Psychometric properties of instruments used in intellectual disabilities

Majid Saleem

Submitted for the award of

Doctorate in Clinical Psychology

The University of Sheffield

September 2016

Declaration

This thesis has been submitted for the award of Doctorate in Clinical Psychology at the University of Sheffield. It has not been submitted for any other qualification or to any other academic institution.

Structure and word count

Literature review	8,034
Including references	9,844
Empirical report	12,237
Including references	13,247
Including references and appendices	14,416
Appendices	1,169
Total word count	24,260
Excluding appendices and references	20,271

Abstract

Literature review: The aim of the current review was to systematically assess the quality of multi-trait measures of psychological distress and psychiatric symptoms, which have been used with adults with intellectual disabilities. Seven multi-trait measures of psychological distress and psychiatric symptoms were identified and their psychometric properties were appraised. Results showed that all multi-trait measures had high levels of internal consistency with only some measures reporting moderate to good levels of inter-rater and test/retest reliability. The assessment of validity of the multi-trait measures was shown to be lacking with number and quality of assessments varying across all of the measures. There was also limited information on acceptability, feasibility and precision of the measures. Findings are discussed in relation to these factors. Methodological limitations and recommendations for clinical practice and future research are outlined.

Empirical report: The aim of the current study was to investigate the factor structure of the WAIS-IV with an intellectual disabilities population in the UK. In addition the relationship between IQ and adaptive functioning was examined. Results showed that the only factor structure that can be reliably applied for this population is a two-factor solution of verbal-performance. It was also found that there is a positive albeit small correlation between full scale IQ and adaptive behaviour composite and a larger correlation between full scale IQ and daily living skills. There were no correlations between full scale IQ and communication or full scale IQ and socialisation These results provide some evidence that the WAIS-IV and the Vineland-II assess different sets of

skills and behaviours. The clinical implications of these findings are discussed alongside future research recommendations.

Acknowledgments

Firstly, I would like to thank my supervisor Prof. Nigel Beail for his support and guidance throughout the entire process. His knowledge in the area of intellectual disabilities is truly astounding and inspiring.

Secondly, I would like to thank my family for their unwavering support over the last three years.

Finally, I would like to thank my wife, without whom none of this would have been possible.

Contents

	Page
Declaration	iv
Word count	v
Abstract	vi
Acknowledgements	vii
Section one: Literature review	
Abstract	2
Introduction	4
Method	9
Results	20
Discussion	29
References	35
Section two: Empirical report	
Abstract	45
Introduction	47
Method	53
Results	63
Discussion	85

References 99

Contents (continued)

Appendices

Appendix A – NHS ethical approval letter 104

Appendix B – R&D ethical approval letter 106

Appendix C – Table showing means and standard 108

deviations of current data and WAIS-IV standardised sub samples

Section one: Literature review

A systematic review of the psychometric properties of multi-trait outcome measures of psychological distress and psychiatric symptoms used in adults with intellectual disabilities

Abstract

Objectives. Appropriate assessment of psychological distress and psychiatric symptoms in adults with intellectual disabilities should be based on validated measurement. Without comprehensive psychometric tools, accuracy of assessment is compromised. The aim of the current review was to systematically assess the quality of multi-trait measures of psychological distress and psychiatric symptoms that have been used with adults with intellectual disabilities.

Method. Three databases (MEDLINE, PsycINFO and Scopus) were searched for journal articles examining the psychometric properties of multi-trait measures of psychological distress and psychiatric symptoms used with people with intellectual disabilities.

Results. Results showed that all multi-trait measures had high levels of internal consistency with only some measures reporting moderate to good levels of inter-rater and test/retest reliability. The assessment of validity of the multi-trait measures was shown to be lacking with number and quality of assessments varying across all of the measures. There was also limited information on acceptability, feasibility and precision of the measures.

Conclusions. The quality appraisal showed that these multi-trait measures of psychological distress and psychiatric symptoms have varying levels of psychometric properties and that there is still work to be done to improve some of their overall quality. The best measures available at the moment appear to be the PASADD and the ADD in terms of reliability and the PASADD in terms of validity and would be recommended for use to clinicians in terms of informant based measures. In terms of self-report measures the BSI appears to have the most reasonable psychometric properties and would be recommended for use.

Practitioner points

- The findings from the current review show that there are a limited number of multi-trait measures of psychological distress and psychiatric symptoms used with adult intellectual disabilities samples compared to other non-intellectually disabled populations.

- The quality appraisal of these measures in the current review showed that these multi-trait measures have varying levels of psychometric properties and that there is still work to be done to improve some of their overall quality.

- The best measures available at the moment appear to be the PASADD and the ADD in terms of reliability and the PASADD in terms of validity and would be recommended for use to clinicians in terms of informant based measures. In terms of self-report measures the BSI appears to have the most reasonable psychometric properties and would be recommended for use.

- Further examination of the psychometric properties of these multi-trait measures would provide a more robust evidence base and more information about the clinical utility of these measures with people with intellectual disabilities.

Approximately three percent of the world's population is affected by a condition referred to as intellectual disability (Asada, Tomiwa, Okada & Itakura, 2010; Cheng & Chen, 2010). Research has shown that there are many genetic conditions, which are known to cause intellectual disability, along with many social and cultural factors that may also intensify this condition (Matson, Smiroldo & Bamburg, 1998; Matson, Kiely & Bamburg, 1997). Individuals with intellectual disabilities have been shown to have deficits in social and adaptive functioning (Lante, Reece & Walkley, 2010; Shin, Shin & Wang, 2010; Wise, Sevick, Romski & Morris, 2010), along with an increased likelihood of epilepsy and challenging behaviours (Lambrechts, Van Den Noortgate, Eeman & Maes 2010; Rose, 2010; Williams, 2010).

Historically it was believed that individuals with intellectual disability were unable to experience emotional and psychological distress like non-intellectually disabled individuals due to deficiencies in ego strength (Deb, Thomas & Bright, 2001). However, many researchers have shown that individuals with intellectual disabilities not only exhibit the same emotional and psychological distress found in the general population, but that this distress is more frequently reported and that they have an increased risk than the general population (Cooper, Smiley, Morrison, Williamson & Allan, 2007; Smiley, 2005; Rush, Bowman, Eidman, Toole & Mortenson, 2004). The prevalence of comorbid psychological distress and psychiatric symptoms differs broadly depending on the research study, with estimates ranging from 10 to 71 %, depending on the diagnostic criteria used and the population included in the study (Cooper, Smiley, Morrison, Williamson, & Allan, 2007; Smiley, 2005; Deb, Thomas & Bright, 2001). Furthermore, the location of where these studies were carried out may contribute to the difference in prevalence rates, with inpatient populations reportedly having a higher rate of psychological distress and psychiatric symptoms than those receiving treatment in the

community (Cooper & Bailey, 2001; Deb, Thomas & Bright, 2001; Trower, Treadwell & Bhaumik, 1998).

There have been a number of disorders of psychological distress/psychiatric symptoms, which have been identified as being most prevalent in individuals with intellectual disabilities. These include major depressive disorder, anxiety disorders, psychotic disorders, attention deficit/hyperactivity disorders and autistic spectrum disorders (Hastings, Beck, Daley & Hill, 2005; Deb, Thomas & Bright, 2001; McGrother, Hauck, Bhaumik, Thorp & Taub, 1996). There is also evidence to suggest that in intellectually disabled individuals the occurrence of one disorder of psychological distress/psychiatric symptoms can increase the risk of developing other disorders of psychological distress/psychiatric symptoms (Kovacs, Paulauskas, Gatsonis & Richards, 1988). Thus highlighting the importance of investigating the presence of multiple psychological disorders/psychiatric symptoms in people with intellectual disabilities.

Although emotional and psychological distress within people with intellectual disabilities is now receiving increasing attention and there is an agreement amongst clinicians that many individuals with intellectual disabilities need assistance from mental health services, there are still challenges in identifying people in this population who are experiencing emotional and psychological distress and many are not referred to the appropriate mental health services (Myrbakk & von Tetzchner, 2008). One possible reason why emotional and psychological distress may not be easily identified in this population is because of the limited verbal skills of individuals, as indications to emotional and psychological distress are usually revealed by a person's thoughts, feelings and beliefs (Mybrakk & von Tetzchner, 2008). This is particularly true in individuals who suffer from severe or profound intellectual disabilities who have difficulties in describing their thoughts and feelings to people around them. Therefore,

informant measures that use caregivers or members of the family are important to help identify signs and symptoms of psychological distress (Taylor, Hatton, Dixon & Douglas, 2004; Deb, Thomas & Bright, 2001).

Several multi-trait measures of psychological distress and psychiatric symptoms have been developed specifically for use with people who have intellectual disabilities. Additionally, some multi-trait measures, which have been developed for the general population, have been adapted to be used with people with intellectual disabilities. The rationale for the development of some tools has been to screen for a range of difficulties to advise in matters of health care insurance in the USA (e.g. Psychiatric Instrument for Mentally Retarded Adults (PIMRA). The Psychiatric Assessment Schedule for Adults who have Developmental Disabilities (PASADD) was developed to screen for a range of mental health problems due to the lack of trained clinicians in services to do this. Furthermore, due to the high comorbidity found in intellectual disability populations referred to mental health services, outcome studies have also used multi-trait measures to evaluate treatment outcomes. The Standard Symptom Checklist (SCL-90-R) has been used to evaluate the outcomes for psychodynamic psychotherapy (Beail, Kellett, Newman & Warden, 2007; Beail, Warden, Morsley & Newman, 2005) and the Brief Symptom Inventory (BSI) has been used to evaluate treatment outcomes in a controlled trail of CBT for adult who have intellectual disabilities (Lindsay et al., 2015).

Despite the development of special checklists and use of general population measures to help clinicians in identifying individuals with intellectual disabilities who need to be referred for a psychiatric and psychological evaluation there is no ‘gold standard’ with few studies where these different checklists have been compared (Myrbakk & von Tetzchner, 2008). There has also been no comprehensive systematic review exclusively focusing on psychometric properties of the range of multi-trait measures of psychological distress and psychiatric symptoms used with people who

have intellectual disabilities. Previous reviews have focused on specific measures that assess both global psychopathology and more discrete domains of psychopathology, and specifically on anxiety and depression measures (Matson, Belva, Hattier & Matson, 2012; Hermans, van der Pas & Evenhuis, 2011; Hermans & Evenhuis, 2010; Myrbakk & von Tetzchner, 2008). These reviews found that measures varied in their design in that they were designed for either the general population or specifically for people with intellectual disabilities. The target population for the measures differed with measures being specifically designed for mild, moderate and severe intellectual disabilities. There was also a large variation in the psychometric properties of the measures. None of these reviews mentioned the use of a quality appraisal tool to assess the quality of the measures.

The current review will appraise the quality of multi-trait measures of psychological distress and psychiatric symptoms used with adults with intellectual disabilities. For each measure the reliabilities, validities, and other psychometric properties will be evaluated. Only measures that assess psychological and psychiatric symptoms will be included with measures being both self-report and informant based. Measures, which have been designed to assess behaviour disturbances of an apparently non-psychiatric nature, have been excluded, along with measures, which have been specifically designed to assess autism.

One of the initial steps in examining whether a measure is appropriate for use is confirming that it has sound psychometric properties (Kraus & Castonguy, 2010). These properties are examined through assessing the reliability and validity of the measure (Beck, Steer & Garbin, 1988). The reliability of a measure refers to its ability to produce a similar result from the same respondent in regular conditions (Field, 2013). This is usually tested through test-retest reliability (amount of which scores are consistent over periods of time), inter-rater reliability (measures of two or more

assessment scorers or observers) and internal consistency (consistency of results across items in a test) (Fitzpatrick, Davey, Buxton & Jones, 1998). The validity of a measure refers to its ability to measure what it is supposed to measure (Rose & Sullivan, 1996). The validity of a measure is assessed through construct validity (the extent a measure actually measures the construct it claims), concurrent validity (when a measure is administered alongside a pre-existing measure and they two are correlated) and discriminant validity (when a measure has low resemblance with another measure that represents another) (Fitzpatrick, Davey, Buxton & Jones, 1998).

Aims of the review

The aim of the current review was to systematically assess the quality of multi-trait measures of psychological distress and psychiatric symptoms that have been used with adults with intellectual disabilities. This aim was achieved by adhering to the following process:

1. To systematically identify multi-trait measures of psychological distress and psychiatric symptoms used with people who have intellectual disabilities, which have been used in published peer reviewed research studies
2. To identify the key papers that report the development/psychometric assessment of the identified multi-trait measures of psychological distress and psychiatric symptoms used with adults with intellectual disabilities
3. To use Fitzpatrick, Davey, Buxton and Jones (1998) criteria to assess the quality of the multi-trait measures of psychological distress and psychiatric symptoms used with adults with intellectual disabilities

Method

Search Strategy

The initial approach involved searches of three different electronic databases, which included 'MEDLINE', 'PsycINFO' and 'Scopus' to help identify research articles that had evaluated the efficacy of multi-trait measures of psychological distress and psychiatric symptoms, utilised with adults with intellectual disabilities. This search was undertaken using a number of keywords in various combinations as main subject headings or text words in titles, abstracts and main bodies to identify relevant articles. The search strategy contained keywords for the terms 'psych* assess*' and 'outcome* measur*' returned 4,234,078 and 3,239,187 articles respectively. To limit the search to include only the population under investigation, keywords for the terms 'intellectual disabilit*', 'learning disabilit*', 'developmental disabilit*', 'mental retardation', 'mental handicap', mental deficiency' and 'low IQ' were applied, which returned 275,897 articles. The terms describing the population were then combined with 'psych* assess*' and 'outcome* measure*', returning 690 articles.

Inclusion and exclusion criteria

The next step was to then inspect articles for relevance to the current review by applying inclusion and exclusion criteria. Articles were included if i) they were published in English language peer-reviewed journals; ii) the target population was over the age of 18; iii) they utilised a multi-trait measure of psychological distress and psychiatric symptoms and iv) the multi trait measure assessed at least three traits of psychological distress and psychiatric symptoms.

Articles were excluded if i) they were a duplicate; ii) they used a single trait measure of psychological distress and psychiatric symptoms; iii) participants did not present with mental health difficulties and iv) the articles were reviews.

The application of the inclusion and exclusion criteria led to 670 articles being excluded and removed from the review. The strategy therefore returned 20 relevant articles (see Figure 1). The 20 articles that satisfied the inclusion criteria were examined to help ascertain the multi-trait measures used (see Table 1). The primary articles concerned with the measure development and validation within the target population were then identified. These articles reported the development of seven different measures, which have been used to assess psychological distress and psychiatric symptoms with people with intellectual disabilities (see Table 2).

Figure 1 Flow diagram of search strategy

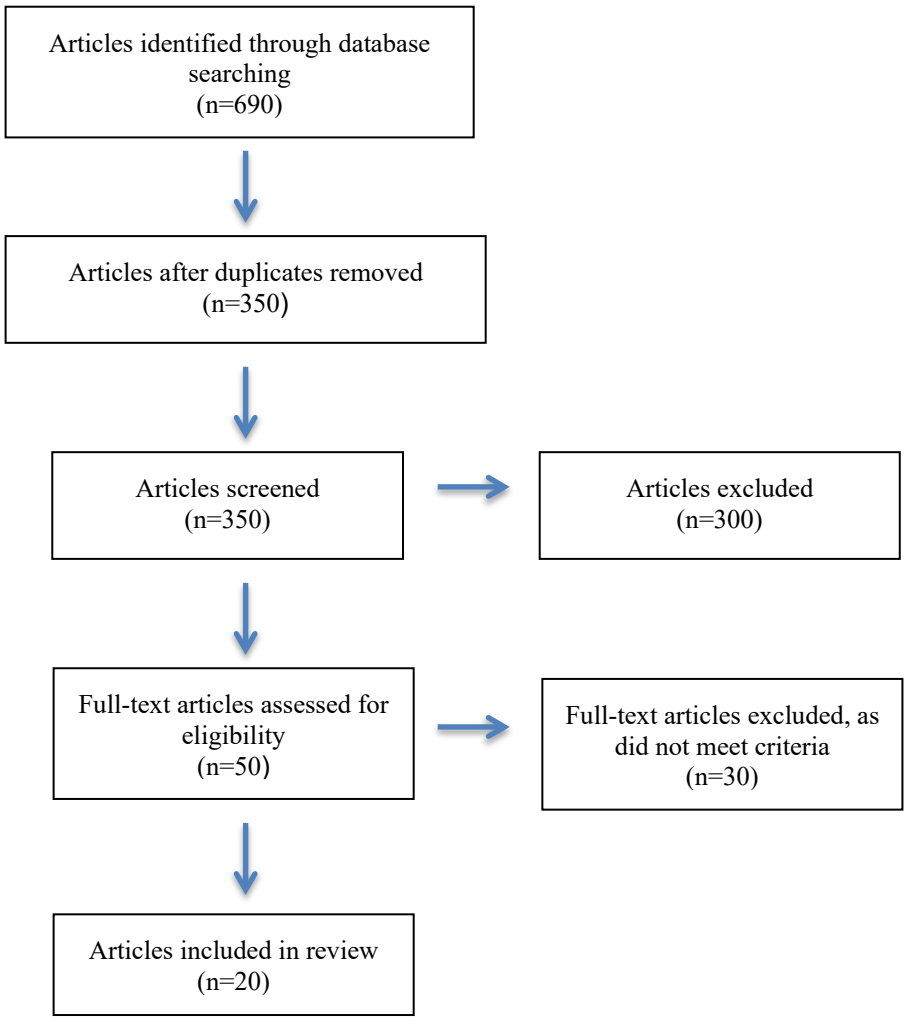


Table 1

Studies that have evaluated psychometric properties of multi-trait measures in intellectual disabilities

Author and Year	Design	Setting	Sample	Multi-trait measure(s)
Beail, Mitchell, Vissides and Jackson (2015)	Cross-sectional	Community	64 male 45 female	Mini-Psychiatric Assessment Schedule for Adults who have Developmental Disabilities (PASADD) Brief Symptom Inventory (BSI)
Gonzalez-Gordon et al., (2002)	Cross-sectional	Community	69 male 11 female	Mini-Psychiatric Assessment Schedule for Adults who have Developmental Disabilities (PASADD)
Gustafsson and Sonnander (2002)	Cross-sectional	Community	73 male 61 female	REISS Screen for Maladaptive Behaviour (RSMB)
Gustafsson and Sonnander (2005)	Cross-sectional	Residential	40 male 31 female	Psychopathology Inventory for Mentally Retarded Adults (PIMRA)
Hove and Havik (2008)	Cross-sectional	Community Residential	300 male 293 female	Psychopathology checklists for Adults with Intellectual Disability (P-AID)
Kellett, Beail, Newman and Mosley (1999)	Cross sectional	Community	60 male 27 female	Symptom Checklist – 90-R (SCL-90-R)
Kellett, Beail, Newman and Frankish (2003)	Cross-sectional	Community	139 male 61 female	Brief Symptom Inventory (BSI)

