, as longer periods increase the likelihood of external influence on the scores. As such, valid test-retest analysis could not be completed as confounding factors such as the commencement of therapy between timepoints was likely.

The validation procedure was based on a framework proposed by Boateng et al., (2018) whilst the abbreviation procedure was based on the framework provided by Barkham et al., (2013). Alternative analyses were considered such as Item Response Theory (IRT) which is a complex statistical modelling technique requiring specific statistical software and large sample sizes (Yang & Kao, 2014) at the cost of independence. The decision to not use these approaches could be considered a limitation.

Missing data is expected with practice-based data (Marino et al., 2021), which can provide limitations. Littles MCAR test showed data to be ‘missing completely at random’. This suggests the probability of items being missed is equal for all observations, thus deemed to provide unbiased results and further imputation methods supposed not necessary (Marino et al., 2021; Pederson et al., 2017). Missing data was more noticeable in the reporting of

demographics particularly ethnicity, level of ID, and IQ. The sample appears to consist of predominately white individuals with a mild to moderate ID, however, due to high levels of missing data for these domains, this cannot be certain. Based on statistics from the Department of Levelling up, this appears to reflect the population of the northern town it was pooled. The survey identified 95.5% of the town's population identified as white, 1% identified as black, and 1% as mixed/multiple ethnic groups (GOV.UK, 2024). Regardless, this may not be reflective of populations with ID nationally/internationally thus identifying issues of generalisability.

The lack of PPI is a weakness. While professionals in the field were consulted, service users and carers were not. This should especially be considered in the context of the measure being a PROM. While a service evaluation was completed on the acceptability of the PTOS (N. Beail, personal communication, July 26th, 2022), this has not been published and was undertaken after the PTOS was developed. An important step in the continued development of the OWLS-ID and OWLS-Mini is to assess its acceptability to people with ID. One important question in addition to their thoughts on the items, is how they would like it to be used in practice. This may also have important implications for figure 6. Patients' views will be assessed when trialling the OWLS measures in practice to ensure the measure is appropriate for the population it intends to serve. This could be done via questionnaire, interviews or verbal feedback from patients following use of the questionnaire.

Implications

The OWLS-ID and OWLS-Mini could be implemented in practice for a variety of reasons, such as to assess individuals entering a service, track changes over time, assess levels of distress, or guide treatment decisions. These uses are evidenced by implementation of the earlier version of the measure. As seen in Table 13 and Figure 5, the OWLS-ID and

OWLS-Mini have different characteristics allowing them to be applied in clinic to their strengths. The OWLS-ID is a more comprehensive measure of distress and wellbeing and may be best implemented at initial assessment stages. Whereas the OWLS-Mini can identify the presence of distress in future sessions to reduce burden while monitoring progress. The two measures can therefore support each other in clinical utility. Research in populations with ID often takes place in clinical practice, therefore, the implementation of routine outcome monitoring is beneficial for both clinical and research purposes (Beail, 2017).

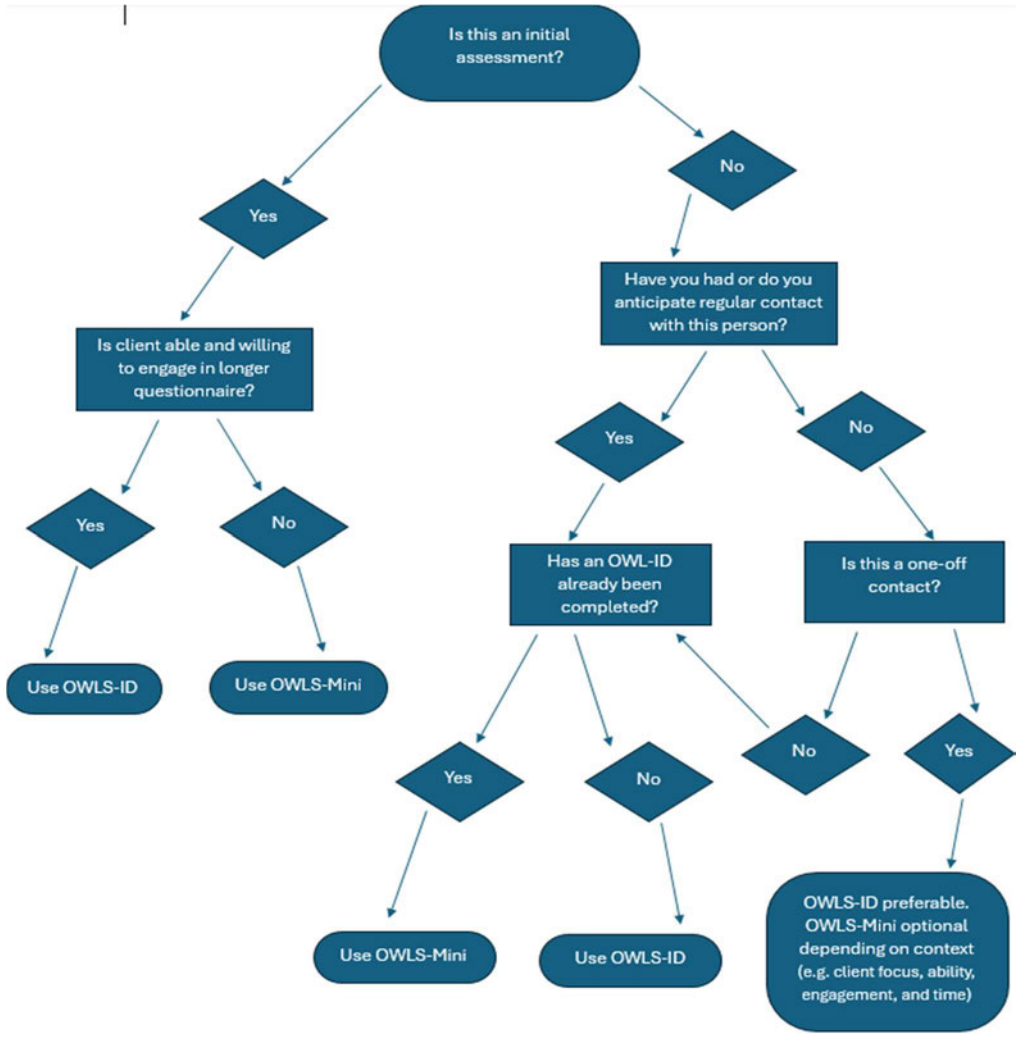
Table 13

Characteristics of the OWLS-ID versus the OWLS-Mini

OWLS-ID (27-item)	OWLS-Mini (10-item)
Fully validated on a large data set.	Partially validated.
Good psychometric properties evident.	Good psychometric properties evident.
Can be time consuming to complete.	Is quick to administer in clinic.
Provides greater accuracy in distress scores.	Identifies the presence of distress.
Can discriminate distress by gender.	Clinical threshold not affected by gender.
Provides more in-depth information.	Provides overview. OWLS-ID needed for more detailed information.
Ideal for use in initial appointments	Ideal as a between session measure

Figure 6

Flowchart Depicting Decision Making Processes when Choosing an Outcome Measure



Future Research

Test-retest reliability to assess stability over time is recommended for both OWLS measures, this would provide evidence of the final remaining domain of Fitzpatrick's (1998) criteria (Responsiveness). Moreover, administration in non-clinical populations could be conducted to assess the measures' ability to discriminate between clinical and non-clinical presentations. In the current sample, 56% of participants scored above the clinical threshold on the OWLS measures, which is expected to be lower in a non-clinical sample.

Development of population norms of the wellbeing indices is recommended. At present the wellbeing score provides the clinician a balance of strengths and difficulties a

client is experiencing. However, the addition of population norms would provide an illustration of how an individual's wellbeing score compares to that of other people in the population, adding contextual meaning to the value.

Conclusion

The OWLS-ID is currently the most validated outcome measure specifically developed for individuals with ID. While the OWLS-Mini is the first abbreviated measure of wellbeing and distress in ID. Both measures show good psychometric properties including construct validity, concurrent validity, internal reliability, sensitivity, and specificity. The OWLS-ID and OWLS-Mini should be used in clinical practice in place of the previous PTOS-ID-II. Further research investigating the measures responsiveness over time is recommended as well as feedback from patients using the measure.

References

Adawi, M., Zerbetto, R., Re, T. S., Bisharat, B., Mahamid, M., Amital, H., Del Puente, G., & Bragazzi, N. L. (2019). Psychometric properties of the Brief Symptom Inventory in nomophobic subjects: insights from preliminary confirmatory factor, exploratory factor, and clustering analyses in a sample of healthy Italian volunteers. *Psychology Research and Behavior Management*, *12*, 145–154. https://doi.org/10.2147/PRBM.S173282

Akoglu, H. (2018). User's guide to correlation coefficients. *Turkish Journal of Emergency Medicine*, *18*(3), 91-93. https://doi.org/10.1016/j.tjem.2018.08.001

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

Barkham, M., Bewick, B., Mullin, T., Gilbody, S., Connell, J., Cahill, J., Mellor-Clark, J., Richards, D., Unsworth, G., & Evans, C. (2013). The core-10: A short measure of psychological distress for routine use in the psychological therapies. *Counselling and Psychotherapy Research*, *13*(1), 3–13. https://doi.org/10.1080/14733145.2012.729069

Barkham, M., Margison, F., Leach, C., Lucock, M., Mellor-Clark, J., Evans, C., Benson, L., Connell, J., Audin, K., & McGrath, G. (2001). Service profiling and outcomes benchmarking using the CORE-OM: Toward practice-based evidence in the psychological therapies. *Journal of Consulting and Clinical Psychology*, *69*(2), 184–196. https://doi.org/10.1037//0022-006x.69.2.184

Beail, N. (2016). *Psychological therapies and people who have intellectual disabilities.* Leicester: British Psychological Society.

Beail, N. (2017). Psychological therapies and people who have ID: A report from the Royal College of Psychiatrists and British Psychological Society. *Advances in Mental Health and Intellectual Disabilities*, *11*(1), 24–26. https://doi.org/10.1108/amhid-09-2016-0021

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),* Seattle Club Conference, Kendal, UK.

