

THE PSYCHOMETRIC ASSESSMENT OF OFFENDERS WITH AN
INTELLECTUAL DISABILITY

by

MARK KELLS

A thesis submitted to
The University of Birmingham
For the degree of
Doctor of Philosophy

Department of Psychology
Faculty of Science
University of Birmingham
August 2011

UNIVERSITY OF
BIRMINGHAM

University of Birmingham Research Archive

e-theses repository

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

ABSTRACT

Two psychometric assessments, developed for use with mainstream offenders of normal intelligence, measuring impulsivity and locus of control respectively, were adapted for use with offenders with an intellectual disability (ID). The language and sentence structures used were simplified, and the content of questions was changed to provide contexts that were familiar to detained offenders with ID. Data generated from the responses of 47 male offenders with ID indicated that both tools demonstrated good levels of reliability, in terms of internal consistency, test re-test reliability and correlations with other tools measuring related concepts. Both instruments were found not to differentiate detained offenders with ID ($n=47$) from two control groups of non-offenders with ID ($n=2 \times 46$). In addition, an examination of the utility of the impulsivity tool in predicting institutional aggression using receiver operating characteristic (ROC) analysis, was examined. Two measures commonly used for risk assessment, the PCL-R (Hare, 2003) and the HCR-20 (Webster, Douglas, Eaves & Hart, 1997), were also included in this study. The results indicated that the impulsivity measure did not predict institutional aggression. However, the HCR-20 and its sub-scales and the PCL-R and its factors possessed moderate and strong predictive ability respectively, particularly in relation to physical aggression.

CONTENTS

List of Tables and Figures

Acknowledgements

Abstract

Chapter 1 Background to the thesis 1

- 1.1. Defining intellectual disability 1**
- 1.2. Features of intellectual disability 3**
 - 1.2.1. Principles of communication 4**
- 1.3. Intellectual disability and crime 5**
 - 1.3.1. Locus of control and criminal behaviour 15**
 - 1.3.2. Impulsivity and criminal behaviour 18**
- 1.4. Assessment in intellectual disability function 22**
 - 1.4.1. Psychometric assessments of intellectual disability 23**
- 1.5. Reliability and validity measurement 26**
 - 1.5.1. Retest method 27**
 - 1.5.2. Alternative-forms method 27**
 - 1.5.3. Split-halves method 28**
 - 1.5.4. Internal consistency method 28**
 - 1.5.5. Validity of assessments 28**
 - 1.5.6. Criterion-related validity 29**
 - 1.5.7. Content validity 29**
 - 1.5.8. Construct validity 30**
 - 1.5.9. The measurement of validity in samples with ID 30**

1.6. Samples described in the thesis	31
1.6.1. Institutionalised offender sample	31
1.6.2. Institutionalised non-offender sample	34
1.6.3. Non-institutionalised non-offender sample	35
1.7. Data collection	35
1.8. Structure of thesis	36
1.9. Conclusions	36

Chapter 2 A review of response biases implicated in the psychometric assessment of

people with an intellectual disability	38
2.0. Introduction	38
2.1. Acquiescence	38
2.2. Negative response bias	40
2.3. Suggestibility	41
2.3.1. Critique of suggestibility research	44
2.4. Types of assessment	44
2.4.1. Either/or questions	44
2.4.2. Open-ended questions	45
2.4.3. Multiple-choice questions	77
2.4.4. Pictorial assessments	48
2.4.5. Repetition	48
2.4.6. Likert-scales	51
2.4.6.1. Consistency of Likert-scales	53
2.4.6.2. Reliability of self-report using Likert-scales	54
2.4.6.3. Conclusions regarding the use of Likert-scales	56

2.5. Possible solutions to alleviate response biases 58

2.6. Conclusions 60

**Chapter 3 A review examining psychometric tools used in the assessment of offenders
with an intellectual disability 62**

3.0. Introduction 62

3.1. Assessment of anger and aggression 62

3.2. Assessment of fire-setting 68

3.3. Assessment of sexually inappropriate behaviour 69

3.3.1. Pro-offending attitudes 69

3.3.2. Socio-affective functioning 73

3.3.3. Self-management problems 78

3.4. Measuring locus of control 80

3.5. Measuring empathy 82

3.6. Measuring impulsivity 83

3.7. Measures of mental health 83

3.8. Discussion and conclusion 85

**Chapter 4 A revision of the Nowicki-Strickland Internal-External locus of control scale
for offenders with an intellectual disability 87**

4.0. Introduction 87

4.1. Locus of control and intellectual disability 88

4.2. The measurement of locus of control 92

4.3. Aims of research 95

4.3.1. Hypotheses 96

4.4. Methods	97
4.4.1. Sample: study 1 and 2a	97
4.4.2. Sample: study 2b	97
4.4.3. Procedure study 1	97
4.4.4. Procedure study 2a	100
4.4.5. Procedure study 2b	101
4.5. Results	101
4.5.1. Study 1: Creation of the M-ANSIE	101
4.5.2. Study 2a: Testing the psychometric properties of the M-ANSIE	109
4.5.3. Study 2b: Factor analysis of the M-ANSIE	112
4.6. Discussion	115
4.7. Conclusions	119

Chapter 5 A revision of the Eysenck impulsivity questionnaire: an assessment of the validity of self-report impulsivity in a population of offenders with intellectual disability 123

5.0. Introduction	123
5.1. Impulsivity	124
5.2. Impulsivity as a personality characteristic	125
5.3. The validity of the psychometric assessment of impulsivity	132
5.4. Impulsivity in offenders with an intellectual disability	133
5.5. Psychological assessment of people with an intellectual disability	134
5.6. Aims of research	134
5.6.1. Hypotheses	134
5.7. Methods	135

5.7.1. Sample: study 1 and 2a	135
5.7.2. Sample: study 2b	135
5.7.3. Measures	135
5.7.4. Procedure: study 1	138
5.7.5. Procedure: study 2a	139
5.7.6. Procedure: study 2b	140
5.8. Results	140
5.8.1. Study 1: Creation of the I7i-R	140
5.8.2. Study 2a: Testing the psychometric properties of the I7i-R	147
5.8.3. Study 2b: Factor analysis of the I7i-R	151
5.9. Discussion	154
5.10 Conclusion	161

Chapter 6 The role of impulsivity in the risk assessment of offenders with an intellectual disability in a medium secure hospital setting 165

6.0. Introduction	165
6.1. Risk assessment in offenders with an intellectual disability	165
6.2. Critique of risk assessment in ID	171
6.3. The identification of risk factors in samples with ID	171
6.4. Dynamic risk factors in offenders with an intellectual disability	173
6.5. Psychopathy	175
6.6. Personality development in people with an intellectual disability	183
6.7. The application of mainstream personality theory	184
6.8. Personality disorder in people with an intellectual disability	186
6.9. Aims of research	190

6.9.1. Hypotheses	191
6.10. Methods	192
6.10.1. Sample	192
6.10.2. Measures	192
6.10.3. Procedure	195
6.11. Results	196
6.11.1. Multiple regression	198
6.11.2. Receiver operator characteristics analysis	199
6.12. Discussion	206
6.13. Conclusion	211

Chapter 7 A comparison of self-report impulsivity and locus of control between institutionalised offenders with ID, institutionalised non-offenders with ID and non-institutionalised non-offenders with an intellectual disability 213

7.0. Introduction	213
7.1. Locus of control	213
7.2. Locus of control and criminal behaviour	214
7.3. Locus of control and other criminogenic factors	214
7.4. Locus of control and institutionalisation	215
7.5. Locus of control and offenders in prison	217
7.6. Impulsivity and criminal behaviour	218
7.7. Aims of research	220
7.7.1. Hypotheses	220
7.8. Methods	221
7.8.1. Sample	221

7.8.2. Design	221
7.8.3. Measures	221
7.8.4. Procedure	221
7.9. Results	221
7.10. Discussion	222
7.11. Conclusion	226
Chapter 8 Thesis conclusions	228
References	235

List of Tables and Figures

Figures

Figure 4.1. Error Bar Graph Showing Mean Flesch Reading Ease Statistics for the ANSIE and the M-ANSIE **103**

Figure 4.2. Error Bar Graph Showing Mean Item Flesch-Kincaid Grade Level Statistic for the ANSIE and the M-ANSIE **104**

Figure 4.3. Error Bar Graph Showing Mean Participant Ratings for Ease of Understanding the ANSIE and M-ANSIE **105**

Figure 4.4. Error Bar Graph Illustrating Mean Participant Ratings for Ease of Answering the ANSIE and the M-ANSIE **106**

Figure 4.5. Error Bar Graph Illustrating the Mean Number of Items Identified as Difficult to Understand on the ANSIE and the M-ANSIE **107**

Figure 4.6. Error Bar Graph Illustrating the Mean Number of Items Identified as Difficult to Answer on the ANSIE and the M-ANSIE 108

Figure 4.7. Scatterplot Illustrating the Correlation between the Participants' Scores on the ANSIE and the M-ANSIE 110

Figure 4.8. Scree Plot of Eigenvalues for the Principal Components Factor Analysis of the M-ANSIE 113

Figure 5.1. Error Bar Graph Showing Mean Flesch Reading Ease Statistic for the I7i and the I7i-R 142

Figure 5.2. Error Bar Graph Showing Mean Item Flesch-Kincaid Grade Level Statistic for the I7i and the I7i-R 142

Figure 5.3. Error Bar Graph Showing Mean Participant Ratings for Ease of Understanding the I7i and I7i-R 143

Figure 5.4. Error Bar Graph Illustrating Mean Participant Ratings for Ease of Answering the I7i and the I7i-R 144

Figure 5.5. Error Bar Graph Illustrating the Mean Number of Items Identified as Difficult to Understand on the I7i and the I7i-R 145

Figure 5.6. Error Bar Graph Illustrating the Mean Number of Items Identified as Difficult to Answer on the I7i and the I7i-R 146

Figure 5.7. Scatterplot Illustrating the Correlation between the Participant's Scores on the I7i and the I7i-R 149

Figure 5.8. Scree Plot of Eigenvalues for the Principal Components Factor Analysis of the I7i-R 152

Figure 6.1 ROC Curves for Verbal/Property Aggression as Predicted by the I7i, I7i-R and IQ. 202

Figure 6.2 ROC Curves for Verbal/Property Aggression as Predicted by the PCL-R, Factor 1, Factor 2 and 13-Item Total. **202**

Figure 6.3 ROC Curves for Verbal/Property Violence as Predicted by the HCR-20, H-subscale, C-subscale and R-subscale **203**

Figure 6.4 ROC Curves for Physical Violence as Predicted by the I7i, I7i-R and IQ. **203**

Figure 6.5 ROC Curves for Physical Violence as Predicted by the PCL-R, Factor 1, Factor 2 and 13-Item Total **204**

Figure 6.6 ROC Curves for Physical Violence as Predicted by the HCR-20, H-subscale, C-subscale and R-subscale. **204**

Figure 6.7 ROC Curves for High-Risk Violence as Predicted by the I7i, I7i-R and IQ. **205**

Figure 6.8 ROC Curves for High-Risk Violence as Predicted by the PCL-R, Factor 1, Factor 2 and 13-Item Total. **205**

Figure 6.9 ROC Curves for High-Risk Violence as Predicted by the HCR-20, H-subscale, C-subscale and R-subscale. **206**

Tables

Table 1.1 Qualitative Descriptions of WAIS-III Full Scale IQ Scores **2**

Table 1.2 Guidelines for Interpreting the Size of Correlations **27**

Table 1.3 Mean Age, IQ and Length of Time as Resident on the Unit **34**

Table 4.1 Ages Appropriate to American School Grades **98**

Table 4.2 Examples of the Original and Modified ANSIE Items **99**

Table 4.3 Cohen's (1988) Categorisation of Effect Sizes **101**

Table 4.4 Readability Statistics for the ANSIE and the M-ANSIE **102**

Table 4.5 The Mean Participant Ratings for Ease of Understanding and Ease of Answering the ANSIE and M-ANSIE **105**

Table 4.6 Mean Number of Items Identified as Either Difficult to Understand or Difficult

to Answer on the ANSIE and M-ANSIE 106

- Table 4.8** The Percentage of Externally Oriented Items Endorsed by Yes and No Responses 111
- Table 4.9** Eigenvalues and Contribution to the Overall Variance Explained by the 8 Factors 114
- Table 4.10** Summary of the Item Content of the Identified Factors 114
- Table 5.1** The Primary Traits Comprising the Super-Factors of N, P and E 126
- Table 5.2** Summary of Item Content of the Factors of the I7i 129
- Table 5.3** Examples of the Original and I7i-R Items 138
- Table 5.4** Readability Statistics for the I7i and the I7i-R 141
- Table 5.5** The mean participant ratings for ease of understanding and ease of answering the I7i and I7i-R 143
- Table 5.6** Mean Number of Items Identified as Either Difficult to Understand or Difficult to Answer on the I7i and I7i-R 145
- Table 5.7** The Percentage of Items Endorsed as Impulsive by Yes and No Responses 147
- Table 5.8** Means Scores and Standard Deviations Obtained on the I7i and I7i-R 148
- Table 5.9** Spearman Correlations of Scores on Self-Report Measures of Impulsivity 150
- Table 5.10** Spearman correlations for the self-report measures of impulsivity and the BADS Total and subscales 150
- Table 5.11** Eigenvalues and Contribution to the Overall Variance Explained by the 6 Factors 153
- Table 5.12** Summary of the Item Content of the Identified Factors 154
- Table 6.1** PCL-R 2-Factor, 4-Facet, 20-item Model 176
- Table 6.2** Internal Consistency (Cronbach's Alpha) for PCL-R Total, Factor and Facet

Scores (Morrissey et al, 2005). **178**

Table 6.3 Biserial Correlations Between PCL-R 20- and 13-Item Totals, PCL-R Factors, PCL-R Facets, and Positive and Negative Progress (from Morrissey et al., 2007) **180**

Table 6.4 Predictive Accuracy of the PCL-R and HCR-20 in Relation to Positive and Negative Progress Measures: ROC Analysis (from Morrissey, Mooney, et al., 2007) **181**

Table 6.5 The Factors and Facets of the Five Factor Model (McCrae, & Costa, 1999) **186**

Table 6.6 Means and Standard Deviations for PCL-R 20- and 13-Item Totals, PCL-R Factors, HCR-20 Total and H-, C-, and R-Subscales, and Aggression Types 1, 2 and 3. **196**

Table 6.7 Spearman's rho correlations for the independent variables with institutional behaviour exhibited over a 3 month period. **198**

Table 6.8 ROC Analyses: ROCs and 95% confidence intervals for three types of aggression occurring within 3 months of assessment. **201**

Table 7.1 Mean scores and standard deviations for total scores on the M-ANSIE and I7i-R and in Institutionalised Offenders, Institutionalised Non-Offenders and Non-Institutionalised Non-Offenders. **222**

CHAPTER ONE: Background to the Thesis

The aim of this thesis was to advance our understanding of offending and the assessment of offending behaviour in people with intellectual disability (ID). In addition, an attempt was made to generate findings that would be of practical use by those employed in the criminal justice system (CJS). This has been done by demonstrating the adaptation of two valid and reliable tools measuring self-report locus of control (LOC) and impulsivity respectively. This thesis has attempted to add to our empirical knowledge of response biases and the adaptation of assessment tools to minimise the threat of these response biases, and thus to improve the reliability and validity of self-report measures used with offenders with ID. Empirical knowledge has also been gained in relation to the predictive accuracy of two tools used in risk assessment, the Psychopathy Checklist - Revised (PCL-R: Hare, 2003) and the Historical Clinical Risk-20 (HCR-20: Webster, Douglas, Eaves & Hart, 1997), developed in relation to mainstream non-ID offenders, when used with offenders with ID. The thesis has also sought to clarify whether two factors associated with offending in mainstream non-ID samples, external orientation of LOC and impulsivity, are also elevated in a sample of offenders with ID.

This chapter consists of a definition of ID and an outline of some the deficits associated with ID that may impact upon self-report assessment in people with ID. The literature relating to the prevalence of people with ID in the JCS is then critically reviewed and summarised. The evidence relating LOC and impulsivity to offending behaviour is also critically reviewed and summarised. The need to assess the reliability and validity of measures is discussed and material relating to different forms of reliability and validity is presented. The actual composition of samples in studies concerning people with ID is critically reviewed. The sample for this thesis and the data collection process are then described. The structure of the thesis is then outlined.

1.1. Defining Intellectual Disability

Intellectual Disability (ID) is defined as a significant impairment of intelligence *and* social functioning acquired before adulthood. The three core criteria within this definition have gained widespread acceptance across professional boundaries in both the UK and America (American Association on Mental Retardation, 1992; American Psychiatric Association, 2000; British Psychological Society, 2001; Royal College of Psychiatry, 2001; World Health Organisation, 1992), these criteria are:

- Significant impairment of intellectual functioning (IQ <70);
- Significant impairment of adaptive/social functioning (significantly reduced functioning in at least two of the following areas: communication; self-care; home living, social/interpersonal skills; use of community resources; self-direction; functional academic skills, work, health and safety);
- Age of onset before adulthood (< 18years old).

Significant impairment of intellectual functioning is determined through the administration of a reliable, valid and properly standardised psychometric assessment, such as the Wechsler Adult Intelligence Scales, currently its fourth edition (WAIS-IV: Wechsler, 2010). The WAIS-III was used for all assessments of intellectual functioning in this thesis. The mean IQ score on the WAIS instruments is 100 with a standard deviation of 15. About two thirds of all scores on the WAIS fall within one standard deviation of the mean (IQ range= 85-115), and about 95% fall within two standard deviations of the mean (IQ range=70-130) (Wechsler, 1998). Significant impairment in intellectual functioning is defined as an IQ below the lower boundary of two standard deviations from the mean, represented by an IQ of less than 70.

Impairment of adaptive/social functioning is determined by the use of one of several scales designed for that purpose, the best known of which is the Vineland Adaptive Behaviour Scale

(VABS: Sparrow, Balla & Cicchetti, 1984) and are recommended by the British Psychological Society (BPS) (2001) and the World Health Organisation (WHO) (1992).

Table 1.1 shows the qualitative descriptions of the WAIS-III full IQ scores.

Table 1.1. Qualitative Descriptions of WAIS-III Full Scale IQ Scores

IQ Score	Classification	Theoretical Percentage of population within IQ range
130 and above	Very Superior	2.2
120-129	Superior	6.7
110-119	High Average	16.1
90-109	Average	50
80-89	Low Average	16.1
70-79	Borderline	6.7
69 and below	Extremely Low	2.2

1.2. Features of ID

The performance of an individual with ID during an assessment will be mediated by motivational and cognitive processes (Beirne-Smith, Patton & Kim, 2006). Motivational factors include self-determination, mediated by locus of control (which will be discussed in detail in Chapter 4), outer-directedness and expectancy of failure (Luthar & Zigler, 1988; Yando & Zigler, 1971; Zigler, 1961; Zigler, Bennett-Gates, Hodapp & Henrich, 2002; Zigler & Burack, 1989), as discussed in Chapter 6. Cognitive processes include attention (Krupski, 1980; Krupski, 1985), memory (Brown, 1972; Brown, Campion & Murphy, 1974; Brown & Scott, 1971) and communication difficulties (Abbeduto & Nuccio, 1991). *Locus of control* refers to the extent that people view events in their life as controlled by themselves or others. People with an external locus of control tend to view life's events as controlled by others (Nowicki, 1976). Outer-directedness refers to a problem-solving strategy that relies upon

external cues for guidance (Balla & Zigler, 1979). Evidence suggests that people with ID have an external locus of control (LOC) (Wehmeyer & Palmer, 1997) (see Chapter 4) and tend to adopt outer-directedness as a problem-solving strategy (Balla & Zigler, 1979) (See Chapter 6). Research has also demonstrated that people with ID have a high expectancy of failure (Cummins & Das, 1980; Schloss, Alper & Jayne, 1994). The detrimental effects of repeated failures are consistent with all three of these motivational themes (Zigler et al., 2002). It would seem reasonable to assume therefore, that the net result of these motivational difficulties is that many people with ID may be prone to under-perform in a variety of social and academic contexts. Evidence presented in Chapter 2 suggests that people with ID do not always perform well using a variety of questioning techniques when compared to peers without ID (Gudjonsson & Henry, 2003)

Deficits in attention (Krupski, 1980; Krupski, 1985), short-term memory (Brown, 1972; Brown et al., 1974; Brown & Scott, 1971) and long-term memory (Cantor & Ryan, 1962), and receptive (Abbeduto et al, 1989; Abbeduto & Nuccio, 1991; Bartel, Bryen & Keehn, 1973; Cromer, 1972, 1974) and expressive language (Dunn & Dunn, 1981), have been reported in populations with ID when compared with people of normal IQ. However, these are not global deficits. Participants with ID perform as well as participants with normal IQ on simple tasks testing attention (Blackwell, McIntyre & Murray, 1983; Krupski, 1985), memory (Brown, 1972) and language skills (Abbeduto et al., 1989). However, as tasks become more complex performance deteriorates. For example, the Span or Apprehension Task is designed to measure selective attention (Krupski, 1987). In this task, participants have to spot target letters displayed on a screen when embedded within a series of interference letters. When embedded within 0, 2 or 4 interference letters no differences between participants with or without ID were observed. However, when the interference

series extends to eight letters or more, participants with ID perform significantly worse than matched participants of normal IQ (Blackwell et al., 1983).

However, there are problems associated with these studies, which include: small sample sizes (Abbeduto et al., 1989; Abbeduto & Nuccio, 1991; Brown, 1972); lack of specificity regarding the reporting of IQ for ID groups (Blackwell et al., 1983; Krupski, 1980; Krupski, 1985); and (possibly due to the developmental nature of this work) an over-representation of studies including children as participants (Abbeduto et al., 1989; Abbeduto & Nuccio, 1991; Blackwell et al., 1983; Krupski, 1980; Krupski, 1985). These limitations mean that caution should be exercised when generalising from the results. Regardless, the conclusions of studies overwhelmingly indicate deficits in these areas associated with ID.

These motivational and cognitive difficulties are likely to impact upon the ability of people with ID to answer questions. For example, attention deficits are likely to make any interview process more taxing and place higher demands upon the cognitive mechanisms of a person with ID. Long-term memory difficulties may hinder the recall of events. Short-term memory difficulties may prevent the person from holding the question in their memory while they decide upon an appropriate response. Difficulties with receptive language mean that an individual with ID may be unable to understand questions, particularly as those questions become more complex. Difficulties with expressive language may prevent an individual with ID from being able to express their actual attitudes or beliefs.

1.2.1. Principles of Communication

In view of these difficulties, Mencap (2010) has produced guidelines on how to provide written information to people with intellectual disabilities. However, the same principles apply to spoken communication. The basic principles are:

- give some thought to the questions you will need to ask before the interview starts;

- keep the language clear and simple;
- try to keep to one idea per sentence;
- keep sentences short;
- try to use the same word for the same thing;
- find out what the person's own words are for specific things (this can often be culturally based);
- be supportive.

In addition, Bull (2010) summarises the factors that interviewers need to consider when speaking to vulnerable witnesses, such as people with ID, in order to gain a valid account of an incident. It is likely that these factors will help to facilitate valid responding in wider contexts. Interviewers need to:

- speak more slowly;
- allow extra time to enable the person with intellectual disabilities to take in what is being said;
- allow time for the person with intellectual disabilities to think about how they are going to answer the question;
- do not rush the questions;
- avoid interrupting;
- be patient.

The interviewing procedures used in gathering data for this thesis conformed to these guidelines. When interviewing an individual as part of an assessment it is important that the information gained is reliable and valid.

1.3. ID and Crime

Ascertaining the true nature and prevalence of offending in ID populations is a complex and problematic venture not least because research does not occur within a socio-cultural vacuum. Pressures associated with public concern about the threat presented by certain groups, political policy, the wider socio-cultural context, the imprecise nature of definitions and measurement of ID, attrition and filtering at all levels of the criminal justice system and the consequent ambivalence of crime statistic are active within the processes and the implications of research (Day, 1994; Holland, 2004; Holland, Clare & Mukhopadhyay, 2002; Lyall, Holland, & Collins, 1995a; Murphy & Mason, 1999).

In addition, those research studies considering the prevalence of offending in ID populations are methodologically flawed. It is not within the scope of this paper to consider all of those methodological difficulties (for a discussion of these issues see Holland, 2004) but rather to provide a flavour of the findings. As attrition occurs at many points between the commissioning of a potentially illegal act and conviction, the end-point of the criminal justice process, it is important at what stage of the CJS process sampling takes place (Holland, 2004). Consequently, it is likely that prevalence will be lower in the later stages of the CJS due to diversion policies (Simpson & Hogg, 2001).

Some studies have highlighted the differential responses of professionals in the criminal justice system when dealing with people with ID (Cant & Standen, 2007; McAfee, Cockeram & Wolfe, 2001) and some have also included the interface between the CJS and community services (Fryson, 2005; Keilty & Connelly, 2001; Lyall, Holland, Collins & Styles, 1995; McBrien & Murphy, 2006). Several of these studies highlight the reluctance of care workers or teachers to report crimes to the police, citing reasons such as fear of blame or investigation of their own practices and the *labelling* of the perpetrator (Fryson, 2005; Keilty & Connelly, 2001; Lyall, Holland, Collins & Styles, 1995). For example, McBrien and

Murphy (2006) reported that only 48% of carers included in their study would report theft, 68% would report an assault and 83% would report a rape. Fryson (2005) reported that 88% of schools had experienced pupils behaving in a sexually inappropriate way, in 65% of schools this occurred at least once a term and in 19% of schools it occurred on a weekly basis. Of the incidents reported 58% involved public masturbation and 15% involved actual or attempted bodily penetration. In response to incidents, 54% of schools had approached social services but only 23% had approached the police.