Author and Year	Design	Setting	Sample	Multi-trait measure(s)
Kellett, Beail, Newman and Hawes (2004)	Cross-sectional	Community	223 male 123 female	Brief Symptom Inventory (BSI)
Masi, Brovedani, Mucci and Favilla (2002)	Cross-sectional	Community	29 male 21 female	Psychopathology Inventory for Mentally Retarded Adults (PIMRA)
Matson and Bamburg (1998)	Cross-sectional	Community	52 male 49 female	Assessment of Dual Diagnosis (ADD)
Matson, Kazdin and Senatore (1984)	Cross-sectional	Community	54 male 56 female	Psychopathology Inventory for Mentally Retarded Adults (PIMRA)
Moss et al., (1993)	Cross-sectional	Community	25 participants	Mini-Psychiatric Assessment Schedule for Adults who have Developmental Disabilities (PASADD)
Moss, Prosser, Ibbotson and Goldberg (1996)	Cross-sectional	Community	100 participants	Mini-Psychiatric Assessment Schedule for Adults who have Developmental Disabilities (PASADD)
Moss et al., (1997)	Cross-sectional	Community	55 male 40 female	Mini-Psychiatric Assessment Schedule for Adults who have Developmental Disabilities (PASADD)

Author and Year	Design	Setting	Sample	Multi-trait measure(s)
Moss et al., (1998)	Study 1- Cross-sectional Study 2 – Cross-sectional	Study 1 - Community Study 2 – Community Hospital	Study 1 - 201 participants Study 2 – 66 participants	Mini-Psychiatric Assessment Schedule for Adults who have Developmental Disabilities (PASADD)
Myrbakk and von Tatzchner (2008)	Cross-sectional	Community	76 male 66 female	Assessment of Dual Diagnosis (ADD)
Prosser et al., (1998)	Cross-sectional	Community	42 male 26 female	Mini-Psychiatric Assessment Schedule for Adults who have Developmental Disabilities (PASADD)
Ramirez and Lukenbill (2008)	Cross-sectional	Community Residential	82 male 55 female	Psychopathology Inventory for Mentally Retarded Adults (PIMRA)
Sturmey et al., (2005)	Cross-sectional	Community	140 male 86 female	Mini-Psychiatric Assessment Schedule for Adults who have Developmental Disabilities (PASADD)
Wieland, Wardenaar, Fontein and Zitman (2012)	Cross-sectional	Community	88 male 136 female	Brief Symptom Inventory (BSI)

Table 2

Multi-trait measures identified for quality appraisal

Measure	Areas assessed	Number of items	Response scale	Population designed for
Assessment of Dual Diagnosis (ADD) Matson and Bamburg (1998)	Mania Depression Anxiety PTSD Substance abuse Somatoform disorders Dementia Conduct disorder Pervasive developmental disorder Schizophrenia Personality disorders Eating disorders Sexual disorders	79 item	3-point Likert scale	Intellectual disabilities
Brief Symptom Inventory (BSI) Derogatis (1993)	Somatization Obsessive compulsive Interpersonal sensitivity Depression Anxiety Hostility Phobic anxiety Paranoid ideation Psychoticism	53 item	5-point Likert scale	General

Measure	Areas assessed	Number of items	Response scale	Population designed for
Psychiatric Assessment Schedule for Adults who have Developmental Disabilities (PASADD) Moss et al (1993)	Organic psychosis Schizophrenia/schizoaffective disorder Autism Depression Anxiety/phobias Mania/hypomania Personality disorder Obsessional disorder Dementia	66 item	4-point Likert scale	Intellectual disabilities
Psychopathology checklists for Adults with Intellectual Disability (P-AID) Hove and Havik (2008)	Psychosis Depression Mania Anxiety OCD Problem behaviour	218 item	6-point Likert scale	Intellectual disabilities
Psychopathology Inventory for Mentally Retarded Adults (PIMRA) Matson, Kazdin and Senatore (1984)	Schizophrenia Affective disorder Psychosexual disorder Adjustment disorder Anxiety disorder Somatoform disorder Personality disorder Inappropriate mental adjustment	56 item	Yes/No	Intellectual disabilities

Measure	Areas assessed	Number of items	Response scale	Population designed for
Reiss Screen for Maladaptive Behaviour (RSMB) Reiss (1988)	Aggressive behaviour Autism Psychosis Paranoia Depression behavioural signs Depression physical signs Dependent personality disorder Avoidant personality disorder	38 item	3-point Likert scale	Intellectual disabilities
Symptom Checklist -90-R (SCL-90-R) Derogatis, Lipman and Covi (1973)	Somatisation Obsessive-compulsive Interpersonal Depression Anxiety Hostility Phobic Paranoid Psychotic	90 item	5-point Likert scale	General

Quality appraisal

In order for a measure to be considered desirable a number of suitable characteristics have been identified (Fitzpatrick, Davey, Buxton & Jones, 1998). These characteristics have been described further below (see Table 3). Additionally, to appraise the quality of these characteristics a rating tool has been developed (Cahill et al., 2008). For the purposes of the current review an adapted version of this rating tool was used to assess the quality of the multi-trait measures used. In this adapted version inter-rater reliability was included to measure the degree of agreement amongst raters of

the measures. The ‘Responsiveness’ criterion was excluded because it does not apply to the purpose of the current review. The ‘Interpretability’ criterion was combined with the ‘Precision’ criterion to assess the quality of the multi-trait measures. This adapted version is summarised below (see Table 4).

Table 3

Fitzpatrick, Davey, Buxton and Jones (1998) criteria taken form Cahill et al., (2008)

Criterion	Definition
Validity	The degree to which a measure genuinely measures the concept that it purposes to measure
Responsiveness	Addresses the question: does the instrument detect changes over time that matter to the patient?
Acceptability	Is the measure acceptable to users?
Feasibility	Is the measure easy to administer and process?
Precision	How precise is the measure?
Interpretability	How interpretable are the scores of the measure

Table 4

Criteria to assess quality of the multi-trait measures (based on Cahill et al., 2008)

Criterion	Definition
Reliability	
Internal consistency	Stability of scores across items that comprise the assessment e.g. measuring whether the items of a scale are measuring the same thing. This usually measured by Cronbach's alpha
Inter-rater reliability	Measures of two or more assessment scorers or observers. This type of reliability is usually not applicable in self-report measures unless the measure was rated by an interviewer/researcher
Test-retest reliability	Degree to which test results are consistent or stable at different time points of administration
Validity	
Construct validity	Extent to which an assessment accurately measures a construct or trait i.e. the degree to which operationalisations of a construct actually measure what the theory states it does
Concurrent validity	A new measure is administered together with a pre-existing one and the two are correlated
Convergent validity	A measure converges with other indications of the same concept that it theoretically should be similar to
Discriminant validity	A measure exhibits low levels of resemblance with a measure that theoretically it should not be similar to
Acceptability	Is the measure acceptable for users? Practicality of administration Time taken to complete Length of instrument Translations Access by ethnic minorities Reading age

Criterion	Definition
Feasibility	Is the measure easy to administer and process? Cost and burden to administrative staff
	Electronic scanning options
	Scoring systems
	Training package
	Training manual
	Support from measure developers
Precision	Interpretability
	Normative data

Analysis

In order to critically evaluate the multi-trait measures information was extracted from the articles identified. Coding instructions were applied to assess the quality of each measure (see Table 5). This coding provided an approximation for each of the six criteria described above.

Table 5

Instructions for coding for the quality assessment of multi-trait measures

Fitzpatrick's criteria	Coding	Explanation
Reliability	Adequate	>0.7
	Partial	>0.5 <0.7
	Inadequate	<0.5
	Unknown	Reliability not supported
Validity	Adequate	Reports >3 types of validity tests
	Partial	Report 2 types of validity tests
	Inadequate	Report 1 validity test
	Unknown	Validity estimates not supplied
Acceptability	Adequate	All of the components described
	Partial	At least one of the components described
	Inadequate	None of the components described
Feasibility	Adequate	All of the components described
	Partial	At least one of components described
	Inadequate	None of the components described
Precision	Adequate	All of the components described
	Partial	At least one of components described
	Inadequate	None of the components described

Results

Information about the multi-trait measures of psychological distress and psychiatric symptoms assessed in the current review is contained in Table 2. This table shows that there were seven multi-trait measures in total measuring a number of different psychological and psychiatric problems. Six of these included depression and anxiety; five included psychosis; four included personality disorders obsessive compulsive disorder (OCD) and somatoform disorder; three included schizophrenia, paranoia, phobias, hostility and mania; two included autistic spectrum disorder (ASD), interpersonal sensitivity and psycho-sexual disorders and one included dementia, conduct disorder, affective disorders, post-traumatic stress disorder (PTSD), substance abuse, pervasive developmental disorder and eating disorder. In six of the measures a Likert scale was used utilising either a three, four, five or six point scale. One of the measures used a yes/no response. With the exception of two measures (BSI and SCL-90-R) all of the measures were designed specifically for use with people who have intellectual disabilities. The important psychometric properties are detailed in Table 6 along with the quality of each multi-trait measure using Fitzpatrick, Davey, Buxton and Jones (1998) criteria.

Table 6

Quality appraisal of the multi-trait outcome measures

Measure	Reliability			Validity				Acceptability	Feasibility	Precision
	Internal	Inter rater	Test/ retest	Construct	Concurrent	Convergent	Discriminant			
Assessment of Dual Diagnosis (ADD)	0.93	0.98	0.82-1.00	No details	Relationship with Mini-PASADD, and RSMB	No details	No details	Partial– describes time taken to complete and length of instrument	Partial – describes how administered and training required to administer	Partial – describes how data is scored
Brief Symptom Inventory (BSI)	0.93	N/A	No detail	8 factor structure	Significant correlation with PASADD	No details	No details	Partial – describes practicality of administration, time taken to complete, length of instrument and translations	Partial – describes how administered, assisted completion format and training required to administer	Adequate – differences found between populations, how data is scored and cites benchmarks
Psychiatric Assessment Schedule for Adults who have Developmental Disabilities (PASADD)	0.87	0.80	0.60	8 factor structure	Significant correlation with BSI	No details	No details	Partial – describes practicality of administration, time taken to complete, length of instrument and translations	Partial – describes how administered, how data is scored, training required to administer and training manual	Adequate – differences found between populations, how data is scored and cites benchmarks

Measure	Reliability			Validity				Acceptability	Feasibility	Precision
	Internal	Inter rater	Test/ retest	Construct	Concurrent	Convergent	Discriminant			
Psychopathology checklists for Adults with Intellectual Disability (P-AID)	0.89	0.52	No detail	4 factor structure	No details	No details	No details	Partial – describes practicality of administration and length of instrument	Partial – describes how administered, how data is scored	Partial – describes how data is scored
Psychopathology Inventory for Mentally Retarded Adults (PIMRA)	0.85	0.71	0.91	3 factor structure	Correlations between RSMB	No details	No details	Partial – describes length of instrument and translations	Partial – describes how administered, how data is scored, training required to administer	Not addressed
REISS Screen for Maladaptive Behaviour (RSMB)	0.90	0.56	No detail	7 factor structure	No details	No details	No details	Partial – describes length of instrument, time taken to complete and translations	Partial – describes how administered, how data is scored,	Not addressed
Symptom Checklist-90-R (SCL-90-R)	0.86	N/A	0.78-0.90	No details	No details	No details	No details	Partial – describes length of instrument and time taken to complete	Partial – describes how administered, how data is scored, training required to administer and assisted completion	Adequate – differences found between populations, how data is scored and cites benchmarks

Reliability

The measures assessed in the current review all showed adequate levels of internal consistency ($\alpha > 0.7$) ranging between $\alpha = 0.85 – 0.93$. In terms of inter rater reliability all of the measures showed good reliability ($r = 0.52 – 0.98$), with the exception of the BSI and the SCL-90-R which did not report any data as they are self-report measures. In terms of test-retest reliability only four of the measures reported data (ADD, PASADD, PIMRA and SCL-90-R). All four of these measures showed moderate to strong levels of test-retest reliability ranging from 0.60 – 1.00. Overall, the ADD and the PASADD, with a combination of excellent internal consistency, inter-rater reliability and test-retest reliability scores, showed the best levels of reliability.

Validity

The number and quality of validity assessments varied for all of the measures included in the current review. All of the measures had some form of validity assessment conducted (see Table 6). Concurrent validity was only assessed on the ADD, PASADD and PIMRA. The ADD was found to have a relationship with the PASADD and the RSMB. The PASADD was found to have a relationship with the BSI and the PIMRA was found to have a relationship with the RSMB.

Construct validity was assessed on all of the measures with the exception of the ADD and SCL-90-R. As all but two of the measures (BSI and SCL-90-R) were explicitly designed for use with people with intellectually disabilities there was no scope to compare the results of the factor analysis to the general population. The exception was the BSI, which was originally designed for the general population. Interestingly, in the Kellet, Beail, Newman and Hawes (2004) study the factor analysis of the BSI slightly differed with that which was originally conducted on the general population during its development. The BSI, for example had been found to consist of nine primary symptom dimensions with a general adult population (Derogatis, 1993).

However, Kellett, Beail, Newman and Hawes, (2004) found only eight interpretable factors in their analysis with 335 adults with intellectual disabilities. One of the scales was identical (originally OCD) but renamed 'cognitive impairment' because people with intellectual disabilities appeared to report distress in relation to their memory, concentration, decision-making and double-checking and the impact this had on their general functioning. The 'anxiety' and 'somatisation' scale contained all of the original items along with additional items. There were two new scales identified labelled 'anger' and 'suicidal ideation'. There were two scales, which were not replicated 'psychoticism' and 'interpersonal sensitivity'. Most of the items from the 'psychoticism' scale converged with the 'depression' scale and the items from the 'interpersonal sensitivity' scale distributed across three of the new proposed factors. Wieland et al (2012) report that they found the same factor structure as in the general population. However, they did not publish the factor loadings data to support this.

The PASADD was shown to have eight-factors which accounted for 65.3% of the variance (Moss et al., 1998). The majority of the factors were readily interpretable in psychiatric terms. The factors were labelled as depression, restlessness, phobic anxiety, psychosis, hypomania, autistic spectrum, depression and non-specific. There were two factors relating to depression with the distinction being that suicidal thoughts and loss of self esteem were more likely to be reported in people whose level of intellectual disability was mild or moderate with the symptoms, loss of appetite, jumpy, irritable or bad tempered, depressed mood, social withdrawal and loss of interest loading on the other factor having a lower association with developmental level.

The P-AID was found to have four components with eigenvalues exceeding 1.00 explaining 54.9% of the total variance (Hove & Havik, 2008). The first factor was labelled problem behaviours and included, physical aggression, destructive behaviour, self-injuries behaviour, verbal aggression, demanding behaviour, OCD, and

oppositional behaviour. This factor accounted for 17.5% of the total variance. The second factor was labelled anxiety and included agoraphobia, social phobia, panic disorder, generalised anxiety, and specific phobia. This factor explained 14.5% of the total variance. The third factor was labelled severe psychopathology and included, depression, dementia, mania and psychosis and this factor explained 14.5% of the total variance. The fourth factor was harder to characterise and was defined by two types of problem, wandering and sexually inappropriate behaviour and this factor explained 8.2% of the total variance.

The PIMRA was shown to have a three-factor solution. The criteria stipulated for a factor was an eigenvalue of 1.50 or above and items with a loading of 0.35 or above were included in each factor. Furthermore, each factor had to have at least five items to be included (Matson, Kazdin & Senatore, 1984). Items were placed only in the factor where the highest factor loading occurred. The first factor was labelled 'affective' and contained 14 items, the second factor was labelled 'somatoform' and contained five items and the third factor was labelled 'psychosis' and contained five items.

The RSMB was shown to have a seven-factor structure (Reiss, 1988). These seven components had eigenvalues >1.00 and they explained 67% of the total variance. These components were labelled as aggressive behaviour, avoidant behaviour, depression (B), psychosis, dependent, paranoia and depression (P). The variance explained by the seven components was 14.8%, 13.5%, 9.3%, 8.3%, 7.7%, 7.3% and 6.5%.

None of the studies included in the current review included data on either convergent or discriminant validity of any of the measures under consideration.

Acceptability

None of the measures in the current review reached adequate levels of acceptability. However, all achieved partial levels of acceptability, (see Table 6). In

these measures data was reported on the following; practicality of administration (BSI, PASADD and P-AID), time taken to complete (BSI, ADD, PASADD, RSMB and SCL-90-R), length of the instrument (ADD, BSI, PASADD, P-AID, PIMRA, RSMB and SCL-90-R) and translations (PASADD, PIMRA and RSMB).

In terms of amount of time that is required to complete a measure, detailed information was provided for five of the measures (ADD, BSI, PASADD, RSMB and SCL-90-R). The amount of time taken to complete each measure varies (ADD, 20 minutes; BSI, eight-ten minutes; PASADD, 10 minutes; RSMB, six minutes and SCL-90-R, 12-15 minutes). The ADD takes the most time to complete, however, since this is an informant-based measure the length of time to administer may be less significant as the chances of people not being able to tolerate the measure may be reduced. The rest of the measures take a relatively shorter period of time to administer therefore suggesting a higher rate of acceptability for both interviewer and interviewee.

Information was provided on the length of all of the measures in the current review (see table 2). The measures ranged from 38 items (RSMB) to 280 items (P-AID). The P-AID is significantly longer than any of the other measures included in the review, however it has been designed for detailed analyses of several different mental disorders and to fit the diagnostic categories of the "diagnostic criterion for psychiatric disorder for use with adults with learning disabilities/mental retardation" (DC-LD). The use of this measure could be seen as cumbersome and therefore affect the completion rate. Furthermore, given that other measures such as the RSMB and PASADD measure similar constructs and take significantly less time to administer use of these measures may be more appropriate as they are potentially less burdensome.

Information on translations of measures into different languages was provided on three of the measures (PASADD, PIMRA and RSMB). The measures have been translated into a number of different European languages including French, Spanish,

Swedish, German and Dutch. The translation of measures will make them more accessible for non-English speaking users. However, there was only limited information available as to the reliability and validity of these translated measures.

In relation to the populations that the measures may not be appropriate for the psychometric properties of all the measures were evaluated with people with 'mild' to 'moderate' intellectual disability. This therefore raises the question as to whether the measures are suitable for use in people with 'severe' intellectual disabilities? There were also no reports of service user views of the measures in any of the studies.

Feasibility

None of the measures in the current review reached adequate levels of feasibility and this was mainly due to a lack of information reported. However, all of the measures reached partial feasibility (see Table 6). There were descriptions of the instructions in how to administer all of the measures. These instructions included the resources required to administer the measure (paper-and–pencil, CD, computer or online administration), professional status/background required to administer the measures, criteria for interviewees/respondents (informant based measures only), instructions that are required to be provided to the interviewee/ respondents and settings/locations required to administer the measures. Five of the measures described the training required to administer (ADD, BSI, PASADD, PIMRA and SCL-90-R). The training included the attendance of formal training sessions organised by the various developers of the measures, being supervised in the administrations of the measure by an individual who has successfully completed the aforementioned training and accessing online tutorials and documents related to training. Five of the measures also described how the data is scored (P-AID, PASADD, PIMRA, RSMB and SCL-90-R). Information on scoring included the scales used to measure responses (Likert or yes/no), reversing of scores for items, converting scores and access to norms. Two measures described an

assisted completion method (BSI and SCL-90-R) and one measure described the training manual required (PASADD). Both of these measures illustrated the efficiency of an assisted completion format with people with an intellectual disability. This format was shown not to influence the participant's ratings of symptoms and did not compromise with the psychometric properties of the measures.