Beail, N., Kerry, E., Rawlings, G. H., & Vlissides, N. (2023). Psychological therapies outcome scale for people who have intellectual disabilities. *FPID Bulletin: The Bulletin of the Faculty for People with Intellectual Disabilities*, *21*(3), 64–65. https://doi.org/10.53841/bpsfpid.2023.21.3.64

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*(2), 238–246. https://doi.org/10.1037/0033-2909.107.2.238

Boat, T. F., Wu, J. T., Institute Of Medicine (U.S.). Committee To Evaluate The Supplemental Security Income Disability Program For Children With Mental Disorders, & National Research Council (U.S.). Center For Education. (2015). *Mental disorders and disabilities among low-income children*. National Academies Press.

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and Behavioral Research: A Primer. *Frontiers in Public Health*, *6*. https://doi.org/10.3389/fpubh.2018.00149

Bond, L., Carroll, R., Mulryan, N., O'Dwyer, M., O'Connell, J., Monaghan, R., Sheerin, F., McCallion, P., & McCarron, M. (2020). Biopsychosocial factors associated with depression and anxiety in older adults with intellectual disability: results of the wave 3 Intellectual Disability Supplement to The Irish Longitudinal Study on Ageing. *Journal of Intellectual Disability Research*, *64*(5), 368–380. https://doi.org/10.1111/jir.12724

Boone, W. J. (2016). Rasch Analysis for Instrument Development: Why, When, and How?. *CBE Life Sciences Education*, *15*(4). https://doi.org/10.1187/cbe.16-04-0148

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*(4), 320–329. https://doi.org/10.1111/bld.12020

Brosseau-Liard, P. E., & Savalei, V. (2014). Adjusting incremental fit indices for nonnormality. *Multivariate Behavioral Research*, *49*(5), 460-470. https://doi.org/10.1080/00273171.2014.933697

Byrne, B. M. (2010). Structural equation modeling with AMOS: Basic concepts, applications, and programming. New York: Routledge.

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. https://doi.org/10.3310/hta12240

Cain, M. K., Zhang, Z., & Yuan, K. (2017). Univariate and multivariate skewness and kurtosis for measuring nonnormality: Prevalence, influence and estimation. *Behavior Research Methods, 49*(5), 1716–1735. https://doi.org/10.3758/s13428-016-0814-1

Chester, R., Chaplin, E., Tsakanikos, E., McCarthy, J., Bouras, N., & Craig, T. (2013). Gender differences in self-reported symptoms of depression and anxiety in adults with intellectual disabilities. *Advances in Mental Health and Intellectual Disabilities, 7*(4), 191-200. https://doi.org/10.1108/AMHID-03-2013-0025

Churruca, K., Pomare, C., Ellis, L. A., Long, J. C., Henderson, S. B., Murphy, L. E. D., Leahy, C. J., & Braithwaite, J. (2021). Patient-reported outcome measures (PROMs): A review of generic and condition-specific measures and a discussion of trends and issues. *Health Expectations*, *24*(4), 1015–1024. https://doi.org/10.1111/hex.13254

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences (2nd ed.). *Journal of the American Statistical Association*, *84*(408). https://doi.org/10.2307/2290095

Cooper, S. A., Smiley, E., Morrison, J., Williamson, A., & Allan, L. (2007). "An epidemiological investigation of affective disorders with a population-based cohort of 1023 adults with intellectual disabilities", *Psychological Medicine, 37*(6), 873-82. https://doi.org/10.1017/S0033291707009968

Costello, A. B., & Osborne, J. (2019). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research, and Evaluation*, *10*(1), 7. https://doi.org/10.7275/jyj1-4868

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. *The British Journal of Psychiatry: The Journal of Mental Science*, *182*, 347–353. https://doi.org/10.1192/bjp.182.4.347

Davis, C., Kellett, S., & Beail, N. (2009). Utility of the Rosenberg self-esteem scale. *American Journal on Intellectual and Developmental Disabilities*, *114*(3), 172–178. https://doi.org/10.1352/1944-7558-114.3.172

Dekker, J., Koelen, J. A., Peen, J., Schoevers, R. A., & Gijsbers-Van Wijk, C. (2007). "Gender differences in clinical features of depressed outpatients: preliminary evidence for subtyping of depression?", *Women & Health, 46*(4), 19-38. https://doi.org/10.1300/j013v46n04_02

Derogatis, L. R. (1983). *SCL-90-R: Administration, Scoring and Procedures: Manual II.* Baltimore, MD: Clinical Psychometric Research.

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

Emerson, E., & Hatton, C. (2007). Mental health of children and adolescents with intellectual disabilities in Britain. *The British Journal of Psychiatry, 191*, 493–499. https://doi.org/10.1192/bjp.bp.107.038729

Esbensen, A. J., Seltzer, M. M., Greenberg, J. S., & Benson, B. A. (2005). Psychometric Evaluation of a Self-Report Measure of Depression for Individuals With Mental Retardation. *American Journal on Mental Retardation*, *110*(6), 469. https://doi.org/10.1352/0895-8017(2005)110[469:peoasm]2.0.co;2

Evans, C., Mellor-Clark, J., Margison, F., Barkham, M., Audin, K., Connell, J., & McGrath, G. (2000). CORE: Clinical Outcomes in Routine Evaluation. *Journal of Mental Health*, *9*(3), 247–255. http://doi.org/10.1080/jmh.9.3.247.255.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, *4*(3), 272. https://doi.org/10.1037/1082-989X.4.3.272

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 (Winchester, England)*, *2*(14), https://doi.org/10.3310/hta2140

Floyd, F. & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment, 7*, 286-299. https://doi.org/10.1037/1040-3590.7.3.286

George, D., & Mallery, P. (2018). Reliability analysis. In *IBM SPSS statistics 25 step by step* (pp. 249-260). Routledge.

Gerbing, D. W., & Hamilton, J. G. (1996). Viability of exploratory factor analysis as a precursor to confirmatory factor analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, *3*(1), 62–72. https://doi.org/10.1080/10705519609540030

Ghaleb, M., & Yaslioglu, M. (2024). Structural Equation Modeling (SEM) for Social and Behavioral Sciences Studies: Steps Sequence and Explanation. *Journal of Organizational Behavior Review, 6*. 69-108. 3559195 (dergipark.org.tr)

Goretzko, D., Siemund, K., & Sterner, P. (2024). Evaluating Model Fit of Measurement Models in Confirmatory Factor Analysis*. Educational and Psychological Measurement, 84*(1), 123-144. https://doi.org/10.1177/00131644231163813

Gourley, T. R., & Yates, L. (2022). "A review of recent clinical measures that can be used to support psychological therapies with adults with intellectual disabilities", *Advances in Mental Health and Intellectual Disabilities, 16*(3), pp. 157-168. https://doi.org/10.1108/AMHID-01-2022-0005

GOV.UK. (2024). *Local data profile*. Department of Levelling Up, Housing and Communities. Retrieved April 18, 2024, from https://assets.publishing.service.gov.uk/media/6602fbaaf1d3a09b1f32ad74/Barnsley.pdf

Greenacre, M., Groenen, P. J., Hastie, T., d'Enza, A. I., Markos, A., & Tuzhilina, E. (2022). Principal component analysis. *Nature Reviews Methods Primers*, *2*(1), 100. DOI:10.1038/s43586-022-00184-w

Griffin, M. M., & Steinbrecher, T. D. (2013). Large-scale datasets in special education research. In R. M. Hodapp, D. J. Fidler, R. M. Hodapp, & D. J. Fidler (Eds.), International review of research in developmental disabilities (pp. 155-183). Academic Press. https://doi.org/10.1016/B978-0-12-407760-7.00004-9

Guadagnoli, E., & Velicer, W. F. (1988). Relation of sample size to the stability of component patterns. *Psychological Bulletin*, *103*(2), 265. https://doi.org/10.1037/0033-2909.103.2.265

Hair, J. F., Babin, B., Anderson, R., & Black, W. (2018). *Multivariate data analysis* (8th ed.). Pearson.

Hair, J. F., Sarstedt, M., Pieper, T. M., & Ringle, C. M. (2012). The use of partial least squares structural equation modeling in strategic management research: a review of past practices and recommendations for future applications. *Long Range Planning*, *45*(5-6), 320-340. https://doi.org/10.1016/j.lrp.2012.09.008

Hosmer, D. W., & Lemeshow, S. (2000). *Applied logistic regression*. Wiley.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1-55. https://doi.org/10.1080/10705519909540118

Hughes-McCormack, L. A., Rydzewska, E., Henderson, A., MacIntyre, C., Rintoul, J., & Cooper, S. A. (2017). Prevalence of mental health conditions and relationship with general health in a whole-country population of people with intellectual disabilities

compared with the general population. *BJPsych Open*, *3*(5), 243–248. https://doi.org/10.1192/bjpo.bp.117.005462

IBM Corp. Released 2023. IBM SPSS Statistics for Windows, Version 29.0.2.0 Armonk, NY: IBM Corp

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, *39*(1), 31-36. https://doi.org/10.1007/BF02291575

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. https://doi.org/10.1046/j.1468-3148.2003.00152.x

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,* 323–34. https://doi.org/10.1111/j.1468-3148.1999.tb00088.x

Kellett, S., Beail, N., Newman, D. W., & Hawes, A. (2004). The Factor Structure of the brief symptom inventory: Intellectual disability evidence. *Clinical Psychology & Psychotherapy*, *11*(4), 275–281. https://doi.org/10.1002/cpp.410

Kline, R. B. (2011). Principles and practice of structural equation modeling (3rd ed.). The Guilford Press.