In addition, it has been reported that the police are reluctant to investigate crimes committed by people with ID, preferring incidents to be dealt with by care staff (Keilty & Connelly, 2001). However, McBrien and Murphy (2006) in a more recent study reassuringly report that all of the police officers included in their study felt that rape and assault should be reported, while 72% felt that a theft should be reported. McAfee et al. (2001) report police officers differential responses to crime scenarios based upon the presence or absence of ID in the perpetrator or victim of a crime. Officers reported that they were more likely to believe the victim of crime if they had ID and were more likely to provide extra help for a victim with ID. For minor crimes, officers reported stronger action against a perpetrator when the victim had ID but weaker action against a perpetrator with ID. For major crimes officers were less likely to believe the victim when the perpetrator had ID but were *more* likely to rate the crime as serious.

These studies highlight some of the complexity at the interface of the community and the CJS. The failure to report offences or to investigate potential offences will clearly impact upon the number of people with ID progressing into the CJS. It should be noted however, that there are limitations to the studies described above. For example, all include relatively small sample sizes, with the exception of McBrien & Murphy (2006), which included 80 care staff and 65 police officers. In the study citing the figures of most concern (Lyall, Holland,

Collins & Styles, 1995) it is unclear how the managers of care-homes were interviewed, or indeed how many were interviewed, and consequently it is difficult to generalise from the findings. In addition, the reliability of offending rates is not clear in all studies, either because offences are identified by carers rather than actual crime reports (Lyall et al, 1995) or because the definition used, for example, sexually inappropriate behaviour may include behaviours over and above those that are offences (Fryson, 2005). Other studies are based upon paper and pencil exercises in which a number of variables are manipulated (McAfee et al., 2001; McBrien & Murphy, 2006). It is difficult to gauge the extent to which responses to hypothetical scenarios generalise to actual behaviour in reality.

Despite filtering of potentially criminal acts described above, many offences involving suspected offenders with ID are investigated by the police. Data suggest that between 0.9% and 9% of people interviewed in police stations have an IQ below 70 (Irving, 1980; Irving & McKenzie, 1989; Lyall, Holland & Collins, 1995). Lyall et al. (1995) identified 5% of people taken into police custody in a single police station over a three-month period as having ID. Winter, Holland and Collins (1997) screened individuals leaving or entering police custody over a 33 day period. Formal IQ testing identified two people (1%) with IQs below 70. However, there are problems with these studies. Both studies relied upon the use of a screening tool (Clare & Gudjonsson, 1992) to identify potential participants. Clare and Gudjonsson (1992) identified problems with the screening tool. In their study it led to the identification of 5% false positives, where two participants from 39 identified themselves as having reading difficulties or a learning disability which subsequent formal testing indicated was not ID. Equally, it led to the identification of 20% false negatives, where 11 from 54 participants who did have ID, did not identify themselves. The net effect is that the screening tool is likely to result in a net under identification of people with ID. Consequently, the failure of Lyall, Holland and Collins (1995) to complete a formal

IQ assessment is likely to result in the under identification of people with ID. In the Winter et al. (1997) study, only 21 of the 47 people identified as possibly ID agreed to participate in the study and therefore it is possible that some people with ID were not tested, leading to a lower identification rate. In the Lyall, Holland and Collins (1995) study, the screening tool was only used with 25% of those entering custody. No explanation was given for this but this non-random sampling may have affected the results. Some studies have considered the vulnerability of suspects with ID in police custody (Clare & Gudjonsson, 1995; Leggett, Goodman & Dinani, 2007). These studies highlight the lack of understanding of the consequences of police interviewing and the use of interviewing techniques not suited to gaining reliable and valid information.

No British studies could be identified that have attempted to assess the prevalence of ID in relation to court appearances. However, a small scale qualitative study again highlights the different attitudes to dealing with people with ID in these contexts (Cant & Standen, 2007), some of which may again lead to filtering away from formal legal procedures. However, two Australian studies identified 24% and 10% respectively, of people accused of committing offences appearing in four Australian courts as having IQs below 70, and a further 14% and 20% respectively, had IQs in the borderline range (Hayes, 1997; Vanny, Levy, Greenberg & Hayes, 2009). There are several methodological problems with these studies. Both studies adopt a non-random selection method and use samples not representative of the wider Australian population. In addition, IQ was determined using the Matrices Section of the Kaufman Brief Intelligence Test (K-BIT: Kaufman & Kaufman, 1990) in the first study, and the K-BIT in the second, although it is not clear if the full assessment was used. The reliability and validity of this measure and the procedure used to estimate the full-scale IQ is not clear. In addition, it is unclear whether cultural differences of

indigenous Australians, included in the study, may have impacted on their scores, thus underestimating actual intellectual ability.

If conviction occurs offenders with ID may face incarceration. The mean IQ score for convicted prisoners is below average suggesting that many prisoners are intellectually disadvantaged (Birmingham, Mason & Grubin, 1996; Eysenck & Gudjonsson, 1989; Mason & Murphy, 2002). Consequently, one might expect high numbers of offenders with ID in the prison population. However, studies generally report low levels of people with IQs less than 70 in prison populations (Fazel, Xenitidis & Powell, 2008).

For example, Murphy, Harnett and Holland (1995) reported 0%. However, Birmingham, et al. (1996) reported 13% and Crocker, Cote, Toupin and St-Onge (2007), using a Canadian sample, report that 19% of their sample were in the probable ID range, with a further 30% in the borderline range of intelligence. Sondennaa, Rasmussen, Palmstierna and Nottestad (2008) using a Norwegian prison sample, report a prevalence rate of 11% with IQs below 70 and an additional 20% in the borderline range. Holland and Persson (2010) assessed prevalence at 1% using an Australian prison sample. These figures suggest a wide range of prevalence rates in prisons and highlight the extremely large numbers of prisoners in the borderline range. However, methodological problems make comparisons difficult and make generalising from the results problematic. Two of the studies considered men on remand and therefore, were not offenders and were unlikely to be representative of the wider prison population (Birmingham et al., 1996; Murphy et al., 1995). The means of determining IQ was different in every study and the reliability of those measures is questionable in several, using screening tools, not specifying the measure used, or reliance upon official records, without reference to how IQ was actually assessed (Birmingham et al., 1996; Crocker et al., 2007; Holland and Persson, 2010; Murphy et al., 1995). In addition, for a number of reasons, only 38% (Crocker et al., 2007) and 77% (Sondennaa et al., 2008) of the

originally identified samples were used. Consequently, it is possible that some non-random sampling had affected the results in these studies.

Shelton (2006) reported the prevalence of ID in a juvenile justice system for offenders between the ages of 15- and 17-years in the US at 20%. This study included a good sample size, containing 290 (25%) of all male detainees and 60 (100%) of female detainees over a four-month period. However, it is unclear how the level of ID was determined, as the means of determining IQ was not specified and no reference was made to the level of adaptive functioning of the sample. In addition, individuals with IQs as high as 82 were included in the ID group, suggesting that the level of ID was over-estimated, as these IQs actually fall in the low/average range.

Mason and Murphy (2002) report the prevalence of ID in a sample under the supervision of the probation service. Mason & Murphy (2002) report WAIS-R Short-Form (WAIS-R; Wechsler, 1981) and VABS (Sparrow et al., 1984) scores, and define ID as individuals who scored two standard deviations below the mean on *both* assessments. The authors reported a prevalence rate of 6% with IQs of 75 or less (allowing for two standard errors). However, if just the WAIS-R scores were used to determine intellectual functioning the figure rose to 17%. The findings of this study are limited by a small sample size ($n=70$) and a non-random sampling technique, based on the case-load of probation officers who volunteered to take part.

Another research strategy for identifying the prevalence of offending in ID populations is to consider the wider population of people with ID and then attempt to assess how many of them have had contact with the CJS. Studies have identified between 1% and 10% respectively of people known to UK developmental disability services having been in contact with the police or other criminal justice agencies because of alleged offending (Lyall,

Holland & Collins, 1995; McBrien, Hodgetts & Gregory, 2003; McNulty, Kissi-Deborah & Newsome-Davies, 1995; Seaward & Rees, 2001).

Studies have reported that 2%, 10%, 5% and 1% respectively, of their samples had been in contact with the CJS in relation to the perpetration of a crime (Lyall, Holland, Collins & Styles, 1995; McBrien et al., 2003; McNulty et al, 1995; Seaward & Rees, 2001). In addition, Tyrer et al. (2006) reported on a large sample ($n=3062$) drawn from the Leicestershire Learning Disability Register. They reported that 14% people with had been physically aggressive towards others in the preceding two years. In this instance physical aggression was determined from carers responses to a questionnaire routinely completed as part the services assessment process. In relation to 80 (3%) individuals, the level of physical aggression had resulted in serious injury to the other person. In response to this aggression, 17 individuals (1%) had been permanently excluded from day services and only 45 (1.5%) resulted in the police taking or threatening to take formal action. Cooper et al. (2009) used a two year longitudinal study to identify the prevalence of aggressive behaviour within a geographically defined area of Scotland. They report prevalence of aggression (including verbal aggression, destructive aggression and physical aggression) as 10%. In addition, the remission rate, that is to say those who had been violent prior to the start of the study and were violent again during the two year follow-up, was 28%.

The advantage of these studies is that they often included large sample sizes (Lyall, Holland & Collins, 1995 ($n=358$); McBrien et al., 2003 ($n=1326$); Tyrer et al., 2006 ($n=3062$)). However, generalisation from these studies is problematic as they do not utilise representative samples, only including those accessing or known to services. Inevitably the sample only includes those who agree to participate, which may influence prevalence rates if those displaying aggressive behaviours are more reluctant to participate than those not displaying aggressive behaviour. None of the studies assessed ID independently, instead

relying upon the criteria for access to services in the particular locality. Consequently, it is unclear what range of ID is included, although the Lyall, Holland and Collins (1995) study appears to include those in the borderline range. A major limitation of these studies, however, is the identification of contact with the CJS and the identification of offending type behaviour. This information is often gained from service managers or care workers and therefore its validity in terms of actual offending rates is uncertain (Lyall, Holland & Collins, 1995; McBrien et al., 2003; McNulty, et al., 1995; Seaward & Rees, 2001; Tyrer, et al., 2006). Cooper et al. (2009) gathered information from a wide variety of sources in order to tackle this issue, including case-notes, and interviews with participant, carers and family members where possible. However, no evidence of the validity of this procedure was reported. There is also evidence of non-random sampling and other problems with the data sampling procedures. For example, Tyrer et al., (2006) report that only 50% of individuals on the Leicestershire Learning Disability Register have had an interview and that for those that have had interviews there are updates every five to seven years. Consequently, the range of participants that could be included in the study is restricted and the data accessed could be out-dated.

Using a different methodology, Dickson, Emerson and Hatton (2005) conducted a secondary analysis of data relating to self-reported antisocial behaviour, obtained from the 1999 Office for National Statistics (ONS) study of the mental health of children and adolescents in the UK. The authors reported that 16% of the group with ID self-reported being in trouble with the police compared to 8% of the group without ID. This study utilised a large sample size ($n=4174$) of whom ($n=124$), 3% were identified as having ID. However, it is unclear how ID was determined in this analysis and the validity of the self-report is uncertain in relation to the actual commissioning of a crime. Equally, there is evidence of

non-random sampling as information was only available for 79% of the sample with ID and 94% of the sample without ID.

The literature regarding prevalence rates is equivocal (Simpson & Hogg, 2001). Methodological difficulties and procedural differences make it difficult to draw firm conclusions from the available evidence. In an attempt to overcome those methodological problems Hayes, Shackell, Mottram and Lancaster (2007) report on a well controlled study which assessed the prevalence of ID in a large big city local prison. A random sample of 140 prisoners, representing approximately 10% of the prison population, participated in the study. The WAIS III-UK Version (WAIS III (UK), Wechsler, 1999) was used to determine IQ. Vineland Adaptive Behaviour Scales Interview Edition (VABS, Sparrow et al., 1984), completed using self-report, was used to determine the level of adaptive behaviour. On the WAIS-III (UK), 7% of participants achieved standard scores below 70 and a further 24% were in the 70–79 borderline intelligence range. Therefore, a total of 31% showed considerable deficits in intellectual abilities. Some community services include clients with standardized scores of up to 75 on psychometric grounds (Holland, 2004). Consequently, this group was also analysed. Results indicated that 16% of the sample had a score of 75 or less on the WAIS-III. The VABS results indicated that 10% of participants had standard scores below 70 and a further 33% were in the borderline range yielding a total of 43% with serious deficits in adaptive behaviour. In addition, 26% had a score of less than 75. These findings suggest that the number of people with ID in the UK prison is larger than previously determined. However, only four participants (3%) had standard scores below 70 on both the VABS composite standard score and WAIS-III (UK) Full Scale IQ (FSIQ), with thirteen participants (9%) falling below 75 on both instruments. When the borderline groups (<80) on both tests were included, a total of 30 participants (22%) fell below 79 on both assessment instruments. Whilst there are differences in the pattern of distribution between the VABS

and WAIS-III, the two tests correlate significantly ($r=.57$). The mean total scores on the WAIS-III (87.09, SD=12.5) and the VABS (82.51, SD=13.2) were both in the low average range.

This study represents a clear attempt to assess the level of ID in a UK prison with regard to IQ and adaptive functioning. However, even in this reasonably well controlled study the continuing debate over the definition of ID is apparent. Using a strict 70 demarcation, 3% of offenders had ID. However, allowing for the 2 standard errors often utilised in ID services to demarcate admission criteria, 9% scored below 75 on both measures, whereas including the borderline range increased the percentage to 22%. These figures suggest that the number of people with intellectual deficits in UK prisons is higher than previously supposed. However, caution is required in generalising from this study. With only 140 participants, this is a relatively small sample size with which to generalise to the wider prison population. It is also unclear to what extent this sample was representative of the wider prison population; the authors note that black and minority ethnic groups were under-represented. In addition, the sample included those on remand and therefore, strictly speaking, is not entirely composed of offenders as some may not be found guilty. In addition, 39 offenders refused to participate, 18 withdrew during data collection and seven were transferred or released before data collection was completed. It is unclear whether this substantial group, 31.37% of the original 204 identified, had any characteristics in common which might have influenced the results. In addition, the authors did not report on the reliability of the VABS assessments. Without, a reliable assessment of social/adaptive functioning it is possible that researchers will continue to define ID primarily in terms of IQ. However, Talbot and Riley (2007) in a qualitative study noted that some individuals, although aware of their intellectual limitations, had learnt to hide them under a 'cloak of

competence' for fear of discrimination. Consequently, some individuals with ID may refuse to participate in research such as this in order to avoid identification.

It is difficult to draw firm conclusions about the prevalence of people with ID in the CJS. Some differences may be due to sampling at different stages of the CJS, as some authors have suggested (Simpson & Hogg, 2001). However, such a pattern was not obviously apparent in the evidence presented above. This may have been due to the significant methodological differences between studies and particularly differences in the means of determining the level of ID of the sample. Some have sampled during the early stages, such as people detained in police stations (Irving, 1980; Irving & McKenzie, 1989; Lyall, Holland and Collins, 1995) and others at late stages, such as prison populations (Crocker et al., 2007; Hayes et al., 2007; Murphy et al., 1995). Some seek to identify the numbers of people with ID present in offender populations (Hayes et al., 2007) and others seek to identify how many individuals known to ID services have been involved with the CJS (McBrien et al., 2003; McNulty, et al., 1995; McBrien, et al., 2003; Seaward & Rees, 2001; Tyrer, et al., 2006). Many use small sample sizes (Fryson, 2005; Keilty & Connelly, 2001; McBrien & Murphy, 2006) and others large sample sizes (Cooper et al., 2009; Tyrer et al., 2006). Some use abbreviated tests to assess level of ID (Mason & Murphy, 2002), others use screening tools (Lyall, Holland & Collins, 1995; Winter et al., 1997) and others use more robust measures such as the WAIS (Hayes et al., 2007). This is often inversely related to sample size with larger sample sizes often relying upon quicker, less valid measures or the identification of ID by the service being examined (Cooper et al., 2009; Tyrer et al., 2006) and smaller sample sizes utilising more robust measures (Mason & Murphy, 2002). Hayes et al. (2007) may be considered an exception with full WAIS and VABS assessments for a sample of 120 offenders. However, even this sample may be considered small when used as an estimate of the level of ID in the whole prison population, of around 85 000 offenders

(Ministry of Justice, 2010). In addition, few studies incorporate assessment of adaptive functioning in a formal sense, although Hayes et al. (2007) was an exception. However, even in that instance the validity of the VABS was uncertain due to the use of a self-report version, rather than administration to a third-party as designed (Sparrow et al., 1984).

Many researchers and clinicians have expressed opinions regarding the potential over- or under-representation of ID offenders commissioning specific offences (Barron, Hassiotis & Banes, 2004; Hayes 1993; Hayes & Craddock, 1992; Holland, 2004; Lindsay, 2002; Lindsay, O'Brien, Carson, Holland, Taylor, Wheeler., ...Johnston, 2010; Simpson & Hogg, 2001; Walker & McCabe, 1973). Studies conducted in different contexts tend to produce different prevalence rates and different profiles of offences committed (Hogue et al., 2006; Lindsay, Hastings, Griffith & Hayes, 2007). For example, in terms of more recent studies, Dickson et al. (2005) reported that adolescents with ID self-reported higher levels of bullying/threatening others; stealing valuable items from houses/ shops/school; using weapons against others; starting fires; deliberately destroying property; stealing from someone in the street, when compared to those without ID. In contrast, Lindsay, O'Brien, et al. (2010) assessed the offence histories of those in offender services for people with ID and reported that fire setting, theft and road traffic offences do not feature prominently. The most frequent index offences and prior problems related to physical and verbal aggression. Holland and Persson (2010) report that prisoners with an ID were younger than offenders without ID, at the commencement of their current incarceration and at their first term of incarceration. In terms of the most serious offence (MSO) for which prisoners were serving their current term of imprisonment, prisoners with an ID were significantly more likely to have a property offence and significantly less likely to have a drug offence as their MSO than prisoners from the non-ID sample. In terms of those purely in high secure settings, Hogue et al. (2006) reported significantly more convictions for threatening behaviour, assault

occasioning actual bodily harm and criminal damage than those in a community setting and a significantly higher number of other offences than the medium and low security sample. In terms of index offence the high security sample had a significantly higher number of index offences of murder and manslaughter than the other two groups and a higher number of violent offences and weapon use than the community sample. Consequently, it appears that the identification of the types of crimes committed is to some extent determined by the context in which the data collection takes place.

However, evidence suggests that prevalence for sexual offences may be relatively higher in populations of offenders with ID (Murphy et al., 1995; Walker & McCabe, 1973; Winter et al., 1997), although Murphy et al. (1995) and Winter et al. (1997) were studies where the participants were sampled at police stations and consequently the number who went on to be convicted is unknown. Cantor, Blanchard, Robichaud and Christensen (2005), using meta-analysis, reported that when adjustments were made to IQ scores to account for the Flynn effect, that is the increase in population IQ scores over time, sex offenders had significantly lower IQs than non-sexual offenders.

This potential link between IQ and paedophilia has been investigated further (e.g., Blanchard et al., 2007; Cantor et al., 2004). For example, Blanchard et al. (2007), using phallometric assessment, reported that men with a sexual preference for prepubescent males and females had significantly lower IQs than men who had a sexual preference for pubescent males and females, who in turn had significantly lower IQs than men with a sexual preference for adult males and females. While these studies are suggestive of a possible increased preference for prepubescent males and females in ID population, caution is required in interpreting the results. The study did not specifically include men with ID, indeed the group that demonstrated a preference for pre-pubescent males and females had an average IQ in excess of 90. The extension of this preference into the mild and moderate ID range has yet to

be demonstrated. In addition, it is not clear to what extent this study was based on a representative sample. For example, the sample comprised men who had been convicted, charged, been the subject of credible accusations or had self-disclosed criminal sexual behaviour. In addition, an unspecified number of men with no involvement with the CJS had referred themselves due to concerns about their sexual orientation, hyper-sexuality or 'sex addiction,' excessive use of telephone sex lines or massage parlours, clinically obsessive patients with intrusive thoughts about unacceptable sexual behaviour, and patients with paraphilic behaviours like masochism, fetishism, and transvestism. It is unclear how the inclusion of these individuals may have influenced the results. However, this appears to be a line of investigation worth pursuing in samples with ID.

Regardless of the criminal versatility of offenders with ID, it is apparent that they participate in a range of offences (Hodgins, 1992; Hogue et al., 2006; Murphy et al., 1995; Walker & McCabe, 1973; Winter et al., 1997). In addition, it appears that there is a significant over-representation of people within borderline intelligence range represented in the CJS. This group appears to conform with the Prison Reform Trusts No One Knows project, which identified a large number of people (20-30% of people in prison), whose level of intellectual ability significantly impaired their ability to cope with various aspects of the CJS. No One Knows avoids a clear definition of ID in order to consider the needs of all those with thinking and understanding difficulties (Loucks, 2007; Talbot & Riley, 2007). Consequently, it is important that a variety of assessment tools are available to consider the wide range of factors associated with the breadth of offending behaviour in people with ID and those in the borderline intelligence range. The tools currently available will be outlined in Chapter 3. However, the validity of self-report tools used with this population has been questioned due to a range of response biases which are more pronounced in this population. These response biases will be addressed in Chapter 2.

1.3.1 Locus of control and criminal behaviour

Numerous studies have reported a correlation between external locus of control and criminal behaviour (Beck & Ollendick, 1976; Beck-Sander, 1995; Duke & Fenhagen, 1975; Eitzen, 1974; Elenewski, 1975; Martin, 1975; Parrott & Strongman, 1984). For example, Beck-Sander (1995) found extreme external orientation of locus of control in a sample of men convicted of a sexual offence against children and in a sample of men responsible for violent offences against children. While these scores were extremely external ($M=19.46$, $SD=6.1$ and $M=22.00$, $SD=5.8$ respectively) the conclusions drawn can only be tentative due to methodological difficulties. These include a small sample size ($n=36$), how representative the sample was of violent offenders (since several had admitted but not been convicted of such offences), and the lack of a non-offender comparison group.

An externally orientated locus of control is also a consistent finding in studies assessing male and female children and adolescents showing delinquent or problematic behaviour (Beck & Ollendick, 1976; Ducette & Wolk, 1972; Duke & Fenhagen, 1975; Eitzen, 1974; Elenewski, 1975; Martin, 1975; Parrott & Strongman, 1984). For example, Beck and Ollendick (1976) found that delinquent adolescent males were more external in their locus of control orientation than a matched sample of non-delinquents. Duke and Fenhagen (1975), and Elenewski (1974), found similar patterns in adolescent girls. While consistent findings exist across these studies there are a number of methodological problems within them. The primary difficulty is that the majority of them have very small sample sizes, with only the Ducette and Wolk (1972) study having a sample size in excess of 28. In addition, little information was provided regarding the nature of delinquency/problematic behaviour or how the sample had been identified. The exception to the latter point is the study by Duke and Fenhagen (1975), which identified the participants as court referrals to a detention unit. Some studies made modifications to existing measures of locus of control but

did not include details of the modifications or tests of reliability/validity (Eitzen, 1974). It is unclear how this may have influenced his results. There are also questions regarding measures of problematic behaviour. For example, Ducette and Wolk (1972) used a paper and pencil task in order to assess risk taking and persistence. It is unclear to what extent these tasks relate to risk taking and persistence in reality or how they are associated with criminal behaviour.

An externally oriented locus of control is also implicated in violent offending. For example, Wiehe (1987a) found that mothers who had physically abused their children were more likely to possess an external locus of control than non-abusing mothers. However, again, the sample size, 32 mothers on the case load of a child protection agency, is small. In addition, the exact nature of abuse was not specified and no reference was made to whether the behaviour in question amounted to criminal conduct. Despite these studies consistent reports of external locus of control orientation in offending populations the methodological problems outlined limit the conclusions that can be drawn.

Locus of control has been identified as an important predictor of treatment change in offending populations (Bowen & Gilchrist, 2004; Duke & Fenhagen, 1975; Eitzen, 1974; Fisher, Beech & Browne, 1998; Johnson & Berry, 1989; Page & Scalora, 2004; Straub, 1979). Successful treatment has been correlated with a shift from an external to an internal orientation suggesting that an offender owning responsibility for his actions is an important component of treatment change (Fisher et al., 1998). For example, Fisher et al. (1998) compared a sample of successfully treated sexual offenders with a sample of unsuccessfully treated sex offenders. Locus of control orientation was assessed using the Adult Nowicki-Strickland Internal-External scale (ANSIE: Nowicki, 1976). Treatment success was defined as significant overall improvement on a battery of 16 psychometric scales between pre- and post-treatment. Analysis identified a significant change in the successfully treated group

($n=28$) ANSIE score between pre- ($M=13.5$, $SD=6.5$) and post-treatment ($M=10.2$, $SD=5$). However, no significant difference was found on ANSIE scores for the unsuccessfully treated group ($n=24$) between pre- ($M=18.8$, $SD=4.0$) and post-treatment ($M=19.2$, $SD=4.9$). There were also significant differences between ANSIE scores for the successfully treated group and the unsuccessfully treated group at pre- and post-treatment. Thirteen men had an internal ANSIE scores (defined as less than 12) prior to treatment and all of these men appeared in the successfully treated group. In contrast, of 39 men with an external ANSIE scores pre-treatment, only eight men appeared in the successfully treated group. Analysis indicated that locus of control orientation pre-treatment was predictive of treatment outcome. It was also observed that 11 men in the unsuccessfully treated group became more externally orientated by at least half a standard deviation through treatment. However, the results of this study are limited by the relatively small sample size and the definition of successful treatment, as it is not clear to what extent improvement on the psychometric battery is related to actual re-offending.