Precision

The level of precision varied amongst measures, with three measures demonstrating adequate levels, two demonstrating partial levels and two having inadequate levels (see Table 6). The three measures demonstrating adequate levels were the BSI, PASADD and SCL-90-R which all found differences between populations, demonstrated how the data is scored and cited benchmarks to assist in the interpretation of scores. The two measures that demonstrated partial levels of precision (ADD and P-AID) both reported how data is scored. The information provided included data on populations with varying levels of intellectual disabilities, data on populations with mental health problems only and data on 'control/normal' populations for comparison of the data and benchmarks. The two measures that demonstrated inadequate levels were the PIMRA and RSMB as they did not address the issue.

Discussion

The aim of the current review was to systematically evaluate the psychometric properties of multi-trait measures of psychological distress and psychiatric symptoms that have been used with adults with an intellectual disability in clinical practice and research. The number of measures, which were identified for the quality appraisal, were quite small (seven), however there were a number of studies, which examined their psychometric properties (20).

Summary of findings

In summary, all of the measures assessed were shown to have high levels of internal consistency. In terms of inter-rater reliability the informant based measures showed good reliability. In terms of test-retest reliability the ADD, PASADD, PIMRA and SCL-90-R showed moderate to strong levels of test-retest reliability following a two-week period, no data was available for the other measures. Overall, the ADD and the PASADD, with a combination of excellent internal consistency, inter-rater reliability and test-retest reliability scores, showed the best levels of reliability. However, the size of the sample in the original PASADD study was very small (n = 25) which creates difficulties in generalising the conclusions to a broader population. However, there have been further studies using larger sample sizes, which have assessed the psychometric properties of the PASADD showing similar results, which have been included in the current review (Beail, Mitchell, Vissides & Jackson, 2015; Gonzalez-Gordon et al., 2002; Moss, Prosser, Ibbotson & Goldberg 1996; Moss et al., 1997; Moss et al., 1998; Prosser et al., 1998; Sturmey et al., 2005) therefore providing further support for the reliability of the PASADD.

The current review highlighted that the assessment of the validity of the measures available is lacking. The number and quality of the assessments varied across the measures. Only four of the measures were assessed for concurrent validity (ADD,

BSI, PASADD and PIMRA), five of the measures were assessed for construct validity (BSI, P-AID, PASADD, PIMRA and RSMB) and none of the measures were assessed for convergent or discriminant validity. The importance of measuring construct validity especially in measures that have been originally developed for the general population is considered key. In the current study there were only two measures (BSI and SCL-90-R), which have been originally developed for the general population and later adapted for use with individuals with intellectual disabilities and only one of these measures (BSI) has had its construct validity assessed. This measure was found to factor slightly differently with adults with intellectual disabilities compared to the general population (Kellett, Beail, Newman & Hawes, 2004). This finding may suggest that experiences of mental health difficulties in adults with intellectual disabilities may be different to the experiences of the general population. This could mean that multi-trait measures designed to assess psychological distress with adults with intellectual disabilities may not be assessing the constructs we believe they are. This also raises the question about whether clinicians should be using measures, which have not been assessed for construct validity like the SCL-90-R, especially with adults with intellectual disabilities. The factor analysis of the other measures also suggests that the original constructs did not factor as predicted.

The acceptability of all of the multi-trait measures included in the current review within the target population was a further area that was not adequately addressed. The ADD, P-AID, PASADD, PIMRA and RSMB had been explicitly developed for use with adults with intellectual disabilities, however even these measures only reached partial acceptability. The other two measures (BSI and SCL-90-R) had been adapted for use with this population. There was no information provided on any involvement of service users in the adaptation of any of the measures or on service user feedback on how they found completing the measures. This exclusion does not conform to the

modern day initiative for increased involvement of service users in health care provision for intellectual disability services (Department of Health, 2009). The involvement of service users in the development of services, resources and assessment tools is considered crucial in facilitating the delivery of effective provision (Greenhill & Whitehead, 2010; Roberts et al., 2013). Therefore, the inclusion of service user feedback can be considered essential to ensure adequate acceptability of a measure.

A further issue that was identified directly effecting the acceptability of the multi-trait measures included in the current review was that the psychometric properties of all the measures have been assessed on people with 'mild' to 'moderate' intellectual disabilities. This method of assessment raises the question as to whether the uses of these multi-trait measures are acceptable in individuals with 'severe' intellectual disabilities. Therefore, further assessment of the psychometric properties in populations from all ends of the spectrum is required to improve the acceptability of the measure.

There was an absence of information on access by ethnic minorities, which also affects the acceptability of the measures. This lack of information is particularly troubling because there is already recognition of how measures can have bias in terms of inappropriate content, inappropriate standardised samples, examiner and language bias, inequitable social consequences, measurement of different constructs and differential predictive validity (Reynolds & Brown, 1984). This therefore means that any development of future measures has to take this factor into consideration as neglecting this issue could potentially lead to the measure being less valid in ethnic minority groups.

In terms of feasibility there was sufficient information offered on how to complete and administer the measures. However, there was less information available on assisted completion formats and training manuals available for some of the measures. This may be less problematic for those measures, which are well established

and have manuals with detailed administration instructions such as the ADD, BSI, PASADD, RSMB and SCL-90-R. Although the combination of these information sources may be utilised to help with administration, there is always the risk of individuals administering the measures, adapting items for service users as they see fit thus affecting the meaning of the items and the overall validity of the measure.

Finally, in relation to the precision of the measures there was limited information provided for some of the measures on benchmarks and cut offs. This missing information is important as it impedes the ability of assessing the severity of individual's difficulties in relation to a normal population and to assess if individuals with intellectual disabilities experience more pronounced difficulties. In addition there is also a need for the ability to be able to monitor any clinically significant changes in presentation over time or during treatment to be able to assess its effectiveness (Evans et al., 2002).

Limitations

There were a number of methodological limitations of the current review. Although all attempts were made to identify relevant papers for the review a forward citation search was not conducted therefore potentially leading to some papers being omitted during the search. There was no inter-rater reliability conducted in relation to screening and eligibility of the papers included in the review. Articles published in languages other than English were not included in the review therefore potentially resulting in the omission of some key papers.

Methodological limitations of the studies included in the current review were that all of the participants used within the studies had either a mild or moderate intellectual disability. There was no reference to those with more 'severe' intellectual disabilities thus bringing the generalisability of the findings into question in this population. There was a lack of inclusion of participants from diverse ethnic

backgrounds, with the proportion of participants being predominately Caucasian. There was also an overrepresentation of males in all of the studies. Again affecting the generalisability of the findings.

Another limitation was the ability of evaluating the quality of the research assessing the psychometric properties of the multi-trait measures in intellectual disabilities samples. Many of the studies identified were selective in terms of their assessment of reliability and validity with there being no real 'gold standard' of assessment established. There were also inconsistencies in the way acceptability, feasibility and precision were reported in relation to the measures therefore making it difficult to assess quality. A good example of this was that none of the studies reported any information on individuals who did not complete the measures and the reasons behind this, thus making it difficult to understand when the measure may not be suitable for use and in what instances. Another example was that there was a lack of detailed information on training manuals. All of the papers implied that there was a need for training to administer the measures, however no information was offered as to what this training is, how long it would take to complete and what the thresholds are to passing through the training.

Clinical Implications

In terms of clinical implications the findings from the current review show that there are a limited number of multi-trait measures of psychological distress used in adult intellectual disabilities samples compared to other non-intellectually disabled populations. The quality appraisal of these measures in the current review showed that these multi-trait measures have varying levels of psychometric properties and that there is still work to be done to improve some of their overall quality. The best measures available at the moment appear to be the PASADD and the ADD in terms of reliability and the PASADD in terms of validity and would be recommended for use to clinicians

in terms of informant based measures. In terms of self-report measures the BSI that appears to have the most reasonable psychometric properties and would be recommended for use.

Future Research

The results from the current review reveal that there are potentially five key areas that need to be addressed. First, a full evaluation of all elements of reliability and validity in the multi-trait measures included in the current review. Second, a detailed assessment of the acceptance, feasibility and precision of all of the multi-trait measure included in the review. Third, the inclusion of service users from the complete spectrum of intellectual disabilities and not just those with mild or moderate intellectual disabilities when assessing the psychometric properties of the multi-trait measures. Fourth, a more balanced male to female ratio and the inclusion of people from ethnic minorities in any future research assessing the psychometric properties of the multi-trait measure. Fifth, the inclusion of service users in the development of any future multi-trait measures. Overall further examination of the psychometric properties of these multi-trait measures would provide a more robust evidence base and more information about the clinical utility of these measures in people with intellectual disabilities.

References

Asada, M., Tomiwa, M., Okada, M., & Itakura, S. (2010). Atypical verbal communication pattern according to others attention in children with Williams syndrome. *Research in Developmental Disabilities, 31*, 452-458. doi:10.1016/j.ridd.2009.10.010

Beail, N., Kellett, S., Newman, D., & Warden, S. (2007). The does-effect relationship in psychodynamic psychotherapy with people with intellectual disabilities. *Journal of Applied Research in Intellectual Disabilities, 20,* 448-454. doi: 10.1111/j.1468-3148

Beail, N., Mitchell, K., Vlissides, N., & Jackson, T. (2015). Concordance of the mini-Psychiatric assessment schedule for adults who have developmental disabilities (PASADD) and the brief symptom inventory. *Journal of Intellectual Disability Research, 59*, 170-175. doi: 10.1111/jir.12073

Beail, N., Warden, S., Morsely, K., & Newman, D. (2005). A naturalistic evaluation of the effectiveness of psychodynamic psychotherapy with adults with intellectual disabilities. *Journal of Applied Research in Intellectual Disabilities, 18,* 245-251. doi: 10.1111/j.1468-3148

Beck, A.T., Steer, R., Brown, G. (1988). *Beck Depression Inventory – 2nd Edition Manual.* San Antonio, TX: The Psychological Corporation

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,* 1-68. doi:10.3310/hta12240

Cheng, Y., & Chen, S. (2010). Improving social understanding of individuals of

intellectual and developmental disabilities through a 3-D facial expression intervention program. *Research in Developmental Disabilities, 31*, 1434-1442. doi:10.1016/j.ridd.2010.06.015

Cooper, S., & Bailey, N. (2001). Psychiatric disorders amongst adults with learning disabilities: prevalence and relationships to ability level. *Irish Journal of Psychological Medicine, 18*, 45-53. doi:10.1017/s0790966700006315

Cooper, S., Smiley, E., Morrison, J, Williamson, A., & Allan, L. (2007). Mental ill health in adults with learning disabilities: prevalence and associated factors. *British Journal of Psychiatry, 190,* 27-35. doi:10.1192/bjp.bp.106.022483

Deb, S., Thomas, M., & Bright, C. (2001). Mental disorder in adults with intellectual disability. I: prevalence of functional psychiatric illness among a community based population aged between 16 and 64 years. *Journal of Intellectual Disability Research, 45*, 495-505. doi:10.1046/j.1365-2788.2001.00374.x

Department of Health (2009). *NHS psychotherapy services in England: Review of strategic policy:* London: Department of Health

Derogatis, L. (1993). *Brief Symptom Inventory: administration scoring and procedures manual* (3rd ed.). Minneapolis: National Computer Systems.

Derogatis, L., Lipman, J., & Covi, R. (1973). *SCL-90R; Administration, Scoring and Procedures: Manual 11.* Towson, MD: Clinical Psychometric Research

Evans, C., Connell, J., Barkham, M., Margison, F., McGrath, G., Mellor-Clark, J., & Audin, K. (2002). Towards a standardised brief outcome measures: psychometric properties and utility of the CORE-OM. *British Journal of Psychiatry, 180*, 51-60. doi:10.1192/bjp.180.1.51

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

Fitzpatrick, R., Davey, C., Buxton, M., & Jones, D. (1998). Evaluating patient based

outcome measures for use in clinical trials. *Health Technology Assessment, 7*, 286-299. doi:10.4135/9781848608344.n11

Gonzalez-Gordon, R., Salvador-Caulla, L., Romero, C., Gonzalez-Saiz, F., & Romero, D. (2002). Feasibility, reliability, and validity of the Spanish version of the psychiatric assessment schedule for adults with developmental disability: A structured psychiatric interview for intellectual disability. *Journal of Intellectual Disability Research, 46*, 209-217. doi:10.1046/j.1365-2788.2002.00402.x

Greenhill, B., & Whitehead, R. (2010). Promoting service user inclusion in risk assessment and management: a pilot project developing human rights based approach. *British Journal of Learning Disabilties, 39*, 277-283. doi:10.1111/j.1468-3156.2010.00664.x

Gustafsson, C., & Sonnander, K. (2002). Psychometric evaluation of a Swedish version of the Reiss Screen of Maladaptive Behaviour. *Journal of Intellectual Disability Research, 46*, 218-229. doi:10.1046/j.1365-2788.2002.00398.x

Gustafsson, C., & Sonnander, K. (2005). A psychometric evaluation of a Swedish version of the psychopathology inventory for mentally retarded adults (PIMRA). *Research in Developmental Disabilities, 26*, 183-201. doi:10.1016/j.ridd.2004.04.002

Hastings, R., Beck, A., Daley, D., & Hill, C. (2005). Symptoms of ADHD and their correlates in children with intellectual disabilities. *Research in Developmental Disabilities, 26*, 456-468. doi:10.1016/j.ridd.2004.10.003

Hermans, H., & Evenhuis, H. (2010). Characteristics of instruments screening for depression in adults with intellectual disabilities: Systematic review. *Research in Developmental Disabilities, 31*, 1109-1120. doi:10.1016/j.ridd.2010.04.023

Hermans, H., van der Pas, F., & Evenhuis, H. (2011). Instruments assessing anxiety in

adults with intellectual disabilities: A systematic review. *Research in Developmental Disabilities, 32*, 861-870. doi:10.1016/j.ridd.2011.01.034

Hove, O., & Havik, O. (2008). Psychometric properties of psychopathology checklists for adults with intellectual disabilities (P-AID) on a community sample of adults with intellectual disability. *Research in Developmental Disabilities, 29*, 467-482. doi:10.1016/j.ridd.2007.09.002

Kellett, S., Beail, N., Newman, D., & 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. doi:10.1046/j.1468-3148.2003.00152.x

Kellett, S., Beail, N., Newman, D., & Hawes, A. (2004). The factor structure of the brief symptom inventory: intellectual disability evidence. *Clinical Psychology & Psychotherapy, 11*, 275-281. doi:10.1002/cpp.410

Kellett, S., Beail, N., Newman, D., & 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*, 323-334. doi:10.1111/j.1468-3148.1999.tb00088.x

Kovacs, M., Paulauskas, S., Gatsonis, C., & Richards, C. (1988). Depressive disorders in childhood III. A longitudinal study of comorbidity with and risk for conduct disorders. *Journal of Affective Disorders, 15*, 205-217. doi:10.1016/0165-0327(88)90018-3

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

Lambrechts, G., Van Den Noortgate, W., Eeman, L., & Maes, B. (2010). Staff

reactions to challenging behaviour: An observational study. *Research in Developmental Disabilities, 31*, 525-535. doi:10.1016/j.ridd.2009.12.004

Lante, K., Reece, J., & Walkley, J. (2010). Energy expanded by adults with and without intellectual disabilities during activities of daily living. *Research in Developmental Disabilities, 31*, 1380-1389. doi:10.1016/j.ridd.2010.06.022

Lindsay, W., Tinsley, S., Beail, N., Hastings, R., Jahoda, A., Taylor, J & Hutton C. (2015) A preliminary controlled trial of transdiagnostic programme for cognitive therapy with adults with intellectual disabilities. *Journal of Intellectual Disability Research, 59,* 360-369. doi: 10.1111/jir.12145

Masi, G., Brovedani, P., Mucci, M., & Favilla, L. (2002). Assessment of anxiety and depression in adolescents with mental retardation. *Child Psychiatry and Human Developmet, 32*, 227-237. doi:10.1023/a:1017908823046

Matson, J., & Bamburg, J. (1998). Reliability of the assessment of dual diagnosis (ADD). *Research in Developmental Disabilities, 19*, 89-95. doi:10.1016/s0891-4222(97)00031-0

Matson, J., Belva, B., Hattier, M., & Matson, M. (2012). Scaling methods to measure psychopathology in persons with intellectual disabilities. *Research in Developmental Disabilities, 33*, 549-562. doi:10.1016/j.ridd.2011.10.023

Matson, J., Kazdin, A., & Senatore, V. (1984). Psychometric properties of the psychopathology instrument for mentally retarded adults. *Applied Research in Mental Retardation. 5*, 881-889. doi:10.1016/s0270-3092(84)80021-1

Matson, J., Kiely, S., & Bamburg, J. (1997). The effects of stereotypies on adaptive skills as assessed with the DASH-II and the Vineland adaptive behaviour scales. *Research in Developmental Disabilities, 18*, 471-476. doi:10.1016/s0891-4222(97)00023-1

Matson, J., Smiroldo, B., & Bamburg, J. (1998). The relationship of social skills to

psychopathology for individuals with severe or profound mental retardation. *Journal of Intellectual and Developmental Disability, 23*, 137-145. doi:10.1080/13668259800033641

McGrother, C., Hauck, A., Bhaumik, S., Thorp, C., & Taub., N. (1996). Community care for adults with learning disability and their carers: Needs ant outcomes from the Leicester register. *Journal of Intellectual Disability Research, 40*, 183-190. doi:10.1046/j.1365-2788.1996.784784.x

Moss, S., Ibbotson, H., Prosser, D., Goldberg, D., Patel, P., & Simpson N. (1997). Validity of the PAS-ADD for detecting psychiatric symptoms in adults with learning disability (mental retardation). *Social Psychiatry and Psychiatric Epidemiology, 32*, 344-354. doi:10.1007/bf00805440

Moss, S., Patel, P., Prosser, H., Goldberg, D., Simpson, N., Rowe, S., & Lucchino, R. (1993). Psychiatric morbidity on older people with moderate and sever learning disability I: development and reliability of the patient interview (PAS-ADD). *British Journal of Psychiatry, 163*, 471-480. doi:10.1192/bjp.163.4.471

Moss, S., Prosser, H., Ibbotson, B., & Goldberg, D. (1996). Respondent and informant accounts of psychiatric symptoms in a sample of patients with learning disability. *Journal of Intellectual Disability Research, 40*, 457-465. doi:10.1046/j.1365-2788.1996.792792.x

Moss, S., Prosser, H., Costello, H., Simpson, N., Patel, P., Rowe, S., Turner, S., & Hatton, C. (1998). Reliability and validity of the PAS-ADD checklist for detecting psychiatric disorders in adults with intellectual disability. *Journal of Intellectual Disability Research, 42*, 173-183. doi:10.1046/j.1365-2788.1998.00116.x

Myrbakk, E., & von Tetzchner, S. (2008). Screening individuals with intellectual

disability for psychiatric disorders: comparison of four measures. *American Journal on Mental Retardation, 1*, 54-70. doi:10.1352/0895-8017(2008)1132.0.co;2

Prosser, H., Moss, S., Costello, H., Simpson, N., Patel, P., & Rowe, S. (1998). Reliability and validity of the Mini-PAS-ADD for assessing psychiatric disorders in adults with intellectual disability. *Journal of Intellectual Disability Research, 42*, 264-272. doi:10.1046/j.1365-2788.1998.00146.x

Ramirez, S., & Lukenbill, J, (2008). Psychometric properties of the zung self-rating anxiety scale for adults with intellectual disabilities (SAS-ID). *Journal of Developmental and Physical Disabilities, 20*, 215-228. doi:10.1007/s10882-008-9120-x

Reiss, S. (1988). *Test manual for the Reiss screen for maladaptive behaviour.* Orlando: International Diagnostic Systems

Reynolds, C & Brown, R. (1984). *Perspectives on bias in mental testing.* London: Open University Press.