Lee K, Cascella M, Marwaha R. Intellectual Disability. [Updated 2023 Jun 4]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2024 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK547654/

Lunsky, Y. (2003). Depressive symptoms in intellectual disability: does gender play a role?. *Journal of intellectual disability research : JIDR*, *47*(Pt 6), 417–427. https://doi.org/10.1046/j.1365-2788.2003.00516.x

Makrinioti, H., Bush, A., & Griffiths, C. (2019). What are patient-reported outcomes and why they are important: improving studies of preschool wheeze. *Archives of Disease in Childhood - Education & Practice Edition*, *105*(3), 185–188. https://doi.org/10.1136/archdischild-2018-316476

Marino, M., Lucas, J., Latour, E., & Heintzman, J. D. (2021). Missing data in primary care research: importance, implications and approaches. *Family Practice*, *38*(2), 200–203. https://doi.org/10.1093/fampra/cmaa134

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. https://doi.org/10.1111/j.1468-3156.2006.00422.x

Marx, R. G., Menezes, A., Horovitz, L., Jones, E. C., & Warren, R. F. (2003). A comparison of two time intervals for test-retest reliability of Health Status Instruments. *Journal of Clinical Epidemiology*, *56*(8), 730–735. https://doi.org/10.1016/s0895-4356(03)00084-2

McCracken, L. M., & Dhingra, L. (2002). A short version of the Pain Anxiety Symptoms Scale (PASS-20): preliminary development and validity. *Pain Research and Management*, *7*, 45-50. https://doi.org/10.1155/2002/517163

Menard, S. (1995). Applied Logistic Regression Analysis: Sage University Series on Quantitative Applications in the Social Sciences. Thousand Oaks, CA: Sage.

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. https://doi.org/10.1046/j.1365-2788.2003.00457.x

Moss, S. (2019). *MPAS-ID: a wide-spectrum mental health assessment for adults who have limited language or reduced cognitive development: fully compliant with ICD-11 and DSM-5: formerly known as the Mini PAS-ADD*. Pavilion.

Mukaka, M. M. (2012). Statistics corner: A guide to appropriate use of correlation coefficient in medical research. *Malawi Medical Journal: The Journal of Medical Association of Malawi*, *24*(3), 69–71. A guide to appropriate use of Correlation coefficient in medical research - PMC (nih.gov)

Myers, R. H., & Myers, R. H. (1990). *Classical and modern regression with applications* (Vol. 2, p. 488). Belmont, CA: Duxbury press.

Novaco, R. W., & Taylor, J. L. (2004). Assessment of anger and aggression in male offenders with developmental disabilities. *Psychological Assessment*, *16*(1), 42–50. https://doi.org/10.1037/1040-3590.16.1.42

Pashak, T. J., Vanderstelt, B. H., Ling, W. L., & Nelson, O. M. (2022). Psychometrics of the Brief Symptom Inventory in University-Enrolled Emerging Adults: Factor Structure and Clinical Caseness Cutoff. *Psychology*, *13*(10), 1539–1551. https://doi.org/10.4236/psych.2022.1310097

Pedersen, A. B., Mikkelsen, E. M., Cronin-Fenton, D., Kristensen, N. R., Pham, T. M., Pedersen, L., & Petersen, I. (2017). Missing data and multiple imputation in clinical

epidemiological research. *Clinical Epidemiology*, *9*, 157–166. https://doi.org/10.2147/CLEP.S129785

Poutanen, O., Koivisto, A. M., Mattila, A., Joukamaa, M., & Raimo, K. R. (2009), "Gender differences in the symptoms of major depression and in the level of social functioning in public primary care patients", *European Journal of General Practice,15* (3),161-7. https://doi.org/10.3109/13814780903186423

Reiss, S., & Benson, B. A. (1985). Psychosocial correlates of depression in mentally retarded adults: I. Minimal social support and stigmatization. *PubMed*, *89*(4), 331–7. Psychosocial correlates of depression in mentally retarded adults: I. Minimal social support and stigmatization - PubMed (nih.gov)

Rosenberg, M. (1965). *Society and the Adolescent Self-image*. Princeton University Press, Princeton, NJ.

RStudio Team (2020). RStudio: Integrated Development for R. RStudio, PBC, Boston, MA URL http://www.rstudio.com/.

Schwartz, A. E., Kramer, J. M., & PEDI-PRO Youth Team. (2021). Inclusive approaches to developing content valid patient-reported outcome measure response scales for youth with intellectual/developmental disabilities. *British Journal of Learning Disabilities*, *49*(1), 100–110. https://doi.org/10.1111/bld.12346

Skelly, A. (2011). Outcomes measurement in routine practice for learning disability psychology services. *Clinical Psychology & People with Learning Disabilities, 9* (4), 2-6. DOI:10.53841/bpsfpid.2011.9.4.2

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. https://doi.org/10.1038/bjc.1988.20

Spitzer, R. L., Kroenke, K., Williams, J. B., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of Internal Medicine*, *166*(10), 1092–1097. https://doi.org/10.1001/archinte.166.10.1092

Statistics Solutions. (2013). Confirmatory Factor Analysis . Retrieved from https://www.statisticssolutions.com/academic-solutions/resources/directory-of-statistical-analyses/confirmatory-factor-analysis/

Steiger J. H. (1998). A note on multiple sample extensions of the RMSEA fit index. *Structural Equation Modeling: A Multidisciplinary Journal*, 5(4), 411–419. https://doi.org/10.1080/10705519809540115

Stevens, J. (2002). *Applied multivariate statistics for the social sciences* (Vol. 4). Mahwah, NJ: Lawrence Erlbaum Associates.

Tapp, K., Vereenooghe, L., Hewitt, O., Scripps, E., Gray, K. M., & Langdon, P. E. (2023). Psychological therapies for people with intellectual disabilities: An updated systematic review and meta-analysis. *Comprehensive Psychiatry*, *122*, 152-372. https://doi.org/10.1016/j.comppsych.2023.152372

Tinsley, H. E. A., & Tinsley, D. J. (1987). Uses of factor analysis in counselling psychology research. *Journal of Counselling Psychology, 34(4),* 414-424. https://doi.org/10.1037/0022-0167.34.4.414

Vlissides, N., Beail, N., Jackson, T., Williams, K., & Golding, L. (2017). Development and psychometric properties of the Psychological Therapies Outcome Scale (PTOS-ID). *Journal of Intellectual Disability Research, 61,* 549–559. https://doi.org/10.1111/jir.12361

Vlissides, N., Goulding, L., & Beail, N. (2016). *A systematic review of the outcome measures used in psychological therapies with adults with ID*, in Beail, N. (Ed), Psychological

Therapies and People Who Have Intellectual Disabilities, British Psychological Society, Leicester.

von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gøtzsche, P. C., & Vandenbroucke, J. P. (2007). The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement. *Epidemiology*, *18*(6), 800–804. https://doi.org/10.1097/ede.0b013e3181577654

Wechsler, D. (2008). *Wechsler Adult Intelligence Scale* (4th ed.). San Antonio, TX: Pearson Assessment.

Williams, B., Onsman, A., & Brown, T. (2012). A Rasch and factor analysis of a paramedic graduate attribute scale. *Evaluation & the Health Professions*, *35*(2), 148-168. https://doi.org/10.1177/0163278711407314

Willner, P. (2005). The effectiveness of psychotherapeutic interventions for people with learning disabilities: a critical overview. *Journal of Intellectual Disability Research, 49,* 75-85. https://doi.org/10.1111/j.1365-2788.2005.00633.x

Wulandari, D., Sutrisno, S., & Nirwana, M. B. (2021). Mardia's skewness and kurtosis for assessing normality assumption in multivariate regression. Enthusiastic: International *Journal of Applied Statistics and Data Science*, 1-6. DOI:10.20885/enthusiastic.vol1.iss1.art1

Yang, F. M., & Kao, S. T. (2014). Item response theory for measurement validity. *Shanghai Archives of Psychiatry*, *26*(3), 171–177. https://doi.org/10.3969/j.issn.1002-0829.2014.03.010

Appendices

Appendix A

PTOS-ID-II

ID NUMBER

PTOS-ID II

Psychological Therapies Outcome Scale

Intellectual Disabilities

2nd Edition

Name			
Gender	Male	Female	*Stage Completed* **S** Screening **A** Assessment **F** First Therapy Session **D** During Therapy **L** Last Therapy Session Stage ☐
Age			
IQ			
Reason for referral			
Date of completion			Episode of Care ☐
Therapist			

Important – Please Read

Please read each question asking how the individual has felt over the last week.
Ask individual to answer 'yes' or 'no'.
If 'yes', use response scale to assist individual to record frequency of how they have felt.

	OVER THE PAST WEEK...	Not at all	A little bit	Sometimes	A lot
1	Have you been interested in doing things or meeting people?	0	1	2	3
2	Have you felt sad?	0	1	2	3
3	Have you felt angry?	0	1	2	3
4	Have you felt frightened of things or places?	0	1	2	3
5	Have you felt like you can make friends?	0	1	2	3

OVER THE PAST WEEK...		Not at all	A little bit	Some times	A lot
6	Have you felt annoyed?	0	1	2	3
7	Have you felt you are a good person?	0	1	2	3
8	Have you suddenly felt scared?	0	1	2	3
9	Have you felt like smashing things?	0	1	2	3
10	Have you felt anxious?	0	1	2	3
11	Have you been sleeping less than usual? If no, have you been sleeping more than usual?	0	1	2	3
12	Have you been able to cope with problems?	0	1	2	3
13	Have you checked things over and over again?	0	1	2	3
14	Have you looked forward to things?	0	1	2	3
15	Have you been able to stand up for yourself?	0	1	2	3

Jackson, Beail & Vlissides (2011) Psychological Therapies Outcome Scale – ID II

OVER THE PAST WEEK...		Not at all	A little bit	Some times	A lot
16	Have you felt you can do things as well as other people?	0	1	2	3
17	Have you felt faint or dizzy?	0	1	2	3
18	Have you felt like you are no good?	0	1	2	3
19	Have you felt like hitting someone?	0	1	2	3
20	Have you been able to tell people how you feel?	0	1	2	3
21	Have you stayed away from some places or things because you are frightened of them?	0	1	2	3
22	Have you been eating more than usual? If no, have you been eating less than usual?	0	1	2	3
23	Have you had a bad temper?	0	1	2	3
24	Have you felt happy with your life?	0	1	2	3

Jackson, Beail & Vlissides (2011) Psychological Therapies Outcome Scale – ID II

OVER THE PAST WEEK...		Not at all	A little bit	Some times	A lot
25	Have you thought about death or dying?	0	1	2	3
26	Have you felt people love or care about you?	0	1	2	3
27	Have you been able to show other people you love or care about them?	0	1	2	3
28	Have you felt wound up?	0	1	2	3
29	Have you felt happy?	0	1	2	3

PTOS-ID II: Risk Screening

OVER THE PAST WEEK...		Not at all	A little bit	Sometimes	A lot
1	Have you hit another person?	0	1	2	3
2	Have you thought of hurting yourself?	0	1	2	3
3	Have you made plans to end your life?	0	1	2	3
4	Have you hurt yourself?	0	1	2	3
5	Have you threatened another person?	0	1	2	3

PTOS-ID II:
Scoring

Psychological Distress

Q2 +
Q3 +
Q4 +
Q6 +
Q8 +
Q9 +
Q10 +
Q11 +
Q17 +
Q18 +
Q19 +
Q21 +
Q22 +
Q23 +
Q25 +
Q28

Total
/16
=

Positive Well-Being

Q1 +
Q5 +
Q7 +
Q14 +
Q15 +
Q16 +
Q20 +
Q24 +
Q26 +
Q27 +
Q29

Total
/11
=

Q 12 and Q 13 do not contribute to the indexes.