Regardless, measures of locus of control inform evaluations of treatment progress and success (Beech, Fisher & Becket, 1998; Johnson & Berry, 1989). Page and Scalora (2004) reviewed the literature concerning the relationship between locus of control, treatment participation, help seeking and treatment outcome in juvenile offenders. They suggest that a shift from a more externally oriented locus of control to a more internal orientation following treatment may indicate a positive treatment effect. Conversely, a shift to a more external orientation of locus of control may indicate ineffective treatment.

Most of the research with offenders regarding locus of control refers to mainstream non-ID populations. However, Rose, Jenkins, O'Connor, Jones and Felce (2002) reported on a small case study series which considered the effectiveness of a group intervention for men with ID and a history of sexual offending. They found that locus of control orientation

became *more external* following treatment. They suggest that this outcome may result from the treatment's emphasis on the external consequences of sexual offending behaviour. If Page and Scalora's (2004) conclusion, that a more internal orientation following treatment is indicative of treatment success, can be extended to the population with ID then treatment outcomes such as Rose et al. are of significant concern because it would indicate unsuccessful treatment. Alternatively, the relationship between locus of control orientation and treatment success may be different in offender populations with ID. Langdon and Talbot (2006) in a small-scale study of sexual offenders with ID, found that whilst locus of control did not appear to change during treatment the level of cognitive distortions endorsed did reduce. So despite locus of control orientation remaining external following treatment the endorsement of cognitive distortions decreased suggesting some level of treatment success. This suggests that the relationship between locus of control orientation and treatment success will apply differently in samples with ID. However, it should be noted that the sample size in the Langdon and Talbot study was small, containing just 23 sex offenders with ID. Despite, the Rose et al. and Langdon and Talbot (2006) being limited by their small sample size they suggest that the relationship between locus of control and offending behaviour might differ in ID populations when compared with non-ID counterparts. Clearly, the role of locus of control and its relationship to offending by people with ID requires further examination. Consequently, there is a need for a reliable and valid assessment tool for the assessment of locus of control.

1.3.2. Impulsivity and criminal behaviour

Links between impulsivity and criminal behaviour are well established. For example, in juvenile and adolescent populations impulsivity has been found to predict re-offending, both sexual and non-sexual (Prentky, Lee, Knight & Cerce, 1997; Miner, 2002) and the

development of delinquent behaviour (White, Moffitt, Caspi, Bartusch, Needles & Stouthamer-Loeber, 1994). White, Tice, Loeber & Stouthamer-Loeber (2002) found that aggressive offences committed under the influence of alcohol were more likely to be committed by more impulsive adolescent males. Eysenck and McGurk (1980) administered the Impulsivity Questionnaire (version 5) (Eysenck & Eysenck, 1978) to 641 male offenders, aged between 17 and 21 years of age ($M=18.05$ years, $SD=1.29$). Scores were compared with those obtained in a non-offenders population (mean age =26.41, $SD=10.43$) reported by Eysenck and Eysenck (1978). The impulsivity scores of the delinquents ($M=14.32$, $SD=4.24$) were reported to be significantly higher than those of the non-offenders ($M=10.49$, $SD=5.40$). Whilst self-control in children has been positively associated with adolescent social and intellectual competence, impulsiveness has been associated with aggressive and delinquent behaviour (Mischel, Shoda & Rodriguez, 1989). In addition, evidence suggests a relationship between childhood impulsivity and adult criminality. For example, Babinski, Hartsough and Lambert (1999) found that childhood behaviour ratings of hyperactivity-impulsivity (associated with ADHD) predicted likelihood of an arrest as an adult.

Similar links between impulsivity and criminal behaviour have been identified in adult populations. For example, impulsivity has been found to correlate with three forms of offending in adult child molesters (Prentky et al., 1997), alcohol use and violence (Hamberger & Hastings, 1991) and intimate partner violence (Cunradi, Caetano, Clark & Schafer, 1999; Schafer, Caetano & Cunradi, 2004). In addition, impulsivity has been found to differentiate male and female violent and non-violent parolees from controls (Cherek & Lane, 1999; Cherek, Moeller, Dougherty & Rhoades, 1997) and to predict re-offence risk across domains of criminal behaviour (Prentky & Knight, 1991).

For example, Cherek et al. (1997) measured impulsivity using a computerised behavioural test where a response requiring a longer delay received a greater reward than a

response with a shorter delay. Impulsivity was interpreted as the inability to inhibit ones behaviour and therefore select the quick option, and the inability to tolerate a delay of gratification and therefore a tendency not to choose the long delay option. Consequently, choosing the short delay with small reward option was deemed to be an impulsive response. The results indicated that violent offenders selected significantly more impulsive responses than did non-violent offenders. In addition, violent offenders scored significantly higher than the non-violent offenders on the Barratt Impulsivity Scale 11th edition (BIS-11: Patton, Stanford & Barratt, 1995).

These studies together demonstrate that impulsivity is associated with many forms of offending. However, the conclusions of these studies are limited by various methodological problems. For example, most of the studies reported rely upon small sample sizes and lack a suitable non-offending control group (Cherek et al., 1997; Cherek & Lane, 1999; Cunradi et al., 1999; White et al., 1994; White et al., 2002), or include participants without convictions in offender groups or include samples that are not representative of the wider population of offenders with ID (Cherek et al., 1997; Cunradi et al., 1999; Prentky & Knight, 1991). Those few studies with a large sample size still contained methodological problems. For example, in the Eysenck and McGurk (1980) study it is unclear whether the administration of questionnaires was comparable for all participants. In addition, it is unclear to what extent the samples were matched. Of particular concern is the difference in mean age between the delinquents group and the control group. Eysenck, Pearson, Easting and Allsopp (1985) found evidence that impulsivity scores decrease with age, although they did not test the statistical significance of these differences. It may be that the difference found by Eysenck and McGurk was due to the higher mean age in the control group, although the mean score quoted by Eysenck et al. (1978) for the age group between 20 and 30 years is actually lower than the control group reported here, suggesting that this explanation is unlikely.

Most of the above research presented above relates to main-stream, non-ID samples. However, such findings and their implications for risk assessment and treatment initiatives are frequently extrapolated into populations with ID. Two studies have considered impulsivity in offenders with ID (Parry & Lindsay (2003; Snoyman & Aicken, 2011). Parry and Lindsay (2003) used a modified version of the BIS-11. The authors report that the sex offenders with ID self-reported significantly lower impulsivity ($M=33.32$) than the non-sex offenders ($M=49.32$) with ID. The results of this study are limited by the small sample size of 22 sexual offenders, 13 non-sexual offenders and 6 non-offenders. The authors do not report how they altered the wording of items without changing the content and yet provided a question context familiar to participants for more awkward items, for example, ‘Do you feel awkward at talks or presentations?’ or the abstract nature of items such as ‘Do you have irrelevant thoughts when you are thinking?’ and ‘Are you more interested in the present than dreaming about the future?’. Reliability and validity data of the adapted measure were not reported upon. Barratt (1994) reports means for impulsive aggressive prisoners ($M=69.8$) and matched non-offender adults ($M=49.1$). This suggests that the means obtained by Parry and Lindsay (2003) were not elevated, although Barratt’s work was conducted some time ago and therefore further research would be required to clarify this issue.

Snoyman and Aicken (2011) also used an adapted version of the BIS-11, reporting a mean score of 79.48 ($SD=12.61$). The authors reported a significant difference between sex offenders ($M=73.17$) and violent offenders ($M=83.44$) on the BIS-11 Total Score. A comparison of sex offenders to violent offenders using a multiple analysis of co-variance (MANCOVA), with age and IQ scores as covariates, found a significant difference with large effect size on all the BIS scales except the Attentional subscale. Sex offenders self-reported being significantly less impulsive overall, with more planning and self-control, and less motor impulsiveness than violent offenders. Neither ‘age’ nor ‘IQ scores’ were significant

covariates with any BIS scores. Similar results were found when comparing sex offenders with non-sex offenders, with a significant difference, with medium to high effect size, between BIS Total, Attentional and Non-Planning scores. IQ was a covariate in relation to Non-planning impulsivity.

There was no significant relationship between BIS Total and any of the IQ scales, but there was a significant correlation between the Non-planning subscale and Full Scale IQ ($r=.308$) and for Performance IQ ($r=.255$). Pearson's correlations indicated that participants with lower IQs with violent and non-sexual most serious historical offence (MSHO) tended to be more inattentive, show more cognitive instability, and make faster decisions than participants with violent and non-sexual MSHO with higher IQ scores. This pattern is reversed for sex offenders with low cognitive ability. There is therefore a complex interaction between the various subscales of the BIS and different types of offending.

It is also apparent that there is some confusion regarding the most recent version and corresponding normative data for the BIS-11, with two versions of the BIS-11 in circulation, which Snoyman and Aicken (2011) labelled 'a' and 'b'. While BIS 11a is the more commonly used 30 item instrument, Parry and Lindsay (2003) used BIS 11b, a 34 item questionnaire. The scoring system used also differed from 1 to 4 and 0 to 3 respectively. Additionally, in modifying the BIS-11b for people with intellectual disabilities, Parry and Lindsay (2003) changed the items of the BIS 11b from statements to a question format, and Snoyman and Aicken expanded some of the original statements. Snoyman and Aicken's reported good internal reliability but their adaptation of the BIS-11 produced statements that were much longer than the original statements. For example, Item 4 "I have racing thoughts" became "I have racing thoughts (I have lots of things in my head at the same time)". Whether this technique promotes reliable responding requires further examination.

Unfortunately the authors did not report test re-test reliability for their modified version of the BIS-11, nor did they appear to consider the presence of any response bias in their data.

Parry and Lindsay (2003) used a sample from the mild to borderline range of ability, whereas Snoyman and Aicken (2011) apparently incorporated participants from the moderate to borderline range, including participants with IQs as low as 40. It is unclear how this may have affected the results, however, evidence suggests that people with lower IQ are more prone to response bias than those with higher IQs, even in the ID range (Budd, Sigelman & Sigelman, 1981; Sigelman, Budd, Spanhel & Schoenrock, 1981a; Sigelman, Schoenrock, Spanhel, Hromas, Winer, Budd, & Martin, 1980), suggesting that Snoyman and Aicken may have more error in their data. Another problem with the Snoyman and Aicken study is the use of MSHO which was used to determine offence category. In the Parry and Lindsay (2003) study the identification of sexual offender or abuser was based upon a current referral. Snoyman and Aicken used the MSHO in order to increase their sample size by allowing those on remand to be included in the study. It is unclear how this procedure may have affected the results but it may have blurred the boundaries between categories, depending upon the criminal versatility of the participant involved. In addition, Snoyman and Aicken did not utilise a control group of non-offenders with ID and consequently it is unclear whether the level of impulsivity is elevated in this population. While Parry and Lindsay found no significant difference in self-report impulsivity between offenders and non-offenders with ID, the small sample of non-offenders ($n=6$), means that this must be viewed with some caution. In addition, it would appear that there are one or two problems associated with the BIS-11, specifically the ease with which it can be simplified for use with a sample with ID. It would be useful to examine the level of impulsivity in a sample with ID, using a reliable and valid tool.

Clearly, there is little evidence regarding the role of impulsivity in offenders with ID. However, the two studies that do exist suggest that self-report impulsivity varies between offence types. The methodological difficulties associated with these two studies means that it is difficult to determine whether self report impulsivity is elevated in offender samples with ID as it is in offender samples without ID (Cherek & Lane, 1999; Cherek et al., 1997; Cunradi et al., 1999; Hamberger & Hastings, 1991; Prentky & Knight, 1991; Schafer et al., 2004). A study comparing self-report impulsivity in offenders with ID and non-offenders with ID is reported in Chapter 7.

1.4. Assessment in ID function

The psychological assessment of anti-social behaviour in a population of people with ID presents significant challenges for professionals, not least because of the paucity of research with this population. The purpose of an assessment is to understand the behaviour in question and the context in which it occurs in order to minimise the risk of re-occurrence (Clare & Murphy, 1998). A good assessment requires good data collection. It is generally accepted that no single method of data collection is without flaw. Consequently, the use of multiple methods such as interviews, psychometrics, official records and observations from multiple sources including the client, significant others and relevant professionals should be included in any comprehensive assessment (Ireland, 2004). Self-report is a particularly important component of assessment because only an individual can report on their own internal state. In addition, with the majority of offender interventions based upon changing cognitions there is a necessity to identify need and change in this area (Beech et al., 1998). Frequently in forensic work the use of multiple methods and sources of information result in inconsistencies in the content of the data collected. Typically such inconsistencies may be due to denial of culpability or the deflection of responsibility on the part of the person under

assessment (Gudjonsson & Singh, 1989). However, other sources of inconsistent responding may result from response biases, such as acquiescence (Gerjuoy & Winters, 1966; March, 1992; Rosen, Floor & Zisfein, 1974; Sigelman, Budd, Spanhel & Schoenrock, 1981a; Sigelman et al., 1980), recency (March, 1992; Sigelman & Budd, 1986; Sigelman, Budd, Spanhel & Schoenrock, 1981b), nay-saying (Budd et al., 1981; Voelker, Shore, Brown-Moore, Hill, Miller, & Perry, 1990) and suggestibility (Clare & Gudjonsson, 1993; Everington & Fulero, 1999; Tully and Cahill, 1984), apparent in the questioning of individuals with learning disabilities. Such biases are likely to be more problematic in the psychometric assessment of individuals where tools rely upon specific response formats susceptible to such response biases. Evidence regarding response biases will be presented in Chapter 2 of this thesis. In addition, consideration will be given to how such response biases can be overcome. There is clearly a need to assess self-report offence related cognitions in offenders with ID. Studies reporting on the reliability and validity of an adapted version of the ANSIE and the Impulsivity Questionnaire (version seven) (I7i; Eysenck et al., 1985) will be reported in Chapter 4 and Chapter 5 respectively.

1.4.1. Psychometric assessments of ID

At the start of this chapter a definition was provided of ID. However, inspection of the literature relating to samples with ID suggests that this definition is rarely adhered to. The level of intellectual functioning is determined through the administration of a reliable, valid and properly standardised psychometric assessment. The WAIS instruments are considered the *gold standard* because of their high degree of reliability and validity (BPS, 2001). However, different studies rely upon different means of determining IQ. Studies assessing the prevalence of personality disorders in samples with ID, described in Chapter 6, highlight this issue. The measures used include: Terman-Merrill IQ (Craft, 1959); the

Disability Assessment Schedule – Version II (DAS-II; Holmes, Shah & Wing, 1982; Khan, Cowan & Roy, 1997; Lidher, Martin, Jayaprakash & Roy, 2005), Vineland Social Maturity Scale (Survey Form) (Sparrow et al., 1984) and C21 Health Check (Cooper et al., 2007), multidisciplinary decision (Eaton & Menolascino, 1982), case-notes with no specified origin (Day 1985 & 1994; Ballinger & Reid, 1987) and not specified in the study (Jacobson, 1990; Reiss, 1990; Alexander, Crouch, Halstead & Piachaud 2006; Bouras & Drummond, 1992; Deb & Hunter, 1991; Flynn, Matthews & Hollins, 2002; Goldberg, Gitta & Puddephatt, 1995; Alexander, Hiremath, Chester, Green, Gunaratna & Hoare, 2011). Indeed, in this whole area only Lindsay et al. (2009) report the use of the *gold-standard* WAIS.

Some studies have included people with IQs over 69 in samples described as having ID (Deb & Hunter, 1991). This may be based upon sound psychometric grounds (BPS, 2001), for example, incorporating a discrepancy (IQ<70 plus or minus 2 standard errors) equating to an IQ of 74 or below (Ballinger & Reid, 1987; Deb & Hunter, 1991; Flynn et al, 2002). In addition, some services providers assess the overall needs of an individual and will provide services for individuals, particularly those with other developmental disorders, and not adhere to a strict cut-off of IQ<70 (Holland, 2004). Consequently, researchers accessing samples from such services often include those in the borderline and low average range (Alexander et al., 2006; Alexander, Piachaud, Odebiyi & Gangadharan, 2002; Crossland, Burns, Leach & Quinn, 2005; Day, 1985, 1994; Eaton & Menolascino, 1982; Hogue et al., 2006).

There are also significant problems in the definition of adaptive or social functioning. Whilst there are a variety of tools purporting to measure such behaviour there is no single measure with demonstrably superior validity (BPS, 2001). Additionally, in practical terms many of the most established measures are difficult to conduct, particularly for those in secure hospitals or in criminal justice establishments (Murphy & Clare, 1991). The vast

majority of studies purporting to cover populations with ID make no reference whatsoever to the establishment of the level of social functioning (Alexander et al., 2011; Craft et al., 1959; Eaton & Menolascino, 1982; Jacobson, 1990; Reiss, 1990). A few studies do make reference to the level of social functioning of participants. However, there are exceptions; for example Mason and Murphy (2002) and Hayes et al. (2007) report WAIS-R Short-Form (WAIS-R; Wechsler, 1981) and WAIS III (Wechsler, 1998) respectively, as well as the Vineland Adaptive Behaviour Scale (VABS; Sparrow et al., 1984) scores, and define ID as individuals who scored two standard deviations below the mean on *both* assessments.

Clearly, this is a major difficulty for researchers working with people with ID. With no standardised, validated measure of social or adaptive functioning there are little means of truly comparing the populations included in the different studies other than simply using IQ. Presumably, however, professionals working within the various services included in these studies have assessed the specific needs of each individuals referred to their service and decided that those identified needs are best met within the learning disability service in question. As a result the comparison of populations based solely on level of IQ is flawed because it does not allow a true comparison of the individuals or their needs in a holistic sense. Whilst, this is problematic for researchers seeking reliable and valid comparisons, this situation probably matches the nature of ID services in reality (Ballinger & Reid, 1987; Deb & Hunter, 1991; Flynn et al., 2002).

The difficulties outlined above are likely to result in the identification of markedly differing populations. For example, examination of the studies reported in this thesis, indicates that some studies include participants in the severe ID ranges (Sigelman, Budd, Spanhel, & Schoenrock, 1981a, 1981b), whereas others deliberately exclude participants in the moderate and severe range, citing the lack validity of self-report associated with people at the lower end of the ID range (Rojahn, Warren, & Ohringer, 1994). Other studies include

samples predominantly drawn from the mild ID and borderline intelligence ranges (Glenn, Bihm, & Lammers, 2003). Consequently, research samples including individuals defined as having ID could potentially be drawn from the range of IQ 20-84, if the borderline IQ range and the allowance of plus or minus two standard errors is included. However, the ability of individuals, in the severe and profound ID range to participate in self-report research is likely to be limited by the significant communication difficulties associated with that level of functioning (Beirne-Smith et al., 2006). Even if the range is limited to moderate and mild ID and includes those with borderline intelligence the range spans IQ 35-84¹, if the 2 standard errors are included. Consequently, such a wide range of IQ scores is likely to encapsulate a wide range of abilities, and conclusions must therefore make specific reference to the sample included in the study. Probably just as important are the characteristics of the control group. If the control group is drawn from the mainstream non-ID population, with a mean IQ of 100, then it is possible to assess differences in samples with ID and attribute them to differences in intellectual functioning. However, many studies do not include a control group (Bramston & Fogarty, 2000; Budd et al, 1981; Sigelman, Budd, Spanhel & Schoenrock, 1981a, 1981b; Sigelman, Budd, et al., 1982; Sigelman et al., 1980), or include a control group but assume that it represents the wider non-ID population without verification the level of intellectual functioning of the people in that group (Glenn et al., 2003). In addition to these difficulties regarding intellectual functioning, it is likely that individuals with IQs ranging from severe ID to those in the borderline range will display a range of socially adaptive behaviour. However, in research terms this is essentially a dichotomous variable; present or not present (although Hayes et al., 2007, is an exception). The relative level of intellectual functioning identified is not mediated by the assessment of social functioning, in research contexts. However, in clinical contexts the presentation or absence of adaptive behaviour is likely to

¹ In practice the extension of the ID boundary through the incorporation of two standard errors on the WAIS III, to take account of standard error in the instrument, is only used at the higher boundary of the Mild ID and Borderline IQ ranges.

influence the identification of treatment goals and to inform risk assessment and management. The reliability and validity of the measurement of ID is therefore brought into question due to the lack of a valid assessment of adaptive functioning. In no study reported in this thesis was an individual excluded because their IQ was within the appropriate parameters for IQ but their social functioning was considered to be adequate.

These issues are likely to result in the identification of populations with markedly differing characteristics. Indeed, the extent to which any taxonomy of broad categories can truly capture the complexity and heterogeneity of those within it is questionable. Consequently, the extent to which findings can be compared between samples and the applicability of any particular research finding to a generic ID population is a moot point. However, many authors highlight the continuum of deficits associated with ID that do not change dramatically at IQ 70. They suggest that services should be designed to meet the needs of the individual regardless of the specific IQ and argue that services should be available to those in the borderline range where required (Beirne-Smith et al., 2006). In addition, the No One Knows project deliberately avoids a specific definition of ID in order to include all who may have difficulties with thinking and understanding. Consequently, it seems appropriate for researchers to include those in the borderline range where possible but provide details of IQ (mean and range) so that the equivalence of samples can be ascertained. The accuracy of assessment, in this instance of offence related variables, related to the reliability and validity of those assessments. Reliability and validity are inextricably linked to the sample in which assessments are being made. A tool that is reliable and valid for use with one sample is not necessarily reliable and valid with another. Consequently, the next two sections of this chapter will consider the reliability and validity of assessment and the samples used in this thesis.

1.5. Reliability and validity of measurement

Reliability concerns the extent to which an experiment, test or any measuring procedure yields the same results on repeated trials (Carmines & Zeller, 1979). Therefore reliability is a measurement of the tendency to be consistent from measurement to measurement. There are four basic methods for estimating the reliability of empirical measurements (Carmines & Zeller, 1979): the retest method; the alternative-form method; the split-halves method; and the internal consistency method.

1.5.1. Retest method

Perhaps the easiest way to estimate the reliability of empirical measurements is when the same test is given to the same people after a period of time. The consistency of results is determined by obtaining the correlation between the scores on the two administrations. The correlation of the scores is equal to the reliability coefficient. Identical scores on both administrations would lead to a retest reliability coefficient of 1.0, but invariably measurements across time will be less than perfect (Carmines & Zeller, 1979). There are problems associated with the retest method. For example, scores may vary because of changes in opinion or beliefs about the phenomena being measured and not because of reliability of the method of measurement. Another problem with the retest method is reactivity, which refers to sensitisation to the phenomena being measured solely as a result of the first measurement. Another problem that can lead to the over-estimation of reliability is memory. Memory of responses at the first administration is likely to influence responses at subsequent administrations. Consequently, the time delay between administrations is likely to mediate the effect of memory and subsequent administrations. Dancey and Reidy (2002) provide guidelines for interpreting the size of correlations, which can be seen in Table 1.2. Dancey and Reidy's guidelines have been used throughout this thesis.

Table 1.2: Guidelines for Interpreting the Size of Correlations (Dancy & Reidy, 2002)

Correlation Co-efficient	Size of Relationship
0.1-0.39	Weak
0.4-0.69	Moderate
0.7-1.00	Strong

It is also worth noting at this point that Pearson's correlation coefficient r has also been used throughout this thesis as a measure of effect size, as recommended by Field (2009). When used as a measure of effect size, Cohen (1988, 1992) has provided guidelines for the interpretation of the size Pearson's r , which can be seen in Table 4.3 in chapter 4.

1.5.2. Alternative-Forms method

This method is very similar to the retest method except that at the second administration an alternative form of the same test is administered. Where an alternative form does not exist, this can be done by randomly administering items from a single measure across two administrations. This should ensure that the two measures are not systematically different from each other and introduce non-random error. The correlation between scores on the alternative forms provides an estimate of reliability. This method removes the impact of memory from subsequent administration but is not able to differentiate true change from unreliability of the measure.

1.5.3. Split-Halves method

The split-halves method can be conducted on one administration with the same group of people. Here the total set of items is divided into halves and the scores on the halves are correlated to obtain an estimate of reliability. However, the correlation gained does not give a true estimate of the overall measures but an estimate of the two halves. Consequently, a statistical adjustment is required to estimate the reliability of the full measure. In addition,

there is uncertainty about how the measurement should be split in half. For a 10-item scale there are 125 different possible splits (Carmines & Zeller, 1979). Each split will probably lead to different correlations between the two halves which, in turn, will lead to different reliability estimates.

1.5.4. Internal consistency method

The internal consistency method is a means of estimating reliability which does not require the splitting or repeating of items. The most popular measure of internal consistency is given by Cronbach's alpha (α). Cronbach's α can be regarded as an extension of the split-halves method in that it is equal to the average values of the alpha coefficients, calculated from the inter-item correlations, obtained for all possible combinations of items split into two half-tests. A potential difficulty with Cronbach's alpha is that simply adding more items with the same inter-item correlations increases the reliability of the measure. Consequently, a 2-item scale with an average inter-item correlation of .2 has an α of .33. However, a 10-item scale with the same average inter-item correlation has an α of .71. Kline (1999) argues that when dealing with psychological constructs an acceptable level of Chronbach's alpha is .70. Consequently, throughout this thesis this will be used as the level of acceptability.