Roberts, A., Townsend, S., Morris, J., Rushbrooke, E., Greenhill, B., Whitehead, R., Matthews, T., & Golding, L. (2013). Treat me right, treat me equal: using national policy and legislation to create positive changes in local health service for people with intellectual disabilities. *Journal of Applied Research in Intellectual Disabilities, 26*, 14-25. doi:10.1111/jar.12009

Rose, J. (2010). Carer reports of the efficacy of cognitive behavioural interventions for anger. *Research in Developmental Disabilities, 3*, 28-50. doi:10.1016/j.ridd.2010.06.007

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

Rush, K., Bowman, L., Eidman, S., Toole, L., & Mortenson, B. (2004). Assessing

psychopathology in individuals with developmental disabilities. *Behaviour Modification, 28*, 621-637. doi:10.1177/0145445503259830

Shin, C., Shin, C., & Wang, S. (2010). Assisting people with disabilities improves their collaborative pointing efficiency with a multiple cursor dynamic pointing assistive program. *Research in Developmental Disabilities, 31*, 1251-1257. doi:10.1016/j.ridd.2010.07.020

Smiley, E. (2005). Epidemiology of mental health problems in adults with learning disability: an update. *Advances in Psychiatric Treatment, 11*, 214-222. doi:10.1192/apt.11.3.214

Sturmey, P., Newton, J., Cowley, A., Bouras, N., & Holt, G. (2005). The PAS-ADD checklist: independent replication of its psychometric properties in a community sample. *British Journal of Psychiatry, 186*, 319-323. doi:10.1192/bjp.186.4.319

Taylor, J., Hatton, C., Dixon, L., & Douglas, C. (2004). Screening for psychiatric symptoms: PAS-ADD checklist norms for adults with intellectual disabilities. *Journal of Intellectual Disability Research, 48*, 37-41. doi:10.1111/j.1365-2788.2004.00585.x

Trower, T., Treadwell, L., & Bhaumik, S. (1998). Acute in-patient treatment for adults with learning disabilities and mental health problems in a specialised admission unit. *British Journal of Developmental Disabilites, 86*, 20-29. doi:10.1179/096979598799156182

Wieland, J., Wardenaar, K., Fontein, E., & Zitman, F. (2012). Utility of the brief symptom inventory (BSI) in psychiatric outpatients with intellectual disabilities. *Journal of Intellectual Disability Research, 56,* 843-853. doi:10.1111/j.1365-2788.2011.01440.x

Williams, D. (2010). Reducing and eliminating restraint of people with developmental

disabilities and severe behaviour disorders: an overview of recent research. *Research in Developmental Disabilities, 31*, 1142-1148. doi:10.1016/j.ridd.2010.07.014

Wise, J., Sevik, R., Romski, M., & Morris, R. (2010). The relationship between phonological processing skills and word and non-word identification performance in children with mild intellectual disabilities. *Research in Developmental Disabilities, 31*, 1170-1175. doi:10.1016/j.ridd.2010.08.004

Section two: Empirical report

The construct validity of the Wechsler Adult Intelligence Scale-Fourth Edition (WAIS-IV) and its relationship with adaptive behaviour in the assessment and diagnosis of intellectual disabilities in a UK population

Abstract

Objective. The WAIS-IV is a popular tool for the assessment of intelligence. The Vineland-II claims to measure the personal and social skills of individuals from birth through to adulthood. The aim of the current study was to investigate the factor structure of the WAIS-IV with an intellectual disabilities population in the UK, in order to assess whether the factor structure is the same as that of the standardisation sample. In addition the relationship between IQ and adaptive functioning was examined.

Design. The study was cross-sectional in design.

Method. The study involved a factor analysis of the WAIS-IV with 170 individuals with an intellectual disability. Pearson's correlation analysis was used to analyse the relationship between subdomains and composite scores from the adaptive functioning measure and the full-scale IQ scores from the WAIS-IV.

Results. Analysis revealed that the only reliable factor solution was that of the traditional two-factor, verbal-performance dichotomy. It was also found that there is a positive albeit small correlation between full scale IQ and adaptive behaviour composite and a larger correlation between full scale IQ and daily living skills. There were no correlations between full scale IQ and communication or full scale IQ and socialisation.

Conclusions. The findings do not support the use of the four-factor solution in intellectually disabled populations currently until further research has been carried out. These results provide also some evidence that the WAIS-IV and the Vineland-II assess different sets of skills and behaviours and that the importance of assessing adaptive behaviour to help understand the concept of intellectual disability can be seen as crucial.

Practitioner points

- The two-factor verbal-performance dichotomy (which has been removed by the developers) appears to have the most robust structure for the population currently under investigation.

- The conclusion that the four-factor solution may not applicable in intellectual disabilities populations leads to the question of its use in clinical practice.

- More research is required about its utility in intellectual disabilities populations and the suitability of using short-forms of the scale.

- In terms of assessing adaptive behaviour there also appears to be a real need. This assessment gathers information, which is not available from just measuring an individual's intelligence and can provide the diagnosing clinician with a more informed view.

A diagnosis of intellectual disability can have a life changing effect on those receiving it (Whitaker, 2010). On the positive side it can assist in the access to appropriate services and support networks, however, on the negative side can have associated stigma attached to it (Dagnan & Waring, 2004). It is therefore of imperative importance that the diagnosis of an intellectual disability is as precise as possible. Before the development of IQ tests, intellectual disability was described in terms of what is now called adaptive behaviour or social functioning, but it was not until 1959 that it became incorporated into diagnostic criteria. This was more formally embedded in the 2002 Diagnostic criteria of the American Association of Mental Retardation (AAMR) and subsequent definitions of the American Psychological Association (APA) and British Psychological Society (BPS). So, diagnosis is now established using three different criteria; a significant impairment in intellectual functioning (scoring less than 70 on a standardised IQ test), a significant impairment of adaptive functioning (social functioning) with both impairments arising before the age of 18 (American Association on Intellectual & Developmental Disabilities, 2010; American Psychiatric Association, 2013; British Psychological Society, 2015). With the assessment of intellectual ability and adaptive functioning playing such a major role in the diagnosis of an intellectual disability it is only proper that the instruments used to assess them are reliable, valid and suitable for the group that is being tested (Kline, 2000).

Amongst the wide array of available intelligence tests, the family of the Wechsler scales is considered among the most eminent (Hill, Reddon & Jackson, 1985). The Wechsler Adult Intelligence Scales (WAIS) have an extensive research base (Groth-Marnat, Gallagher, Hale & Kaplan, 2000) and have a long history of being the most commonly used instruments to assess adult intelligence both in clinical and non-clinical populations (Jones, van Schaik & Witts, 2006). In particular these tools are

widely used by clinicians in contributing to accurately diagnosing intellectual disabilities in adults (British Psychological Society, 2015).

The most up to date version of the WAIS is the WAIS-IV (Wechsler, 2008). Like its predecessor the WAIS-III (Wechsler, 1997) it boasts strong levels of psychometric properties (Kaufman, Lichtenberger & McLean, 2001; Sattler & Ryan, 1999; Benson, Hulac & Kranzler, 2010; Weiss, Keith, Zhu & Chen, 2013; Wechsler, 2008). In the UK the WAIS-IV is recommended for the assessment of IQ in adults who have intellectual disabilities. It is designed for individual administration, constructed on the basis of the normal distribution and standardised on a representative US sample and UK sub sample. It also has psychometric properties that lie within the range of scientific acceptability, and is based on a multidimensional, hierarchical model of intelligence, producing "not just an overall score but also related index and composite scores" (British Psychological Society, 2015). Factor analytic studies have been conducted, which have produced a four-factor structure in the WAIS-IV (Wechsler, 2008). Using this four factor structure four indexes named, Working Memory, Perceptual Reasoning, Processing Speed and Verbal Comprehension have been constructed (Wechsler, 2008) and are used in the interpretation of the test. These indexes have previously been shown to have a useful clinical application and have been shown to correspond to neuropsychological constructs (Kaufman, 2000) and this has been extended to applications of the WAIS- IV with adults with different levels of intellectual disabilities (Lichtenberger & Kaufman, 2013). However, Loring and Bauer (2010) are of the view that the four-factor structure is less useful in neuropsychology than the previous two-factor structure; but this is no longer calculable and has no associated norms.

Similar to the WAIS-III, the four-factor structure of the WAIS-IV has been generally supported in the general population in the USA and in one study of adults who have intellectual disabilities in the USA (Reynolds, Ingram, Seeley & Newby,

2013). The authors of this study investigated the structure of the WAIS-IV using confirmatory factor analysis and tested for measurement invariance across a group with intellectual disabilities (n=104) and a control group (n=104) matched on age, ethnicity, education and region of the USA. They found that the first-order four-factor structure demonstrated strong factorial invariance, when the standard battery was used and partial strong factorial invariance was supported when all of the subtests were used. These findings supported the use of the WAIS-IV in the assessment of adult intellectual disability. However, carrying out a confirmatory factor analysis on the basis of what had been found in the standardisation sample may have been premature as the same structure was not replicated in adults with intellectual disabilities in two studies carried out in the UK on the WAIS-III (Jones, van Schaik & Witts, 2006; MacLean, McKenzie, Kidd, Murray & Schwannauer, 2011) questioning the validity and utility of the four factor model of the test with this population. Jones, van Schaik and Witts (2006), in their research using exploratory factor analysis, found that only two factors (verbal and performance) generalised to populations of adults with intellectual disabilities leading them to the conclusion that the WAIS-III factor structure in intellectual disabilities populations differed considerably from the factor structure found in the US standardisation sample. MacLean et al (2011), in their study of mild and moderate intellectually impaired samples, used a confirmatory factor analysis technique to assess the goodness of fit of the proposed four-factor model using 13 and 11 subtests. They found that none of the indices used suggested a good fit for the model, indicating a lack of factorial validity and suggesting a lack of measurement invariance of the assessment with people with an intellectual disability. In view of this evidence from the WAIS-III studies in the UK with people with intellectual disabilities, there is need for caution when interpreting the indexes of the WAIS-IV in routine clinical practice.

A possible explanation for these findings is that there is limited representation of people with intellectual disabilities in both the standardisation and sampling procedures of the WAIS-IV. Considering that this group of people are the most likely to undergo intellectual assessment for the diagnosis of an intellectual disability, surprisingly little attention has been paid to the validity of such scales in people in the lower ability range. Further, Lichtenberger and Kaufman (2013) state that their biggest disappointment is the lack of available studies on the WAIS-IV with brand new clinical samples. Similarly, Loring and Bauer (2010) argue that there are insufficient clinical data in neurological populations to ensure the appropriate use of the WAIS-IV structure in neuropsychological evaluations. The same argument applies to people who have intellectual disabilities. Further, the WAIS-III retained the traditional verbal and performance score, but this was removed from the WAIS-IV, in Loring and Bauer's (2010) view, prematurely. They also argue that the removal from the WAIS-IV of items that detect psychomotor slowing would lead to fewer full scale IQ's of 70 or below in neurological populations, as this condition is a core feature of many forms of brain injury. Loring and Bauer (2010) are concerned as to the implications for this in relation to access to disability benefits in the USA. Similarly poor psychomotor skills are a feature of intellectual disability and so this could impact on the sensitivity of the scale in the diagnosis of intellectual disability with similar consequences. Whitaker (2010) has suggested that the psychometric properties of the WAIS-IV are population specific and that low IQ samples are not validly tested using the four-factor structure. However, according to the standardisation fallacy (Jensen, 1987), not having a representation of a subpopulation in a standardisation sample is not indicative that a measure is biased or unfair when used in that subpopulation. Therefore, further studies are needed using the WAIS-IV in intellectual disabilities samples to test the validity of the constructs when applied to these populations.

For a diagnosis of intellectual disability the person also has to have an assessment of social functioning or adaptive behaviour. Difficulties in making reliable and valid assessments of adaptive functioning have led to their being a greater emphasis being placed on intellectual functioning when assessing whether someone has an intellectual disability (British Psychological Society, 2015). Indeed, recent research with children has found that the assessment of adaptive behaviour does not contribute to the diagnosis (Obi, Braun, Baio, Drews-Botsch, Devine & Yeargin-Allsopp, 2011). The authors of this study suggest that IQ assessment alone appears to be more appropriate, but their research was concerned with prevalence and epidemiology rather than individual diagnosis. However, the measurement of adaptive behaviour is important as it contributes to the framework for person-referenced education and habilitation goals. (Tasse et al., 2012). In their review Tasse et al (2012) found four measures of adaptive behaviour that had the appropriate psychometric properties for use in the diagnosis of intellectual disability. Of these the most widely used are the Vineland Adaptive Behaviour Scales II (Sparrow, Balla & Cicchetti, 1984). Their research with 83 typically developing adults aged 17-68 using the Vineland Adaptive Behaviour Scales II found near zero correlations between IQ on the WAIS-III and Vineland adaptive behaviour sub- and total scores, suggesting their independence as constructs (Sparrow, Balla & Cicchetti, 1984). However, the nature of this relationship has not been tested in people with intellectual disabilities despite the now equal weight given to IQ and adaptive behaviour in the diagnosis of intellectual disabilities (American Association on Intellectual & Developmental Disabilities, 2010; American Psychiatric Association, 2013; British Psychological Society, 2015). Thus, if this assessment is not making a contribution to diagnosis, then the burden should be removed from the assessment process for intellectual disability.

Clinical Implications

The WAIS-IV and the Vineland-II are commonly used to determine whether an individual has significant deficits in intellectual and social functioning. In addition to age they provide the other two accepted criteria for the diagnosis of having an intellectual disability. Results from intellectual and social functioning assessments can be used in the context of mental health legislation regarding mental impairment, legal decision making and planning interventions (Murray, McKenzie & Lindsay, 2003). Therefore, it is imperative to ensure that the psychometric properties of the scale are the same for a population with low IQ and intellectual disabilities as for the population upon which it is normed. Similarly it is important to test that the additional required assessment of adaptive functioning is contributing to the diagnosis.

Aim

The aim of the current study was to investigate some psychometric properties of the WAIS-IV and its relationship to the Vineland-II when used to diagnose intellectual disability in the UK. The analysis focused on the internal consistency and factor structure of the WAIS-IV with people who have intellectual disabilities in the UK, in order to assess whether these are similar to those of the US standardisation sample. This was achieved through examining whether a four-factor solution could be extracted using exploratory factor analysis. By examining if two and three factor solutions could be extracted and by extracting a one-factor solution. In addition the relationship between IQ and adaptive functioning was examined.

Method

Design

The study was cross-sectional in design and used quantitative methodology.

Participants

The WAIS-IV data was collected from 170 individuals who have had at one point in time been diagnosed as having intellectual disability following a diagnostic assessment at a Community Intellectual Disabilities service in a town in the north of England. Of these 147 also had a completed assessment on the Vineland Adaptive Behaviour Scales II. The assessments were undertaken by an assistant, trainee or qualified clinical psychologist working within the service. An experienced Consultant Clinical Psychologist who had been involved in the WAIS-III standardisation study for the UK norms trained all assessors in the administration of the WAIS-IV and Vineland II. Administrators undertook independent administrations when they had demonstrated competence under observation. A qualified psychologist checked all administration by trainees and graduate psychologists.

The assessments were completed as part of routine clinical practice for a variety of reasons, including, informing interventions and support requirements, as part of capacity assessments and to determine eligibility for receiving intellectual disabilities services. Participants were all residents in the local area at the time of the assessments. Data was included for all individuals for whom there was service case files and met all three diagnostic criteria for an intellectual disability (IQ <70, concurrent deficits in adaptive functioning and both having an onset before the age of 18). Participant data was omitted from the final analysis if any of the values from the WAIS-IV subtests were missing. The table below (Table 1) shows all demographic data for the participants.

Table 1

Demographic Data

	WAIS-IV data only	WAIS-IV and Vineland II data
Age		
Mean (SD)	27.67 (11.91)	27.01 (11.91)
IQ		
Mean (SD)	50.01 (7.16)	58.96 (7.35)
Gender		
Male, n, (%)	96 (56.47)	87 (59.20)
Female, n, (%)	74 (43.52)	60 (40.80)
Ethnicity		
Caucasian, n, (%)	170 (100)	147 (100)

Exclusion criteria

Data was excluded if the participant had an acquired brain injury.

Sample size

There are a number of opinions regarding what constitutes an adequate number of participants for an effective factor analysis. These include:

1) The number of participants ≥ 5 x number of variables (Lewis, 1995)
2) N – n – 1 ≥ 50 (where N = number of participants and n = number of variables (Lawley & Maxwell, 1971)
3) A ratio of at least 2:1 participants to variables (Kline, 1994)

The sample size of the current study (n=170) therefore satisfies all three of these criteria.

Instruments

Wechsler Adult Intelligence Scale – Fourth Edition (WAIS-IV).

The WAIS-IV is an individually administered intelligence test for individuals aged between the ages of 16 to 90 years. In the clinical service from which the data was

collected all subtests were carried out in the same order each time following the standardised set of instructions in the WAIS-IV manual. These instructions included information about how to introduce and conduct each subtest. There were also instructions in relation to the timing of subtests. Due to the detailed nature of the administration instructions, inter-rater reliability was very high.

The WAIS-IV includes 15 subtests (10 core subtests and 5 supplemental subtests) with only the 10 core subtests contributing to the full scale IQ. In the service from which the data were collected only the core subtests were administered. Four index scales (Verbal Comprehension, Perceptual Reasoning, Working Memory and Processing Speed) are derived from the 10 core subtests. The specific subtests contributing to each index are shown in Table 2. The scaled scores across the ten core sub-tests were combined to derive the full scale IQ.

Table 2

Subtests of the WAIS-IV

Subtest	Proposed abilities measured
Verbal Comprehension Index	
Vocabulary (core)	Word knowledge
Similarities (core)	Verbal concept formation and abstract thinking
Information (core)	General knowledge acquired from culture
Comprehension (supplementary)	Knowledge of abstract social conventions
Perceptual Reasoning Index	
Block design (core)	Visuospatial problem solving, nonverbal concept formation and construct ability
Matrix reasoning (core)	Nonverbal abstract problem solving, serial reasoning and spatial reasoning
Visual puzzles (core)	Spatial reasoning
Picture completion (supplementary)	Ability to differentiate visual details
Figure weights (supplementary)	Quantitative and analogical reasoning
Working Memory Index	
Digit span (core)	Attention, concentration and mental control
Arithmetic (core)	Concentration whilst conducting mental mathematical problems
Letter number sequencing (supplementary)	Attention, concentration and mental control
Processing Speed Index	
Symbol search (core)	Perceptual discrimination, speed, accuracy and sustained attention
Coding (core)	Visual scanning, visual tracking, handwriting speed and paired associate learning
Cancellation (supplementary)	Visual perceptual speed

Reliability of the WAIS-IV.