Jackson, Beail & Vlissides (2011) Psychological Therapies Outcome Scale – ID II

Appendix B

Permission to Access Service Database and Ethical Approval.

B (i) Ethical Approval

The University Of Sheffield.

Downloaded: 19/01/2024
Approved: 18/01/2024

Emily Kerry
Registration number: 210154917
Psychology
Programme: Doctorate of Clinical Psychology

Dear Emily

PROJECT TITLE: Validation and Abbreviation of the Psychological Therapies Outcome Scale (PTOS-II) Outcome Measure
APPLICATION: Reference Number 058517

This letter confirms that you have signed a University Research Ethics Committee-approved self-declaration to confirm that your research will involve only existing research, clinical or other data that has been robustly anonymised. You have judged it to be unlikely that this project would cause offence to those who originally provided the data, should they become aware of it.

As such, on behalf of the University Research Ethics Committee, I can confirm that your project can go ahead on the basis of this self-declaration.

If during the course of the project you need to deviate significantly from the above-approved documentation please inform me since full ethical review may be required.

Yours sincerely

Department Of Psychology Research Ethics Committee
Departmental Ethics Administrator

B(ii) Letter of Access to Service Database

Dear Emily,

This letter should be presented to each participating organisation before you commence your research at that site: South West Yorkshire Partnership NHS Foundation Trust.

In accepting this letter, each participating organisation confirms your right of access to conduct research through their organisation for the purpose and on the terms and conditions set out below. This right of access commences on 21 March 2023 and ends on 31 May 2024 unless terminated earlier in accordance with the clauses below.

As an existing NHS employee, you do not require an additional honorary research contract with the participating organisation(s). The organisation(s) is/are satisfied that the research activities that you will undertake in the organisation(s) are commensurate with the activities you undertake for your employer. Your employer is fully responsible for ensuring such checks as are necessary have been carried out. Your employer has confirmed in writing to this organisation that the necessary pre-engagement checks are in place in accordance with

the role you plan to carry out in the organisation(s). Evidence of checks should be available on request to South West Yorkshire Partnership NHS Foundation Trust.

You have a right of access to conduct such research as confirmed in writing in the letter of permission for research from this organisation. Please note that you cannot start the research until the Principal
Investigator for the research project has received a letter from us giving the organisation(s) permission to conduct the project.

You are considered to be a legal visitor to South West Yorkshire Partnership NHS Foundation Trust premises. You are not entitled to any form of payment or access to other benefits provided by South West Yorkshire Partnership NHS Foundation Trust or this organisation to employees and this letter does not give rise to any other relationship between you South West Yorkshire Partnership NHS Foundation Trust or this organisation, in particular that of an employee.

While undertaking research through South West Yorkshire Partnership NHS Foundation Trust you will remain accountable to your employer, but you are required to follow the reasonable instructions of your nominated manager in each organisation or those given on her/his behalf in relation to the terms of this right of access.

Where any third-party claim is made, whether or not legal proceedings are issued, arising out of or in connection with your right of access, you are required to co-operate fully with any investigation by [Insert organisation] or this organisation in connection with any such claim and to give all such assistance as may reasonably be required regarding the conduct of any legal proceedings.

You must act in accordance with South West Yorkshire Partnership NHS Foundation Trust policies and procedures, which are available to you upon request, and the Research Governance Framework.

You are required to co-operate with South West Yorkshire Partnership NHS Foundation Trust in discharging its duties under the Health and Safety at Work etc Act 1974 and other health and safety legislation and to take reasonable care for the health and safety of yourself and others while on South West Yorkshire Partnership NHS Foundation Trust premises. Although you are not a contract holder, you must observe the same standards of care and propriety in dealing with patients, staff, visitors, equipment and premises as is expected of a contract holder and you must act appropriately, responsibly and professionally at all times.

If you have a physical or mental health condition or disability which may affect your research role and which might require special adjustments to your role, if you have not already done so, you must notify your employer and each participating prior to commencing your research role at each site.

You are required to ensure that all information regarding patients or staff remains secure and *strictly confidential* at all times. You must ensure that you understand and comply with the requirements of the NHS Confidentiality Code of Practice and the Data Protection Act 2018. Furthermore, you should be aware that under the Act, unauthorised disclosure of information is an offence and such disclosures may lead to prosecution.

The organisation(s) will not indemnify you against any liability incurred as a result of any breach of confidentiality or breach of the Data Protection Act 2018. Any breach of the Data Protection Act 2018 may result in legal action against you and/or your substantive employer.

You should ensure that, where you are issued with an identity or security card, a bleep number, email or library account, keys or protective clothing, these are returned upon termination of this arrangement.
Please also ensure that while on the premises you wear your ID badge at all times or are able to prove your identity if challenged. Please note that the organisation(s) accept no responsibility for damage to or loss of personal property.

This letter may be revoked and your right to attend the organisation(s) terminated at any time either by giving seven days' written notice to you or immediately without any notice if you are in breach of any of the terms or conditions described in this letter or if you commit any act that we reasonably consider to amount to serious misconduct or to be disruptive and/or prejudicial to the interests and/or business of the organisation(s) or if you are convicted of any criminal offence. You must not undertake regulated activity if you are barred from such work. If you are barred from working with adults or children, this letter of access is immediately terminated. Your employer will immediately withdraw you from undertaking this or any other regulated activity and you MUST stop undertaking any regulated activity immediately.

Your substantive employer is responsible for your conduct during this research project and may in the circumstances described above instigate disciplinary action against you.

If your circumstances change in relation to your health, criminal record, professional registration or suitability to work with adults or children, or any other aspect that may impact on your suitability to conduct research, or your role in research changes, you must inform the organisation that employs you through its normal procedures. You must also inform the nominated manager in each participating organisation.

Yours sincerely

Dr Wajid Khan Research & Development Manager

B (iii) HRA correspondence showing NHS ethics is not required

Dear Miss Kerry,

REC Reference: 23/HRA/0497
IRAS ID: 324291
Study Title: Validation and Abbreviation of the PTOS-II outcome measure

Thank you for submitting the above referenced study which was booked for REC review.

Upon a check, this application does not require ethical review for the following reasons:

1. The research is limited to secondary use of information previously collected in the course of normal care (without an intention to use it for research at the time of collection) **and** the dthe patients or service users are not identifiable to the research team in carrying out the research.

This application will still require study wide review, and the study wide reviewer will be in touch.

Kind regards

Rachel Katzenellenbogen
Approvals Specialist

Appendix C

Completed STROBE Checklist.

	Item No	Recommendation
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found
Introduction		
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported
Objectives	3	State specific objectives, including any prespecified hypotheses
Methods		
Study design	4	Present key elements of study design early in the paper
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case **N/A**
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group
Bias	9	Describe any efforts to address potential sources of bias
Study size	10	Explain how the study size was arrived at
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding
		(*b*) Describe any methods used to examine subgroups and interactions
		(*c*) Explain how missing data were addressed
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy
		(*e*) Describe any sensitivity analyses

Continued on next page

Results		
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed
		(b) Give reasons for non-participation at each stage
		(c) Consider use of a flow diagram
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders
		(b) Indicate number of participants with missing data for each variable of interest
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount) N/A
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time
		Case-control study—Report numbers in each exposure category, or summary measures of exposure
		Cross-sectional study—Report numbers of outcome events or summary measures
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included
		(*b*) Report category boundaries when continuous variables were categorized N/A
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period N/A
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses
Discussion		
Key results	18	Summarise key results with reference to study objectives
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence
Generalisability	21	Discuss the generalisability (external validity) of the study results
Other information		
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based N/A

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

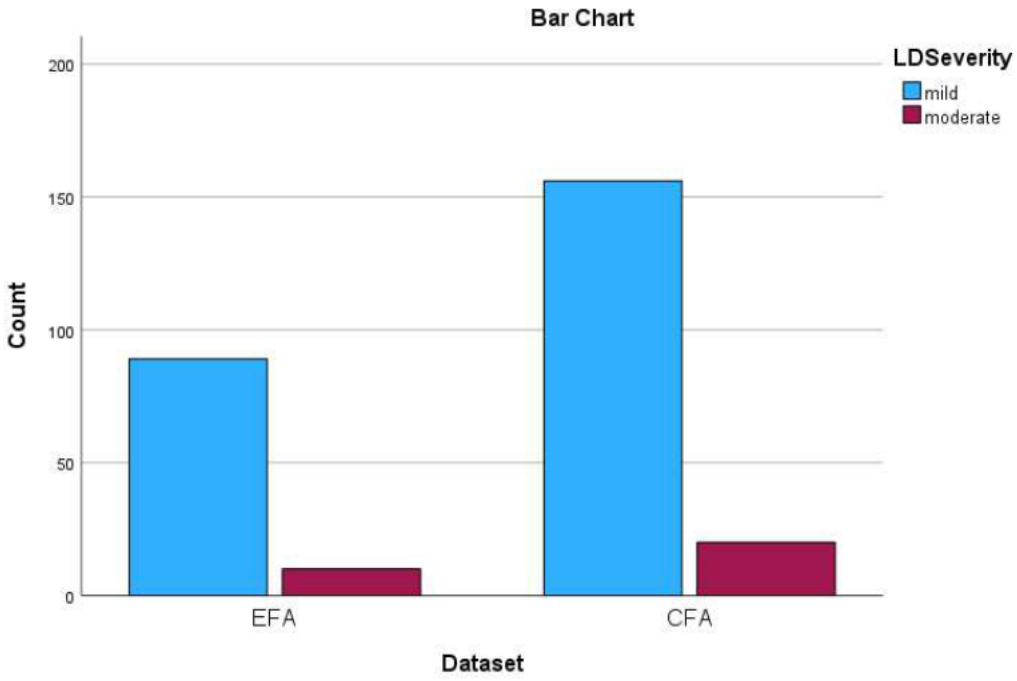
Appendix D

ANOVA, Chi-Square, and Descriptive Data Comparing the Different Datasets

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Gender	Between Groups	.001	1	.001	.002	.961
	Within Groups	96.763	396	.244		
	Total	96.764	397			
Ethnicity	Between Groups	.044	1	.044	.856	.358
	Within Groups	3.905	76	.051		
	Total	3.949	77			
LDSeverity	Between Groups	.010	1	.010	.103	.748
	Within Groups	26.717	273	.098		
	Total	26.727	274			