1.5.5. Validity of assessments

Validity is defined as the extent to which any measuring instrument measures what it is intended to measure (Carmines & Zeller, 1979). However, strictly speaking it is actually an interpretation of data arising from a specified procedure (Cronbach, 1971). This is because an instrument can be valid for measuring one kind of phenomenon but invalid for measuring another phenomenon. Consequently, one measures the instrument in relation to the purpose for which it is being used. There are three basic types of validity which can be used to assess

the extent to which an instrument measures what it purports to measure (Carmines & Zeller, 1979): criterion-related validity; content validity; and construct validity.

1.5.6. Criterion-related validity

Criterion related validity is sometimes referred to as predictive validity. Criterion related validity refers to the extent that an instrument estimates a form of behaviour that is external to the measuring instrument itself (Carmines & Zeller, 1979). Consequently, criterion related validity for an assessment purported to measure anger, would be that it differentiates those who express anger inappropriately from those who do not. The means of expressing the degree of correspondence between measurement and criterion is usually expressed through the size of their correlation. A difficulty with criterion related validity in social sciences is that not all psychological phenomena have an obvious criterion. For example, it is difficult to determine a direct criterion for self-esteem. Indeed the more abstract the phenomena the more problematic identifying a criterion becomes.

1.5.7. Content validity

Content validity refers to the extent that an empirical measure represents a specific domain of content (Carmines & Zeller, 1979). Consequently, a content-valid measure of Novaco's (2003) concept of anger should include cognitive, arousal and behavioural components. Therefore, a researcher must be able to specify the full domain of content, be able to select a representative set of items to represent the full content domain, since including everything from that domain is likely to be unfeasible, and they must format a means of gathering relevant information e.g. questionnaire. Consequently, in terms of questionnaire construction items must be constructed that reflect the meaning of each component of a domain. Problems with the establishment of content validity in social sciences relate to difficulty defining the

content domain of abstract concepts which, different researcher may disagree on. In addition, it is often impossible to sample content directly. Consequently, items are developed that *reflect* the content of a given theoretical concept. There is no consensus regarding a criterion for determining the extent to which a measure has attained content validity (Carmines & Zeller, 1979).

1.5.8. Construct validity

Construct validity is concerned with the extent to which a particular measure relates to other measures consistent with theoretically derived hypotheses concerning the concepts, or constructs, that are being measured (Carmines & Zeller, 1979). Clearly, this form of validity is theory dependent because it is impossible to assess construct validity without the ability to generate theoretical predictions, which directly lead to empirical tests. One of the problems associated with construct validity is that it is not established by confirmation of a single, or indeed several, predictions. Construct validity requires a pattern of consistent findings, reported by different researchers, using different theoretical structures across a number of different studies (Carmines & Zeller, 1979).

1.5.9. The measurement of validity in samples with ID

In many of the early studies reporting on the validity of self-report in populations with ID used criterion related validity. The criterion most commonly employed was the extent to which self-report matched that reported by an informant who knows the person well (Voelker et al., 1990; Lovett & Harris, 1987; Sigelman & Budd, 1986; Sigelman, Budd, et al., 1982; Sigelman, Budd, Spanhel, & Scheonrock., 1981a, 1981b; Sigelman et al., 1980; Sigelman, Scheonrock, et al., 1981). Using this technique the validity of self-report in ID populations has been demonstrated (Lovett & Harris, 1987; Voelker et al., 1990; Kabzems, 1985).

However, discrepancies between self-reports and informant observations have also been found. For example, informants appear more likely to rate individuals with ID as more problematic or angrier than the individuals themselves (Voelker et al., 1990; Benson & Ivins, 1992; Lally, 1993). Lewis and Morrissey (2010) reported no significant relationships between corresponding scales (thought disorder, depression and low self-esteem) of the Emotional Problems Scales (EPS) (Prout & Strohmer, 1991) Self Report Inventory (EPS-SRI) and an informant Behaviour Rating Scale (EPS-BRS), apart from the EPS-SRI and EPS-BRS scales measuring anxiety ($r = .38$).

Different authors have reported various explanations for such findings. Voelker et al. (1990) suggest that the discrepancies were due to a socially desirable response bias expressed by individuals with ID. Benson and Ivins (1992) suggest that fatigue or denial in the individual with ID may have been responsible. Lally (1993) however, argues that staff members were equally likely to have over-estimated the difficulties they encounter because these behaviours may induce stress which may bias their judgement.

The extent to which discrepancies invalidate responses is a moot point. It may be argued that individuals are providing more accurate insights into their attitudes and beliefs which cannot be accessed otherwise (Gollay, Freedman, Wyngeerden, & Kurtz, 1978). In addition, the high turnover of staff in residential settings may actually reduce the validity of informant reports (Crocker, 1989). More recently however, studies have begun to provide evidence of construct validity negating the need for informant ratings and providing construct and criterion related validity that will be reported later in chapter 2 (Alder & Lindsay, 2007; Bramston & Fogarty, 2000; Dagnan & Sandhu, 1999; Glenn et al., 2003; Kazdin, Matson & Senatore, 1983; Kellett, Beail, Newman, & Frankish, 2003; Lindsay et al., 2009; Lindsay & Skene, 2007; Masi, Brovedani, Mucci & Favilla, 2002; Mindham & Espie, 2003; Nezu et al., 1995; Rojahn et al., 1994).

1.6. Samples described in the thesis

1.6.1. Institutionalised offender sample: Group 1

The experimental group for this research comprised 47 adult males detained in a regional medium secure unit for men with ID or another developmental disability who have committed an offence or displayed behaviour indicating that there was a serious risk of an offence being committed. The unit's catchment area included Norfolk, Suffolk, Bedford and Cambridgeshire. In addition, some residents were housed from out-of-area, usually when beds were not available in the home area. The unit is divided into an eight-bed admissions ward, 2x 8-bedded progression wards and a four-bed pre-discharge bungalow. The data was collected between January 2001 and March 2008 and all the men who were residents during that time were considered for inclusion in this research. In total, 55 men were considered as possible participants during this period. One man was excluded from the study because during his stay on the unit it became apparent that his cognitive deficits and behaviour were largely influenced by a frontal lobe injury acquired during his adolescence. This man was transferred to an appropriate brain injury facility. Three men with high IQs and a diagnosis of Asperger's syndrome were excluded from the study because their characteristics and responses suggested that they differed from the rest of the population. One man was excluded from the study because he was non-verbal and at the time of data collection was only willing to communicate on subjects associated with meeting his needs. Three men refused to take part in the research. All three of these men were residents on the unit for a three month assessment period pending court disposals. All three refused to participate in that assessment process as well as the research. One man refused to participate in the research for an 18-month period after admission to the unit during which time he remained unsettled. Following medication changes and a period of engagement in psychological therapy he approached the researcher and asked to participate. Several men, who had agreed

to participate in the research, refused to complete particular components usually because they were angry about an incident that had occurred on the unit. In most instances these men approached the researcher within a few days asking to complete the task. On occasions when the individuals had not approached the researcher, they were approached by the researcher and asked if they wanted to continue to participate. In all instances, once the problem causing the anger had been resolved the individuals were eager to re-engage with the research process. Four men in the population had IQs in the low-average range and displayed some more significant characteristics associated with the autistic spectrum but without a formal diagnosis of Asperger's syndrome. Whilst having the four highest IQs in the population their general level of functioning was not obviously different from the majority of residents and therefore they were included in the study. Consequently, 47 men agreed to participate in this study following the process of gaining informed consent.

These men were detained under the Mental Health Act (Mental Health Act,1983) and had committed a range of offences. Five men were detained with an index offence for a sexual offence against an adult, seven for sexual offences against male children, six for sexual offences against female children, nine for arson, eleven for violent offences, five for acquisitive offences (with histories of persistent offending) and two for violence resulting in death. Two men were detained without an index offence. One of these men admitted to committing numerous acquisitive crimes during a period of vagrancy associated with poor mental health to fund alcohol, drugs and gambling. He had come to the clinic some years before as a pre-court diversion. The second man had apparently come to the clinic for a period of assessment following a request for respite for his mother. During the assessment behaviours of concern came to light regarding control and violence directed towards his mother. This man displayed extremely high levels of violence within the clinic and remained under section without a formal court disposal during the data collection.

Before individuals were approached discussions were held with staff teams to explain the nature and purpose of the research. Staff information sheets were given out to all staff and question and answer sessions were held. The chief researcher was based on the unit and regularly discussed research progress in staff meetings and with individuals on a formal and informal basis. Specific discussions were held with the named nurses of potential participants to explain the research, answer questions and to gain their support for the informed consent process. All named nurses received the staff information sheet so that they could discuss relevant issues with their clients and answer questions on a day-to-day basis. All potential participants were interviewed in the presence of their named nurses in a private but informal lounge setting. The nature of the research was explained, including the time commitment involved and the nature of the tasks that each individual would be required to complete. Depending upon the needs of the individual the amount of information given at any one time varied. Some individuals were able to understand and discuss the research in its entirety whereas with others the information related solely to the next task. However, over the course of the whole data collection process essentially the same information was given to all participants. It was also explained to some individuals, where necessary that some of the information required was held in their existing files and that the consent process would allow the researcher to use that information rather than complete the tasks again. The participant was given an easy read information sheet (See Appendix 1) to take away with them. The opportunity was given to ask questions. Checks were made that the person understood the instructions by asking questions throughout this process. They were instructed that they didn't need to make a decision immediately about whether they wanted to participate or not but could go away and think about it for a few days. It was also made clear to clients that they did not have to participate and that there would be no negative consequences or sanctions imposed if they didn't want to take part. It was also stressed that even if they

consented to participate in the research they could withdraw that consent at any time, even after data collection had started and they were in the middle of completing a questionnaire or assessment. Appointments were arranged for at least seven days in the future and potential participants were encouraged to think about whether they wanted to participate or not, to discuss it with their named nurse if they wanted or to speak to the chief researcher as most of them saw him on the unit on a daily basis.

After a period of at least seven but no more than ten days potential participants were re-interviewed. Checks were made on their understanding of the research and what was required of them. In all instances potential participants were able to give at least a rudimentary explanation of what the research is about (“It’s about the control I have in my life”) and the task involved (“You’re going to ask me some questions about it”). Again the opportunity was given to ask questions. The participant was then asked if they were willing to participate in the research. It was again stated that even after the individual had signed to consent to participate in the research that they could withdraw that consent at any time without any negative consequences or sanctions. Individuals willing to participate were asked to sign an east-read consent form which was first read through to them.

Table 1.3: Mean Age, IQ and Length of Time as Resident on the Unit (*n*=47)

	Mean	SD	Range
Age	31.5 years	12 years	18.5 – 62.3
IQ	70.55	8.03	51 - 84
Length of time detained on the unit	7.8 years	8.2 years	0 – 30.3 years

The mean age, IQ and length of time on the unit at the commencement of data collection are show in table 1.1. It should be noted that this sample includes those up to an IQ of 84, which falls in the low average range. This is to take account of the error

measurement inherent in the WAIS and allows leeway of two standard errors. The data for the age of the population was somewhat skewed by a cohort of older offenders with index offences displaying high levels of sexual violence directed at children. In fact, over 50% of the population was under 25 years of age and over 70% of the population was under 30 years of age. Equally, the length of time individuals had been detained under section was also skewed by this same cohort. For example 5 men had been detained under section for between 25 and 30 years respectively. The sample was predominantly white with 44 white English and 3 participants described themselves as black English.

It should be noted that the average IQ of the sample included in this study was high for a sample with ID, at 70.55. This was because 25 men in the sample had an IQ in excess of 70. Of these 10 had IQs between 70 and 75 and were therefore within two standard errors of the formal ID range. A further 11 men had IQs between 75 and 80, and therefore fell in the borderline ID range. Finally, 4 men had IQs between 80 and 84 and therefore fell within two standard errors of the borderline cut-off. All of these men had been admitted to a specialist forensic service for men with ID, which included the admission criteria of an IQ in the ID or borderline range. Consequently, all of these men were deemed suitable to be approached to participate in this study.

No participants displayed active symptoms of mental illness during the data collection process.

1.6.2. Institutionalised non-offender sample: Group 2

The institutional type sample comprised 46 adult males, who were residents in institutional or institutional type accommodation. Thirty-one of the men were residents in a large hospital for people with ID before it closed down. Fifteen were residents in community homes for people with ID. For these fifteen individuals, the community homes were identified, by

psychologists working in the local NHS Community Service for people with ID, as predominantly the same type of environment as the hospital for people with ID. These community homes tended to be large and operated in a structured way; providing all meals, laundry and structured activities during the day and evenings. In addition, locked door policies were in operation to prevent members of the public entering without supervision. The present author approached individual community home managers with the intention of recruiting participants for the research. The research was described to staff working in the home and requests made for them to approach men who they thought would be willing and able to participate in the study. Inclusion and exclusion criteria included: no active mental illness; no previous offending history; aged between 18 and 65 years of age. Once identified appointments were made with the chief researcher and the same informed consent process was conducted as with the detained forensic sample.

1.6.3. Non-institutionalised non-offender sample: Group 3

The community sample comprised 46 men living in the community. Twenty-seven were living independently or semi-independently in flats or supported-living flats or shared houses in Norfolk and nineteen were living with their parents. The majority of these men were recruited through the Football in the Community Project run through Norwich City Football Club, providing regular football tournaments for people with ID or through a local Wednesday Club football team. A further fourteen were recruited through (Norwich) City College which provided a variety of educational and vocational courses for people with ID. All football teams were attached to specific organisations and therefore contact was initially made to these organisations with requests to approach the men individually. No organisation refused permission to approach the men with whom they were working. The men were

approached individually and the same informed consent process was conducted as with the other two groups

1.7. Data Collection

Data was collected between January 2001 and March 2008. At the start of data collection the unit had a static population, most of whom had been residents for some years, with little or no progression towards discharge. For this initial group data was collected by a range of different individuals. Questionnaires such as the ANSIE, M-ANSIE, I-7, I-7R and the BIS were completed by a range of individuals including psychologists, trainee psychologists and nurse therapists. Standardised procedures were followed for their completion and completion by non-qualified staff was always supervised by a qualified psychologist.

The PCL-Rs were completed by 3 doctors and a psychologist. All were appropriately trained on the instrument. Internal reliability for the total scale and two main factor scores was calculated using Cronbach's alpha. The Total ($\alpha=.83$), Factor 1 ($\alpha=.86$) and Factor 2 ($\alpha=.76$) were all acceptable (Kline, 1999). Inter-rater reliability was assessed by the re-rating 19 cases, using the same methodology, and analysed to ascertain intra-class correlation coefficients and was found to be very good ($r=.92$).

The HCR-20s were completed by one of two doctors or by a psychologist. All were appropriately trained on the instrument. Internal reliability for the total scale and three sub-scale scores was calculated using Cronbach's alpha. The Total scale ($\alpha=.87$), Historical scale ($\alpha=.86$), Clinical scale ($\alpha=.84$) and Risk Management scale ($\alpha=.80$) were all acceptable (Kline, 1999). Inter-rater reliability was assessed by re-rating 19 cases, using the same methodology, and analysed to ascertain the intraclass correlation coefficient, which was found to be very good ($r=.90$).

The BADS were all completed by a psychologist. Internal reliability for the total scale and six sub-scale scores was calculated using Cronbach's alpha. The BADS Total ($\alpha=.71$), Rule Shift Cards ($\alpha=.78$), Key Search ($\alpha=.74$), Action Programme ($\alpha=.70$) and Zoo map ($\alpha=.70$) were all acceptable. However, Temporal Judgement ($\alpha=.64$) and Modified Six Elements ($\alpha=.48$) were unacceptable. Many of the men tested on the BADS did not appear to understand the requirements of the Modified Six Elements Task and this probably reflects the poor internal reliability. Due to the nature of the BADS it was not possible to conduct inter-rater reliability assessments.

1.8. Structure of the thesis

In Chapter 2, the literature relating to the expression of the responses biases of acquiescence, nay-saying, recency and suggestibility are critically reviewed and summarised. In Chapter 3, the literature relating to the use of psychometric tools used in the assessment of offenders with ID is critically reviewed and summarised. Chapter 4 presents a study assessing the reliability of an adapted version of the ANSIE (Nowicki, 1976). Chapter 5 presents a study assessing the reliability and validity of an adapted version of the I7i (Eysenck et al, 1985). In Chapter 6 the literature relating to personality disorder and risk assessment in samples with ID is critically reviewed and summarised. This is followed by a study which assessed the ability of the I7i, I7i-R, PCL-R, Factor 1, Factor 2, 13-Item Total and HCR-20 and its subscales to predict institutional violence. Chapter 7 presents a study which assessed whether self-report scores on the M-ANSIE and the I7i-R differentiated offenders with ID from non-offenders with ID. Chapter 8 summarises the thesis conclusions.

1.9. Conclusions

The formal definition of ID requires a significant impairment of intellectual functioning (IQ <70) and a significant impairment of adaptive/social functioning with an age of onset before adulthood (< 18years old). However this was contrasted to the samples used in studies of people with ID, which often included those in the borderline intelligence range (Alexander et al., 2006; Alexander et al., 2002; Crossland et al., 2005; Day 1984, 1985, 1994; Eaton & Menolascino, 1982; Hogue et al., 2006). A review of the literature on the prevalence of people with ID in the CJS, suggests that when those in the borderline intelligence are included they are significantly over-represented. The Prison Reform Trust's *No One Knows* project, deliberately avoided clear definitions of ID, in an attempt to consider the needs of all those who find some activities that involve thinking or understanding difficult. In addition, the Forensic Service of NHS Trust where the research was conducted also provides a service for those in the borderline range. Consequently, the sample for this thesis included those in the borderline range. In addition, the evidence suggested that people with ID commit a wide variety of offences and consequently there is a need for a range of assessment tools to assess risk and need in this population. A critical review of the literature relating to LOC and impulsivity suggests that little is known about these factors in offenders with ID. The reliable and valid assessment of these factors in people with ID requires the development of a tool or the adaptation of an existing tool in order to facilitate reliable and valid responding. The reliability and validity of the new measure must then be established.

CHAPTER TWO: A REVIEW OF RESPONSE BIASES IMPLICATED IN THE PSYCHOMETRIC ASSESSMENT OF PEOPLE WITH AN INTELLECTUAL DISABILITY

2. Introduction

This review seeks to outline the role of response biases implicated in the psychometric assessment of offenders with ID. This is appropriate to this thesis because response biases can reduce the validity of self-report assessment, particularly in people with ID. Chapters four, five, six and seven of this thesis report on studies that utilise the self-report assessment of impulsivity and locus of control. Consequently, any threat to the validity of this self-report requires consideration. The response biases of acquiescence (2.1), negative response bias (2.2) and suggestibility (2.3) are identified and explored. The way assessments are framed (i.e., types of questions, yes/no, open-ended, multiple, choice, pictorial, the use of repetition in assessment, and Likert approaches) are then discussed. Possible solutions to alleviate response biases in the psychometric assessment of offenders with ID are discussed in the chapter.

2.1. Acquiescence

Acquiescence or acquiescence response set is defined as the tendency to respond affirmatively to questions requiring a yes or no response, regardless of their content. Different methods of assessing acquiescence have been developed, including: contradictory paired questions (Budd et al., 1981; Sigelman et al., 1981a; Sigelmam et al., 1981b), *no* questions (Sigelmam et al., 1981b), factual question (Budd et al., 1981; Sigelman, Budd, et al., 1982; Sigelmam et al., 1981b) and ambiguous questions or statements (Gerjuoy, & Winters, 1966; Rosen et al., 1974).

Studies use these techniques to assess acquiescence in slightly different ways. Using contradictory paired questions, the rate of acquiescence is represented by the number of times a participant contradicts themselves by responding yes to both of a contradictory pair of questions, such as: *Are you usually happy?*; *Are you usually sad?* Equally, the acquiescence rate for *no* questions is represented by the number of times a participant answers *yes* to a question when the correct response is obviously *no*. For example, a participant is asked, *Is it raining outside?* when it is clearly not. Using factual questions, such as questions about the participants social activities, such as *Do you play football?*, the acquiescence rate is represented by the number of times a participants yes response does not correspond with observations made by care staff. The acquiescence rate to ambiguous questions or statements is represented by the number of times a participant answers yes to items such as, *Winter is more fun than summer*. Acquiescence rates are normally represented as a percentage of the possible total. The acquiescence rates reported below are referenced to the question type used as it was unclear whether the rates gained using different techniques are equivalent. Acquiescence rates for participants with ID using *contradictory pairs* have been reported at 44% (Sigelman et al., 1981a), 51% (Sigelman et al., 1981b) and 45% (Budd et al., 1981), compared with 4% (Sigelman et al., 1981a) and 8% (Budd et al., 1981) for the opposing bias of answering *no* to both questions. Clare and Gudjonsson (1993) used the Winkler Acquiescence scale (Winkler, Kanouse, & Ware, 1982), a scale containing 24 contradictory question pairs, and reported scores for a group with ID ($M=7.7$, $SD=3.1$) (64%), as significantly higher than the control group of normal intelligence ($M=2.2$, $SD=1.6$) (18%).

Sigelman et al. (1981b) report the acquiescence rate of 48% in relation to *no* questions. Acquiescence rates for *factual questions* have been reported 21% (Sigelman et al., 1981b) and 42% (Budd et al., 1981). Budd et al. (1981) compare this with the opposite form of discrepancy for factual question, where the participants no response does not correspond

with observations made by an informant, which occurred at a rate of 6%. Acquiescence rates to *ambiguous questions* or statements have been reported at 59% (Gerjuoy & Winters, 1966) and 60% (Rosen et al., 1974), showing remarkable similarity. Gerjuoy and Winters (1966) reported that the number of yes responses was significantly higher than the number of no responses, and Rosen et al. (1974) report that the acquiescence rate for people with ID was significantly higher than for a mental age matched control group (48%). Acquiescence leads to a reduction in the validity of responding, producing significantly lower levels of agreement between participants and informants using yes/no question formats (52%) compared with open-question formats (60.1%) (Sigelman, Winer, & Schoenrock, 1982). In this study, the major source of disagreement between participant and informant occurred when participants indicated involvement in a particular sport by responding yes to a yes/no question but the informant indicated that the participant was not involved with that sport which occurred in response to 42% of questions. In contrast, disagreements where a participant indicated that they were not involved in a sport by answering no to a yes/no questions but an informant claimed that they were involved in that sport occurred in response to only 6% of questions. This was a highly significant difference. These findings indicate that acquiescence can lead to significant over-reporting of behaviour.

Within samples with ID, IQ appears to be related to acquiescence (Budd, et al., 1981; Sigelman, et al., 1981b; Sigelman, et al., 1980). Sigelman et al. (1980) report significant negative correlations between IQ and acquiescence for two samples of children with ID ($r_s = -.40$ and $r_s = -.30$) and one sample adults with ID ($r_s = -.31$). Budd et al. (1981) also reported a significant moderate negative correlation ($-.44$) in a sample of children as did Sigelman et al., (1981b) ($r = -.37$). However, acquiescence rates are also related to other variables because within studies the rates of acquiescence vary between questions. For example, for contradictory pairs, response rates for individual items have been reported ranging from 20%

to 83.3% (Sigelman et al., 1981b), from 35.1% to 50% (Budd et al., 1981), from 27.3% to 56% (Sigelman et al., 1981a). Equally, for *no questions*, acquiescence rates ranging from 28% to 73% have been reported (Sigelman et al., 1981b). The lowest acquiescence rates for contradictory pairs were in response to the question, *Do you live in (correct cottage number) right now?* vs. *Do you live in (incorrect cottage number) right now?* Equally, for *no questions*, the lowest acquiescence rate was reported in response to the question, *Right now is it raining outside?* (Sigelman et al., 1981b). The question with the highest acquiescence rate (83%) was in response to a question where the authors report that participants were unlikely to know the distinction: *Here is a picture of (superintendent's name). Is he the superintendent here at the colony?* versus. *Here is a picture of (business manager's name). Is he the superintendent here at the community?* This suggests that simple questions, about concrete subjects where the participant's know the correct answers are likely to lead to low levels of acquiescence. More complex questions where participants may be unable to make the distinction between an incorrect and correct response appear to lead to higher acquiescence rates.

2.2. Negative response bias

Negative response bias (nay-saying) may be regarded as the opposite response bias to acquiescence, namely the tendency to respond "no" to questions (Schuman, & Presser, 1977). The tendency to report negatively is most often reported in response to threatening questions. Threatening questions probe behaviours that are illegal, contra-normative, or generally not discussed in public without tension, or relate to issues of self-preservation (Blair, Sudman, Bradburn, & Stocking, 1977). Closed format threatening questions result in under-reporting and are more sensitive to social desirability factors in mainstream non-ID samples (Sudman, Seymour, Norman, & Bradburn, 1974). However, Blair et al. (1977) reported that long,

open-ended questions, with familiar wording (potentially related to ease of reading) obtained much higher levels of reporting than short, closed, standard questions.

Consistent with these findings, the tendency for ID individuals to negatively acquiesce appears to be stronger than the tendency to acquiesce when questions relate to social taboos (Budd et al., 1981; Voelker et al., 1990). For example, Budd et al., (1981) reported high levels of acquiescence (45%) to contradictory pairs of questions. In contrast, only 8% of participants contradicted themselves by responding 'no' to contradictory pairs. However, when question topic was switched to the rules of the establishment ('Are you allowed to...?' 'Is it against the rules here to...?') the pattern was reversed with significantly more participants contradicting themselves by responding no to both questions (38%) as opposed to those contradicting themselves by responding yes to both questions (17%). The authors report that this pattern of responding was associated with level of IQ in this population ($r=-.24$, $p<.07$). In these instances the behaviours in question were clearly prohibited. Consequently, the "no" response may be interpreted as an expression of disapproval or a desire to present oneself in a socially desirable light by denying any association with taboo subjects or actions (Heal & Sigelman, 1995).