The reliability estimates for internal consistency reported in the WAIS-IV manual used Cronbach's Alpha; these are generally >0.80, with several being >0.90 for subtests. The Cronbach's Alphas for the Full Scale IQ, Verbal Comprehension index, Perceptual Reasoning index and Working Memory index are reported to be >0.90 across the age ranges. The internal reliability for the Processing Speed index based on

test-retest data for 298 people at four age levels (16-29, 30-54, 55-69 and 70-90) ranges from 0.87 to 0.92 (Wechsler, 2008). The WAIS-IV technical manual also reports reliability estimates based on the internal consistency among items for subtests in the intellectual disabilities subgroups (mild severity n=73 and moderate severity n=31) with these estimates being very consistent with the overall standardisation sample, except for the one lower estimate (0.67) reported for the Information subtest for the sample of individuals with a moderate intellectual disability.

Validity of WAIS-IV.

Three different concepts of validity were used to measure the validity of the WAIS-IV. These included content validity, construct validity and criterion-related validity.

The developers of the WAIS-IV (Wechsler, 2008) propose that the content validity of the measure is good and that its subtests, IQ and index scores measure neuropsychological concepts that have been shown to be valid. In order to assess the content validity of the measure systematic reviews of the literature along with consultation with those practitioners who use the measure were undertaken. This investigation allowed the Psychological Corporation to assess whether domains of behaviour are satisfactorily measured (Wechsler, 2008).

The construct validity of the WAIS-IV was measured using an inter-correlation analysis (Campbell & Fiske, 1959). This analysis highlighted that all subtests measure a general intelligence factor (e.g. *g)* and that all of the items on the measure had some degree of association. It was also found that subtests contributing to a specific index had higher correlations with each other than with subtests comprising other scales and that some subtests were more related to *g* than other subtests (Wechsler, 2008).

Finally, criterion-related validity was measured by comparing the WAIS-IV with a number of other measures of intelligence (WAIS-III, WISC-IV, Standard

Progressive Matrices and Stanford-Binet Intelligence Scale – fourth edition) both in terms of side-by-side administration and comparison. Correlation coefficients ranging from 0.68 to 0.93 were found between the WAIS-IV and other relevant measures, therefore suggesting that the criterion-related validity is acceptable.

The Vineland Adaptive Behaviour Scales – Second Edition (Vineland-II)

The Vineland-II claims to measure the personal and social skills of individuals from birth through to adulthood. As adaptive functioning refers to an individual's typical performance of the day-to-day activities required for personal and social sufficiency, these scales assess what a person actually does, rather than what he or she is able to do. Thus they assess different skills and abilities to tests of intellectual functioning and therefore their content is very different. In order to determine the level of an individual's adaptive functioning, someone who is familiar with that individual, such as a parent or caregiver, is asked to describe his or her activities. Those activities are then compared to those of other people the same age to determine which areas are average, above average or in need of social help. Learning about an individual's adaptive functioning helps us to gain a total picture of them. When adaptive functioning information is combined with information about an individual's intelligence, school achievement and physical health, plans can be made to address any special needs that person may have at home, school, or work. For adults the Vineland-II assesses adaptive functioning in three domains: Communication, Daily Living Skills, and Socialisation. It also provides a composite score that summarises the individual's performance across all three domains.

Reliability of the Vineland-II.

Internal reliability tests of both the survey and expanded forms show that, the survey form Pearson coefficients for three age groups ranged from 0.82 to 0.95 for the domains and 0.96 to 0.98 for the Adaptive Behavior Composite. The expanded form

Pearson coefficients range from 0.90 to 0.97 for the domains and 0.98 to 0.99 for the Adaptive Behaviour Composite.

The test-retest reliability for the survey form intra-class correlation coefficients for caregivers of children between the ages of six months and two years ranges from 0.78 to 0.92 for the domains and 0.90 for the Adaptive Behaviour Composite. There is no test-retest reliability for the expanded form.

The inter-rater reliability for the survey form intra-class correlations coefficients, with a mean of eight days between the interviews of caregivers of children ages six months to 18 years, ranges from 0.62 to 0.78 for the domains and is 0.74 for the Adaptive Behaviour Composite. There are no inter-rater reliability tests for the expanded form.

Validity of the Vineland-II.

Analysis of content validity included a literature review and field tests with caregivers. Criterion-related validity was assessed through Pearson correlations between the adaptive behavior composite and the original Vineland unadjusted social quotient, and Silverstein's deviation social quotients among caregivers of children between ages six months and 18 years. Both scores were 0.55. Comparisons between the total of the raw scores for the four domains of the revised Vineland and the original Vineland yielded a correlation of 0.97 in a sample of intellectually disabled clients, and an age-adjusted partial correlation of 0.88 in a sample of hearing-impaired children.

Data collection

Historical and current files stored within the Community Intellectual Disabilities service in a town in the north of England were reviewed by the lead researcher and information was extracted from those files where a WAIS-IV and Vineland-II had been administered and scored. The WAIS-IV was published in 2008; therefore data obtained were from the period 2008-2015. The following data were collected; gender, age,

ethnicity; WAIS-IV subtest scores and full IQ scores; adaptive behaviour scores (Vineland-II) and an identifying code. The identifying code was recorded as the clinician's initials that completed the assessment along with a data set number. The latter can be matched with the clients name and were held separately by the research team until the data were entered into the database and then destroyed. This was to ensure that data could be extracted again if lost during the inputting procedure and also to ensure that entries were not duplicated.

Data analysis

Data were analysed using SPSS. Descriptive data were presented. Preliminary analysis was carried out to make sure that the sample was appropriate for factor analysis (Tabachnick & Fidell, 2001). This analysis included normality and linearity of variables (skewness and kurtosis and the Kolmogorov-Smirnov test). The absences of outliers were examined using standardised scores (univariate outliers) and Mahalanobis distance (multivariate outliers) and multicollinearity and singularity were assessed by the determinant of the correlation matrix. The correlation matrix was examined to evaluate the 'factorability' of the data (Tabachnick & Fidell, 2001). Principal axis factoring (PAF) and principal components analysis (PCA) with the application of oblique rotations were employed as these techniques were used by Jones, van Schaik and Witts, (2006) in their previous study of the WAIS-III.

Correlation analyses (Pearson's) were used to analyse the relationship between subdomains and composite scores from the adaptive functioning measure and the full-scale IQ scores from the WAIS-IV.

Factor analytic techniques used

Principal axis factor (PAF) analysis was conducted, with an application of oblique rotations. The rationale for using this method was because the current study is a replication of a previous study conducted by Jones, van Schaik and Witts (2006) on the

WAIS-III. This technique of factor analysis reflects an adherence to the notion that underlying processes are responsible for the performance of individuals on the subtests, therefore suggesting that the WAIS-IV is a measure of neuropsychological capability.

Principal components analysis (PCA) was also used since there has been limited previous research that has examined the population in the current study. Tabachnick and Fidell (2001) propose PCA as a technique to empirically summarise data, however this technique does not necessarily give a reflection of underlying processes. This technique of analysis has been described as the process of decomposing the original data into a set of variables (Field, 2000). Furthermore, the confidence in a factor solution is increased if the same factor solution is obtained using more than one factor extraction technique (Tabachnick & Fidell, 2001).

For the current research, initial analyses common to all techniques will be described, then PAF and rotations applied will be reported and finally PCA and applied rotations. These results will be appraised to summarise conclusions that may be drawn from the analysis of the data set. Factor loading criteria suggested by Comrey and Lee (1992) will be used to accept or reject factor loadings. They recommend that a variable can be considered as having a 'poor' loading onto a factor if the correlation is at least 0.32, 'fair' if the correlation is at least 0.45, 'good' if the correlation is at least 0.55, 'very good' if the correlation is at least 0.63 and 'excellent' if the correlation is at least 0.71.

Differences between PCA and PAF

A brief outline of the differences between PCA and PAF are that PCA is designed to understand a data set more generally while PAF is designed to understand, support and produce theory. The aim of PCA is to extract maximum variance from a data set with a few orthogonal components and the aim of PAF is to reproduce the correlation matrix with a few orthogonal factors (Tabachnick & Fidell, 2001).

Analysing different aspects of the correlation matrix achieves these aims. In PCA, the positive diagonal in the correlation matrix contains ones and for PAF it contains an estimate of communality. In PCA, all of the variance, including error and unique variance is included in the analysis, while in PAF only common variance is included in the analysis (Tabachnick & Fidell, 2001).

Ethical implications

The main ethical consideration for the current study was one of confidentiality. The issue of confidentiality was addressed by coding all data so that it was not recognisable to anyone not involved in the study. Furthermore, no personal details were revealed at any stage of the study or write-up and full anonymity was maintained. Also by using a retrospective approach consent was not gained from individuals to extract their data for the purposes of research. However, both local National Health Service Research and Ethic Committee (NHS REC) (Appendix A) and Research and Development (Appendix B) department from the host organisation were satisfied that the use of the data for the research was appropriate. NHS REC approval was sought on a proportionate basis and approved.

Results

A data file was constructed and all analyses carried out using SPSS version 21.0.

Descriptive data

The table below (Table 3) shows the raw data for the individual subtests on both the WAIS-IV and Vineland-II across all participants.

Table 3

Raw Data for WAIS-IV and Vineland

Subtest	Mean	SD±	Cronbach's Alpha (α)
WAIS-IV			
Full scale IQ	50.01	7.16	0.77
Block design	4.54	2.17	0.78
Similarities	3.56	1.58	0.77
Digit span	3.71	2.19	0.77
Matrix reasoning	3.61	1.77	0.78
Vocabulary	4.16	1.08	0.77
Arithmetic	3.85	1.67	0.77
Symbol search	3.20	1.91	0.78
Visual puzzles	4.99	1.78	0.78
Information	4.39	1.52	0.77
Coding	2.50	1.62	0.77
Vineland-II			
Adaptive behaviour	46.42	15.19	0.69
Communication	40.87	17.37	0.73
Daily living skills	52.78	15.23	0.72
Socialisation	50.63	16.97	0.71

Note. $\alpha \geq 0.9$ (excellent), $0.9 > \alpha \geq 0.8$ (good), $0.8 > \alpha \geq 0.7$ (acceptable), $0.7 > \alpha \geq 0.6$ (questionable), $0.6 > \alpha \geq 0.5$ (poor), $0.5 > \alpha$ (unacceptable)

From the analysis we can see that there is an acceptable level of internal consistency for both of the instruments under investigation with this specific sample.

Preliminary analysis

Before the factor analysis was carried out on the data, various screening procedures were conducted to prepare for major analysis (Tabachnick & Fidell, 2001).

Normality of variables

According to Tabachnick and Fidell (2001), standardised scores should be calculated (z scores) for skewness and kurtosis to assess normality, as well as the distribution of the data using histograms. The standardised skewness and kurtosis values for each subtest of the WAIS-IV were calculated and shown in Table 4. Field (2000) suggests that values above the criterion of 1.96 signify a distribution that departs from a normal distribution.

Table 4

Skewness and Kurtosis z Scores for Each Subtest

Subtest	Skewness z score	Kurtosis z score
Block design	0.42	0.33
Similarities	0.48	0.45
Digit span	0.29	-0.89
Matrix reasoning	0.88	**2.43**
Vocabulary	0.90	1.79
Arithmetic	0.11	0.14
Symbol search	1.08	**2.09**
Visual puzzles	0.43	1.12
Information	1.60	**5.41**
Coding	0.89	-0.15

Note. Bold type denotes a value > 1.96

Visual inspection of the histograms in combination with the information in the table above yields the conclusion that not all of the distribution of the data is normal. However, given that this is a clinical population, a normal distribution may not be

expected and this does not necessarily degrade the factor analysis (Tabachnick & Fidell, 2001).

A further analysis to assess the normality of the dataset was also conducted. The Kolmogorov-Smirnov test was carried out to assess the significance of the deviations from normality as described above. This analysis indicated that the distribution of all of the WAIS-IV subtests differed significantly from a normal one. This deviation might be expected because of the limited range of scores anticipated from a population with an intellectual disability. In some cases, transformation of the data might be considered, so a log transformation was performed on the original data and subsequently reanalysed using PAF and PCA.

Linearity of variables

Inspecting bivariate scatterplots, which were produced for all variables, assessed the linearity of variables. These bivariate scatterplots suggested that some of the subtests are correlated in a linear way while some are not. Therefore, the linearity of all sets of subtests could not be confirmed. While the assumption of linearity is usually a pre-requisite for factor analysis, transformation of the data was not considered for the same reason as stated above.

Absence of univariate outliers

Standardised scores for all cases were calculated. The number of cases is obtained by multiplying the number of variables with itself (which in this case is 10 x 10). Therefore, 100 cases were inspected and within these three were found to have standardised scores in excess of 3.29 that may identify them as potential outliers. However, a few scores of this magnitude may be expected given the size of the data set (Tabachnick & Fidell, 2001). This indicates that it is appropriate to carry out a factor analysis with the current data set.

Absence of multivariate outliers

Mahalanobis distance was calculated for all cases. None of the cases exceeded the critical value of 34.52, therefore showing that no multivariate outliers were present in the current dataset. This again confirmed that it was appropriate to proceed with the factor analysis.

Multicollinearity and singularity

The determinant of the correlation matrix 0.01336, which indicates the multicollinearity (i.e. variables that are highly correlated) and singularity (i.e. variables which were perfectly correlated) were not present in the current data set.

Factorability

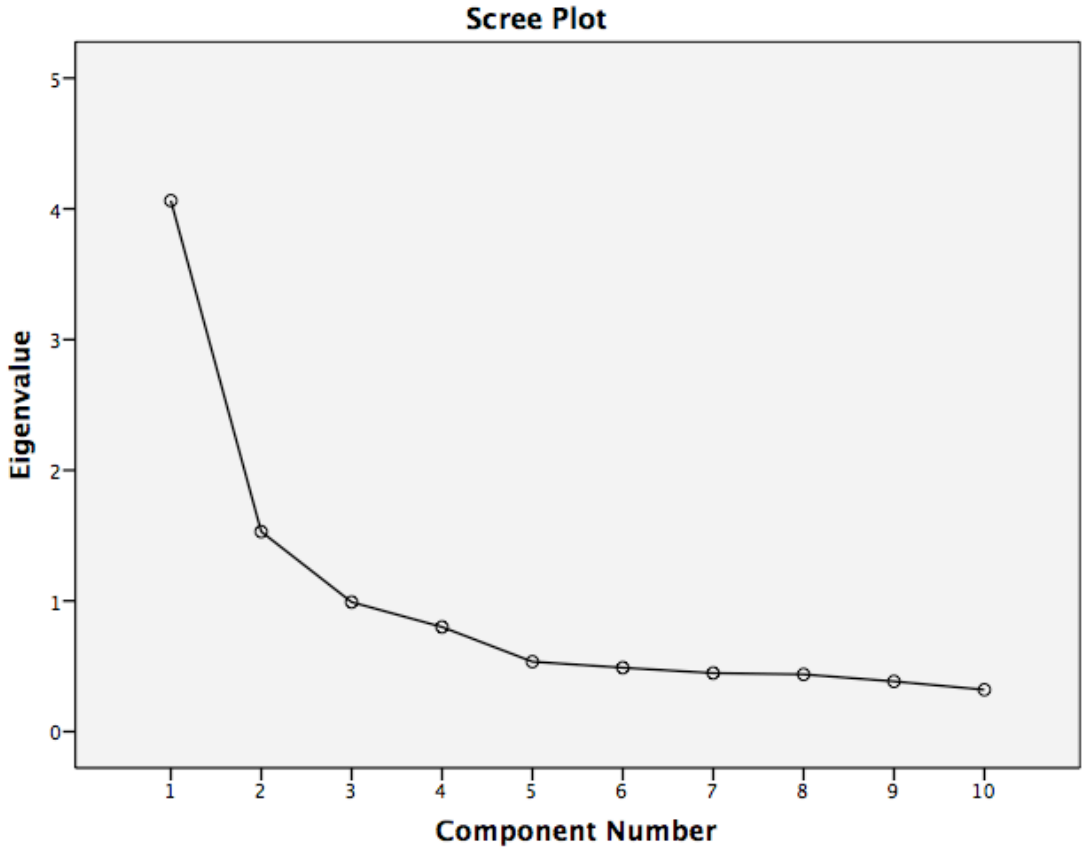
Tabachnick and Fidell (2001) advise evaluating the 'factorability' of the data. This is done through examining the correlation matrix, which should show several correlations at least over 0.3. Examination of the correlation matrix generated by the current data set showed that at least 70% of the correlations in the matrix were at least 0.3. Furthermore, Kaiser and Rice's (1974) Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was 0.86 indicating this value as being 'great' in terms of acceptance (Field, 2000). Again this can be considered as high sampling adequacy and good preconditions for factor analysis.

Factor analysis

Number of factors.

The scree plots produced for both PAF and PCA were identical as were the eigenvalues (see Figure 1). A visual examination of this scree plot suggested that the two factors should be extracted. Using the eigenvalues ≥ 1 criterion, two factors were also indicated. Analyses of the log transformed data found no differences in the factor structures with virtually identical scree plots and a two-factor solution being extracted.

Figure 1 Scree Plot



Previous research with the WAIS-III, WAIS-R and WAIS has found two, three and four factor solutions being extracted. Therefore, in the current study further analyses were carried out with three and four factor solutions specified a priori. The eigenvalues and percentage of explained variance for the first four factors are shown in Table 5.

Table 5

Eigenvalues and Explained Variance Factors 1 – 4

Factor	Eigenvalue	% variance explained
I	4.06	40.62
II	1.53	15.30
III	0.99	9.91
IV	0.80	8.00

Principal axis factoring (PAF).

A one-factor solution was extracted; the factor matrix is shown in Table 6.

Table 6

Factor Matrix for One Factor Solution (PAF)

Variable (sub-test)	Factor I
Vocabulary	**0.55**
Coding	**0.59**
Similarities	**0.57**
Block design	**0.49**
Arithmetic	**0.71**
Matrix reasoning	**0.58**
Digit span	**0.60**
Information	**0.52**
Visual puzzles	**0.60**
Symbol search	**0.57**

Note. Bold type denotes a value > 0.45

All of the loadings in the factor matrix exceeded the criterion suggested by Comrey and Lee (1992) of 0.45 for what they consider a 'fair' loading. As there was only one factor there was no need to apply a rotation to the solution.

The next step was to extract pattern loadings for two, three and four factor solutions using PAF with an oblique rotation (direct Oblimin, delta = 0). This type of rotation was chosen because it was used by the Psychological Corporation in the development of the WAIS-IV. Oblique rotation does not require resultant factors to be independent and reflects possible correlations between factors.

The pattern loadings for two, three and four factor solutions are shown in tables 7, 8 and 10 respectively.

Table 7

Factor Pattern Loadings for Two Factor Solution (PAF) with Oblique Rotation

Variable (sub-test)	Two factor solution	
	I	II
Vocabulary	0.01	**-0.69**
Coding	**0.54**	-0.11
Similarities	0.00	**-0.73**
Block design	**0.61**	0.05
Arithmetic	*0.39*	**-0.42**
Matrix reasoning	**0.65**	-0.00
Digit span	*0.32*	**-0.47**
Information	-0.09	**-0.73**
Visual puzzles	**0.74**	0.05
Symbol search	**0.65**	0.00

Note. Bold type denotes loadings that exceed 0.45; italic type indicates loadings that exceed 0.32

The correlation between factor I and II was -0.47, indicating that shared variance between them is equal to 47%.