Comparison by Severity of LD:



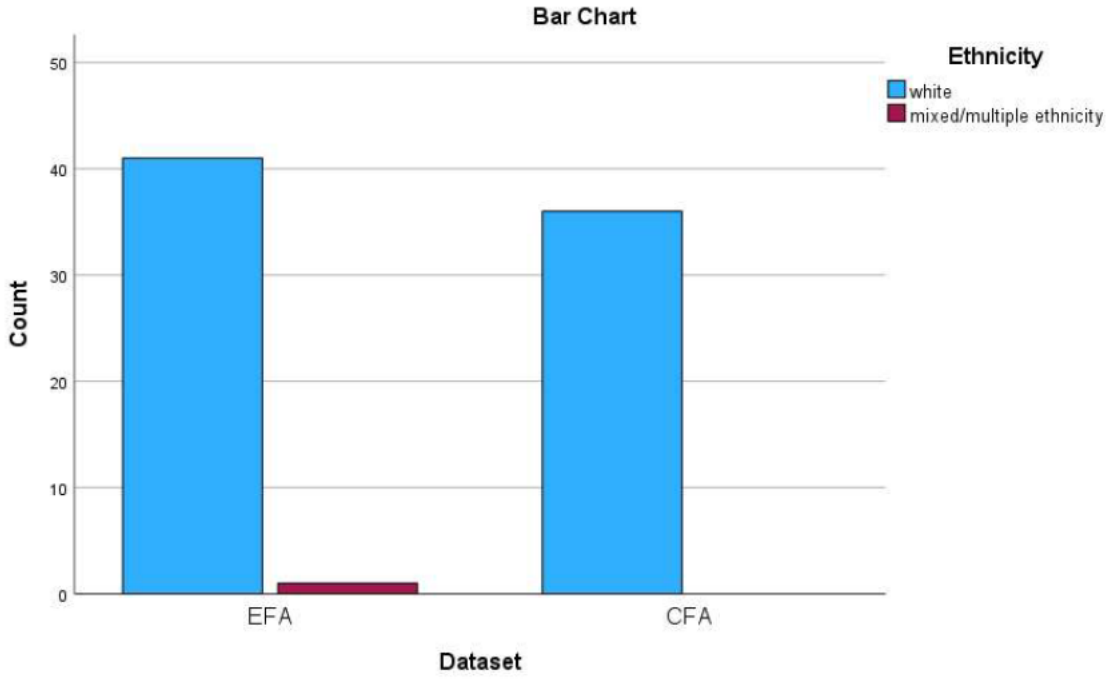
Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	.104[a]	1	.747		
Continuity Correction[b]	.015	1	.904		

Likelihood Ratio	.105	1	.746		
Fisher's Exact Test				.842	.458
Linear-by-Linear Association	.104	1	.748		
N of Valid Cases	275				

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 10.80.
b. Computed only for a 2x2 table

Comparison by Ethnicity:

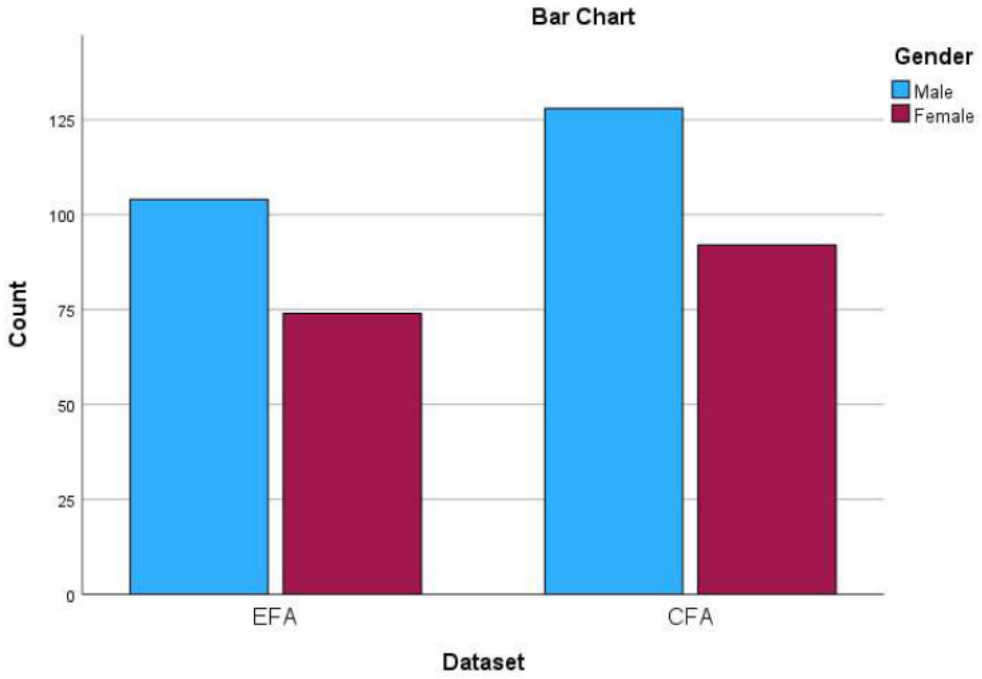


Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	.868[a]	1	.351		
Continuity Correction[b]	.000	1	1.000		
Likelihood Ratio	1.249	1	.264		
Fisher's Exact Test				1.000	.538
Linear-by-Linear Association	.857	1	.355		
N of Valid Cases	78				

a. 2 cells (50.0%) have expected count less than 5. The minimum expected count is .46.
b. Computed only for a 2x2 table

Comparison by Gender:



Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	.002[a]	1	.961		
Continuity Correction[b]	.000	1	1.000		
Likelihood Ratio	.002	1	.961		
Fisher's Exact Test				1.000	.521
Linear-by-Linear Association	.002	1	.961		
N of Valid Cases	398				

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 74.24.

b. Computed only for a 2x2 table

Appendix E

Correlation Matrix from First PCA

Correlations

Question		Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19	Q20	Q21	Q22	Q23	Q24	Q25	Q26	Q27	Q28	Q29
1	Correlation Coefficient	--																												
	Sig. (2-tailed)																													
	N	182																												
2	Correlation Coefficient	-.352**	--																											
	Sig. (2-tailed)	0.000																												
	N	182	184																											
3	Correlation Coefficient	-.169*	.546**	--																										
	Sig. (2-tailed)	0.023	0.000																											
	N	182	184	184																										
4	Correlation Coefficient	-.224**	.406**	.443**	--																									
	Sig. (2-tailed)	0.002	0.000	0.000																										
	N	181	183	183	183																									

5	Correlation Coefficient	.574**	-.280**	-.153*	-.291**	--																								
	Sig. (2-tailed)	0.000	0.000	0.039	0.000																									
	N	181	183	183	182	183																								
6	Correlation Coefficient	-0.144	.497**	.623**	.331**	-.171*	--																							
	Sig. (2-tailed)	0.052	0.000	0.000	0.000	0.021																								
	N	182	184	184	183	183	184																							
7	Correlation Coefficient	.158*	-.299**	-.266**	-.400**	.324**	-.296**	--																						
	Sig. (2-tailed)	0.034	0.000	0.000	0.000	0.000	0.000																							
	N	182	184	184	183	183	184	184																						
Q8	Correlation Coefficient	-.162*	.446**	.490**	.649**	-0.136	.373**	-.332**	--																					
	Sig. (2-tailed)	0.029	0.000	0.000	0.000	0.067	0.000	0.000																						
	N	181	183	183	182	182	183	183	183																					
9	Correlation Coefficient	-.246**	.380**	.511**	.172*	-.245**	.410**	-.211**	.190*	--																				
	Sig. (2-tailed)	0.001	0.000	0.000	0.020	0.001	0.000	0.004	0.010																					
	N	182	184	184	183	183	184	184	183	184																				
10	Correlation	-.183*	.503**	.601**	.507**	-.227**	.498**	-.383**	.565**	.405**	--																			

	Coefficient																													
	Sig. (2-tailed)	0.014	0.000	0.000	0.000	0.002	0.000	0.000	0.000	0.000																				
	N	182	184	184	183	183	184	184	183	184	184																			
11	Correlation Coefficient	-.203**	.328**	.401**	.304**	-.198**	.383**	-0.122	.294**	.257**	.307**	--																		
	Sig. (2-tailed)	0.006	0.000	0.000	0.000	0.007	0.000	0.099	0.000	0.000	0.000																			
	N	181	183	183	182	182	183	183	182	183	183	183																		
12	Correlation Coefficient	.227**	-.257**	-.293**	-.272**	.217**	-.234**	.234**	-.304**	-.190**	-.249**	-0.136	--																		
	Sig. (2-tailed)	0.002	0.000	0.000	0.000	0.003	0.001	0.001	0.000	0.010	0.001	0.066																			
	N	181	183	183	182	182	183	183	182	183	183	182	183																	
13	Correlation Coefficient	-0.073	.267**	.316**	.340**	-0.030	.387**	-.173*	.321**	.183*	.313**	.159*	-0.035	--																
	Sig. (2-tailed)	0.327	0.000	0.000	0.000	0.690	0.000	0.019	0.000	0.013	0.000	0.032	0.637																	
	N	181	183	183	182	182	183	183	182	183	183	182	183	183																
14	Correlation Coefficient	.539**	-.346**	-.228**	-.233**	.485**	-.273**	.378**	-0.088	-.299**	-.260**	-.194**	.218**	-0.133	--															
	Sig. (2-tailed)	0.000	0.000	0.002	0.002	0.000	0.000	0.000	0.234	0.000	0.000	0.008	0.003	0.072																
	N	182	184	184	183	183	184	184	183	184	184	183	183	183	184															
15	Correlation Coefficient	.269**	-0.107	-.163*	-.256**	.367**	-0.124	.385**	-.168*	-0.128	-.213**	-0.073	.291**	-0.132	.311**	--														
	Sig. (2-tailed)																													
	N																													