2.3. Suggestibility

In his work with adolescents and adults, Gudjonsson (1983, 2003) has identified and labelled two distinct and relatively independent types of response bias, referred to as Yield and Shift. The former refers to the tendency of interviewees to give in to leading questions, whereas the latter is more related to the ability to cope with interrogative pressure, such as negative feedback and repeated questioning (Gudjonsson, 2003). Gudjonsson developed an instrument, the Gudjonsson Suggestibility Scales (GSS), as a means of identifying people who were particularly susceptible to giving erroneous accounts of events when subjected to

questioning. There are two scales (GSS1, & GSS2) but the formats are identical, containing a narrative paragraph describing an event, followed by a series of 20 questions about that narrative. Of these questions, 15 are describes as leading, in that they ask the respondent to answer questions that introduce erroneous information not contained in the original narrative. The number of times a respondent gives in to leading questions is scored as Yield 1. When these answers have been given the respondent is given “negative feedback” by being told firmly and clearly “You have made a number of errors. It is therefore necessary to go through the questions once more, and this time try to be more accurate”. The questioning process is repeated and the number of times the respondent gives in to leading questions is scored as Yield 2. In addition, the number of occasions the respondent changes answers on the two iterations is scored as Shift. Total suggestibility scores are determined by adding Yield 1 and Shift.

Research indicates that intellectual functioning is correlated with total suggestibility in participants with ID (Clare, & Gudjonsson, 1993; Everington, & Fulero, 1999; Gudjonsson, & Clare, 1995; Henry, & Gudjonsson, 1999; Henry, & Gudjonsson, 2003; Milne, Clare, & Bull, 2002; Tully, & Cahill, 1984). It appears that the two components of suggestibility, Yield and Shift, are both mediated by similar factors (Gudjonsson, 2003). Yield however appears to differentiate between ID and non-ID groups better than Shift (Cardone, & Dent, 1996; Clare, & Gudjonsson, 1993; Gudjonsson, & Clare, 1995; Gudjonsson, & Henry, 2003; Henry, & Gudjonsson, 2003; Milne et al., 2002). For example, Clare and Gudjonsson (1993) reported significant difference between Yield scores but not Shift between a group with ID and a control group without. Gudjonsson and Henry (2003) reported significant increases in all component scores of suggestibility (except Shift for adult participants), from the normal IQ group, through the mild ID group, to the moderate ID

group. These findings indicate that Shift does not appear to vary with IQ, in adult populations with ID.

Henry and Gudjonsson (2003) also reported that Actual Yield (AY) and Actual Shift (AS) were significantly higher in participants with ID than both mental aged matched and chronological age matched control groups, following observation of a staged incident. Actual Yield and Actual Shift were terms used to differentiate between Yield and Shift in relation to questioning regarding an event that was actually observed as opposed to a narrative passage, such as that used in the GSS. A regression analysis indicated that GSS Yield was the strongest predictor of participants' *actual* suggestibility during both interviews. In the immediate interview, GSS Yield accounted for 28% of the variance, and in the delayed interview, this value rose to 45%. GSS Shift did not significantly predict any suggestibility behaviours during either interview. In contrast, Everington and Fulero (1999) reported that Yield and Shift were elevated in adults with ID. However, there is little support for their findings and there are a number of problems with the study. The authors used a modified and abbreviated version of the GSS 1 but did not provide any assessment of its reliability or validity. No assessment of IQ was conducted for either participants with ID or those in the control group, who were assumed to be of average intelligence. In addition, this study used a very small sample size ($n=18$).

Suggestibility in ID populations appears to be primarily due to poor memory (Cardone, & Dent, 1996; Clare, & Gudjonsson, 1993; Everington, & Fulero, 1999; Gudjonsson, 1983; Gudjonsson, 1988; Gudjonsson, 2003; Gudjonsson, & Clare, 1995; Gudjonsson, & Henry, 2003; Henry, & Gudjonsson, 1999; Henry, & Gudjonsson, 2003; Milne et al., 2002; Tully, & Cahill, 1984). For example, in an early study, Gudjonsson (1983) reported highly significant moderate correlations between immediate and delayed recall and Yield and Shift scores. Cardone and Dent (1996) reported that significant

increases in immediate and delayed recall, which resulted from the augmentation of the narrative through the use of photographs, was also associated with a significant reduction in Yield scores. Consequently, it appears that improved memory for an event may lead to a reduction in suggestibility. Gudjonsson and Henry (2003), in the study described above with child and adult participants, reported that there were highly significant differences between mild and moderate ID groups and a control group of normal IQ, on scores for immediate recall on the GSS 2 in the expected direction. An ANOVA analysis, using Immediate Recall as a covariate, reduced the significance of differences between groups for Total Suggestibility and Yield, resulting in the Yield score no longer being significantly different between groups for child participants. The effect of Immediate Recall on Shift was not significant. These results indicate the importance of memory on suggestibility. However, Total Suggestibility remained significantly different across groups of adults and children with and without ID, as did Yield across the adult groups. In addition, Shift was not influenced by Immediate Recall scores in the adult groups. The results indicate that other factors are also involved in the expression of suggestibility.

Beail (2002) has cautioned against the use of the GSS with individuals with ID due to their poor performance on the narrative recall and the effect that this has on subsequent suggestibility scores obtained. Beail argues that recall of the GSS narrative taps into semantic memory, rather than episodic and autobiographical memory that would influence recall of an experienced event. Several studies have assessed the suggestibility in people with ID following more naturalistic situations, where the narrative passage is replaced by a video (Milne et al, 2002), an actual event (Willner, 2008, 2011) or a staged event (Henry & Gudjonsson, 2003). Willner (2008, 2011) reported that recall for actual events was no different to recall for narrative passages. However, recognition, as judged by a choice between an actual and a false component of the passage or event, improved significantly

when questioning related to an experienced event as opposed to a narrative passage. In addition, the level of suggestibility was significantly lower in the condition with the experienced event as opposed to the narrative passage. However, this study did not contain a control group of normal intelligence and therefore it was not possible to determine whether suggestibility remained elevated in the group with ID. In addition, these studies utilised small sample sizes ($n=8$). Those studies that have included a control group have continued to find significantly increased suggestibility in people with ID. Milne et al. (2002) used a naturalistic adaptation of the GSS and reported that participants with ID scored significantly higher than the matched sample of normal IQ on all components of the suggestibility except shift, for which there was no significant difference between groups. This indicates that suggestibility is not solely expressed in response to a narrative passage. In addition, the results confirm the lack of relationship between IQ and Shift. Recall scores for the film had a medium size negative correlation with Total Suggestibility ($r=-.40$), Yield 1 ($r=-.61$) and Yield 2 ($r=-.64$) all of which were highly significant. There was no significant correlation between recall score and Shift.

Several studies have considered the relationship between suggestibility and acquiescence (Clare & Gudjonsson, 1995; Gudjonsson, 1986; Gudjonsson, 1990; Gudjonsson & Clare, 1995). All of these studies used using Winkler's acquiescence scale (Winkler et al., 1982). Clare and Gudjonsson (1995) reported significantly higher scores for Total Suggestibility, Immediate and Delayed Recall, Yield 1, Yield 2 and Acquiescence in a sample with ID when compared to a sample without ID. Gudjonsson (1986) reported weak to moderate positive correlations ($r=.32-.42$) between acquiescence, Total Suggestibility, Yield 1 and Yield 2 respectively. Gudjonsson and Clare (1995) reported that acquiescence had weak positive correlations ($r=.27$ and $.21$) with Yield 1 and Yield 2 and weak to moderate negative correlations ($r=-.37$ and $-.40$) with Immediate and Delayed Recall. In contrast

Gudjonsson (1990) did not find a significant relationship between acquiescence and Total Suggestibility. Consequently, it appears that there is a weak relationship between suggestibility and acquiescence, and that memory function is also implicated. It seems feasible that a study could test memory for specific items and compare suggestibility and acquiescence on those items for which there was memory recall compared for those where there was not.

2.3.1. Critique of suggestibility research

Much of the research on suggestibility conducted with people with ID is limited by the small samples sizes used ($n=8-54$) (Cardone & Dent, 1996; Clare & Gudjonsson, 1993; Everington & Fulero, 1999; Gudjonsson, 1983; Gudjonsson, 1986; Henry & Gudjonsson, 1999; Henry & Gudjonsson, 2007; Milne et al., 2002; White & Willner, 2005; Willner, 2008, 2011). In addition, comparisons between studies are problematic because of differences in the methodologies used and differences in the samples. For example, some studies use samples of children with ID (Henry & Gudjonsson, 1999; Henry & Gudjonsson, 2007), some use adults with ID adults (Cardone & Dent, 1996; Clare & Gudjonsson, 1993; Gudjonsson & Clare, 1995; White & Willner, 2005), some use offenders with ID (Clare & Gudjonsson, 1995; Everington & Fulero, 1999), and other adults without ID (Gudjonsson, 1983). Several studies in which WAIS measures were reported did not report full IQ details (mean and range) (Clare & Gudjonsson, 1993; Cordone & Dent, 1996; Gudjonsson, 1986) and one study used people accessing day services for people with ID without investigating IQ (White & Willner, 2005). In addition, several studies incorporate amendments to the GSS such as: including photographs (Cardone & Dent, 1996); using alternative narrative passages (White & Willner, 2005), using a video-taped incident (Milne et al., 2002), or using staged incidents in familiar settings (Henry & Gudjonsson, 1999; Henry & Gudjonsson, 2006). Nonetheless,

the evidence appears to consistently indicate increased suggestibility in samples with ID. I will now examine the types of approaches used in assessments.

2.4. Types of assessment

2.4.1. Either/or questions

Acquiescent response bias can be overcome by presenting questions that require a choice between two options, such as either/or questions (Sigelman et al., 1981a). Sigelman et al. (1981a) reported no significant difference between the response rates for yes/no questions (68%) and either/or questions (63%). However, there appeared to be an increase in the validity of responding for either/or questions, demonstrated by greater consistency of responses to contradictory pairs with either/or questions (86%) as opposed to yes/no questions (53%). In addition, examination of the inconsistencies demonstrates an apparent reduction in acquiescence. The authors report that 44% of contradictory pairs of yes/no questions were answered inconsistently due to both questions being answered with a *yes* compared to only 4% where contradiction was due to responding to both questions with a *no*. However, either/or questions may also produce a systematic response bias, with people with ID tending to contradict themselves by choosing the second option on both occasions (13%) as opposed to the first (1%) of the two options presented, however, this trend was not significant (Sigelman et al., 1981a). In contrast, March (1992) reported a significant difference between the number of second options chosen (69%) as opposed to first options (31%) when people with ID were questioned about everyday likes and dislikes. Equally, Loper and Reeve (1983) reported that participants with ID were significantly more likely to choose the second option on an either/or scale for the Intellectual Achievement responsibility Questionnaire (Crandall, Katkovsky, & Crandall, 1996).

In a more realistic situation Milne et al. (2001), reported above, using a visual adaptation of the GSS, reported that participant with and without ID were significantly more likely to select the second option on questions offering a choice of two false options, and participants with ID were significantly more likely to answer such questions because they were significantly less like to give resistant responses or don't know responses.

2.4.2. Open-ended questions

The utility of open-ended questions has been demonstrated in many studies involving people with ID (Budd et al., 1981; Lovett & Harris, 1987; Sigelman et al., 1982; Sugg, 1987).

However, evidence indicates that open-ended questions may result in lower response rates than yes/no questions. Sigelman et al. (1982) report that response rates to yes/no questions (86%) were significantly different to the response rates for open-ended questions (55%).

This indicates that in response to open ended questions, people with ID may under-report behaviour or, in response to yes/no questions, possibly due to acquiescent responding, over-report behaviour.

Responses to open-ended questions appear be more valid than responses to yes/no questions (Sigelman et al., 1982; Voelker et al.,1990) and multiple-choice questions (Sigelman et al., 1982). Sigelman et al. (1982) report that the validity of responses to open-ended questions, in terms of agreement with an informant, was significantly higher with open-ended questions (60%) than it was with yes/no questions (52%). This suggests that in response to yes/no questions participants with ID are more likely, possibly through the process of acquiescence, to report behaviour that is untrue. Although, if the lack of correspondence with informants is accepted as a criterion for validity, these figures also suggest that there is threat to the validity of responses to open-ended questions as well but to a lesser degree. Consistent with this interpretation, evidence suggests that the differences in

response rates to yes/no and open-ended questions may not be solely due to over-reporting because of acquiescence. Sigelman et al. (1982) report that in response to 32% of open-ended questions participants failed to mention participation in a sport that informants stated that they were involved in. This suggests that open-ended questions may lead to under-reporting in people with ID however; this tendency was not significantly different to other combinations of disagreement, indicating that the major influence is probably over reporting to yes/no questions. A similar finding was reported in relation to multiple choice questions in the same study, where the response rate to multiple choice questions (94%) was significantly different to the response rate for open-ended questions (74%) but the increase in responsiveness was not associated with an increase in validity (Sigelman et al., 1982).

Several studies also demonstrate the validity of open-ended questions (Lovett & Harris, 1987; Voelker et al., 1990). Lovett and Harris (1987) report that the use of an open-question produced additional information, following the completion of a questionnaire about community living skills. The authors report that the validity was provided by correspondence between self-report responses and responses to a questionnaire completed by significant others. Unfortunately, no statistical analysis was reported but the authors state that both procedures produced the same relative ranking on categories, with the same three categories rated as most important according to both self-report and informant rating. Voelker et al. (1990) reported promising criterion validity for a self-report adaptation of the VABS (Sparrow et al., 1984). The use of open-ended questions with frequent questions to probe responses, produced scores for the Adaptive Behaviour Composite and the three domains contributing to this score (Communication, Daily Living Skills and Socialisation) that showed no significant differences with scores produced by informants, thus providing some validity for this procedure with people with ID. However, self-report and informant scores on the Maladaptive Behaviour Domain were significantly different, with participants

reporting less maladaptive behaviours than informants. This may be due to social desirability effects with participants reluctant to report maladaptive behaviours but this was not investigated. Consistent with this study, Sigelman et al. (1982) reported that open-ended questions with a probe significantly increased response rates to questions about social activities. However, the authors report that the level of agreement with an informant was greater with open-questions alone (77%) rather than with probes (71%) but this difference was not significant. Cardone and Dent (1996) in a study investigating suggestibility, reported that the use of specific questions produced a significantly more complete and accurate recall for a narrative passage.

These studies indicate that people with ID are able to respond to open-questions in a valid way. This is possibly because, in the case of Lovett and Harris (1987) and Voelker et al. (1990) response biases may interfere less because the questions are about concrete, objective, real-life activities that place minimal demands upon cognitive abilities. However, the poor response rate to open questions (Sigelman et al., 1980) suggests that those receptive and expressive communication difficulties associated with ID may limit the applicability of such techniques to all within such populations.

2.4.3. Multiple-choice questions

Individuals with ID may have significant difficulty remembering a choice of possible responses and consequently some display a tendency to repeat the last phrase of the interviewer which will result in the selection of a greater proportion of later possibilities (Sigelman et al., 1982). Sigelman et al. (1982) reported that the response rate to multiple choice questions (94%) was significantly different to open-ended questions (74%). In addition, they reported that the use of pictures to represent response choices on multiple-choice questions increased the response rate to 100%. However, this increase was not

statistically significant, probably due to the generally high level of responding to multiple-choice questions. Despite the significantly higher response rates associated with multiple-choice questions, the validity of the responses, as measured by agreement with informant responses, was not significantly different between groups (open 85%, verbal 86% and pictorial 77%), with the trend suggesting that the increase in response rate with pictorial presentation is countered by a reduction in validity. A possible interpretation of this would be that pointing at a picture makes the answering process easier in practical term, in that a response can be signified simply by pointing at a picture. However, the use of pictures does not aid the more conceptual or abstract decision making processes underlying the decision, so pictures can facilitate participants providing meaningless responses. However, the authors found no evidence of participants selecting the last option in either multiple-choice condition.

2.4.4. Pictorial assessments

The use of line drawings and photographs has been demonstrated to increase the response rate to multiple-choice and either/or questions but not to yes/no questions (March, 1992; Sigelman & Budd, 1986), possibly due to ceiling effects with yes/no questions. For example, for either or questions, March (1992) reported that the response rate for pictorial question (95%) was significantly different to the verbal format (82%). Sigelman and Budd (1986) reported remarkably similar response rates for pictorial (96%) and verbal formats (82%) which was a significant difference. In addition, both studies reported a significant reduction in the selection of the second option response bias. March (1992) reported a significant reduction in second options chosen for preference questions from 69% to 56% when pictures were used, and for factual questions a reduction from 59% to 46% when pictures were used. However, there did not appear to be any increase in the validity of responding, as measured by agreement between participant and informants, regarding the participants' preferences.

However, in response to factual questions, in which an obvious answer existed (e.g., Do you sleep on a table or a bed?), there was a significant difference between the number of correct answers given in the pictorial format (69%) than in the verbal format (46%).

2.4.5. Repetition

The use of repetition can be beneficial or detrimental to the reliability and validity of self-report depending upon the circumstances of its use. There is evidence that familiarisation with the questioning procedure through repetition and correction using the question and answer format can improve reliability and validity of responding (Bromley, Emerson, & Cummings, 1998; Glenn et al., 2003; Lunsky & Benson, 1997; Lunsky & Benson, 2001; Powell, 2003). For example, studies reporting the use of repetitive training and familiarisation components report high levels of internal consistency ($\alpha \leq .85$) for the particular scale utilised in the study (Glenn et al., 2003; Lunsky & Benson, 1997; Lunsky & Benson, 2001; Powell, 2003). In addition, Sigelman et al. (1981a) reported that a repeat interview after a two-week interval increased responsiveness to yes/no questions from 77% to 81%. Unfortunately, the authors did not report on the significance of this difference.

Bromley, Emerson, and Cummings (1998) provide a model for training and familiarisation that has been adopted by others (Glenn et al., 2003; Powell, 2003). Bromley et al. (1998) evaluated participants ability to: (1) distinguish between a big and a small wooden block; (2) distinguish between a big, a medium sized and a small wooden block; (3) place the blocks in ascending order; (4) place the blocks at appropriate points along a visual scale (the pain ruler), depicting increasing and decreasing levels; and (5) to manipulate a marker on the scale to demonstrate where the three blocks should go (in the absence of the three blocks).

Participants were given three chances following the correction of errors. The authors note the importance of repetition and correction at every stage of this process. If the participant still failed the task after three attempts the task was discontinued. Using the pain ruler and blocks, participants were then asked to judge, the levels of mild, moderate or severe pain, depicted by actors in series of three photographs. This training session was repeated 7-14 days later. In addition, participants were prepared for the identification of pain locations by pointing to a specified part of their own body and then pointing to the location on a body-map. Corrections were made and the participant given four opportunities to complete the task. After the fourth failure the task was discontinued. Participants were then shown a series photographs depicting simulated painful situations and were asked to locate the site of pain depicted in the picture on to the body-map and to estimate the level of pain depicted. In the actual study participants were asked to judge the intensity and location of pain depicted in a series of 52 test photographs. The authors report 7-14 day test-retest stability of pain location ($r=.86, p<.001$) and intensity ($r=.74, p<.001$) indicating that the measure has good reliability. However, evidence for the validity of the scale was weaker. Participants with ID rated pain as more intense for 100% of mild stimuli and 36% of severe stimuli. Overall, judgements of pain intensity were logically consistent on only 65% of trials for participants with ID compared to 100% of trials for the control group.

In contrast to this, research examining the use of the Gudjonsson Suggestibility Scales (GGS; Gudjonsson, 1983, 1992), suggests that the immediate repetition of questions can lead to reduced validity of responding because it induces people to change their previous response to questions (Gudjonsson, 2003). However, evidence suggests that while it is true that receiving negative feedback followed by the repetition of questions does induce people with ID to change their responses, the tendency to do this is no different to that in people of normal intelligence (Clare & Gudjonsson, 1993; Everington & Fulero, 1999; Gudjonsson,

2003; Gudjonsson & Clare, 1995; Henry & Gudjonsson, 1999; Henry & Clare, 2003). Rosen et al. (1974) provide an interesting distinction between the repetition of process as opposed to the repetition of question content. Rosen et al. delivered two assessments of acquiescence separated by several other tasks. Acquiescence was represented by the number of yes responses to 21 ambiguous questions or statements, such as *Winter is more fun than summer*. For the second administration the questions are reversed so that the process is repeated but the question is not. Consequently, for the second administration the statement would be *Summer is more fun than winter*, with each question forming a pair with its opposite. The number of yes responses on both iterations was recorded as was the number of occasions that the participant gave contradictory responses to those given during the first administration, during the second administration. During the first administration participants with ID acquiesced significantly more often than a mental-age control group. In addition, participants with ID contradicted themselves more often by giving the same response to both questions in a pair than did the control group. However, during the second administration the acquiescence rate fell and was not significantly different to the controls. This suggests that it is the familiarisation with the questioning process that overcomes the threat of response bias.

It is difficult to judge the impact of the training and familiarisation processes utilised in studies using Likert-scales, as outlined in Bromley et al. (1998), because no control group was incorporated in the research design. In addition, it would be unwise to generalise from these studies due to the small samples used (range $n=20-54$) (Bromley et al., 1998; Glenn et al., 2003; Powell, 2003; Rosen et al., 1974; Sigelman et al., 1981). However, the results suggest that the repetition of the questioning process increases the reliability and validity of responding but future research may seek to clarify this relationship.

The evidence presented suggests that people with ID are prone to response biases when questioned using a range of different question formats. However, there are a number of

methodological problems related to these studies. All of the studies reported by Sigelman and colleagues (Budd et al, 1981; Sigelman et al., 1981a, 1981b; Sigelman et al., 1982; Sigelman et al., 1980), draw their samples from the same cohort of people with ID: 58 adults living in an institution; 52 children living in an institution; and 57 children living in the community.

However, it is not always clear how many of this total is taking part in each task reported. Where samples sizes are reported they can be small ($n=16-55$) (Sigelman et al., 1981a). In addition, some of the analyses are based upon very few questions. For example, contradictory pairs are often based on responses to just two pairs of questions (Sigelman et al., 1981a) so answering just one question *incorrectly* can potentially have a large impact on the outcome. In addition, the majority of studies did not include a control group without ID (Budd et al., 1981; Sigelman et al., 1981a, 1981b; Sigelman et al., 1982; Sigelman et al., 1980), so actual differences in participants with and without ID were not demonstrated.

Also, the results are seldom presented systematically, which makes it difficult to make comparisons between studies. Of the other studies cited, most have relatively small samples sizes: $n=60$ - Gerjuoy and Winters (1966); $n=20$ - Gudjonsson (1986); $n=48$ - Lovett and Harris (1987); $n=15$ - March (1992); $n=24$ - Rosen et al. (1974). In addition, many of the studies do not report IQ mean or IQ range (Budd et al., 1981; Guerjuoy & Winters, 1966; Lovett & Harris, 1987; Sigelman et al., 1981a,1981b; Sigelman et al., 1982; Sigelman et al., 1980) and consequently it is difficult to make comparisons between the studies.

2.4.6. Likert-scales

Perhaps in recognition of the threat of response bias associated with yes/no and either/or questions, and the tendency to under-report in response to open-ended questions, a wide range of studies with participants with ID, have incorporated Likert-scales. Evidence

indicates that Likert-scales result in lower response rates than yes/no (68% to 85%), either/or (63% to 66%) and open-ended (51% to 85%) formats (Sigelman et al., 1981, 1982b). These differences appear to be associated with the level of ID in participants, with participants drawn from mild ID and borderline IQ range, reporting higher response rates. For example, 100% response rates have been reported for studies with participants from the mild (Baker & Bramston, 1997), mild to borderline IQ range (Glenn, Bihm and Lammers, 2003; Lindsay & Lees, 2003). This is a tentative conclusion, however, as there are methodological difficulties and differences between studies which may have influenced the results, which will be reported later.

However, this relationship between level of IQ and response rate is also mediated by the wording of the response format. For example, Beck, Carlson, Russell and Brownfield (1987) ($n=26$) found that participants were able to complete the Depression Self-Rating Scale (DSS; Birlerson, 1981) but were unable to complete the Beck Depression Inventory (BDI; Beck, Ward, Mendalim, Mock, & Erbaugh, 1981). The DSS uses brief descriptors of the type normally associated with Likert-scales, whereas the BDI uses four self-description response alternatives, with each question presenting a new set of response alternatives. The DSS allows the individual to generalise having learned to use the response format but the BDI requires participants to select from subtle differences unique to specific attitudes or behaviours. All of the other studies outlined in this section include brief descriptors which remain the same for all items which may help to avoid this difficulty. However, this study suggests that the relative ease or difficulty of the response format will affect whether questions will be answered and hence the reliability and validity of the instrument. Lindsay and Lees (2003) modified the BDI items and response format in an attempt to over-come potential difficulties associated with the subtle changes in the BDI items while retaining the meaning of the original items. They reported strong a test re-test correlation for the amended

scale ($r=.88$) although this was with a small sample ($n=32$). Lindsay and Skene (2007), using a modified version of the BDI-II reported high internal consistency ($\alpha=.90$) using a larger sample size ($n=108$).