Table 8

Factor Pattern Loadings for Three Factor Solution (PAF) with Oblique Rotation

Variable (sub-test)	Three factor solution		
	I	II	III
Vocabulary	-0.01	**0.70**	0.01
Coding	0.26	0.03	*-0.39*
Similarities	-0.08	**0.76**	0.06
Block design	-0.10	-0.00	**0.83**
Arithmetic	**0.45**	*0.41*	0.10
Matrix reasoning	0.07	0.06	**0.64**
Digit span	0.23	*0.38*	0.12
Information	-0.00	**0.77**	-0.09
Visual puzzles	*0.32*	0.00	**0.46**
Symbol search	0.29	-0.02	*0.38*

Note. Bold type denotes loadings that exceed 0.45; italic type indicates loadings that exceed 0.32

Correlations between the three factors are shown in table 9 below. An examination of this table indicates that the shared variance between them is equal to or less than 65%.

Table 9

Factor Correlations Matrix

	I	II	III
I	1.00	0.41	0.49
II	0.41	1.00	0.36
III	0.49	0.36	1.00

Table 10

Factor Pattern Loadings for Four Factor Solution (PAF) with Oblique Rotation

Variable (sub-test)	Four factor solution			
	I	II	III	IV
Vocabulary	**0.81**	-0.01	0.02	0.17
Coding	0.07	**-0.88**	-0.10	-0.11
Similarities	**-0.63**	-0.25	0.07	-0.14
Block design	-0.09	0.01	0.22	*-0.33*
Arithmetic	0.21	*-0.36*	0.07	-0.29
Matrix reasoning	0.09	**-0.47**	0.23	-0.02
Digit span	*0.33*	-0.01	**0.47**	-0.22
Information	*-0.32*	**-0.60**	-0.07	*-0.32*
Visual puzzles	0.31	0.00	*0.35*	-0.06
Symbol search	**0.63**	-0.02	0.16	0.03

Note. Bold type denotes loadings that exceed 0.45; italic type indicates loadings that exceed 0.32

Correlations between the four factors are shown in table 11 below. An examination of this table indicates that the shared variance between them is equal to or less than 73%.

Table 11

Factor Correlations Matrix

	I	II	III	IV
I	1.00	-0.32	0.47	-0.40
II	-0.32	1.00	-0.28	0.49
III	0.47	-0.28	1.00	-0.34
IV	-0.40	0.49	-0.34	1.00

The correlations between all factors for two, three and four factor solutions after applying oblique rotations were examined. For all factor solutions, some correlations

between factors were detected above Comrey and Lee's (1992) factor loading criterion of 0.32, which suggests that at least 10% of the variance was shared between these factors. In instances such as this Tabachnick and Fidell (2001) suggest that there is no need to carry out an orthogonal rotation, since there is a lack of evidence to suggest that resultant factors will be independent.

Principal Components Analysis (PCA).

A one factor solution was extracted; the factor matrix is shown in table 12.

Table 12

Factor Matrix for One Factor Solution (PCA)

Variable (sub-test)	Factor I
Vocabulary	**0.61**
Coding	**0.64**
Similarities	**0.63**
Block design	**0.55**
Arithmetic	**0.74**
Matrix reasoning	**0.64**
Digit span	**0.65**
Information	**0.58**
Visual puzzles	**0.65**
Symbol search	**0.63**

Note. Bold type denotes a value > 0.45

All of the loadings in the factor matrix exceeded the criteria suggested by Comrey and Lee (1992) of 0.45 for what they consider a 'fair' loading. As there was only one factor there was no need to apply a rotation to the solution.

The next step was to extract pattern loadings for two, three and four factor solutions using PCA with an oblique rotation (direct Oblimin, delta = 0).

The pattern loadings for two, three and four factor solutions are shown in tables 13, 14 and 16 respectively.

Table 13

Factor Pattern Loadings for Two Factor Solution (PCA) with Oblique Rotation

Variable (sub-test)	Two factor solution	
	I	II
Vocabulary	-0.01	**0.79**
Coding	**0.60**	0.14
Similarities	-0.01	**0.81**
Block design	**0.73**	-0.09
Arithmetic	*0.40*	**0.49**
Matrix reasoning	**0.73**	0.00
Digit span	*0.32*	**0.46**
Information	-0.11	**0.86**
Visual puzzles	**0.79**	-0.03
Symbol search	**0.73**	-0.00

Note. Bold type denotes loadings that exceed 0.45; italic type indicates loadings that exceed 0.32

The correlation between factor I and II is 0.53 indicating that the shared variance between them is equal to 53%.

Table 14

Factor Pattern Loadings for Three Factor Solution (PCA) with Oblique Rotation

Variable (sub-test)	Three factor solution		
	I	II	III
Vocabulary	**-0.53**	0.09	*0.34*
Coding	0.20	**0.51**	-0.09
Similarities	**-0.48**	0.22	*0.39*
Block design	-0.11	0.00	**0.91**
Arithmetic	**0.48**	0.13	0.05
Matrix reasoning	0.04	0.06	0.29
Digit span	**0.48**	0.31	**0.46**
Information	**-0.51**	0.24	**-0.50**
Visual puzzles	0.28	**0.53**	0.15
Symbol search	0.26	**-0.48**	0.26

Note. Bold type denotes loadings that exceed 0.45; italic type indicates loadings that exceed 0.32

Correlations between the three factors are shown in table 15 below. An examination of this table indicates that the shared variance between them is equal to or less than 66%.

Table 15

Factor Correlations Matrix

	I	II	III
I	1.00	0.33	0.40
II	0.33	1.00	0.27
III	0.40	0.27	1.00

Table 16

Factor Pattern Loadings for Four Factor Solution (PCA) with Oblique Rotation

Variable (sub-test)	Four factor solution			
	I	II	III	IV
Vocabulary	*0.33*	0.21	0.03	0.19
Coding	**0.85**	**0.46**	-0.11	-0.16
Similarities	-0.08	**-0.78**	**0.49**	-0.09
Block design	-0.12	-0.01	0.25	*-0.35*
Arithmetic	**0.45**	-0.13	0.05	-0.17
Matrix reasoning	0.06	-0.08	**0.78**	**-0.46**
Digit span	0.01	-0.00	0.07	**-0.89**
Information	*-0.34*	**-0.73**	-0.10	-0.23
Visual puzzles	0.28	-0.02	*0.33*	-0.07
Symbol search	**0.82**	-0.01	*-0.33*	-0.16

Note. Bold type denotes loadings that exceed 0.45; italic type indicates loadings that exceed 0.32

Correlations between the four factors are shown in table 17 below. An examination of this table indicates that the shared variance between them is equal to or less than 73%.

Table 17

Factor Correlations Matrix

	I	II	III	IV
I	1.00	0.25	0.36	0.31
II	0.25	1.00	0.28	0.39
III	0.36	0.28	1.00	-0.27
IV	-0.31	-0.39	-0.27	1.00

The correlations between all factors for two, three and four factor solutions after applying oblique rotations were examined. For all factor solutions, some correlations

between factors were detected above Comrey and Lee's (1992) factor loading criteria of 0.32, which suggests that at least 10% of the variance was shared between these factors. In instances such as this Tabachnick and Fidell (2001) suggest that there is no need to carry out an orthogonal rotation, since there is a lack of evidence to suggest that resultant factors will be independent.

Current study PAF, PCA and WAIS-IV development four-factor solutions

The four-factor solutions extracted through the PAF and PCA analysis were compared to the four-factor solution extracted from the WAIS-IV development sample, as shown in table 18. This table shows how specific tests vary across populations in their loadings.

Table 18

Factor Pattern Loadings for Four-Factor Solution (PAF, PCA and WAIS-IV Development Sample

Variable (sub-test)	PAF four factor solution				PCA four factor solution				WAIS-IV development four factor solution			
	I	II	III	IV	I	II	III	IV	I	II	III	IV
Vocabulary	0.81								0.46			
Coding		0.88			0.85	0.46						0.55
Similarities	0.63					0.78	0.49		0.58			
Block design										0.62		
Arithmetic					0.45						0.60	
Matrix reasoning		0.47					0.78	0.46		0.68		
Digit span			0.47					0.89			0.65	
Information		0.60				0.73			0.59			
Visual puzzles										0.66		
Symbol search	0.63				0.82							0.63

Table 18 shows major differences between the four-factor solution in the WAIS-IV development data and those found in this study. The four factor solutions in this study are difficult to interpret and lack any theoretical rationale. Unfortunately direct

comparisons of loading were not possible as cross loading are not reported in the WAIS-IV technical manual.

Appraisal of factor analysis results

Two factor solutions.

A two factor solution was indicated by both the scree plot and the criterion of an eigenvalue ≥ 1. Analysis using both PAF and PCA with the application of oblique rotations extracted the same factor structure as seen in table 16. As suggested by Comrey and Lee's (1992) criterion, individual variables were assigned to factors if the loading was ≥ 0.32 as a minimum. However all factor loadings exceeded 0.45.

The factors in this structure were representative of what have historically been referred to as the verbal and performance sections of the WAIS-IV. They have therefore been labelled 'verbal' and 'performance' (see Table 19).

Table 19

Two Factor Solution PAF and PCA, Factor Loading Criterion > 0.45

Factor	
Verbal	Performance
Vocabulary	Coding
Similarities	Block design
Arithmetic	Matrix reasoning
Digit span	Visual puzzles
Information	Symbol search

The production of this factor structure using both PAF and PCA techniques with the application of oblique rotations suggests a robust factor structure. The internal consistency of this factor structure was first assessed by calculating Cronbach's coefficient alpha for the verbal and performance factors identified. According to Nunnally (1978), a minimal level of 0.7 indicates reliability. The alpha coefficient for

the verbal factor was calculated as 0.806 and for the performance factor it was 0.771, which indicates that both of the factors have a high internal consistency.

Three factor solutions.

A three factor solution was also extracted from both the PFA and PCA. However, a different factor configuration was evidenced in each analysis. The factor structures are displayed in tables 20 and 21. Factor structures with applied loading criteria of 0.45 and 0.32 have both been illustrated to highlight the relative instability of a three factor solution.

Table 20

Three Factor Solution (PAF)

Factor		
I	II	III
>0.45		
Arithmetic	Vocabulary	Block design
	Similarities	Matrix reasoning
	Information	Visual puzzles
>0.32		
Arithmetic	Vocabulary	Block design
Visual puzzles	Similarities	Matrix reasoning
	Information	Visual puzzles
	Arithmetic	Coding
	Digit span	Symbol search

Table 21

Three Factor Solution (PCA)

Factor		
I	II	III
>0.45		
Similarities	Coding	Block design
Information	Visual puzzles	Digit span
Arithmetic	Symbol search	Information
Digit span		
Vocabulary		
>0.32		
Similarities	Coding	Block design
Information	Symbol search	Digit span
Arithmetic	Visual puzzles	Information
Digit span		Similarities
Vocabulary		Vocabulary

The three factor solution shows that there are some similarities and differences produced by PAF and PCA. In both types of analysis there is a strong factor, which contains, Similarities, Information, Arithmetic, Digit Span and Vocabulary (factor II in PAF and factor I in PCA). This extracted factor includes subtests, which are traditionally associated with verbal abilities (Similarities, Information and Vocabulary) along with two subtests, which typically assess working memory (Arithmetic and Digit span).

The remaining two factors, which were extracted, were not synonymous in the PAF and PCA techniques used and both seemed to produce a mixture of items that loaded onto the working memory, perceptual organisation and processing speed indexes on the standardised sample. These two factors were not labelled, as they were unable to

be viewed as discreet factors due to the differences in their structure when extracted using different techniques and loading criterion. It was also evident from the loading criterion that when a lower threshold was utilised some of the subtests loaded onto more than one factor. This cross loading indicates that a three factor structure is not robust.

Four factor solutions.

The Psychological Corporation proposed that a four factor solution is the most reflective of underlying cognitive processes as assessed by the WAIS-IV and that it is the most useful in clinical practice. However, the four factor solutions extracted in this study were not easily interpretable using both PAF and PCA (see tables 22 and 23). Similar to the three factor solutions, factor structure as shown using two different loading criterions of 0.45 and 0.32.

Table 22

Four Factor Solution (PAF)

Factor			
I	II	III	IV
>0.45			
Vocabulary	Coding	Digit span	
Similarities	Matrix reasoning		
Symbol search	Information		
>0.32			
Vocabulary	Coding	Digit span	Block design
Similarities	Matrix reasoning	Visual puzzles	Information
Symbol search	Information		
Digit span	Arithmetic		
Information			

Table 23

Four Factor Solution (PCA)

Factor			
I	II	III	IV
>0.45			
Coding	Coding	Similarities	Matrix reasoning
Arithmetic	Similarities	Matrix reasoning	Digit span
Symbol search	Information		
>0.32			
Coding	Coding	Similarities	Matrix reasoning
Arithmetic	Similarities	Matrix reasoning	Digit span
Symbol search	Information	Visual puzzles	Block design
Vocabulary		Symbol search	
Information			

Each of the four factor solutions illustrated above reveal a different factor structure. Therefore, this makes the interpretability of them difficult, as they do not appear to represent stable concepts.

The criterion set out by the scree plot and eigenvalues of ≥1 were not achieved by either the three or four factor solutions. However, both were still set as priori in order to investigate their structure and neither of them provided solutions that were easily interpretable due to the presence of complex variables (Tabachnick & Fidell, 2001). This resulted in it not being possible to obtain a simple structure from three or four factor solutions. These results will be discussed further in the next chapter along with the clinical and theoretical implications.

Relationship between IQ and adaptive behaviour

A Pearson's product-moment correlation coefficient was computed to assess the relationship between the adaptive behaviour composite scores and subdomains (daily living skills, socialisation and communication) from the adaptive functioning measure and the full-scale IQ scores from the WAIS-IV.

There was a positive correlation between the full scale IQ scores and the adaptive behaviour composite scores [r = 0.17, n = 147, p = 0.03]. There was a positive correlation between full-scale IQ scores and the daily living skills scores [r = 0.22, n = 147, p = 0.005]. There was no correlation between full-scale IQ scores and socialisation scores [r = 0.11, n = 147, p = 0.17] and there was no correlation between full-scale IQ scores and communication scores [r = 0.12, n = 147, p = 0.13]. Scatter plots summarising the results are shown below (figures, 2, 3, 4 and 5).

Figure 2 Scatterplot showing relationship between full scale IQ scores and adaptive behaviour composite scores

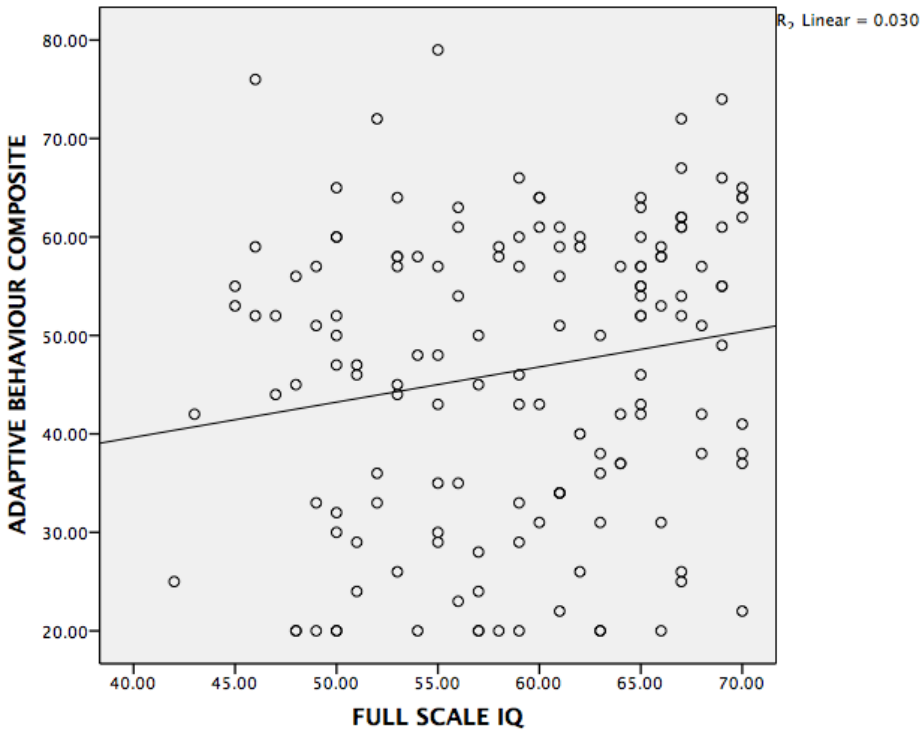


Figure 3 Scatterplot showing relationship between full scale IQ scores and daily living skills scores

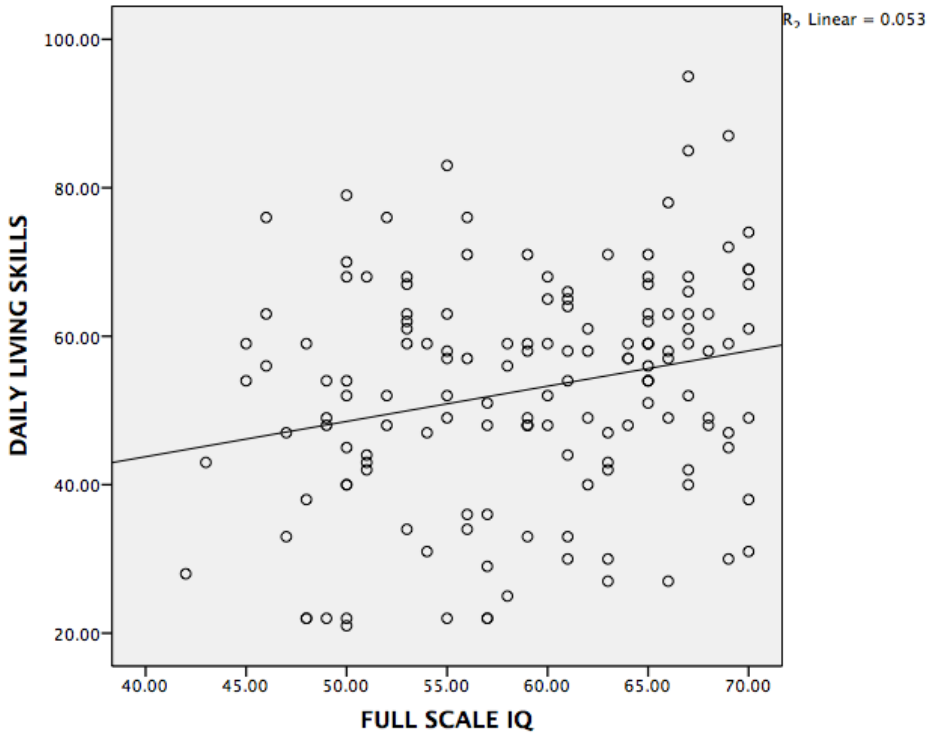


Figure 4 Scatterplot showing relationship between full scale IQ and socialisation scores