	Sig. (2-tailed)	0.000	0.150	0.028	0.001	0.000	0.096	0.000	0.024	0.086	0.004	0.329	0.000	0.076	0.000																
	N	180	182	182	181	181	182	182	181	182	182	182	181	181	182	182															
16	Correlation Coefficient	.412**	-.264**	-.238*	-.318**	.354**	-.199**	.358**	-.215**	-.311**	-.274**	-.190*	.202**	-.152*	.407**	.287**	--															
	Sig. (2-tailed)	0.000	0.000	0.001	0.000	0.000	0.007	0.000	0.004	0.000	0.000	0.011	0.006	0.041	0.000	0.000															
	N	181	182	182	181	181	182	182	181	182	182	181	181	181	182	180	182														
17	Correlation Coefficient	-.252**	.412**	.356**	.454**	-.290**	.415**	-.345**	.475**	0.138	.338**	.271**	-.226**	.257**	-.282**	-.251**	-.226**	--														
	Sig. (2-tailed)	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.062	0.000	0.000	0.002	0.000	0.000	0.001	0.002														
	N	182	184	184	183	183	184	184	183	184	184	183	183	183	184	182	182	184													
18	Correlation Coefficient	-.359**	.572**	.482**	.426**	-.260**	.410**	-.356**	.426**	.387**	.429**	.275**	-.231**	.294**	-.397**	-.208**	-.406**	.411**	--													
	Sig. (2-tailed)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.000	0.000	0.005	0.000	0.000													
	N	182	183	183	182	182	183	183	182	183	183	182	182	182	183	181	181	183	183												
19	Correlation Coefficient	-.173*	.274**	.362**	0.063	-.184*	.304**	-0.126	.187*	.360**	.249**	.195**	-.321**	0.126	-.158*	-0.142	-.264**	.329**	.320**	--												
	Sig. (2-tailed)	0.020	0.000	0.000	0.396	0.013	0.000	0.088	0.011	0.000	0.001	0.008	0.000	0.089	0.033	0.056	0.000	0.000	0.000												
	N	182	184	184	183	183	184	184	183	184	184	183	183	183	184	182	182	184	183	184											
20	Correlation Coefficient	.323**	-0.139	-.168*	-0.136	.367**	-.151*	.251**	-0.045	-.204**	-0.063	-.198**	.241**	-0.006	.309**	.332**	.349**	-0.080	-.217**	-.191**	--										
	Sig. (2-tailed)	0.000	0.060	0.023	0.066	0.000	0.041	0.001	0.550	0.006	0.398	0.007	0.001	0.938	0.000	0.000	0.000	0.281	0.003	0.010											

	N	182	184	184	183	183	184	184	183	184	184	183	183	183	184	182	182	184	183	184	184									
21	Correlation Coefficient	-.208**	.301**	.282**	.544**	-.262**	.310**	-.345**	.473**	.263**	.418**	0.105	-0.091	.337**	-.249**	-.293**	-.242**	.402**	.339**	-0.068	0.068	--								
	Sig. (2-tailed)	0.005	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.161	0.223	0.000	0.001	0.000	0.001	0.000	0.000	0.361	0.362									
	N	181	182	182	182	181	182	182	181	182	182	181	181	181	182	180	180	182	182	182	182	182								
22	Correlation Coefficient	-0.037	0.117	.193**	.178*	-0.117	.208**	-0.029	.256**	0.057	.193**	.388**	-0.005	0.010	-0.011	-0.026	-0.029	0.116	0.119	-0.074	-.152*	0.115	--							
	Sig. (2-tailed)	0.625	0.116	0.009	0.017	0.116	0.005	0.696	0.001	0.446	0.009	0.000	0.945	0.892	0.887	0.730	0.704	0.119	0.112	0.321	0.041	0.124								
	N	179	181	181	181	180	181	181	180	181	181	181	180	180	181	180	179	181	180	181	181	180	181							
23	Correlation Coefficient	-0.117	.377**	.646**	.355**	-.171*	.515**	-.260**	.432**	.516**	.524**	.415**	-.372**	.278**	-.177*	-0.070	-.247**	.287**	.319**	.382**	-.168*	.162*	0.130	--							
	Sig. (2-tailed)	0.117	0.000	0.000	0.000	0.021	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.016	0.351	0.001	0.000	0.000	0.000	0.023	0.029	0.082							
	N	181	183	183	182	182	183	183	182	183	183	182	182	182	183	181	181	183	182	183	183	181	180	183						
24	Correlation Coefficient	.486**	-.543**	-.385**	-.354**	.492**	-.392**	.480**	-.283**	-.395**	-.412**	-.246**	.318**	-.166*	.580**	.366**	.399**	-.388**	-.526**	-.272**	.349**	-.267**	-0.102	-.326**	--					
	Sig. (2-tailed)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.025	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.174	0.000						
	N	181	183	183	182	182	183	183	182	183	183	182	182	182	183	181	181	183	182	183	183	181	180	183	183					
25	Correlation Coefficient	-.347**	.494**	.400**	.337**	-.347**	.292**	-.356**	.329**	.445**	.389**	.248**	-.283**	.267**	-.376**	-.277**	-.328**	.383**	.473**	.365**	-.275**	.261**	0.050	.331**	-.544**	--					
	Sig. (2-tailed)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.502	0.000	0.000					
	N	181	183	183	182	182	183	183	182	183	183	182	182	182	183	181	181	183	182	183	183	181	180	183	183	183				

26	Correlation Coefficient	.462**	-.489**	-.336**	-.309**	.500**	-.285**	.467**	-.248**	-.394**	-.363**	-.176*	0.130	-0.124	.592**	.323**	.425**	-.310**	-.511**	-.281**	.321**	-.241**	-0.117	-.261**	.634**	-.517**	--				
	Sig. (2-tailed)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.018	0.081	0.097	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.119	0.000	0.000	0.000				
	N	180	182	182	181	181	182	182	181	182	182	181	181	181	182	180	180	182	181	182	182	180	179	182	182	182	182			
27	Correlation Coefficient	.366**	-.405**	-.264**	-.259**	.537**	-.256**	.366**	-.246**	-.343**	-.247**	-.270**	.209**	-0.086	.411**	.260**	.362**	-.282**	-.377**	-.249**	.419**	-.168*	-0.050	-.291**	.580**	-.406**	.496**	--			
	Sig. (2-tailed)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.001	0.000	0.005	0.252	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.024	0.506	0.000	0.000	0.000	0.000				
	N	180	182	182	181	181	182	182	181	182	182	181	181	181	182	180	180	182	181	182	182	180	179	182	182	182	182	182			
28	Correlation Coefficient	-.179*	.496**	.588**	.416**	-.199**	.603**	-.310**	.426**	.478**	.534**	.333**	-.329**	.265**	-.290**	-0.088	-.252**	.301**	.443**	.336**	-0.069	.304**	.165*	.596**	-.396**	.371**	-.306**	-.249**	--		
	Sig. (2-tailed)	0.016	0.000	0.000	0.000	0.007	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.239	0.001	0.000	0.000	0.000	0.358	0.000	0.028	0.000	0.000	0.000	0.000	0.001			
	N	179	181	181	180	180	181	181	180	181	181	180	180	180	181	179	179	181	180	181	181	179	178	181	181	181	181	181	181	
29	Correlation Coefficient	.442**	-.496**	-.311**	-.280**	.545**	-.364**	.571**	-.246**	-.308**	-.334**	-.182*	.201**	-0.116	.607**	.334**	.429**	-.351**	-.511**	-.223**	.366**	-.249**	-0.069	-.270**	.775**	-.483**	.696**	.603**	-.325**	--	
	Sig. (2-tailed)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.014	0.007	0.119	0.000	0.000	0.000	0.000	0.000	0.003	0.000	0.001	0.359	0.000	0.000	0.000	0.000	0.000	0.000		
	N	180	182	182	181	181	182	182	181	182	182	181	181	181	182	180	180	182	181	182	182	180	179	182	182	182	182	182	181	182

Appendix F

Summary of Factor Loadings from Initial PCA

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.900
Bartlett's Test of Sphericity	Approx. Chi-Square	2511.210
	df	406
	Sig.	<.001