Most of the studies examined did not report on the presence of any response bias apparent in their data. However, several studies report that participants tended to chose the most positive response alternative (Hartley & MacLean, 2005; Verri et al., 1999), a tendency that may be associated with level of ID. For example, several studies use subjects from the mild ID to borderline intelligence range (Alder & Lindsay, 2007; Fogarty, Bramston and Cummins, 1997; Harley & MacLean, 2005). They report 0%, 0% and 3% levels of response bias respectively. In contrast, those studies using a sample with moderate ID, reported higher levels of response bias (Bramston, Fogarty, & Cummins, 1999; Sigelman & Budd, 1986; Verri et al., 1999). For example, Verri et al. (1999) excluded 37% of their Italian sample and 27% of their Australian sample because they chose the most positive response option on every occasion. None of their control groups containing people without ID were excluded because of response bias. The Verri et al. study suggests that response bias may be more pronounced in participants with moderate ID; however more research is required before firm conclusions can be drawn.

Modifications have been made to process of self-report using Likert-scales in samples with ID, including:

(1) the addition of pictures to response options (Alder & Lindsay, 2007; Kellett et al., 2003; Lindsay et al., 2009; Lindsay & Lees, 2003; Steptoe, Lindsay, Forrest, & Power, 2006);

(2) the rephrasing of questions during interview (Fogarty et al., 1997; Masi et al., 2002; Mindham & Espie, 2003; Novaco & Taylor, 2004; Snoyman & Aicken, 2011);

(3) using clarifying questions to check the validity of responding (Baker & Bramston, 1997; Bramston & Fogarty, 2000; Fogarty et al., 1997; Mindham & Espie, 2003);

(4) and training and familiarisation procedures (Cummins et al., 1997; Glenn et al., 2003, Lindsay & Lees, 2003; Lunskey & Benson, 2001; Snoyman & Aicken, 2011).

It appears that expanding upon item content and /or paraphrasing items increases the reliability and validity of responding to Likert-scales. For example, researchers followed each question on a Likert-scale with the phrase “Tell me more about that” which they state was a means of checking the validity of responses and enabling the participant to formulate an appropriate response (Bramston & Fogarty, 2000; Fogarty et al., 1997; Mindham & Espie, 2003). All of these studies report high internal consistencies for the scales included ($\alpha=.89$, $\alpha=.86$ and $\alpha=.96$ respectively).

In addition, pictures have been used to improve the reliability and validity of responses (Adler & Lindsay, 2007; Baker & Bramston, 1997; Sigelman, Budd et al., 1982; Sigelman & Budd, 1986; Rojahn et al., 1994; Steptoe et al., 2006; Snoyman & Aicken, 2011). For example Rohan et al. (1994) and Sigelman and Budd (1986) both used response formats represented by smiling and frowning faces on a three-point and four-point Likert-scales respectively. For these studies, response rates increased in the picture format from 76% to 100% (Rohan et al., 1994) and from 52% to 69% (Sigelman & Budd, 1986). It is noticeable that the response rate from a sample drawn from people with severe, moderate and mild ID (Sigelman & Budd, 1986) is still lower than the response rate from a sample drawn from people in the mild to moderate ID range (Rohan et al., 1994), even when pictures are used. In addition a number of authors (i.e., Adler & Lindsay, 2007; Baker and Bramston, 1997; Bramston & Fogarty, 2000; Lindsay, et al., 2009; Lindsay & Lees, 2003; Snoyman & Aicken, 2011; Steptoe et al., 2006) all used pictorial aids and reported 100% response rates. The evidence presented here suggests that pictorial aids, used in conjunction with Likert-scales, increases the response rate to the scales used.

2.4.6.1. Consistency of Likert-scales

The following studies report internal consistency using Cronbach's alpha and test-retest correlational coefficients: Kellett et al. (2003) report internal consistencies ($\alpha=.68-.85$) for the symptom dimensions of the Brief Symptoms Inventory (Derogatis, 1993) and test-retest ($r=.68-.91$). Mindham and Espie (2003) report internal consistency ($\alpha=.96$) and test-retest ($r=.95$) for the Glasgow Anxiety Scale for people with an ID (GAS-ID: Mindham & Espie, 2003). Payne and Jahoda (2004) report internal consistency ($\alpha=.78$) and test-retest ($r=.90$) within a four-week interval for the Glasgow Social Self-efficacy Scale (GSSES). Kellett, Beail, and Newman (2005) report internal consistencies for the subscales of the Inventory of Interpersonal Problems-32 (IIP-32: Barkham, Hardy, & Startup, 1996) ($\alpha=.82-.85$). Rojahn et al. (1994) report ($\alpha=.90$) and test-retest ($r=.63$) after an 11-week interval for the Self-Report Depression Scale (SRDQ) (Reynolds & Baker, 1988). These findings are very encouraging with the majority of Cronbach's alphas reported at acceptable levels (Kline, 1999) and test-retest reporting strong associations across all measures. Lindsay et al. (2009) report strong internal consistency ($\alpha=.84$) and test re-test ($r=.79$) for the Social Problem Solving Inventory – Revised (SPSI-R, D'Zurilla, Nezu, & Maydeu-Olivares, 2002). Lindsay and Lees (2003) report strong test re-test reliability over a one week interval for the BAI ($r=.91$) and the BDI ($r=.88$).

A number of studies report the internal consistency of the Likert-scale used in the study (e.g., Adler & Lindsay, 2007; Bramston & Fogarty, 2000; Bramston & Miochic, 2001; Emerson, 2005; Glenn et al., 2003; Kellett et al., 2003; Lindsay, et al., 2009; Lindsay & Lees, 2003; Mindham & Espie, 2003; Payne & Jahoda, 2004; Rojahn et al., 1994; Snoyman & Aicken, 2011; Verri et al., 1999). While, Adler and Lindsay (2007) report ($\alpha=.91$) for the Dundee Provocation Inventory (DPI: Lindsay, 2000). Bramston and Fogarty (2000) report ($\alpha=.89$) for the State Trait Anger Expression Inventory (STAXI; Spielberger, 1988), and

($\alpha=.86$) for the Life-stress Inventory (LI). Bramston and Miochie (2001) report ($\alpha=.73$) for the (LI). Emerson (2005) reports ($\alpha=.61$) for the Strengths and Difficulty Questionnaire (SDQ; Goodman, 1997). Glenn et al. (2003) report ($\alpha=.92$) for the Beck Anxiety Inventory (BAI; Beck & Steer, 1990), ($\alpha=.92$) for the Reynolds Child Depression Scale (Reynolds, 1989), ($\alpha=.97$) for the Automatic Thoughts Questionnaire (ATQ; Hollon & Kendall, 1980), and ($\alpha=.94$) for the Cognitions Checklist (CCL; Beck, Brown, Steer, Eidelson, & Riskind, 1987). Lindsay and Skene (2007) reported good internal consistency for the BAI ($\alpha=.91$) and the BDI-II ($\alpha=.90$). Snoyman and Aicken (2011) report strong internal consistency ($\alpha=.80$) for a modified version of the BIS-11. It can be seen that two studies incorporated the Life-stress Inventory (Bramston & Fogarty, 2000; Bramston & Miochie, 2001). However, Bramston and Fogarty (2000) used a sample with mild and moderate ID, whereas Bramston and Miochie (2001) included individuals with mild and moderate disabilities including ID, visual disabilities and physical disabilities. The Cronbach's alpha reported in this study is for the whole sample and consequently it is difficult to make comparisons between the two studies.

2.4.6.2. Reliability of self-report using Likert scales

A number of studies have provided evidence regarding the validity of self-report using Likert-scales in samples with ID (Alder & Lindsay, 2007; Bramston & Fogarty, 2000; Dagnan & Sandhu, 1999; Glenn et al., 2003; Kazdin et al., 1983; Kellett et al., 2003; Lindsay et al., 2009; Lindsay & Skene, 2007; Masi et al., 2002; Mindham & Espie, 2003; Nezu et al., 1995; Rojahn et al., 1994). In relation to the BAI and the BDI, Lindsay and Skene (2007) reported a factor structure for both tools that conformed entirely to factor solutions identified in mainstream, non-ID samples. In addition, the authors reported that anxiety referrals in

their sample scored significantly higher BAI and depression referral scored significantly higher on the BDI-II.

Adler and Lindsay (2007) reported moderate and strong correlations respectively between the DPI and the Novaco Anger Scale (NAS: Novaco, 2003) ($r=.57$) and the NAS Provocation Inventory (NAS-PI: Novaco, 2003) ($r=.77$). Dagnan and Sandhu (1999) report a moderate correlation ($r=.57$) between negative self-esteem, as measured on the Rosenberg Self-esteem Scale (Rosenberg, Schooler, & Schoenbach, 1989), and the achievement subscale of the Social Comparison Scale (Gilbert & Allen, 1994), and a moderate correlation ($r=.41$) between negative self-esteem and the Zung Depression Scale (Zung, 1965). These results were generally consistent with results from studies with participants without ID (Swallow & Kuiper, 1988) thus providing some validity for the self-report measures used. Kellet, Beail, Newman and Hawes (2004) reported an eight factor structure for the BSI which appeared to have high face validity, with six of the original nine scales broadly reproduced. The large sample size used in this study ($n=335$) suggests that this is likely to be a robust factor structure. Masi et al. (2002) reported moderate correlations between the Zung Self-Rating Anxiety Scale (ZAS: Zung, 1971) and informant ratings on the Anxiety subscale of the Psychopathology Instrument for Mentally Retarded Adults (PIMRA; Matson, Senatore, & Kazdin, 1984) ($r=.48$) and the Anxiety and Depression Subscale of the Child Behaviour Checklist (CBCL: Achenbach, 1983) ($r=.43$). However, they reported no correlations between the Zung Self-Rating Depression Scale (Zung, 1965) and informant ratings on the PIMRA Depression Subscale or the Anxiety and Depression subscale of the CGCL. The findings with regard to the ZAS are somewhat at odds with results reported by Lindsay and Michie (1988), who reported that the ZAS, with modified items to aid understanding, but using the 5-point Likert-scale resulted in unacceptable split-half reliability ($r=.12$). Masi et al. (2002) do not report on the reliability of the tools in their study. Emerson (2005) reported

significant weak to moderate correlations ($r=.30-.49$), between child self-report and parental report, on six of the seven subscales of the SDQ, compared with seven from seven for children without ID ($r=.30-.46$).

Bramston and Fogarty (2000) report a weak correlation ($r=.19$) between self-report Children's Depression Inventory (CDI: Kovacs, 1985) and an informant completed version. Glenn et al. (2003) reported moderate to strong correlations ($r=.76-.92$) between the BAI, RCDS, ATQ and the CCL, thus matching the links between anxiety and depression found in populations with normal IQ (Beck et al., 1987). Kellett et al. (2003) reported that the BSI significantly differentiated a clinical group from a control group on eight of the nine symptoms dimensions and two of the three global indices of distress. Similar trends in the obsessive-compulsive symptom dimension and the positive symptoms distress index failed to reach significance. In addition, BSI total correlated ($r=.15-.45$) with the IIP-32 and its subscales (Kellet et al., 2005). Mindham and Espie (2003) reported a strong correlation between the GAS-ID and the BAI ($r=.72$). Payne and Jahoda (2004) report a moderate correlation ($r=.42$) between the GSSSES and Sherer's Social Self-Efficacy Scale (Sherer et al., 1982). Benson and Ivins (1992) reported weak correlations between self-report scores on the Birleson Depressive Self-Rating Scale (Birleson, 1981) and informant ratings ($r=.26$), but no correlation between self-report scores on the Children's Inventory of Anger (CIA: Finch et al., 1983) and informant ratings (see Chapter 3). Rojahn et al. (1994) report that there were no association between classifications derived from the SRDQ and a standard diagnostic interview (The Diagnostic Interview for Children and Adolescents; Reich, Shayka, & Taibleson, 1992) and an informant rating scale for Maladaptive Behaviour (The Reiss Screen; Reiss, 1988). Lindsay et al. (2009) conducted a factor analysis on data from 132 participants with ID on the SPSI-R. The authors report that the four factor solution identified was entirely consistent with those of D'Zurilla et al. (2002).

2.4.6.3. Conclusions regarding the use of Likert scales

It appears from the evidence presented above that a significant amount of work incorporating self-report measures utilising Likert-scales is being conducted and that this is producing reliable and valid measures for use in samples with ID. Good examples are the BAI and the BDI. The BAI has high internal consistency, as reported by Glenn et al. (2003) ($\alpha=.92$) and Lindsay and Skene (2007) ($\alpha=.91$). Lindsay and Lees (2003) report strong test re-test reliability over a one week interval for the BAI ($r=.91$). Lindsay and Skene reported a factor structure entirely consistent with mainstream non-ID samples and reported that referrals for anxiety scores significantly higher than controls. Similarly, with the BDI, Lindsay and Skene (2007) reported good internal consistency ($\alpha=.90$) and Lindsay and Lees report strong test re-test reliability over a one week interval ($r=.88$). Lindsay and Skene reported a factor structure that conformed entirely with factor solutions identified in mainstream, non-ID samples. In addition, the authors reported depression referrals scored significantly higher on the BDI-II. Glenn et al. (2003) reported moderate to strong correlations ($r=.76-.92$) between the BAI, RCDS, ATQ and the CCL, thus matching the links between anxiety and depression found in populations with normal IQ (Beck et al., 1987), a finding supported by Lindsay and Skene who reported a moderate correlation between the BAI and the BDI-II. This evidence indicates that with appropriate modifications, reliable and valid self-report measures can be produced for use with samples with ID.

Two studies incorporating the same self-report measure suggest that there may be differences in validity depending upon the level of ID. Kazdin et al. (1983) ($n=110$), using a sample comprising near 50% of people with moderate to severe ID, report a low correlation ($r=.10$) between the BDI-R, rated on a four-point rating scale, and the depression scale of the Psychopathology Instrument for Mentally Retarded Adults (PIMRA-D). In contrast, Nezu *et*

al. (1995), using a sample of similar size ($n=107$), but comprising only adults with mild ID, found a medium correlation ($r=.53$) between the same instruments.

The results from studies that report reliability and validity data for the use of Likert-scales for self-report assessment in people with ID are generally encouraging. However, it is difficult to draw firm conclusions because of methodological difficulties and differences between the studies. Several studies include small sample sizes (range $n=3-53$) (Bramston & Miochie, 2001; Dagnan & Sandhu, 1999; Glenn et al., 2003; Lindsay & Lees, 2003; Mindham & Espie, 2003; Payne & Jahoda, 2004; Rojahn et al., 1994; Singh, Lancioni, Winton, Singh, Adkins & Singh, in press). In addition, some studies included participants with borderline IQ (Glenn et al., 2003; Lindsay & Lees, 2003; Payne & Jahoda, 2004; Verri et al., 1999) and low average IQ (Alder & Lindsay, 2007; Payne & Jahoda, 2004). Several studies did not report the level of ID of the sample used, often relying upon a service for people with ID to identify a sample (Dagnan & Sandju, 1999; Emerson, 2005; Verri et al., 1999). Several studies excluded participants because of concerns about the reliability and validity of their self-report (Glenn et al., 2003; Payne & Jahoda, 2004; Rojahn et al., 1994) which will have resulted in higher response rates and is likely to have improved the assessment of reliability and validity in those studies. In addition, the complexity of the Likert-scales used varied between studies from 3-point (Bramston & Fogarty; Rojahn et al., 1994), 4-point (Baker & Bramston, 1997; Bramston & Fogarty, 2000; Bramston et al., 1999; Emerson, 2005; Glenn et al., 2003; Lindsay & Lees, 2003; Singh et al., in press; Snoyman, & Aicken, 2011), 5-point (Baker & Bramston, 1997; Bramston, & Miochie, 2001; Fogarty et al., 1997; Kellett et al., 2003), 6-point (Dagnan & Sandhu, 1999), 7-point (Steptoe et al., 2006) to 14-points (Payne & Jahoda, 2004). It is unclear if the number of options available affects the complexity of the assessment and in turn impacts upon reliability and validity.

Several studies have incorporated training and familiarisation procedures in attempts to improve reliability and validity of responding (Bromley, Emerson, & Cummins, 1998; Glenn et al., 1994; Lindsay & Lees, 2003; Lunsky & Benson, 2001; Lunsky, Emery, & Benson, 2002; Verri et al., 1999). It is difficult to judge the impact of these procedures because no control groups without the training and familiarisation procedures were included in the studies. In addition, some studies also included visual aids representing the different points on the Likert-scale, for example, using smiling or frowning faces (Dagnan & Dandhu, 1999; Glenn et al., 2003; Lindsay & Lees, 2003; Minham & Espie, 2003; Steptoe et al., 2006; Verri et al., 1999). Again, the effects of these visual aids are difficult to assess because no control group was included without the use of visual aids with which to make comparisons. In addition, the differences between studies make it difficult to draw conclusions from those that did include these components and those that did not.

These findings are encouraging, indicating that Likert-scales can be used to obtain reliable and valid self-report assessments for people with ID. Future research could seek to clarify which combination of scale length, pictures, rephrasing of questions, clarifying questions and training and familiarisation processes can be used to optimise performance.

2.5. Possible solutions to alleviate response biases in the assessment of offenders with ID

In relation to the self-report assessment of individuals with ID, the evidence suggests that there are a number of response biases that represent a threat to the reliability and validity of those assessments. A number of authors have suggested amendments to assessments that should aid reliable and valid assessment in people with ID. The evidence presented above indicates that if an individual knows the answer to a question then response biases are far less likely to occur (Sigelman et al., 1981b; Heal & Sigelman, 1995). Knowing the answer to a question is of course dependent upon understanding the question being asked. This has

particular implications for people with ID who may have language and communication deficits (Mackinnon, Bailey, & Pink, 2004). Consequently, when asking questions of people with ID the use of simple language will obviously make questions easier to understand and therefore, answer. Care should be taken to avoid sophisticated vocabulary, jargon, unexplained technical terms, figurative language and colloquialisms (Prosser & Bromley, 1998).

Equally, the use of short and simple sentence structures and grammar is likely to facilitate understanding and therefore valid responding (Prosser & Bromley, 1998). Consequently, single clause questions are preferable to difficult propositions and complex linguistics. Using active verbs in the present tense rather than passive verbs regarding possible future actions is likely to focus on concrete rather than abstract propositions. The readability of questions, in terms of sentence structure and the vocabulary used can be measured through the use of standard software programmes (Microsoft Corporation, 1993). These include the Flesch Reading Ease (Flesch, 1948) and Flesch-Kincaid Grade Level (Kincaid, Fishburne, Rogers, & Chissom, 1975), which both rely upon the same core measures of average number of syllables per word and average number of words per sentence but are weighted differently. These statistics provide reading level estimates based on various percentages of the respective normative samples that answered comprehension questions regarding sets of standard passages. The Flesch Reading Ease formula generates a rating between 0 and 100, where high scores indicate material that is easier to read. The Flesch-Kincaid formula translates the 0-100 score into an American grade level reading ability, which makes it easier to interpret the readability level of books and texts. The US grade can then be converted into age equivalent reading ability for the UK (See the table in Chapter 4).

Simple questions about concrete issues in the life of a person with ID produce the most reliable and valid responses (Heal & Sigelman, 1995; Sigelman et al., 1981b). The question of sentence structure is perhaps particularly pertinent to the assessment of factors where social desirability of responding may be a factor (Paulus, 1984); such as the assessment of attitudes and beliefs associated with a range of offending behaviours (Saunders, 1991). Questionnaires assessing such areas frequently include reversed items in order to avoid response sets (Eysenck & Eysenck, 1977; Nowicki & Duke, 1974a). This often results in the inclusion of questions containing double negatives. Such questions are likely to present significant difficulty for someone with ID because the sentence structure is not simple. Failure to understand the question or to understand the question but fail to understand which response signifies a true representation of one's attitudes and beliefs increases the bias of response sets (Heal & Sigelman, 1995; Sigelman et al., 1981a). Consequently, it is reasonable, if not essential, that such items are removed from test instruments for use with people with ID (Prosser & Bromley, 1998).

The use of pictures appears to aid understand and consequently improves response rates (Adler & Lindsay, 2007; Baker & Bramston, 1997; Sigelman et al., 1982a, Sigelman & Budd, 1986; Rojahn et al., 1994; Snoyman & Aicken, 2011; Steptoe et al., 2006). Use in conjunction with other adaptations also appears to improve reliability and validity of responding (Lindsay & Lees, 2003; Lindsay & Skene, 2007; Lunsy & Benson, 2001; Snoyman & Aicken, 2011) but research has not yet demonstrated which amendments make the meaningful difference.

Most of the evidence presented indicates the prevalence of response biases in ID populations. However, this is clearly an over-simplification. Response biases appear to vary with IQ within the ID population (Budd et al., 1981; Gudjonsson & Henry, 2003). Also if we consider the most prevalent response bias of acquiescence there are those within the ID

population who do not acquiesce (Heal & Sigelman, 1995). The adoption of modification to existing tools outlined above will increase an individual's ability to participate in the assessment process and thus increase the reliability and validity of responding. While, the psychometric assessment of individuals with ID may be problematic the evidence suggests that such assessment can reveal useful and valid information regarding the interviewee. Clearly however, interviewers have an obligation to ensure that the threat of response biases associated with this population is minimised in the assessment procedure.

2.6. Conclusions

Evidence suggests that there are threats to the reliability and validity of self-report in people with ID due to the influence of response biases. However, evidence also suggests that the adaptation of tools and procedures can lead to reliable and valid self-report. Consequently, it appears that response biases such as acquiescence can be overcome to some extent by controlling question format, question content and response format. The question format should be simple with short and familiar words and short straightforward sentences. The content of questions should be concrete where possible and relate to concepts familiar to the population in questions. Finally, response format should be simple so it is clear which response represents the individual's actual attitudes or beliefs.

Talbot and Riley (2007) conclude that there are large numbers of people in prison have intellectual difficulties that represent a significant problem that is likely to interfere with an individual's ability to cope with many aspects of the Criminal Justice System. Part of the process of providing an appropriate service that meets the needs of these individuals, involves the development of new, or the adaptation of existing, assessment tools to ensure that they are reliable and a valid means of assessing risk and need in this population. Once adapted it is necessary to demonstrate the reliability and validity of the new tool with a sample of the

target population. As previously outlined, *reliability* can be determined through the assessment of internal consistency and test-retest reliability. *Validity* can be determined concurrently, through correlation with another measure that is theoretically linked to the variable in question, providing construct validity, or in relation to a criterion, or an important form of behaviour that is external to the instrument itself. For the purposes of this thesis the construct validity for adapted assessment tools will be assessed through the correlation of the revised instrument with the instrument in its original format. In relation to criterion validity, measures of impulsivity and risk assessment will be compared with the criterion of aggressive institutional behaviour.

Chapter 3 will consider the reliability and validity of psychometric tools available for the assessment of people with ID in forensic settings.

CHAPTER THREE: A REVIEW EXAMINING PSYCHOMETRIC TOOLS USED IN THE ASSESSMENT OF OFFENDERS WITH AN INTELLECTUAL DISABILITY

3.0. Introduction

In recent years, researchers have begun to consider those factors associated with offending in ID populations. This paper aims to review this evidence and identify those psychometric measures used with this population. Initially, evidence relating to the general topics of anger, fire-setting and sexually inappropriate behaviour is considered. Then, the specific criminogenic factors of empathy, locus of control and impulsivity are examined.

Criminogenic factors are those that are directly linked to crime (Ministry of Justice, 2010).

Some assessment tools have been specifically developed for people with ID while others have been adapted from mainstream assessments. For example, in the assessment of pro-offending attitudes related to sexually inappropriate behaviour, Broxholme and Lindsay (2003) developed the Questionnaire on Attitudes Consistent with Sex Offending (QACSO) specifically for use with people with ID. In contrast, Kolton, Boer, and Boer (2001) adapted the Abel and Becker Cognition Scale (ABCS; Abel, Becker & Cunningham-Rathner, 1984) for use with offenders with ID by simplifying the language and sentence structure of the original scale.

Evidence presented in Chapter 1 demonstrated that people with ID are convicted of (Barron et al., 2004; Hayes 1993; Hayes & Craddock, 1992; Lindsay, 2002; Lindsay, Elliott & Astell, 2004; Simpson & Hogg, 2001; Walker & McCabe, 1973) and self-report (Dickson et al., 2005) involvement in a wide range of offences. Consequently it is important that a range of tools are available to assess risk and need in relation to these offences. Also, as outlined in Chapter one it is important that the tools used are reliable and valid, so that the assessments are meaningful and can be used to accurately assess risk and need. Chapter 6

contains a critical review and summary of risk assessment tools used with offenders with ID.

This chapter critically reviews the literature relating to tools used to assess psychological factors associated with offending in samples with ID.

3.1. Assessment of anger and aggression

Of all areas of forensic research involving individuals with ID perhaps the most developed is that concerned with the assessment and treatment of anger and aggression. A number of studies have investigated the experience of anger and the expression of aggression in people with ID (Alder & Lindsay, 2007; Benson & Ivins, 1992; Benson, Rice & Miranti, 1986; Black & Novaco, 1993; Hagiliassis, Gulbenkoglou, Di Marco, Young and Hudson, 2005; Howells, Rogers & Wilcock, 2000; Jackson & Altman, 1996; Jahoda, Pert, Squire & Trower, 1998; King, Lancaster, Wynne, Nettleton & Davis, 1999; Lindsay, Allan, MacLeod, Smart & Smith, 2003; Lindsay, Allan, Parry, MacLeod, Cottrell, Overend and Smith, 2004; Lindsay, Overend, Allan, Williams & Black, 1998; MacMahon, Jahoda, Espie & Broomfield, 2006; Moore, Adams, Elsworth & Lewis, 1997; Novaco & Taylor, 2004; Pert & Jahoda, 2008; Rose & Gerson, 2009; Rose, Loftus, Flint & Carey, 2005; Rose, O'Brien & Rose, 2009; Rose & West, 1999; Rose, West & Clifford, 2000; Rossiter, Hunnisett & Pulsford, 1998; Taylor, Novaco, Gillmer & Thorne, 2002; Taylor, Novaco, Guinan & Street, 2004; Walker & Cheseldine, 1997).