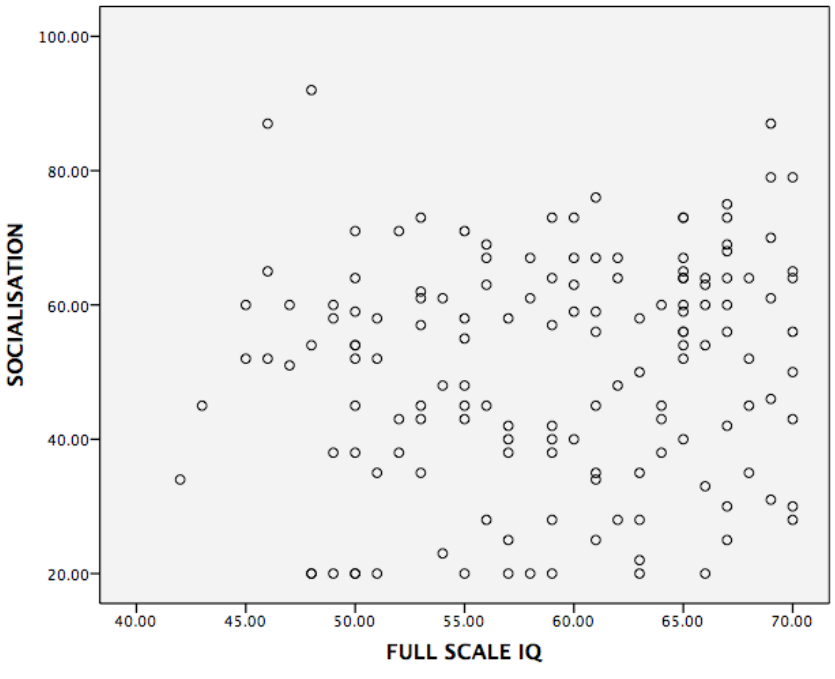
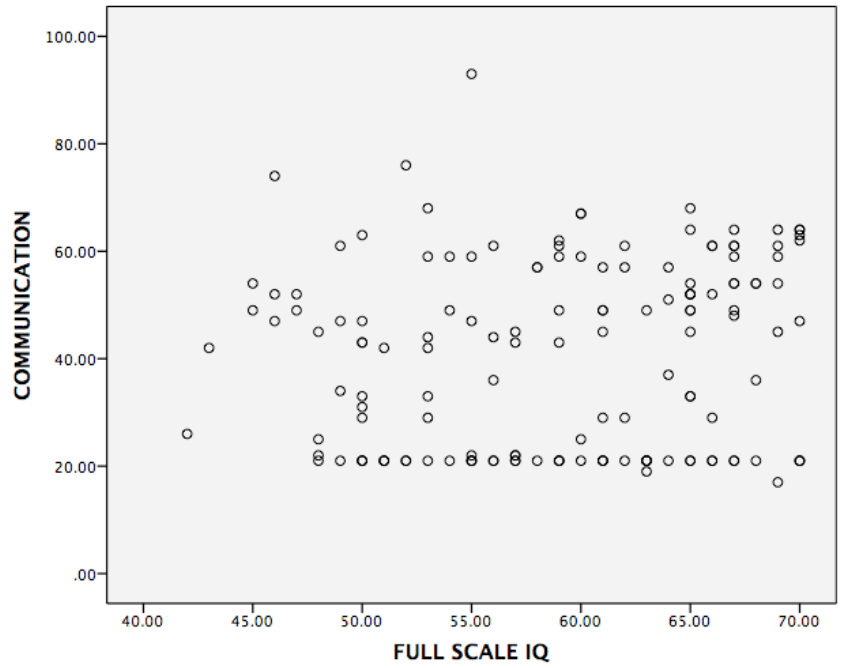


Figure 5 Scatterplot showing relationship between full scale IQ and communication scores



Discussion

This section will discuss the findings in the current study and its implications. The limitations of the research will be considered, along with suggestions for future research and theoretical implications.

Like its predecessors, the WAIS-IV is one of the most frequently used tools in the assessment of intelligence. The WAIS-IV is used for a range of purposes, including the measuring of neuropsychological functioning and impairment to gate-keeping access to services. The earlier versions of the WAIS have all been subject to numerous factor analytic studies in order to help develop the theory about intelligence through describing human intellectual abilities and allowing the formation and confirmation of theoretical hypothesis (Hill, Redden & Jackson, 1985).

In relation to individuals with intellectual disabilities, the WAIS-IV is a well-established assessment tool. It provides service providers with one part of the information needed in the categorisation process along with assessment of adaptive

functioning. It is particularly important in the assessment of individuals with a less severe impairment where classification is not straightforward.

The aim of the current study was to explore the factor structure of the WAIS-IV with a sample of individuals identified as having intellectual disabilities in the UK. This was achieved through investigating the validity of the four-factor model as proposed by the developers of the WAIS-IV (Wechsler, 2008). By looking at two and three factor solutions as these have been extracted in previous studies of the WAIS, WAIS-R and WAIS-III. Finally, by extracting a one-factor solution as previous research has compared this one factor solution with the theoretical concept of *g*. However, caution needs to be taken when examining the concept of *g,* as there is no single definition of this concept within the field of intelligence as a whole. In addition we also examined the relationship between IQ and adaptive functioning.

The current research was carried out on a sample of 170 individuals (96 males and 74 females) identified as having intellectual disabilities. The mean age of the participants was 27.67 (SD 11.91). The mean full scale IQ of the sample was 50.01 (SD 7.16) and the mean adaptive behaviour composite was 46.42 (SD 15.19). All participants were of Caucasian ethnicity, which was representative of the community where the sample was obtained. The high alpha levels confirmed the reliability of the 10 sub-tests with people with intellectual disability. The mean and standard deviations for the subtests and Full Scale IQ for the intellectual disabilities sample in this study are virtually the same as those reported for the mild intellectual disabilities sample in the United States; means for subtest being around 4.0 and standard deviations around 2.0. These are in sharp contrast to the means and standard deviation's for those reported for the typically developing group in the UK. For the subtests means these are reported to be around 10 and the standard deviation around 3.0 as in the US standardisation sample (see Appendix C). Such differences have implications for the calculation of Standard

Errors of Measurement (SEM) on which Critical Values of Difference Scores are calculated for the Full Scale score, Index scores and Sub-test scores. However, SEM scores could not be compared as these are not reported in the WAIS IV manuals (Wechsler, 2008, Wechsler, 2008b)

Factor analysis

The developers of the WAIS-IV have identified a four-factor solution as reflecting theoretical constructs and argue that scores based on this solution lead to a more meaningful way to interpret an individual's performance. These four factors are included as index scores and are purported to reflect verbal comprehension, perceptual reasoning, working memory and processing speed.

Since the publication of the WAIS-IV there have been two exploratory factor analytic studies (which is the technique used in the current study) that have been published looking at the factor structure within populations of young adults aged 16-19 (Canivez & Watkins, 2010a) and from young adults and adults aged 16-90 (Canivez & Watkins, 2010b) using the Wechsler norming sample. Both of these studies did not replicate the four-factor model. They found that different extraction criteria (i.e. scree test, standard error of the scree, eigenvalue 1≥, Horn's parallel analysis and minimum average partial) indicated either one or two factors when 10 subtests were analysed (Canivez & Watkins, 2010a; Canivez & Watkins, 2010b) and one to four factors when 15 subtests were analysed (Canivez & Watkins 2010b). The only indication that the four-factor model could be replicated in both studies was when the Schmid-Leiman procedure was applied. However the authors of the study suggested that caution is required when using the WAIS-IV and that interpretation should be limited to measuring general intelligence as it accounts for the highest proportion of common variance amongst participants.

Interestingly in the present study two factors were indicated as the best solution by both the scree plot and using the eigenvalue ≥1 criterion. Three and four factor solutions were also specified a priori, but they both revealed poor validity, due to a number of complex variables in each solution. To assess the concept of *g* a one-factor solution was also specified. These solutions will be discussed further.

One factor solution.

Extraction of the one-factor solution using both PAF and PCA showed relatively high factor loadings for all 10 subtests. This finding supports the theory of a construct of *g*, which is that all intellectual ability has one underlying factor.

Two factor solution.

Similarly a relatively strong two-factor solution was extracted using both PAF and PCA. These two factors showed a clear verbal-performance difference (even though this distinction has been removed and replaced with index scores in the WAIS-IV), with all verbal subtests loading highly on one factor and all performance subtests loading highly on the second. Internal consistency measures also indicated that the structure was robust. The consistency of this factor structure in both extraction techniques suggests that it is favourably reliable for use in both clinical and research settings.

Previous factor analytic research has shown this verbal-performance dichotomy in the WAIS, WAIS-R and WAIS-III (Hill et al., 1985; Leckliter, Matarazzo & Silverstein, 1986; Jones, van Schaik & Witts, 2006). Indeed, research by Hill et al, (1985) suggests that empirical support is highest for the two-factor solution and that the weaker three and four factor solutions are a product of the desire of clinicians to have something more clinically meaningful.

Three- factor solution.

The extraction of three factors from the data was shown not to be stable across the two different methods employed. One explanation for this was that the criterion of

acceptance of factor loadings had an effect on the structure, particularly when the lower value of 0.32 was used. Using this criterion many of the variables were shown to be complex, with loadings on more than one factor.

If the higher criterion value of 0.45 was used then some tenuous findings can be examined. Firstly, for both PAF and PCA, one factor could be extracted (factor II in PAF and factor I in PCA) which could both be considered to describe mainly verbal abilities. However the structure of these two factors was not identical. In both PAF and PCA the items included Similarities, Information and Vocabulary with Digit Span and Arithmetic included in the PCA technique. Some of these items load onto the Verbal Comprehension Index (Vocabulary, Similarities and Information) and the Working Memory Index (Digit Span and Arithmetic).

Additionally, a factor with items mainly describing 'performance' abilities could be extracted using the PAF technique (factor III). The structure of this factor included Block Design, Matrix Reasoning and Visual Puzzles. These items would load onto the Perceptual Reasoning Index. Interestingly no factor describing 'performance' was extracted using the PCA technique thus questioning the stability of the findings.

As mentioned earlier PCA did not extract a second factor when a three-factor solution was specified a priori unlike PAF. However there was no evidence of a third factor from the PAF analysis as it only contained one item (Arithmetic).

A verbal-performance distinction is offered from these rather tenuous findings. Jones, van Schaik and Witts, (2006) in their factor analytic study of the WAIS-III in a low IQ population found that in their three factor solution both PAF and PCA techniques produced a second factor describing 'performance' abilities. The composition of these factors for both PAF and PCA included Coding, Picture Arrangement, Symbol Search and Picture Completion; Block Design was included in the PAF structure only. These items (Picture Completion) would load onto the

Perceptual Organisation Index (which is now known as the Perceptual Reasoning Index) and the Processing Speed Index (Coding and Symbol Search). However they questioned the stability of their findings and recommended extreme caution when interpreting their results as their initial acceptance criterion (scree plot and eigenvalues) did not suggest that a three-factor solution could be extracted from their data.

Four-factor Solution.

The factor solutions extracted using both PAF and PCA when a four-factor solution was set a priori were both unclear. If a loading criterion of 0.32 is used, no defined structure can be identified with most of the variables in both solutions being complex, loading two or more factors. If a criterion of 0.45 was used then a verbal factor (factor I in PAF and factor II in PCA) can be precariously interpreted (containing the Similarities, Vocabulary and Symbol Search subtests in the PAF analysis and Similarities, Information and Digit Symbol subtests in the PCA analysis). However this interpretation has very limited corroboration. No other factor was easily interpreted. None of the solutions extracted bore a resemblance to the four-factor structure of the WAIS-IV indexes.

Although there is potential clinical utility in the four-factor solution in intellectual disabilities samples unfortunately there was little evidence in the current study to support this notion. However, future research may be able to clarify this picture.

Concluding observations regarding factor solutions

The aim of the current study was to evaluate the reliability of the four-factor solution of the WAIS-IV as applied to an intellectual disabilities population. This study has shown that utilising a four-factor solution in order to interpret the WAIS-IV in intellectual disabilities samples may not be appropriate, since there is no clear support for this model. This therefore implies that the neuropsychological utility of the test may

be reduced for such populations. The findings from the current study indicate that the only factor structure that can be reliably applied for this population is a two-factor solution of verbal-performance.

The evidence for this verbal-performance distinction has been demonstrated on a number of occasions in previous versions of the WAIS. In the construction of the WAIS-IV this dichotomy was removed and replaced by the index scores as it was thought that this would be a more precise measure of cognitive abilities (Wechsler, 2008).

Previous neuropsychological research supports the verbal-performance distinction, however, modern-day models of intelligence have highlighted the influence of other factors in particular the role of executive functioning in intellectual abilities (Wechsler, 2008). Distinctive factors describing these intellectual abilities have not been reliably extracted from the data in the current study and there may be a number of explanations for this. Individuals with intellectual disabilities are not a homogeneous group in terms of the underlying reasons for poor performance on intellectual tests and hence the processes by which tasks such as those assessed by the WAIS-IV may be different. For instances intellectual disabilities have numerous causative factors, including genetic predisposition, brain injury at birth or during childhood and due to the normal distribution of intellectual ability in the population. As a result, there may be variation within the sample in terms of executive functioning abilities such as working memory and processing speed. Research by Anderson (1998) has implied that individuals with distinct aetiologies may perform differently on a particular ‘cognitive module’ and that this difference has an explicit effect on the approach taken by different groups of individuals. One could therefore speculate that the current sample under investigation may not have adopted a homogenous approach to the subtests in terms of executive functioning abilities used in tasks.

Furthermore, there are those who object to the current diagnostic criteria for intellectual disabilities, particularly the use of IQ testing. Rather than IQ being seen as a measure of intelligence it has been argued that it is a statistical construct and that the index scores that underline its hierarchical model have little meaning outside of the conceptual structure of the WAIS-IV (Hare, 2015). Currently alternative developments for the diagnosis and classification of intellectual disabilities are being put forward, based on developments in neuroscience, genetics and behavioural phenotyping (McPartland & Pelphre, 2012; Rapin, 2014) and endophenotyping (Bilder, Howe & Sabb, 2013). However, the World Health Organisation has concerns about the current validity of these approaches and also notes that there are ethical and social issues that need to be considered. Joyce (2015) has argued that the alternatives being put forward could be seen as a return to reductionist approaches to intellectual disability, as such approaches show that a person has a genetic or other disorder, but that does not always mean that they also have an intellectual disability. In view of the recent developments in alternative approaches it seems likely that this debate will continue.

Therefore, the findings from the current study do not provide a convincing answer to the question of how many factors best describe the fundamental intellectual performance in an intellectual disabilities sample. It is however, a useful exploration of the factor structure of the WAIS-IV in people with intellectual disabilities and leads one to the conclusion that further research is required with this population before a four-factor solution can be considered a reliable measure of cognitive functioning.

Relationship between Vineland and WAIS-IV

In terms of assessing intellectual disability the adaptive functioning construct fulfils four essential functions. Firstly, it helps to define intellectual disability operationally. Secondly, it helps to assess whether the individual has significant limitations as expressed in conceptual, social and practical adaptive skills. Thirdly, it

helps to provide a framework for mapping the development of adaptive skills and establishing rehabilitation goals. Finally, it involves a critical dimension in a multidimensional understanding of human functioning (Tasse et al., 2012).

In the current study it was found that there is a positive albeit small correlation between full scale IQ and adaptive behaviour composite and a larger correlation between full scale IQ and daily living skills. There were no correlations between full scale IQ and communication or full scale IQ and socialisation. These results are somewhat similar to those reported by the developers of the Vineland-II who compared full scale IQ scores from the WAIS-III to the different domains on the Vineland-II (Sparrow, Cicchetti & Balla, 2005). These results provide some evidence that the WAIS-IV and the Vineland-II assess different sets of skills and behaviours and that it is therefore important that both are used in the assessment and diagnoses of intellectual disabilities even if the scores are consistently lower than IQ scores.

However, these results are not consistent with the findings of Obi et al (2011) who suggested that in their population of eight year old children, use of adaptive behaviour scales was not necessary and that IQ scores were sufficient as the sole criterion for intellectual disability case status. However, this study was examining prevalence rates of intellectual disability and not individual diagnosis therefore caution needs to be taken when extrapolating these findings. Therefore, there could be an argument that the importance of assessing adaptive behaviour to help understand the concept of intellectual disability in particular when diagnosing a mild intellectual disability can be seen as crucial. There also appears to be a need for a better understanding of the construct of adaptive behaviour as this will be central to its evolution as a construct.

Limitations of the current study

There are a number of limitations that must be taken into consideration when assessing the outcome of the current study. Firstly, although the administration of the WAIS-IV for this study was carried out following training with an experienced clinical psychologist, no fidelity checks were made. This was out with the scope of the study as these data were collected in routine clinical practice. Secondly, the sample was exclusively Caucasian therefore making it difficult to generalise findings to an ethnically diverse population. However, the geographical area from which the sample was obtained is considered to be almost predominately Caucasian and is a true reflection of the population within that region. Third, the usual assumptions of normality and linearity of the subtests were not fulfilled; however, log transformation of the data produced virtually identical results. Despite this the results should be interpreted with some caution, although this does not suggest complete degradation of the analysis employed (Tabachnick & Fidell, 2001). However, a violation of these assumptions is not surprising given the nature of the sample under investigation, as clinical populations may not be expected to be homogenous in nature. Fourth, no information regarding organic or genetic conditions was collected, therefore not allowing the examination of whether this may have had an effect on the factor solutions extracted. Fifthly, no information was collected on the length of time it took clinicians to administer the WAIS-IV. There may be an argument that those who completed that assessment in more than one session due to a lack of time, resource or attention span may display particular characteristics that make them different from the population as a whole. Sixthly, in order to diagnose an individual as having a learning disability, it must be clear that the individual has an IQ <70 and impairment in adaptive functioning. However, in the current study there were a number of participants (23) for whom there was no data on adaptive functioning, as the assessment was not completed for various reasons. These people were therefore not included in the analysis examining the

relationship between the Vineland-II and the WAIS-IV. As mentioned before this population who did not complete the assessment of adaptive functioning may display particular characteristics that make them different from the population as a whole therefore having a bearing on the outcomes.

Subjectivity of factor analysis.

The technique of factor analysis is complex and requires interpretation, which can therefore make it susceptible to subjectivity. There is also a potential for there being an unlimited number of mathematically correct solutions, which can lead to a lack of agreement concerning the best solution (Kline, 1991). However, in order to overcome this, a simple structure can be adopted where a rotation is applied which produces factors with a few high loadings and many zero or near zero loadings. This technique should propose the simplest and most precise solution and should be the one accepted (Kline, 1991). This concept of 'simple structure' has been utilised in the current study. Furthermore, when the use of different extraction techniques produces the same factor solution confidence in the validity of the factor structure increases (Tabachnick & Fidell, 2001). The current study utilises two common factor extraction techniques for this reason.

Implications for clinical practice

The findings of this study along with the previous factor analytic studies of the WAIS-III questions the validity of the four-factor solution in the interpretation of administrations with people who have intellectual disabilities. It questions usefulness of the four-factor theory that has developed from the WAIS-III and the abandonment of the two-factor model with people who have intellectual disabilities. The assertion in the WAIS-IV manual that the Verbal Comprehension Index and the Perceptual Reasoning Index should be substituted for the Verbal IQ and Performance IQ in clinical decision making (Wechsler, 2008) does not appear to hold for people who have intellectual

disabilities. Loring and Bauer (2010), commented that this assertion was premature for neurological testing and go on to state that the VIQ and PIQ are more sensitive to brain impairment in non-focal brain disease. Thus it may be the case that this also applies to assessments of people who have intellectual disabilities.