Communalities

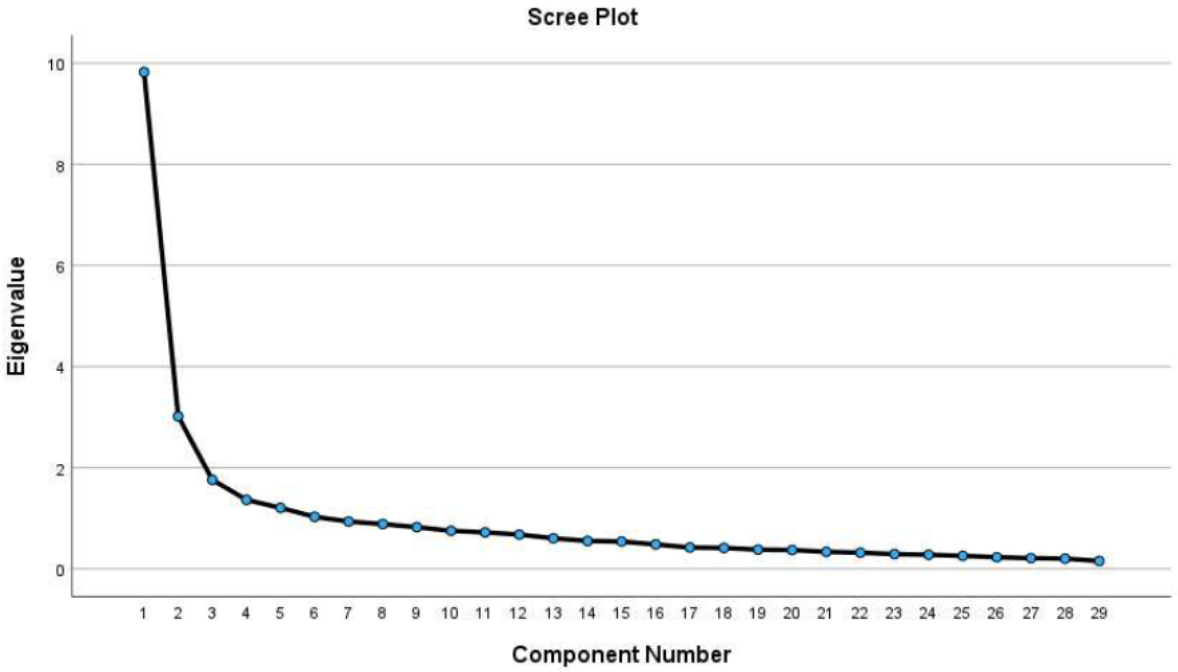
	Initial	Extraction
Q1	1.000	.543
Q2	1.000	.640
Q3	1.000	.714
Q4	1.000	.746
Q5	1.000	.604
Q6	1.000	.601
Q7	1.000	.497
Q8	1.000	.701
Q9	1.000	.699
Q10	1.000	.629
Q11	1.000	.609
Q12	1.000	.673
Q13	1.000	.404
Q14	1.000	.647
Q15	1.000	.659
Q16	1.000	.417
Q17	1.000	.653
Q18	1.000	.588
Q19	1.000	.627
Q20	1.000	.601
Q21	1.000	.708
Q22	1.000	.680
Q23	1.000	.678
Q24	1.000	.712
Q25	1.000	.515
Q26	1.000	.729
Q27	1.000	.552
Q28	1.000	.629
Q29	1.000	.738

Extraction Method: Principal Component Analysis.

Total Variance Explained

	Initial Eigenvalues			Extraction Sums of Squared Loadings		
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	9.824	33.875	33.875	9.824	33.875	33.875
2	3.015	10.397	44.272	3.015	10.397	44.272
3	1.760	6.067	50.340	1.760	6.067	50.340
4	1.364	4.702	55.042	1.364	4.702	55.042
5	1.206	4.157	59.199	1.206	4.157	59.199
6	1.027	3.542	62.741	1.027	3.542	62.741
7	.934	3.221	65.962			
8	.884	3.047	69.009			
9	.822	2.836	71.845			
10	.751	2.589	74.434			
11	.720	2.481	76.915			
12	.676	2.331	79.246			
13	.601	2.074	81.320			
14	.550	1.897	83.217			
15	.538	1.856	85.073			
16	.481	1.660	86.732			
17	.421	1.451	88.183			
18	.412	1.420	89.603			
19	.378	1.305	90.908			
20	.371	1.278	92.186			
21	.335	1.157	93.343			
22	.318	1.096	94.439			
23	.290	1.001	95.439			
24	.277	.954	96.394			
25	.255	.879	97.273			
26	.228	.785	98.058			
27	.210	.723	98.781			
28	.200	.690	99.471			
29	.153	.529	100.000			

Extraction Method: Principal Component Analysis.



Component Matrix[a]

	Component					
	1	2	3	4	5	6
Q1	-.546	.436	.016	.129	-.196	.015
Q2	.691	.165	.069	.127	.299	.159
Q3	.658	.468	.229	.014	-.041	-.085
Q4	.643	.272	-.448	-.219	-.082	.043
Q5	-.558	.461	.078	.267	-.039	.031
Q6	.596	.433	.159	.002	.059	-.172
Q7	-.584	.188	.295	-.087	.158	.039
Q8	.577	.405	-.317	-.130	-.053	.289
Q9	.548	.119	.417	.221	-.078	-.394
Q10	.671	.388	-.095	.046	-.045	-.124
Q11	.438	.245	.267	-.479	.213	.109
Q12	-.417	.016	-.083	.108	.628	-.293
Q13	.364	.363	-.251	.190	-.009	-.201
Q14	-.655	.413	.002	-.019	-.114	.182
Q15	-.415	.364	.263	.145	.480	.183
Q16	-.563	.218	.042	.028	.132	.180
Q17	.589	.084	-.261	.057	.054	.474
Q18	.706	.039	-.017	.198	.193	.108
Q19	.446	.027	.463	.228	-.226	.331
Q20	-.364	.431	-.289	.376	.218	.106

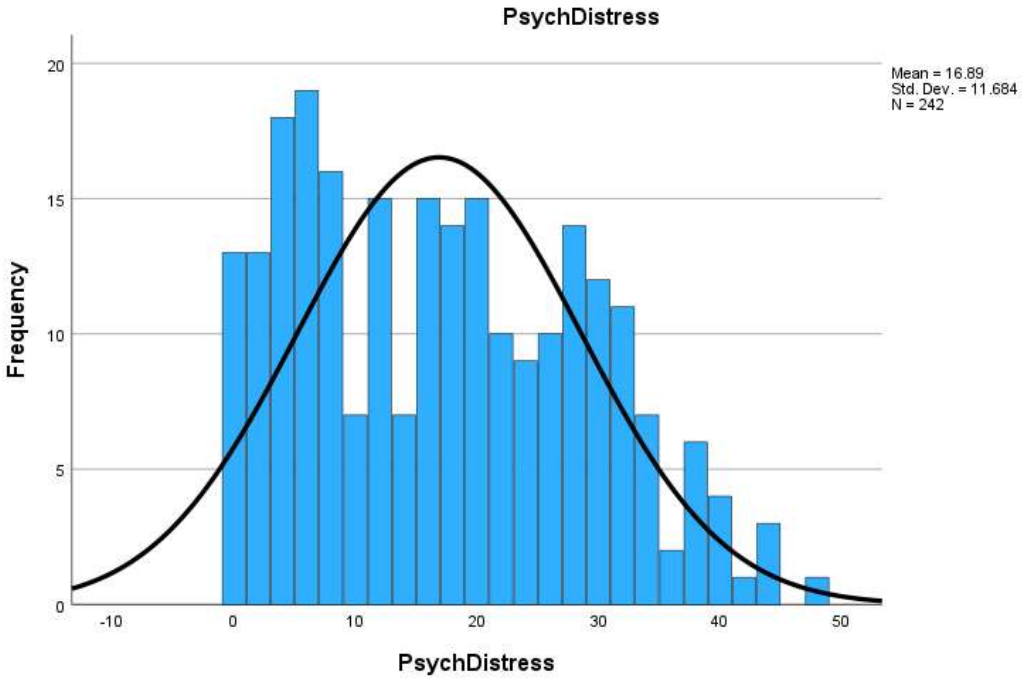
Q21	.488	.228	-.594	-.003	.010	-.256
Q22	.174	.264	.096	-.715	.239	-.051
Q23	.598	.408	.308	-.015	-.242	-.033
Q24	-.768	.321	-.054	-.090	-.076	-.037
Q25	.678	-.095	-.029	.178	.075	.092
Q26	-.717	.388	-.040	-.164	-.185	.045
Q27	-.617	.350	-.161	.076	-.075	-.104
Q28	.636	.418	.178	.045	-.014	-.125
Q29	-.736	.395	.010	-.124	-.145	-.065

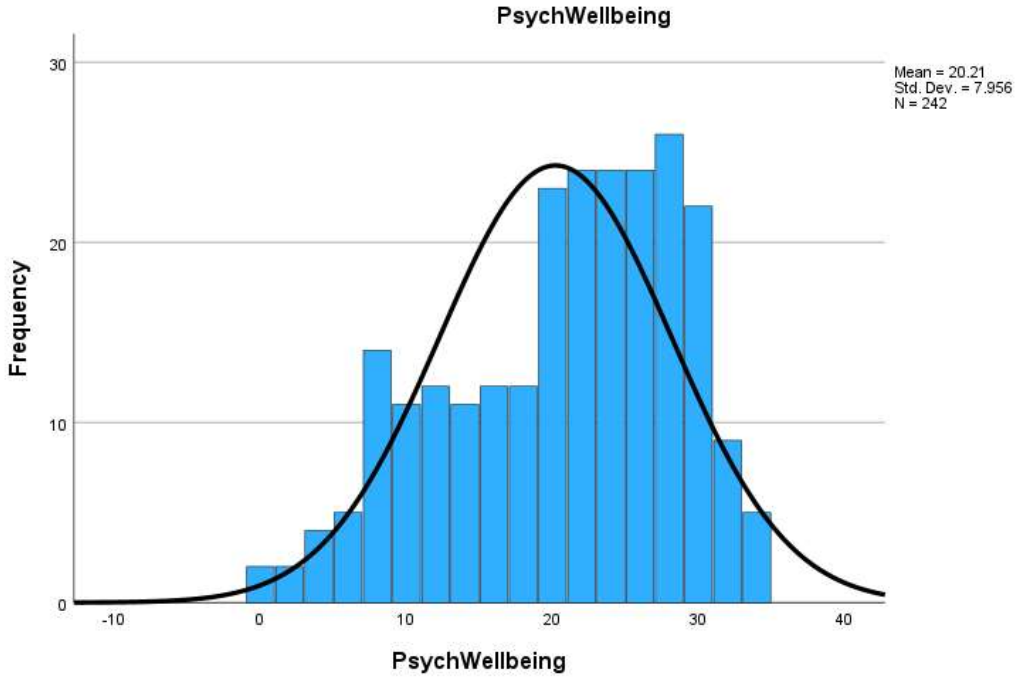
Extraction Method: Principal Component Analysis.