The majority of these studies use some form of self-report assessment, with just one exception (Jackson & Altman, 1996). The self-report measures used fall into seven broad categories: anger (Alder & Lindsay, 2007; Benson & Ivins, 1992; Benson et al., 1986; Hagiliassis et al., 2005; King et al., 1999; Lindsay, Allan, MacLeod, Smart & Smith, 2003; Lindsay, Allan, Parry, MacLeod, Cottrell, Overend, & Smith, 2004; MacMahon et al., 2006; Novaco & Taylor, 2004; Rose & Gerson, 2009; Rose, O'Brien & Rose, 2009; Rose & West,

1999; Rose et al., 2000; Rose et al., 2005; Taylor et al., 2004); provocation (Alder & Lindsay, 2007; Black & Novaco, 1993; Lindsay et al., 1998; Lindsay et al., 2004; MacMahon et al., 2006; Novaco & Taylor, 2004 Taylor et al., 2002; Taylor et al., 2004; Walker & Cheseldine, 1997); problem-solving/assertiveness (Basquill, Nezu, Nezu & Klein, 2004; Benson et al., 1986; Black & Novaco, 1993); mental health (Benson & Ivins, 1992; Black & Novaco, 1993; Howells et al., 2000); self-concept/self-esteem (Benson & Ivins, 1992; Howells et al., 2000; King et al., 1999); threat appraisal (Jahoda, Pert, Squire, & Trower, 1998; Pert & Jahoda, 2008); and impulsivity (Novaco & Taylor, 2004).

The two most frequently used self-report measures of anger in populations with ID are the Novaco Anger Scale (NAS: Novaco, 2003) and the Anger Inventory (AI: Benson et al., 1986). The AI has been used in a number of studies (Benson et al., 1986; Benson & Ivins, 1992; King et al., 1999; Rose et al., 2000; Rose et al., 2005; Rose & West, 1999; Rose & Gerson, 2009; Rose, O'Brien & Rose, 2009; Rossiter et al., 1998). Benson et al (1986) report a moderate correlation ($r=.62$) on test re-test, over an unspecified interval, which remains the only report of the instrument's reliability. Four studies report pre-treatment means (range 98.11-103.8) which are fairly similar (Rose et al., 2000; Rose et al., 2005; Rose & Gerson, 2009; Rose, O'Brien & Rose, 2009). Three studies report post-treatment means (range 83.9-93.7) which are all lower than the pre-treatment scores. Indeed, several studies have reported significantly lower mean scores from pre- to post-treatment (King et al., 1990; Rose et al., 2000; Rose et al., 2005; Rose et al., 2009; Rose & West, 1999) suggesting that the AI is sensitive to treatment change. There is some variation between the mean scores reported at pre- and post-treatment. The range of means for pre-treatment vary by 5.69 and the range of means post-treatment vary by 9.8 suggesting that some treatment interventions may be more successful in reducing AI scores than others. The largest reduction in AI scores pre-post treatment was 14.4 (Rose et al., 2009). However, it is difficult to draw conclusions

from these findings because the authors do not provide any details of participants' IQs. Therefore it is possible that differences in self-report scores are due to differences in the cognitive ability of group participants.

The NAS has been used in several studies (Alder & Lindsay, 2007; Hagiliassi et al., 2005; MacMahon et al., 2006; Novaco & Taylor, 2004; Taylor et al., 2004). Novaco and Taylor (2004) reported good internal consistency ($\alpha=.92$), which remains the only report of the instrument's reliability. Novaco and Taylor (2004) also reported moderate correlations between the NAS Total and subscale scores and the number of hospital assaults (range $r=.31-.41$), and between the NAS Cognitive subscale and staff ratings of anger ($r=.34$). Equally, Alder and Lindsay (2007) reported a moderate correlation between NAS Total and the Dundee Provocation Inventory (DPI) ($r=.57$), providing more evidence of the validity of NAS scores in this population. Hagiliassi et al (2005) also reported on the NAS but the authors scored it in an idiosyncratic manner, with lower scores representing higher levels of anger. It is therefore difficult to make comparisons between scores in this study and those using the traditional scoring method. However, Hagiliassi et al (2005) report significant increases in scores post-treatment and at a four-month follow-up, providing some validity for the NAS, and suggesting that it is sensitive to changes in anger levels as a result of treatment. In addition, the correlations reported in the studies above provide some evidence of the validity of scores on the NAS, as it appears to be measuring a similar construct to related measures.

A third tool used in the assessment of anger in populations with ID is the State-Trait Anger Expression Inventory (STAXI: Spielberger, 1996), which was used in a couple of studies (Novaco & Talyor, 2004; Taylor et al., 2004). Taylor et al. (2004) report good internal consistency for State ($\alpha=.87$) and Trait ($\alpha=.86$) dimensions. In addition, evidence of

the validity of scores on the STAXI was provided by weak but significant correlations between subscales and number of hospital assaults ($r=.33-.37$).

Two provocation inventories are reported in the literature: the Provocation Inventory (PI: Novaco, 1988) and the DPI (Alder & Lindsay, 2007). The PI has been used in a number of studies (Alder & Lindsay, 2007; Black & Novaco, 1993; MacMahon et al., 2006; Novaco & Taylor, 2004; Taylor et al., 2004; Walker & Cheseldine, 1997). Novaco and Taylor (2004) reported good internal consistency ($\alpha=.92$) for the PI, which remains the only report of the instrument's reliability with a sample with ID. Taylor et al (2004) reported a significant reduction in mean score post-treatment compared to pre-treatment, suggesting that the PI is sensitive to changes in responses to provocative situations resulting from treatment. Alder and Lindsay (2007) reported a strong correlation between scores on the PI and scores on the DPI, providing more evidence of the PI's validity, with scores similar to those for an instrument measuring the same construct.

Lindsay et al. (2004) reported adequate internal consistency for the DPI ($\alpha=.79$) compared with Alders and Lindsay (2007) who reported very good internal consistency ($\alpha=.91$). Lindsay et al. (2004) used a relatively small sample size ($n=33$) compared with Alders and Lindsay ($n=114$), suggesting that this is likely to be a more robust assessment of internal reliability. However, Alder and Lindsay (2007) included participants into the low average range. Lindsay et al. (2004) did not report the range of IQs on their study, so it is possible that a discrepancy between the ability of the relative samples has affected the internal consistency scores. Moderate to high inter-item correlations ($r\geq.40$) and item to total correlations ($r=.48-.72$) have been reported for the DPI (Alder & Lindsay, 2007). In addition, using exploratory factor analysis with an oblique rotation, Alder and Lindsay (2007) identified a five-factor structure accounting for 63.76% of the total variance. The identified factors related to themes of: threat to self-esteem; locus of control; resentment; frustration;

and rejection. Lindsay et al. (2004) reported that the DPI had a moderate correlation with anger provoking role-plays and that mean scores were significantly lower post-treatment compared with pre-treatment. The assessment of the psychometric properties of the DPI is well advanced compared to other measures used in samples with ID. The tool demonstrates good reliability, structural qualities and promising validity.

Taylor, Novaco, Guinan and Street (2004) developed the Imaginal Provocation Test (IPT). The authors report acceptable levels of internal consistency for the four subscales ($\alpha=.70-.80$). The concurrent validity of the IPT was examined through the correlation of its indices with the STAXI, NAS and PI. The authors report that overall the pattern of correlations obtained provided confirmation of the concurrent validity of the IPT indices. The highest correlations were between NAS Total and subscales and IPT Behavioural Reaction ($r= -.52-.63$) and IPT Anger Composite ($r=-.51-.62$).

In terms of problem-solving and assertiveness, only two self-report tools have been used in samples with ID and problems with anger and aggression. Benson et al. (1986) report on the Conflicting Situations Test (CST: Saylor, Benson & Einhaus, 1985) and Black and Novaco (1983) report on the Problem-Solving Assessment (PSA: Castles & Glass, 1986). The CST uses open questions, the responses to which are subsequently allocated to pre-defined categories. Benson et al. report good inter-rater reliability for the scoring ($r=.94$). The Black and Novaco (1983) study is a single case study. The authors report no data for the PSA. There is currently no reliable and valid assessment of problem-solving skills and assertiveness that has been used with a sample with ID who have problems with anger and aggression

Self-concept or self-esteem has been investigated in a sample with ID by King et al. (1999), using the Coopersmith Self-esteem Inventory (Coopersmith, 1981) and Howells et al. (2000) using the Rosenberg Self-esteem Inventory (Rosenberg, 1979). Both reported

significant increases in mean scores following treatment, indicating improvements in self-report self-esteem following treatment. Black and Novaco (1993) also assessed self-esteem, using the Self-esteem Inventory (O'Malley & Bachman, 1979), but did not report the results in their single case-study. It is not possible to draw firm conclusions from these studies due to the small sample sizes involved.

Two studies using self-report measures have considered the role of threat appraisal in the expression of aggressive behaviour in samples with ID (Jahoda et al., 1998; Pert & Jahoda, 2008). Jahoda et al. (1998) used a word-stem task and Pert and Jahoda (2008) used the Social Goals and Strategies for Conflict (SGASC) assessment. The SGASC uses hypothetical conflict situations and ambiguous situations, asking *What would you do?* and *Why would you do it?* The open-ended responses produced to the *What?* question are then coded into categories of: seek revenge; show strength; seek a fair outcome; avoid conflict. The open-ended responses produced to the *Why?* question are coded into three response categories: passive, assertive and aggressive. The authors report good inter-rater reliability of the coding using a second blind rater ($\kappa=.71-.88$). Participants who had recently displayed aggressive incidents gave more aggressive responses to the word stem task. On the SGASC aggressive participants tended to view submissive responses as having more negative outcomes than did non-aggressive participants, however, this difference was not significant.

It is difficult to draw firm conclusions from the evidence presented on the self-report measures used in the assessment of anger and aggression in samples with ID. Some of the evidence reported relates to single-case studies (Black & Novaco, 1993; Jackson & Altman, 1996; MacMahon et al., 2006), series of case studies (Lindsay et al., 1998) and small group studies (Hagiliassis et al., 2005; Howells et al., 2000; King et al., 1999; Lindsay et al., 2003; Moore et al., 1997; Rose & West, 1999; Rossiter et al., 1998; Taylor et al., 2002). Only six studies from a possible 21 have samples in excess of 48 (Alder & Lindsay, 2007; Benson et

al., 1986; Benson & Ivins, 1992; Novaco & Taylor, 2004; Rose et al., 2005; Rose & Gerson, 2009). The generally small sample sizes mean that it is difficult to generalise from the research as a whole.

In addition, there are differences in the intellectual ability of the samples used in the studies. For example, Benson et al. (1986) report that their sample ($n=54$) were from the mild and moderate ID range. Whereas, Taylor et al. (2004) report that their sample ($n=48$) were from the moderate to low-average range of ability. Consequently, some differences between studies may be due to differences in IQ. However, the problems associated with the level of intellectual functioning of the sample are more fundamental in nature. Only eight studies from a possible 21 reported using a valid and reliable measure of IQ, such as the WAIS-R, WAIS-II or WASI (Howells et al., 2000; Jackson & Altman, 1996; Lindsay et al., 1998, 2003; Lindsay et al., 2004; MacMahon et al., 2006; Novaco & Taylor, 2004; Taylor et al., 2004). Only six studies (Alder & Lindsay, 2007; Howells et al., 2000; Lindsay et al., 1998, 2003; Taylor et al., 2002, 2004) from a possible 21 reported the participants' range of IQs. Eight studies made no reference to the level of participants' intellectual functioning at all.

What is apparent from the research above is the lack of well researched, reliable and valid measures of anger and aggression for use with participants with ID. The AI (Benson et al., 1986) has demonstrated moderate test re-test reliability ($r=.62$) and good internal consistency ($\alpha=.93$). There is some evidence for its validity in terms of significant reductions in mean scores post treatment in two studies with reasonable sample sizes ($n=86$ & $n=57$) (Rose et al., 2005; Rose & Gerson, 2009). However, its validity has yet to be established against a more direct criterion of anger and aggression such as assaultive incidents. The NAS (Novaco, 2003) has also demonstrated good internal consistency ($\alpha=.92$) (Novaco & Taylor, 2004). In addition, scores on the NAS subscales showed low to moderate correlations with

hospital assaults ($r=.34-.43$) in a study with a good sample size ($n=110$) (Novaco & Taylor, 2004). Equally, the PI (Black & Novaco, 1993) has good internal consistency ($\alpha=.92$) and moderate test re-test reliability, over a two-six month interval ($r=.57$) in a study with a good sample size ($n=104$) (Novaco & Taylor, 2004). In addition, there is some evidence supporting its validity, in that scores on the PI have been shown to differentiate patients who are assaultive from those who are not (Novaco & Taylor, 2004). The DPI (Lindsay et al., 2004) has demonstrated good internal consistency ($\alpha=.91$) and moderate to high inter-item ($r\geq.40$) and item-total ($r=.48-.72$) correlations. Evidence for the validity of the tool is provided by a strong correlation with the PI ($r=.75$), suggesting that these provocation inventories are indeed measuring the same construct. In addition, the DPI had a moderate correlation with anger-provoking role plays ($r=.55$). The DPI mean score has shown a significant reduction post-treatment, although the sample size in this study was modest ($n=33$) (Lindsay et al., 2004). In addition, Alders and Lindsay (2007) have demonstrated an interpretable 5-factor solution accounting for 63% of the variance. Consequently, the psychometric properties of the DPI are more advanced than for the other self-report measures for people with ID and problems with anger and aggression.

3.2. Assessment of fire-setting

There are a small number of studies which include self-report measures in the assessment of fire-setting which include samples with ID (Brown, Johnson & Peedie, 2002; Clare, Murphy, Cox & Chaplin, 1992; Murphy & Clare, 1996; Rice & Chaplin, 1979; Taylor, Robertson, Thorne, Belshaw & Watson, 2006; Taylor, Thorne, Robertson & Avery, 2002; Taylor, Thorne & Slavkin, 2004). All of these studies represent single case-studies or single group interventions and consequently all use small samples ($n=1-14$). The self-report measures used in these studies include fire specific questionnaires and those assessing anger, self-

esteem and depression. Fire specific questionnaires include the Fire-setting Assessment Schedule (FAS: Murphy & Clare, 1996), the Fire Interest Rating Scale (FIRS: Murphy & Clare, 1996), and the Fire Attitude Scale (FAS* : Muckley, 1997). Other measures used include the NAS (Novaco, 2003), Culture Free Self-Esteem Inventory-2nd edition (CFSEI-2: Battle, 1992) and the Beck Depression Inventory (BDI: Beck & Beck, 1972).

The evidence relating to fire-setting behaviour in samples with ID is scant. Consequently, there is little data relating to self-report measures in this population. The studies included in this section are all small in scale and therefore it is difficult to draw meaningful conclusions. No studies report reliability statistics for any of the fire-specific measures listed above. In addition, most of the studies do not report whether mean scores changed significantly following treatment (Brown et al., 2002; Clare et al., 1992; Murphy & Clare, 1996; Rice & Chaplin, 1979). There are a number of exceptions where changes to mean scores are considered pre- and post-treatment (Taylor et al., 2002, 2004, 2006). However, the evidence from these studies is not consistent. For example, Taylor et al (2002), reported significant improvements in scores on the FIRS and the FAS* post-treatment. However, Taylor et al (2004, 2006) reported no significant change in mean scores on the FIRS and the FAS* following treatment. It is difficult to evaluate these different findings because of the small sample sizes used in these studies ($n=4-14$ respectively).

No reliability data have been reported for any of the non-fire-specific self-report measures used in relation to fire-setters with ID. Taylor et al (2004) reported improvements in NAS, CFSEI-2 and BDI scores for four, two and three clients respectively but these improvements did not reach significant levels. From the evidence presented above it must be concluded that at present there are no reliable or valid measures for the assessment of fire-setting in samples with ID.

3.3. Assessment of inappropriate sexual behaviour

The psychometric assessment of sexual offenders with ID can be divided into four broad areas of treatment need: deviant sexual interests; pro-offending attitudes; socio-affective functioning; and self management problems (Keeling, Beech & Rose 2007a). The current author has been unable to identify any new or adapted measures of deviant sexual arousal for sexual offenders with ID. Consequently, attention will be focussed on the remaining three areas.

3.3.1. Pro-offending attitudes

Psychometrics used in the assessment of pro-offending attitudes in samples with ID include the Questionnaire on Attitudes Consistent with Sex Offending (QACSO: Broxholme & Lindsay, 2003), Abel and Becker Cognition Scale (ABCS; Abel, Becker & Cunningham-Rathner, 1984), The Sex Offenders Opinion Test (SOOT: Bray 1997), the Criminal Sentiments Scale (CSS: Gendreau, Grant, Leipziger & Collins, 1979), the Adapted Victim Empathy Task (Offending Behaviour Programmes Unit (OBPU), HM Prison Service (HMPS), 1996), and the Sexual Offenders Self Appraisal Scale (SOSAS; Bray & Foreshaw, 1996). The psychometric most frequently used to assess pro-offending attitudes in sex offenders with ID is the QACSO (Broxholme & Lindsay, 2003; Langdon, Maxted, Murphy & SOTSEC-ID Group, 2007; Lindsay, Marshall, Neilson, Quinn & Smith, 1998; Lindsay, Neilson, Morrison and Smith, 1998; Lindsay & Smith, 1998; Lindsay, Steptoe & Beech, 2008; Lindsay, Whitefield & Carson, 2007; Murphy, Powell, Guzman and Hays, 2007; Sex Offender Treatment Services Collaborative – Intellectual Disabilities (SOTSEC-ID), 2010). The QACSO possesses good internal consistency for the total scale ($\alpha=.95$) and for the subscales ($\alpha=.68-.86$) and has demonstrated good test re-test reliability over a one-month interval for the overall scale ($r=.95$) and the subscales ($r=.56-.90$) (Lindsay et al., 2007). In

addition, it possesses good inter-subscale correlations ($r_s=.41-.91$) and item-to-total correlations ($r \geq .35$) (Lindsay et al., 2007). Lindsay et al. (2008) also report high levels of agreement (96.8%) between independent raters of QACSO responses ($\kappa=.94$). Evidence of the validity of scores on the QACSO comes from its ability to differentiate groups of sex offenders with ID, non-sex offenders with ID and non-offenders with ID (Lindsay et al., 2007). Equally, scores on the QACSO have been shown to be significantly lower following treatment (Keeling, Rose & Beech, 2006; Murphy et al., 2007; SOTSEC-ID, 2010) and at follow-up (SOTSEC-ID, 2010), indicating that it is sensitive to treatment change. However, scores on the QACSO failed to predict re-offending (SOTSEC-ID, 2010). These findings indicate that the QACSO is a well researched instrument, with evidence supporting the reliability of the measure and the validity of the scores obtained from sex-offenders with ID.

The SOSAS, has been used in several studies (Langdon et al., 2007; Murphy et al., 2007; SOTSEC-ID, 2010; Williams, Wakling & Webster, 2007). Williams et al. (2007) reported acceptable internal consistency for the overall scale ($\alpha=.76$), but only after five items with negative corrected item-total correlations were removed from the analysis. In addition, a PCA revealed a two-factor solution, accounting for 39.9% of the variance in the data. Denial of responsibility was the theme for Factor 1, which displayed reasonable internal consistency ($\alpha=.74$). Factor 2 had a theme of acceptance of a problem but had a barely acceptable level of internal consistency ($\alpha=.70$) (Kline, 1999). The SOTSEC-ID. (2010) and Williams et al. (2007) reported a significant reduction in mean scores following treatment, with a large effect size (Williams et al, 2007), with the difference being maintained at six-month follow-up (SOTSEC-ID, 2010). Murphy et al. (2007) reported no significant change to scores following treatment. However, the SOTSEC-ID (2010) and Williams et al. used larger sample sizes ($n=46$ & $n=211$, respectively). Consequently, it is likely that the results from these studies will be more robust than those associated with the smaller sample size

($n=8$) from the Murphy et al. (2007) study. However, scores on the SOSAS did not predict re-offending (SOTSEC-ID, 2010).

Other assessments of pro-offending attitudes used with sex offenders with ID include the Adapted-ABCS (Kolton, Boer, & Boer, 2001) and an adapted version of the CSS (Keeling et al., 2007b). Little information of the reliability and validity of these tools, when used with sex offenders with ID is available. Kolton et al. (2001) report acceptable internal consistency and adequate item-to-total correlations for the adapted-ABCS, but they do not report the values. Keeling, Rose and Beech (2006) reported significantly lower scores on the Adapted-ABCS post-treatment suggesting that it is sensitive to treatment change, although this was with a small sample ($n=11$). Keeling, Rose, and Beech (2007b) did not report the alpha for the overall CSS. However, only one of the two adapted CSS subscales, *Identification with Criminal Others* (IOC), correlated with its equivalent on the original scale ($r=.60$) but this scale did not possess a good level of internal consistency ($\alpha=.26$). The other subscale, *Attitude Towards the Law* (ALPC), had acceptable internal consistency ($\alpha=.79$) but was not significantly correlated with its equivalent on the original scale. Consequently, the CSS does not appear to be reliable measure of pro-offending attitudes in sex-offenders with ID. In addition, there is no evidence of the validity of the scores obtained on these tools with this population.

Williams et al. (2007) reported that the SOOT had good internal consistency ($\alpha=.82$) with a good sample size ($n=211$) of male offenders with IQs below 80. This alpha was only achieved after deleting one item with a very low item-total correlation. A PCA identified a two-factor solution which accounted for 36.5% of the variance. The factor 1, representing, *Deceitful women and children* and factor 2 *Children, sex and the law* both displayed good levels of internal consistency: Factor 1, $\alpha=.80$ ($n=211$), and Factor 2, $\alpha=.83$ ($n=211$). In addition, scores on factor 2 were significantly higher for child molesters when compared to

rapists pre-treatment. However, post-treatment there was no significant difference in scores between child molesters and rapists. Williams et al. (2007) reported significantly lower scores on the SOOT post treatment, with large effect sizes for the SOOT total and factor 1, and a small effect size for factor 2. These findings suggest that the SOOT can differentiate between child molesters and rapists and is sensitive to treatment change.

Williams et al. (2007) also reported on the Adapted Victim Empathy Task (OBPU, HMPS, 1996). This scale was an adaptation of the Victim Empathy Questionnaire (Bowers, Mann, & Thornton, 1995). The authors reported acceptable internal consistency ($\alpha=.75$) (Kline, 1999). A principal components analysis (PCA) found that one component explained the data adequately, which accounted for 36.9% of the variance. Due to the small number of items in this scale, nine-items, a factor analysis was not performed.

There are a range of methodological difficulties associated with the studies mentioned which mean that caution must be used in interpreting the results. For example, it was previously mentioned that the QACSO and the SOSAS did not predict re-offending (SOTSEC-ID, 2010). However, the sample size in this study was limited ($n=46$) and consequently the re-offending rates were very low; three participants re-offended during the intervention period and four participants during the follow-up which will limit the predictive ability of the instruments. In addition, as the sample was drawn from both community and institutionalised settings it is likely that there was a differential opportunity to re-offend, with potentially more of the high-risk men still detained and consequentially having less opportunity to re-offend. This again may have affected the predictive ability of the instruments.

Many of the studies reported include small sample sizes, with eight including samples of 22 or less (Dukes & McGuire, 2009; Hays et al., 2007; Jacoub & Hall, 2008; Keeling et al., 2007; Lindsay & Smith, 1998; Lindsay, Marshall, Neilson, Quinn, & Smith, 1998;

Lindsay, Neilson, Morrison, & Smith, 1998; Murphy et al., 2007). Consequently, caution must be exercised when drawing conclusions from much of this work. However, unlike the work on anger, aggression and fire-setting there have been a significant number of studies reporting findings with good sample sizes, with thirteen including sample sizes of 46 or over (Broxholme & Lindsay, 2003; Galea, Butler, & Iacono, 2004; Keeling et al., 2007; Kennedy et al., 1999; Kolton et al., 2001; Lindsay et al., 1992; Lindsay et al., 2007; Lindsay et al., 2008; SOTSEC-ID, 2010; Steptoe et al., 2006; Talbot & Langdon, 2006; Williams et al., 2007).