This two-factor solution finding is unfortunate as the superior utility provided by the four indexes has the potential to offer greater understanding of the intellectual functioning in intellectual disabilities populations. However, these findings do not suggest that WAIS-IV is not valid with people who have ID. It is interpretations of the four factors that are not valid. Unfortunately the manual does not enable computation of the VIQ and PIQ and this needs to be addressed in future norming of the WAIS.

In terms of assessing adaptive behaviour, the study showed that this is unrelated to IQ score in people who have intellectual disabilities and therefore providing a unique dimension to the diagnostic criteria of intellectual disability. This assessment gathers information, which is not available from just measuring an individual's intelligence and can provide the diagnosing clinician with a more informed view.

Implications for future research

In terms of future research, the next logical step could potentially be to apply confirmatory factor analysis techniques as this approach is considered to be theory testing as opposed to theory generating (Stevens, 1966). Such an analysis was carried out on a USA sample of people who have intellectual disabilities (Reynolds, Ingram, Seeley & Newby 2013) and confirmed the four-factor solution. However, this was premature, as the hypothesis generating stage had not been carried out. Clinicians and researchers need to take into account the contradictory findings across the two exploratory and two confirmatory factor analyses of the WAIS-III and IV. It would also be useful for research to examine the factor structure of the WAIS-IV in larger scale populations with specific conditions and syndromes to ascertain if the factor structure is

similar to that found in the current study or if it is closer to that which is proposed by the developers. For example Styck, and Watkins (2016) similarly found that the four-factor model of the Wechsler Intelligence Scale for Children-IV provided an inadequate explanation of the data.

The WAIS-V is under development and there is a call for qualified professionals to participate in the field research that is currently taking place (Psychological Corporation, 2016). In view of the wide scale use of the WAIS in evaluations of intellectual disabilities the test producers should pay more attention to the psychometric properties of the scale with people who have intellectual disabilities. In the development of the WAIS-IV they included the results of administrations for 13 special groups of which two had mild and moderate intellectual disability respectively. Their aim as stated in the manual was to provide initial evidence of the clinical validity with these groups (Wechsler, 2008b). This was also carried out for the WAIS-III, but not previous versions. In the development of the WAIS-V it would be appropriate to go beyond what they have previously done. i.e. only report means and standard deviations for a subgroup of people who have intellectual disabilities. This study would suggest that larger samples of people who have intellectual disabilities; sufficient for testing the factor structure are obtained. It may also be appropriate to introduce fidelity checks into the development programme of the WAIS-V with the general population and special groups to ensure the accuracy of the examinations. This study used previously collected data and so such checks could not be made.

Conclusions

The current study aimed to explore the factor structure of the WAIS-IV with an intellectual disabilities population. This was done in order to examine the validity of the four factor structure as proposed by the developers of the assessment tool in an

intellectual disabilities population. Research examining the factor structure with earlier versions of the scale (WAIS, WAIS-R and WAIS-III) have shown the one, two, three and four factor solutions may be reliably extracted for a variety of populations. In the current study it was found that only one and two factors could be reliably extracted from the factor analysis. Therefore, this finding does not support the use of the four factor solution for the population currently under investigation until further research has been carried out. In addition the relationship between IQ and adaptive functioning and the contribution of adaptive functioning assessment to the diagnosis was examined. It was found that there is a positive albeit small correlation between full scale IQ and adaptive behaviour composite and a larger correlation between full scale IQ and daily living skills. There were no correlations between full scale IQ and communication or full scale IQ and socialisation. These results provide some evidence that the WAIS-IV and the Vineland-II assess different sets of skills and behaviours. Therefore, there could be an argument that the importance of assessing adaptive behaviour to help understand the concept of intellectual disability in particular when diagnosing a mild intellectual disability can be seen as crucial.

It is hoped that future research will be carried out with an attempt made to avoid the limitations of the current study for both clinical and academic purposes.

References

American Association on Intellectual and Developmental Disabilities. (2010). *Intellectual disability: Diagnosis, classification and systems of supports* (11th ed.). Washington: AAIDD.

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental Disorders* (5th ed.). Washington: American Psychiatric Association

Anderson, M. (1998). Mental retardation general intelligence and modularity. *Learning and Individual Differences, 10*, 159-178. doi:10.1016/s1041-6080(99)80128-9

Benson, N., Hulac, D., & Kranzler, J. (2010). Independent examination of the Wechsler Adult Intelligence Scale Fourth Edition (WAIS-IV): What does the WAIS-IV measure? *Psychological Assessment, 22,* 121-130. doi: 10.1037/a0017767

Bilder, R., Howe, A., & Sabb, F. (2013). Multilevel models from biology to psychology: Mission impossible? *Journal of Abnormal Psychology, 122,* 917-927

British Psychological Society, (2015). *Guidance on the assessment and diagnosis of intellectual disabilities in adulthood.* Leicester: British Psychological Society

Campbell, D.T., & Fiske, D.W. (1959). Convergent and discriminant validation by the multi-trait multi-method matrix. *Psychological Bulletin, 56*, 81-105. doi:10.1037/h0046016

Canivez, G.L., & Watkins, M.W. (2010a). Exploratory and higher order factor analysis of the Wechsler Adult Intelligence Scale Fourth Edition (WAIS-IV) adolescent subsample. *School Psychology Quarterly, 25*, 223-235. doi: 10.1037/a0022046

Canivez, G.L., & Watkins, M.W. (2010b). Investigation of the factor structure of

the Wechsler Adult Intelligence Scale Fourth Edition (WAIS-IV): Exploratory and higher order factor analysis. *School Psychology Quarterly, 22*, 827-836. doi: 10.1037/a0020429

Comrey, A.L., & Lee, H.B. (1992). *A first course in factor analysis.* Hillsdale, NJ: Erlbaum

Dagnan, D., & Warning, M. (2004). Linking stigma to psychological distress: Testing a social cognitive model of the experiences of people with intellectual disabilities. *Clinical Psychology and Psychotherapy, 11,* 247-254. doi:10.1002/cpp.413

Field, A. (2000). *Discovering statistics using SPSS for windows: Advanced techniques for the beginner.* London: Sage

Groth-Marnat, G., Gallagher, R., Hale, J., & Kaplan E. (2000). The Wechsler Intelligence Scales. In Groth-Marnat, G. (Ed.) *Neuropsychological Assessment in Clinical Practice,* John Wiley & Sons: New York

Hare, D. (2015). Letter: Response to faculty guidance on assessment and diagnosis of intellectual disability. *The Bulletin of the Faculty for People with Intellectual Disabilites, 13,* 22-24

Hill, T., Reddon, J., & Jackson, N. (1985). The factor structure of the Wechsler scales: A brief review. *Clinical Psychology Review. 5,* 287-306. doi: 10.1016/0272-7358(85)90009-1

Jensen, A.R. (1987). *Historical foundations of educational psychology.* New York: Plenum

Jones, J., van Schaik, P., & Witts, P. (2006). A factor analysis of the Wechsler Adult Intelligence Scale 3rd Edition (WAIS-III) in a low IQ sample. *British Journal of Clinical Psychology, 45*, 145-152. doi: 10.1348/014466505x29134

Joyce, T. (2015). Response to Dr Hare's letter. *The Bulletin of the Faculty for People with Intellectual Disabilites, 13,* 25-26

Kaiser, H.F., & Rice, J. (1974). Little Jiffy, Mark IV. *Educational and Psychological Measurement, 34*, 111-117. doi:10.1177/001316447403400115

Kaufman, A. (2000). Tests of Intelligence. In R.J. Sternberg (Ed.), *Handbook of intelligence.* Cambridge: Cambridge University Press

Kaufman, A., Lichtenberger, E., & McLean M. (2001). Two and three factor solutions of the WAIS-III. *Assessment, 8,* 267-280

Kline, P. (1991). *Intelligence. The psychometric view.* London: Routledge

Kline, P. (1994). *An easy guide to factor analysis.* London: Routledge

Kline, P. (2000). *A psychometrics primer.* London: Free Association Books

Lawley, D., & Maxwell, A. (1971). *Factor analysis as a statistical method* (2nd ed.). London: Butterworths

Leckliter, I.N., Matarazzo, J.D., & Silverstein, A.B. (1986). A literature review of the factor analytic studies of the WAIS-R. *Journal of Clinical Psychology, 42*, 332-342. doi: 10.1002/1097-4679

Lewis, J. (1995). IBM computer usability satisfaction questionnaire: Psychometric evaluation and instructions for use. *International Journal of Human Computer Interaction. 7,* 57-78. doi:10/1080/10447319509526110

Lichtenberger, E.O., & Kaufman, A.S. (2013). *Essentials of WAIS-IV Assessment:* 2nd ed.). Hoboken, NJ: Wiley.

Loring, D., & Bauer, R. (2010). Testing the limits: Cautions and concerns regarding the new Wechsler IQ and memory scales. *Neurology, 74,* 685-690

MacLean, H., McKenzie, K., Kidd, G., Murray, A., & Schwannauer, M. (2011).

Measurement invariance in the assessment of people with an intellectual disability. *Research in Developmental Disabilities, 32,* 1081-1085. doi:10.1016/j.ridd.2001.01.022

McPartland, J., & Pelphre, K. (2012). The implications of social neuroscience for social disability. *Journal of Autism and Developmental Disorders, 42,* 1256-1262

Murray, G., McKenzie, K., & Lindsay, W. (2003). A pilot study of the use of the WAIS-III with people with a learning disability. *Clinical Psychology, 28*, 17-21

Nunnally, J.C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill

Obi, O., Braun, K.V.N., Baio, J., Drews-Botsch, C., Devine O., & Yeargin-Allsopp, M. (2011). Effect of incorporating adaptive functioning scores on the prevalence of intellectual disability. *American Journal on Intellectual and Developmental Disabilities, 118*, 156-176. doi:10.1352/1944-7558-116.5.360

Psychological Corporation. (2016). Retrieved from http://www.pearsonclinical.com/field-research/current-opportunities/wais-5.html

Rapin, I. (2014). Classifications of behaviourally defined disorders: Biology versus the DSM. *Journal of Autism and Developmental Disorders, 44,* 2661-2666

Reynolds, M., Ingram, P., Seeley, J., & Newby, K. (2013). Investigating the structure and invariance of the Wechsler Adult Intelligence Scales, Fourth Edition in a sample of adults with intellectual disabilities. *Research in Developmental Disabilities, 10,* 3235-3245

Sattler, J., & Ryan, J. (1999). *Assessment of children, revised: WAIS-III supplement.* La Mesa, CA: Jerome Sattler Publishing

Sparrow, S.S., Cicchetti, D.V., & Balla, D.A. (2005). *Vineland II: Vineland Adaptive Behaviour Scales* (2nd ed.). Minneapolis: Pearson Assessments

Stevens, J. (1996). *Applied multivariate statistics for the social sciences* (3rd

ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Styck, K., & Watkins, M. (2016). Structural validity of the WISC-IV for students with learning disabilities. *Journal of Learning Disabilities, 49*, 216-224. doi:10.1177/0022219414539565

Tasse, M.J., Schalock, R.L., Balboni, G., Bersani, H., Borthwick-Duffy, S.A., Spreat, S., Thissen, D., Widaman, K.F., & Zhang, D. (2012). The construct of adaptive behaviour: Its conceptualisation, measurement, and use in the field of intellectual disability. *American Journal on Intellectual and Developmental Disabilities, 117*, 291-303. doi: 10.1352/1944-7558-117.4.291

Tabachnick, B., & Fidell, L. (2001). *Using multivariate statistics* (4th ed.). Needham Heights, MA: Allyn and Bacon

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale-third edition.* San Antonio, TX: Harcourt Assessment

Wechsler, D. (2008a). *Wechsler Adult Intelligence Scale-fourth edition: Technical and interpretive manual.* San Antonio, TX: Pearson.

Wechsler, D. (2008b). *Wechsler Adult Intelligence Scale-fourth edition: Technical and interpretive manual.* San Antonio, TX: Pearson.

Weiss, L., Keith, T., Zhu, J., & Chen, H. (2013). WAIS-IV and clinical validation of the four and five factor interpretative approaches. *Journal of Psychoeducational Assessment, 31,* 91-113. doi:10.1177/0734282913478030

Whitaker, S. (2010). Error in estimation of intellectual ability in the low range using the WISC-IV and WAIS-III. *Personality and Individual Differences, 48,* 517-521. doi:10.1016/j.paid.2009.11.017

Appendix A

NHS
Health Research Authority

National Research Ethics Service

NRES Committee London - Hampstead
Barlow House
3rd Floor
4 Minshull Street
Manchester
M1 3DZ

Telephone: 0161 625 7815
Fax: 0161 625 7299

13 February 2015

Dr Majid M Saleem
Trainee Clinical Psychologist
Sheffield Health and Scocial Care NHS Trust
Clinical Psychology Unit Department of Psychology
University of Sheffield
Western Bank
S10 2TN

Dear Dr Saleem

Study title: **The reliability and validity of the Wechsler Adult Intelligence Scale-Fourth Edition (WAIS-IV) and its concordance with adaptive behaviour in the assessment and diagnosis of intellectual disabilities**
REC reference: **15/LO/0313**
IRAS project ID: **172911**

Thank you for your email of 11 February 2015. I can confirm the REC has received the documents listed below and that these comply with the approval conditions detailed in our letter dated 11 February 2015.

Documents received

The documents received were as follows:

Document	*Version*	*Date*
Other [email confirmation of data storage]		11 February 2015

Approved documents

The final list of approved documentation for the study is therefore as follows:

Document	*Version*	*Date*
Covering letter on headed paper		30 January 2015
Other [email confirmation of data storage]		11 February 2015
REC Application Form	3.5	05 February 2015
Referee's report or other scientific critique report		29 January 2015

Research protocol or project proposal	2	11 December 2014
Summary CV for Chief Investigator (CI) [Majid Saleem]		
Summary CV for supervisor (student research) [Nigel Beail]		

You should ensure that the sponsor has a copy of the final documentation for the study. It is the sponsor's responsibility to ensure that the documentation is made available to R&D offices at all participating sites.

15/LO/0313 **Please quote this number on all correspondence**

Yours sincerely

[redacted]

Dr Ashley Totenhofer
REC Manager

E-mail: nrescommittee.london-hampstead@nhs.net

Copy to: Dr Andrew Thompson – University of Sheffield

Mrs Helen Carter - South West Yorkshire Partnership NHS Foundation Trust

Professor Nigel Beail – University of Sheffield

Appendix B

With all of us in mind

South West Yorkshire Partnership NHS
NHS Foundation Trust

10th March 2015

Dr Majid M Saleem
Trainee Clinical Psychologist
Clinical Psychology Unit
Clinical Psychology Department
University of Sheffield
Western Bank
S10 2TN

Research Department
Ward 2
Castleford & Normanton District Hospital
Lumley Street
Castleford
WF10 5LT

Tel: 01977 605285
Fax: 01977 605298
Ref: Approval Letter
Email address: research@swyt.nhs.uk

Dear Dr Saleem

Re: The Reliability and Validity of the WAIS-IV and its Concordance with Adaptive Behaviour in the Assessment and Diagnosis of Intellectual Disabilities

REC Reference: 15/LO/0313

Following the recent review of the above project I am pleased to inform you that the above project complies with Research Governance standards, and NHS Permission has been granted on behalf of Trust management. We now have all the relevant documentation relating to the above project. As such your project may now begin within South West Yorkshire NHS Foundation Trust.

The final list of documents reviewed and approved is as follows:

Document	Version	Date
Research Protocol	2	11 December 2014
Scientific Approval University of Sheffield		29 January 2015
Indemnity University of Sheffield		29 January 2015
Sponsor Confirmation University of Sheffield		16 February 2015
CV – M Saleen & Professor N Beail		

This approval is granted subject to the following conditions:

- You must comply with the terms of your approval. Failure to do this will lead to permission to carry out this project being withdrawn. If you make any substantive
- changes to your protocol you must inform us immediately.
- You must comply with the procedures on project monitoring and audit[1].

- You must comply with the guidelines laid out in the Research Governance Framework for Health and Social Care[2](RGF). Failure to do this could lead to permission to carry out this research being withdrawn

- You must comply with any other relevant guidelines including the Data Protection Act, The Health and Safety Act and local Trust Policies and Guidelines
- If you encounter any problems during your research you must inform your Sponsor and us immediately to seek appropriate advice or assistance.

Research projects will be added to any formal Department of Health research register.

Please note that suspected misconduct or fraud should be reported, in the first instance, to local Counter Fraud Specialists for this Trust. R&D staff are also mandated to do this in line with requirements of the RGF.

Adverse incidents relating to the research procedures and/or SUSARs (suspected unexpected serious adverse reactions) should be reported, in line with the protocol requirements, using **Trust incident reporting procedures in the first instance and to the chief investigator**[3].
They should **also** be reported to:

- The R&D Department
- the Research Ethics Committee that gave approval for the study (if applicable)
- other related regulatory bodies as appropriate.

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 (http://www.dh.gov.uk/assetRoot/04/06/92/54/04069254.pdf) and the Data Protection Act 1998. Furthermore you should be aware that under the Act, unauthorised disclosure of information is an offence and such disclosures may lead to prosecution.

Changes to the agreed documents MUST be approved by in line with guidance from the Integrated Research Applications System (IRAS), before any changes in documents can be implemented. Details of changes and copies of revised documents, with appropriate version control, must be provided to the R&D Office. Advice on how to undertake this process can be obtained from R&D.

Projects sponsored by organisations other than the Trusts are reminded of those organisations obligations as defined in the Research Governance Framework, and the requirements to inform all organisations of any non-compliance with that framework or other relevant regulations discovered during the course of the research project.

The research sponsor or the Chief Investigator, or the local Principal Investigator, may take appropriate urgent safety measures in order to protect research participants against any immediate hazard to their health or safety.

The R&D office should be notified that such measures have been taken. The notification should also include the reasons why the measures were taken and the plan for further action.

---

[2] Details from:
http://www.dh.gov.uk/PublicationsAndStatistics/Publications/PublicationsPolicyAndGuidance/PublicationsPolicyAndGuidanceArticle/fs/en?CONTENT_ID=4108962&chk=Wde1Tv

[3] SUSARS – this must be within 24 hours of the discovery of the SUSAR incident

Appendix C

Table showing means and standard deviations for current study and WAIS-IV standardised sub samples

	Current study (n=170)		WAIS USA (mild severity, n=73)		WAIS USA (moderate severity, n=31)		WAIS (UK validation sample, n=270)	
Subtest	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Block design	4.5	2.1	4.0	1.9	1.9	1.1	10.65	3.10
Similarities	3.5	1.5	3.8	1.3	2.0	1.1	9.72	2.86
Digit span	3.7	2.1	2.8	1.9	1.3	0.7	10.47	2.77
Matrix reasoning	3.6	1.7	3.6	2.0	2.0	1.2	10.81	3.15
Vocabulary	4.1	1.1	4.1	1.1	3.1	0.9	10.74	3.20
Arithmetic	3.8	1.6	3.6	1.5	1.9	1.3	10.93	3.10
Symbol search	3.2	1.9	3.5	2.4	1.7	1.6	9.92	3.02
Visual puzzles	4.9	1.7	4.7	1.6	3.4	1.3	10.85	3.43
Information	4.3	1.5	4.3	1.3	3.0	1.0	10.92	2.89
Coding	2.5	1.6	3.2	2.3	1.6	1.1	10.14	2.88
Full scale IQ	50.0	7.1	58.5	7.5	48.2	4.7	103.34	14.67