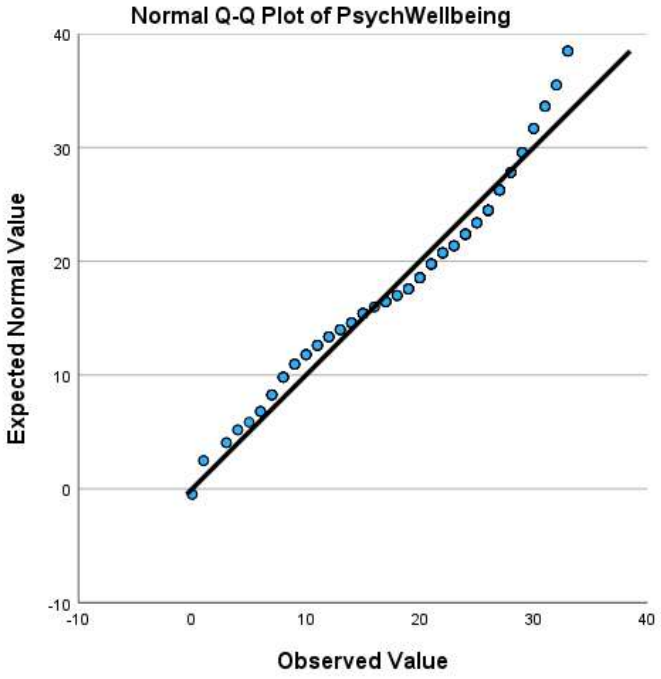
a. 6 components extracted.

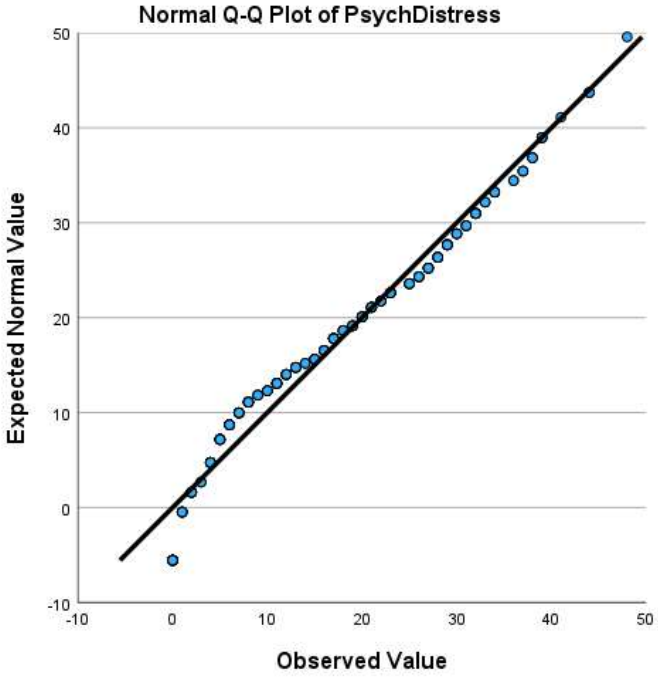
Appendix G

Histograms and Q-Q plots to Test Normality






Normal Q-Q Plot of PsychWellbeing
Expected Normal Value
Observed Value
40
30
20
10
0
-10
-10
0
10
20
30
40


Normal Q-Q Plot of PsychDistress
Expected Normal Value
Observed Value
50
40
30
20
10
0
-10
-10
0
10
20
30
40
50

Appendix H

Correlational Analyses Wellbeing and Distress

		PsychWellbeing	Q1	Q5	Q7	Q14	Q15	Q16	Q20	Q24	Q26	Q27	Q29
PsychWellbeing	Pearson Correlation	--											
	N	184											
Q1	Pearson Correlation	.693**	--										
	Sig. (2-tailed)	<.001											
	N	182	182										
Q5	Pearson Correlation	.733**	.585**	--									
	Sig. (2-tailed)	<.001	<.001										
	N	183	181	183									
Q7	Pearson Correlation	.617**	.192**	.326**	--								
	Sig. (2-tailed)	<.001	.009	<.001									
	N	184	182	183	184								
Q14	Pearson Correlation	.758**	.590**	.507**	.406**	--							
	Sig. (2-tailed)	<.001	<.001	<.001	<.001								
	N	184	182	183	184	184							
Q15	Pearson Correlation	.570**	.269**	.366**	.372**	.332**	--						
	Sig. (2-tailed)	<.001	<.001	<.001	<.001	<.001							
	N	182	180	181	182	182	182						
Q16	Pearson Correlation	.629**	.417**	.355**	.367**	.427**	.288**	--					
	Sig. (2-tailed)	<.001	<.001	<.001	<.001	<.001	<.001						
	N	182	181	181	182	182	180	182					
Q20	Pearson Correlation	.587**	.332**	.379**	.246**	.346**	.339**	.347**	--				
	Sig. (2-tailed)	<.001	<.001	<.001	<.001	<.001	<.001	<.001					
	N	184	182	183	184	184	182	182	184				
Q24	Pearson Correlation	.811**	.512**	.508**	.499**	.602**	.382**	.408**	.366**	--			
	Sig. (2-tailed)	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001				
	N	183	181	182	183	183	181	181	183	183			
Q26	Pearson Correlation	.770**	.525**	.498**	.485**	.614**	.324**	.436**	.326**	.628**	--		
	Sig. (2-tailed)	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001			
	N	182	180	181	182	182	180	180	182	182	182		
Q27	Pearson Correlation	.698**	.406**	.538**	.361**	.421**	.257**	.369**	.426**	.559**	.493**	--	
	Sig. (2-tailed)	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001		
	N	182	180	181	182	182	180	180	182	182	182	182	
Q29	Pearson Correlation	.823**	.467**	.546**	.583**	.615**	.342**	.431**	.355**	.783**	.687**	.557**	--
	Sig. (2-tailed)	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	
	N	182	180	181	182	182	180	180	182	182	182	182	182

**. Correlation is significant at the 0.01 level (2-tailed).

Correlations

		Q2	Q3	Q6	Q9	Q11	Q18	Q19	Q23	Q25	Q28	Q4	Q8	Q10	Q13	Q17	Q21	PsychDistres
Q2	Pearson Correlation	--																
	N	184																
Q3	Pearson Correlation	.542**	--															
	Sig. (2-tailed)	<.001																
	N	184	184															
Q6	Pearson Correlation	.498**	.627**	--														
	Sig. (2-tailed)	<.001	<.001															
	N	184	184	184														
Q9	Pearson Correlation	.378**	.528**	.405**	--													
	Sig. (2-tailed)	<.001	<.001	<.001														
	N	184	184	184	184													
Q11	Pearson Correlation	.325**	.408**	.385**	.240**	--												
	Sig. (2-tailed)	<.001	<.001	<.001	.001													
	N	183	183	183	183	183												
Q18	Pearson Correlation	.573**	.489**	.393**	.394**	.271**	--											
	Sig. (2-tailed)	<.001	<.001	<.001	<.001	<.001												
	N	183	183	183	183	182	183											
Q19	Pearson Correlation	.317**	.376**	.318**	.401**	.223**	.319**	--										
	Sig. (2-tailed)	<.001	<.001	<.001	<.001	.002	<.001											
	N	184	184	184	184	183	183	184										
Q23	Pearson Correlation	.382**	.642**	.497**	.510**	.395**	.314**	.378**	--									
	Sig. (2-tailed)	<.001	<.001	<.001	<.001	<.001	<.001	<.001										
	N	183	183	183	183	182	182	183	183									
Q25	Pearson Correlation	.478**	.383**	.274**	.398**	.251**	.498**	.278**	.356**	--								
	Sig. (2-tailed)	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001									
	N	183	183	183	183	182	182	183	183	183								
Q28	Pearson Correlation	.492**	.583**	.606**	.463**	.324**	.434**	.312**	.583**	.361**	--							
	Sig. (2-tailed)	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001								
	N	181	181	181	181	180	180	181	181	181	181							
Q4	Pearson Correlation	.394**	.455**	.328**	.174*	.305**	.428**	.069	.374**	.375**	.418**	--						
	Sig. (2-tailed)	<.001	<.001	<.001	.018	<.001	<.001	.353	<.001	<.001	<.001							
	N	183	183	183	183	182	182	183	182	182	180	183						
Q8	Pearson Correlation	.443**	.485**	.351**	.196**	.289**	.427**	.181*	.445**	.368**	.424**	.628**	--					
	Sig. (2-tailed)	<.001	<.001	<.001	.008	<.001	<.001	.014	<.001	<.001	<.001	<.001						
	N	183	183	183	183	182	182	183	182	182	180	182	183					
Q10	Pearson Correlation	.501**	.607**	.509**	.404**	.303**	.431**	.251**	.528**	.411**	.523**	.521**	.565**	--				
	Sig. (2-tailed)	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001					
	N	184	184	184	184	183	183	184	183	183	181	183	183	184				
Q13	Pearson Correlation	.270**	.310**	.396**	.166*	.150*	.296**	.136	.270**	.288**	.270**	.359**	.307**	.312**	--			
	Sig. (2-tailed)	<.001	<.001	<.001	.025	.044	<.001	.067	<.001	<.001	<.001	<.001	<.001	<.001				
	N	183	183	183	183	182	182	183	182	182	180	182	182	183	183			
Q17	Pearson Correlation	.416**	.351**	.390**	.142	.287**	.419**	.378**	.307**	.400**	.306**	.436**	.461**	.329**	.272**	--		
	Sig. (2-tailed)	<.001	<.001	<.001	.055	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001			
	N	184	184	184	184	183	183	184	183	183	181	183	183	184	183	184		
Q21	Pearson Correlation	.296**	.278**	.313**	.248**	.114	.319**	-.058	.163*	.244**	.308**	.554**	.446**	.418**	.344**	.369**	--	
	Sig. (2-tailed)	<.001	<.001	<.001	<.001	.128	<.001	.437	.029	<.001	<.001	<.001	<.001	<.001	<.001	<.001		
	N	182	182	182	182	181	182	182	181	181	179	182	181	182	181	182	182	
PsychDistress	Pearson Correlation	.709**	.795**	.708**	.588**	.549**	.671**	.455**	.702**	.596**	.735**	.671**	.689**	.756**	.414**	.584**	.516**	--
	Sig. (2-tailed)	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	
	N	184	184	184	184	183	183	184	183	183	181	183	183	184	183	184	182	184