Another problem which makes it difficult to draw conclusions and make comparisons between studies is the inconsistent reporting of the intellectual ability of the sample included in the study. Several studies make no reference to IQ (Jacoub & Hall, 2008; Kolton et al., 2001; Lindsay, Belshaw, Culross, Stains, & Michie, 1992; Lockhart, Guerin, Shanahan & Coyle, 2010) and others do not report the range of IQs (Galea et al., 2004; Hays et al., 2007; Lindsay & Smith, 1998; Michie, Lindsay, Martin, & Grieve, 2006; Steptoe et al., 2006; Talbot & Langdon, 2006). Only a minority of studies make reference to which assessment tool was used to determine IQ and report the range as well (Keeling et al., 2007b; Lindsay & Smith, 1998; Lindsay, Neilson et al., 1998; Lindsay et al., 2007; Williams et al. 2007). Only three studies make explicit reference to WAIS (R or III), IQ range and VABS scores (Duke & McGuire, 2009; Murphy et al., 2007; SOTSEC-ID, 2010). These omissions make it difficult to generalise from the results of the study and to make comparisons with other studies. Those studies that do report the IQ range indicate that the IQs of the respective samples vary: 65-78 Broxholme and Lindsay (2003); <70 Forchuk et al (1995); 63-84 Keeling et al. (2007a) ; 55-83 Langdon et al (2007); 64-71 Lindsay, Marshall et al. (1998); 65-71 Lindsay, Neilson et al. (1998); 53-74 Lindsay et al. (2007); 52-83 Murphy et al. (2007) and SOTSEC-ID (2010).

Consequently, the different levels of ability in the samples discussed make comparisons between the studies potentially problematic if scores on different measures vary with IQ.

3.3.2. Socio-affective functioning

An offender's relationships is a specific domain that has received considerable attention with regard to sex-offenders with ID. The majority of studies in this area relate to the experience of relationships, particularly with parents and peers. Psychometrics used in the assessment of socio-affective functioning in sex offenders with ID include the Significant Others Scale (SOS: Power, Champion & Aris, 1988), the Life Experience Checklist (LEC; Ager, 1990), Social Intimacy Scale (MSIS: Miller & Lefcourt, 1982), Relationship Scales Questionnaire (RSQ: Griffin & Bartholomew, 1994), UCLA Loneliness Scale – Revised (UCLA-R: Russell, 1996), and The Relationship Questionnaire (RQ: Bartholomew & Horowitz, 1991).

Of these, the RQ ($\alpha=.95$), MSIS ($\alpha=.89$), and the UCLA-R ($\alpha=.92$) have demonstrated good levels of internal consistency (Keeling et al., 2007a). Williams et al. (2007) reported on an adapted version of the UCLA, called the Adapted Emotional Loneliness Scale (AELS: OBPU, HMPS, n.d.a). Williams et al. reported good internal consistency ($\alpha=.84$) for the AELS and using PCA identified a single component accounting for 29.9% of the variance. In addition, the MSIS had good test-re-test reliability ($r=.93$) and was moderately correlated ($r=.61$) with the UCLA-R (Keeling, Rose & Beech, 2007a). This correlation with the UCLA-R provides some evidence of the validity of scores on the MSIS and the UCLA-R in this sample as both tools assess similar constructs. The RSQ does not have acceptable levels of internal consistency, with only one of its subscales reaching acceptable levels (Avoidant-Fearful $\alpha=.72$). However, significantly higher scores have been reported on the RSQ following treatment (Keeling, Rose & Beech, 2007b), suggesting that it is sensitive to

treatment change. However, this evidence must be viewed with some caution as the sample in this study was small ($n=22$). Williams et al. (2007) reported that scores on the AELS did not significantly change over treatment. No other evidence regarding the reliability or validity of these tools is available.

The Significant Others Scale (SOS: Power, Champion & Aris, 1988) and the Life Experience Checklist (LEC; Ager, 1990) are reported by Steptoe, Lindsay, Forrest & Power (2006). The authors do not report any data regarding the reliability of these tools. However, both measures produced interesting findings; sex offenders scored significantly lower for the leisure and relationship sections of the LEC, and scored lower than non-sex offenders on actual and ideal levels of support from mother and father. Both of these findings suggest a level of social isolation also found in mainstream sex offender populations and perhaps represent areas for further investigation. In addition, the Adapted Self Esteem Questionnaire (ASEQ: OBPU, HMPS, n.d.b.) has been used with sex offenders with ID (Williams et al, 2007). The ASEQ had good internal consistency ($\alpha=.77$) and PCA identified a single factor accounting for 38.4% of the variance. Scores on the AELS were also significantly higher post-treatment, with a medium effect size, suggesting that it is sensitive to treatment change (Williams et al, 2007).

Another area of socio-affective functioning that has received attention in the literature regarding sexual offenders with ID is sexual knowledge. This area is, at least in part, related to the hypothesis of counterfeit deviance (Hingsburger, Griffiths & Quinsey (1991). Counterfeit deviance refers to the expression of behaviour that is undoubtedly sexually deviant but is the result of factors such as lack of sexual knowledge, lack of social understanding, limited opportunities and skills to develop relationships, and sexual naivety. Studies relating to this hypothesis will be discussed later in this chapter. First, the range of tools used to assess sexual knowledge in samples with ID will be considered.

Tools used to assess sexual knowledge in samples with ID include: the Assessment of Sexual Knowledge (ASK; Galea et al., 2004); the Socio-Sexual Knowledge and Attitude Test (SSKAAT; Wish, McCombs & Edmonson, 1979) which has since been revised (SSKAAT-R; Lunsky, Frijters, Griffiths, Watson & Williston, 2007), the Bender Sexual Knowledge Questionnaire (BSKQ; Bender, Aitman, Biggs & Haug, 1983), The SexKen-ID (McCabe, Cummins & Deeks, 1999); and the Sexual Attitudes and Knowledge Scale (SAKS; author unknown).

The SSKAAT and SSKAAT-R have been used in several studies (Lunsky et al., 2007; Michie, Lindsay, Martin & Grieve, 2006). The SSKAAT-R has seven subscales: Anatomy, Women's Bodies, Men's Bodies, Intimacy, Pregnancy, childbirth and child rearing, Birth Control and STDs. In addition, it has a section on Healthy Boundaries and a number of Attitude items addressing issues such as masturbation and pornography. The SSKAAT-R has good internal consistency ($\alpha=.90$), inter-rater reliability ($r=.94$) and test re-test reliability ($r=.95$). In addition, the SSKAAT-R has some evidence supporting the validity of its scores. It was strongly correlated with its predecessor, the SSKAAT and the SexKen-ID, and showed improvements on all sub-scales following treatment (Griffith & Lunsky, 2003).

Galea et al. (2004) developed the ASK, which comprises four sections: Knowledge Section; Attitudes Section; a Quick Knowledge Quiz; and a Problematic Socio-Sexual Behaviours Checklist. The areas covered by the assessment include body parts, public and private activities, masturbation, sexuality, sexual health and legal issues regarding sexuality. The assessment has 124-items comprising open-questions with prompts, which are then coded on a three-point scale. A sub-sample ($n=33$) was blind rated after two administrations of the ASK. Inter-rater reliability was reported at 92% and 95% respectively and test re-test, over a one to two week interval was 83%. Correlations between sections from the two

completions were strong ($r_s=.83-.99$). In addition, using a comparatively good sample size ($n=96$) the authors report good overall internal consistency ($\alpha=.89$).

Talbot and Langdon (2006) used the General Sexual Knowledge Questionnaire (GSKQ), a revised version of the BSKQ. The GSKQ has six sections, covering; physiology, intercourse, pregnancy, contraception, STDs and sexuality. The authors report good internal consistency ($\alpha=.94$). However, one of the sub-scales, physiology pictures, had low internal consistency ($\alpha=.35$), indicating a problem with that scale. Equally, the authors report good split-half reliability ($r=.80$), with most sub-scales showing moderate to strong split-half correlations ($r=.67-.85$). The exception being physiology pictures with a moderate correlation ($r=.43$), again suggesting a problem with that subscale.

McCabe et al. (1999) describe four studies outlining the development of the Sex Ken-ID. The SexKen-ID was designed to assess Sexual Knowledge, Experience, Feelings and Needs in relation to a range of topics; Friendship, Dating and intimacy, Marriage, Body part identification, Sex and sex education, Menstruation, Sexual interaction, Contraception, Pregnancy, abortion, and childbirth, Sexually transmitted diseases, Masturbation, and Homosexuality. With sexual knowledge, Experience, Feelings and Needs assessed in relation to many of these areas the scale has a total of 48 sub-scales. The authors report internal consistency for 28 of the scales ranging from Homosexuality Feelings ($\alpha=.01$) to Body part Knowledge ($\alpha=.96$). Kline (1999) has stated that $\alpha=.70$ is acceptable in the assessment of psychological phenomena. If this criteria is adopted then only 12 of the 28 subscales have acceptable levels of internal consistency. The authors also report test re-test data for 38 scales, of which 33 had significant correlations between administrations (range .41 - .95). The type of correlation was not reported.

The Sexual Consent and Education Assessment (SCEA: Kennedy, 1993) was developed in order to evaluate the effectiveness of individually tailored sex education (Dukes

& McGuire, 2009). Kennedy reported adequate to good levels of internal consistency for the Knowledge-Scale (K) ($\alpha=.89$) and the Safety Practices-Scale (S) ($\alpha=.74$). Equally, test re-test correlations were strong after a three-year interval: K-Scale 84% ($\kappa=.96$); S-Scale 80% ($\kappa=.92$). In addition, Dukes and McGuire (2009) reported that scores on the SCEA were significantly higher following treatment, providing some evidence for the validity of the scores in terms of their sensitivity to treatment change.

Forchuk, Martin and Griffiths (1995) reported on the development of the Sexual Knowledge Interview Schedule (SKIS). The SKIS is a 46-item interview schedule with an abuse scale and knowledge scale. The knowledge scale includes feelings, body parts identification, body parts function and general sexual knowledge. The SKIS is an interview schedule which uses open-questions, which the interviewer then scores against defined categories. The internal consistency for the subscales was good: abuse subscale ($\alpha=.96$); knowledge subscale ($\alpha=.90$); feelings subscale ($\alpha=.92$); body parts identification subscale ($\alpha=.86$); body parts function subscale ($\alpha=.88$); and the general sexual knowledge subscale ($\alpha=.78$). Inter-rater reliability was 95.3% agreement and test-retest reliability, over a three to seven day interval, was 70.1%. There is no other evidence relating to the reliability or validity of this assessment.

The SAKS is a tool with no accredited author that appears to have been developed specifically for people with ID. It consists of 19 pictures with accompanying questions on sexual knowledge and attitudes. The questions cover four areas: understanding relationships; social interactions; sexual awareness; and assertiveness. Using a modest sample ($n=34$) Langdon et al (2007) reported acceptable internal consistency for the scale ($\alpha=.82$).

Broadly speaking the SSKAAT-R, the ASK, the SexKen-ID and the BSKQ cover similar areas of sexual knowledge and attitudes, with the SSKAAT-R, SexKen-ID and the ASK being the most comprehensive. However, this makes them time consuming and

resource intensive to administer and has perhaps led to the use of shorter assessments that are easier to administer, in some studies (SOTSEC-ID, 2010; Talbot & Langdon, 2006). Both the ASK and the SSKAAT-R are comprehensive, using a range of response formats including open-questions that are then coded according to predetermined categories. Consequently they take a minimum of two hours to administer and an hour to score. The SexKen-ID is a 284-item assessment, with yes/no and 5-point Likert-scale response formats. The authors state that the assessment can be administered in approximately one hour. In contrast the SAKS and the BSKQ are abbreviated assessments, that are short and easy to administer and which can be delivered in a time-frame more suited to research. Galea et al. (2004) reported impressive reliability for the ASK but at present no studies of its validity have been conducted. The reliability data for the SSKAAT are equally impressive and scores on the SSKAAT correspond with scores on other measures in relation to sex-offenders and non-sex-offenders (Lockhard, et al., 2010; Lunsky et al., 2007; Michie et al., 2006; Talbot & Langdon, 2006) providing some validity for those scores in this context. The published studies also included larger sample sizes for the ASK ($n=96$) and the SSKAAT-R ($n=276$) and therefore the findings are likely to be more robust. The SexKen-ID has the advantage of being available in a format suitable for those with physical disabilities and those in the mainstream non-ID population and is therefore a sound choice if comparisons between samples are required. However, the SexKen-ID study only included a small sample of people with ID ($n=30$). In addition, most of these studies included people with IQs over 70, including: people in the borderline IQ range (Galea et al., 2004; Lunsky et al., 2007); and people in the low average IQ range (Langdon et al., 2007; Murphy et al., 2007). However, only Murphy et al. (2007) included the assessment of adaptive functioning, by using and reporting VABS scores. In addition, Lunsky et al. (2007) and McCabe et al. (1999) did not specify the means of identifying ID used in their studies. Consequently, despite some good

evidence of reliability it is difficult to determine whether that was due to differences in the tools used or differences in the samples used. Some of these tools have produced promising reliability data but further research is required to assess their validity. In addition, the development of a reliable and valid but concise tool would be of considerable benefit to researchers and clinicians.

As mentioned previously, several studies that incorporate self-report measures of sexual knowledge and socio-sexual skills and attitudes, report findings with a direct bearing on the hypothesis of counterfeit deviance (Lockhard et al., 2010; Lunskey et al., 2007; Michie et al., 2006; Talbot & Langdon, 2006). All of these studies reported evidence contrary to the hypothesis of counterfeit deviance. Michie et al. (2006), using the SSKAAT, reported significantly higher scores on the birth control and masturbation sections for sex offenders with ID compared to non-sex-offenders with ID. Talbot and Langdon (2006), using the BSKQ, reported that there was no significant difference between the sexual knowledge of treated and untreated sex-offenders. However, treated and untreated sex-offenders with ID scored significantly higher than non-offenders with ID. Lunskey et al. (2007), using the SSKAAT-R, reported that more serious sexual offenders (paedophiles, rapists or those who had committed repeated and forced sexual assaults) scored significantly higher on all six subscales, when compared with less deviant offenders (inappropriate touching, public exhibitionism or public masturbation). Lockhart et al. (2010) reported that when adaptive behaviour was controlled as a covariate, a group categorised as displaying sexualised challenging behaviour had significantly higher scores on the Birth Control, STDs and Healthy Socio-sexual Boundaries, then a group categorised as displaying non-sexual challenging behaviour. None of these studies reported any subscales where non-offenders or less deviant offenders scores higher than sexual offenders or more serious sexual offenders respectively. All of these findings suggest that lack of socio-sexual knowledge is not related

to committing sexual offences. The convergence of scores on these different measures provides some evidence of the validity of those measures.

3.3.3. Self-management problems

In terms of self-management problems the Adapted Relapse Prevention Interview (ARPI: OBPU, HMPS, n.d.c.) has been used to assess an individuals' offence specific coping strategies. Also, the Social Problem-Solving Inventory-Revised (SPSI-R: D'Zurilla & Nezu, 1990) has been used to assess more generalized social problem-solving skills. The ARPI was an adapted version of Beckett, Fisher, Mann, and Thornton's (1997) Relapse Prevention Interview. Williams et al. (2007) reported that the ARPI has good internal consistency ($\alpha=.86$) ($n=211$). In addition, an exploratory factor analysis identified only one meaningful factor accounting for 31.4% of the variance. In addition, scores on the ARPI were significantly higher post treatment with a large effect size.

When assessing social problem-solving deficits it is necessary to consider both the component processes involved in problem solving and the outcomes of problem-solving efforts (Nezu, Nezu & Dudek, 1998). Identification of deficits in the component process will dictate where interventions may be usefully targeted. Equally, assessment of the outcomes of problem-solving efforts may help to identify inappropriate or unhelpful goals that direct the individual's problem solving strategies. Interventions can then be directed towards problem identification and solution generation specific to the individual. Two self-report measures are available, each of which fulfills one of these components of the assessment of problem deficits. The Social Problem-Solving Inventory-Revised (SPSI-R: D'Zurilla & Nezu, 1990) was designed to assess the various component processes involved in problem-solving and the Social Problem-Solving Task (Nezu, Nezu & Arian, 1991) can be used to evaluate the actual outcome of problem-solving efforts.

The SPSI-R can be delivered in two formats; the full 52-item assessment or a shorter, 25-item version. The structure of the SPSI-R comprises two constructive and three dysfunctional dimensions. The two constructive dimensions are positive problem orientation and rational, or systematic, problem-solving style. The three dysfunctional dimensions are negative problem orientation, impulsive/careless style, and avoidance style. The SPSI-R utilizes a 5-point (0-4) Likert-scale, with zero indicating '*not at all true of me*' and four indicating '*extremely true of me.*' Lindsay et al. (2009) report on the use of an adapted version of the short form of the SPSI-R. The short form of the SPSI-R was adapted so that items could be more easily understood by individuals with ID. Test re-test reliability was acceptable ($r_s=0.79$) over a two week interval and internal consistency was also good ($\alpha=0.84$), suggesting that the 25 items converge on a single concept. In addition, utilising a sample of 132 people with IQs ranging from 52 to 85 the authors identified a four factor structure entirely consistent with D'Zurilla et al's. (2002) original study. In addition, evidence of the validity of scores on the adapted tool was provided by improvements on three sub-scales following a social problem-solving programme. The results indicated that participants demonstrated a more positive problem solving orientation, a less avoidant style and were less impulsive. The first part of this study was well controlled with data from 132 participants used to investigate the factor structure of the SPSI-R. However, the validity component of this study included only 10 participants from a social problem-solving programme and consequently the evidence is somewhat limited.

The Social Problem-Solving Task (SPST: Nezu et al., 1991) presents five common interpersonal problematic situations commonly faced by people with ID (e.g. making friends and conflict resolution). Participants are asked: *What is the actual problem?*; *Think of as many ways to solve the problem as you can?*; *What are the positive and negative consequences of each idea?*; and, *Which idea do you think will solve the problem?* All

responses are later scored on a five point scale (1=low quality; 5=high quality). The authors report good inter-rater reliability ($r=.83$) and test re-test, over an unspecified interval ($r=.79$). Nezu, et al. (1991) provided some evidence of the validity of scores on the SPST, when they reported on the efficacy of assertiveness and problem-solving training for people with ID. Self-report measures used included the Brief Symptoms Inventory (BSI: Derogatis, 1993), a Social Problem-Solving Task (SPST) and a Subjective Units of Distress Scale (SUDS). The BSI measures psychological symptoms that yield three global indices and nine dimensional scales. The General Severity Index (GSI) is the most sensitive of the global scales and provides an overall assessment of changes in psychiatric symptoms as a function of treatment. The authors reported adequate to good internal consistency across dimensions ($\alpha=.71-.85$) and strong test re-test reliability ($r=.90$) for the GSI. The SUDS was described as an *upset thermometer*; a diagrammatic representation of a thermometer, with marks ranging from 1-10 on which participants rated their feelings of distress for the previous week. Following a problem-solving intervention, and using regression analysis, the authors report that improvement in SPSI scores was related to decreases in BSI and SUDS scores. This indicates that, as expected, improved problem-solving was associated with reduced feelings of psychological distress. In addition, Basquill et al. (2004) reported results using the SPST consistent with previous findings providing validity for the scores. Aggressive participants scored significantly lower in terms of generating positive and negative consequences for potential problem-solving solutions. In addition, aggressive participants generated more aggressive responses to hostile, non-hostile and ambiguous social situations.

The investigation of social problem-solving is proving a promising area for researchers concerned with sexual and aggressive offending in samples with ID. In particular, the adapted version of the SPSI-R provides a reliable tool to assess social problem solving in samples with ID and would appear to be a useful avenue for future research.

However, at present there are very few self-report measures available to assess potential deficits in samples with ID. These are encouraging results and suggest that similar processes are operating in samples with ID as have been investigated in mainstream non-ID samples. However, there are some methodological problems with the studies suggesting that caution is required when interpreting the results. For example, Basquill et al. (2004) identified aggressive participants using the Aggressive Behaviour Scale of the Reiss Screen (Reiss, 1988) and so identification was not based on criminal behaviour or seriously aggressive behaviour. Consequently, it is unclear to what extent the aggressive participants were actually aggressive as this was not measured in an objective sense. In addition, the assessments were based upon verbal tests and yet no assessment was made of verbal abilities or consideration given to potential differences in verbal abilities between groups. Consequently, it is possible that differences between groups are due to differences in verbal abilities. In addition, the most significant weakness of the self-report measures of problem solving, such as the SPSI-R and the SPST, is that they do not objectively measure an individual's social problem-solving skills, but how individuals' view problems and themselves as problem-solvers. It is not clear whether self-appraisal of problem-solving skills reflects actual problem-solving skills or some other underlying construct, and if it does reflect problem-solving skills, whether or not such self-appraisal is accurate.

3.4. Measuring locus of control

The assessment of LOC in non-offender samples with ID is reported in Chapter 4. There is little evidence of the measurement of locus of control in offenders with ID. Rose, Jenkins, O'Connor, Jones and Felce (2002) reporting on a small case study series considering the effectiveness of a group intervention for men with a history of sexual offending reported that locus of control orientation became more external following treatment. Langdon and Talbot

(2006) however, in a small-scale study of sexual offenders with intellectual disabilities, found that locus of control orientation did not change during treatment, remaining externally orientated. Both of these studies use the Adult Nowicki-Strickland Internal-External locus of control scale (ANSIE) (Nowicki & Duke, 1974). In both the Rose et al (2002) and Langdon et al (2006) studies no reliability or validity data were reported in relation to ID offenders specifically.

Goodman, Leggett and Garrett (2007) report on the development of a locus of control scale for people with an intellectual disability. The authors identify two assessment tools measuring locus of control, namely the ANSIE (Nowicki, 1976) and a 17-item scale developed by Craig, Franklin and Andrews (1984). The authors state that both of these measures include three subscales relating to powerful others, internal control and luck or chance. During a pilot study the authors concluded that people with an intellectual disability struggled with the concept of luck or chance, for example, not understanding the symbolism of the four-leaf clover referred to in the ANSIE. Consequently, they developed a 20-item scale with just the two subscales of powerful others and internal control. They developed 16 questions based on concepts from the two scales, simplifying the language used, making them shorter and more relevant to the lives of people with learning disabilities. In addition, they developed 4 cartoon scenarios related to locus of control concepts which were used to illicit an open, unstructured response from participants. This open response was then rated for orientation. The measure was completed by 102 people with ID, 75 of whom were men and 27 were women. Of these, 61 had previously had contact with the criminal justice system (CJS), of which 22 had convictions. The scale showed improvements over the original scales on readability statistics. In addition, the scale showed moderate internal consistency ($\alpha=.53$) and moderate test-retest reliability ($r=0.66$) over an eight week period. The sixteen written questions were subject to a principle components factor analysis with rotation which

generated 7 factors accounting for 65.6% of the variance. The authors report that the two expected subscales of personal control over behaviour and the perceived role of powerful others in determining outcome were identified. No other information about the structure of the factors or how these two subscales were determined was included. No reliability or validity data were presented in relation to the cartoon scenarios. In addition, the results section makes no reference to data from these scenarios being included in the factor analysis. Subsequently, the scale was found to differentiate offenders with a conviction from non-offenders.

The findings of these studies are limited by the sample sizes. Goodman et al (2007) used a larger sample size ($n=102$) but this only included 22 convicted offenders. Clearly, the questionnaire is also limited in that it is only assessing part of the concept of locus of control, with no reference to luck or chance, consequently it is likely to have low content validity. The addition, the level of internal consistency reported by the authors is unacceptably low.

3.5. Measuring empathy

There exists a small body of literature indicating that ID may be associated with difficulties decoding emotional cues beyond general cognitive abilities (Hobson, Ouston & Lee, 1989; Rojahn, Rabold & Schneider, 1995). Indeed, poor facial emotion decoding has been identified in ID populations (Holder & Kirkpatrick, 1991). Rojahn and Warren (1997) have found significant positive correlations between level of intellectual disability, the ability to match facially expressed emotion and psychometrically measured empathy. In this instance empathy was measured using the Appropriate Skills sub-scale of the Social Performance Survey Schedule (Matson, Helsel, Bellack & Senatore (1983). However, the correlations between level of ID and facial emotion expression ($r=.49$) and empathy ($r=.34$) were only moderate. Woodcock and Rose (2007) used the AI (Benson & Ivins, 1992) and reported that

high self-report anger did not impair facial expression recognition or increase the interpretation of facial expressions as more hostile compared to those with low self-report anger. The findings are limited by the small sample size ($n=30$). In addition, high-levels of anger were determined by self-report only. The authors did not appear to assess IQ or the level of adaptive functioning. Inclusion was based upon individuals accessing a social education centre. The test is also cross-sectional in design and is therefore not sensitive to state dependent changes to performance.

Proctor and Beail (2007) examined empathy and theory of mind in a sample of offenders with ID in comparison with a sample of offenders without ID. The tests of empathy used were the Interpersonal Reactivity Index (IRI: Davis, 1983) and the Test of Emotional Perception (TEP: Moffatt, Hanley-Maxwell & Donnellan, 1995). The IRI was modified to simplify the language and concepts used and to eliminate ambiguity. In addition, a slight modification was added to the 5-point Likert-scale, where diagrammatic representation of a cross to signify no and a series of ticks in ascending size order to depict levels of agreement. No reliability and validity data exists for this tool with a sample with ID participants. Theory of mind tests included the Deceptive Box (Smarties) Task (Gopnick & Astington, 1988), the Location Change (Sally-Anne) Task (Wimmer & Perner, 1983) and the Second Order Location Change (Ice Cream Van) Task (Perner & Wimmer, 1985). The results indicated that offenders performed better than non-offenders on the second order theory of mind task and on emotion recognition and required fewer prompts to provide an emotional descriptor. This is an interesting study which set out to define and explore the relationship between empathy and theory of mind in people with ID. The findings of this study bring into question the practice of empathy training for offenders with ID as their skills in this area appear to be superior to those of non-offenders. However, the conclusions drawn must be tentative due to the small sample size, with both groups containing 25 participants.

In addition, the IRI was adapted for this population but reliability and validity data were not reported.

3.6. Measuring impulsivity

Parry and Lindsay (2003) and Snoyman and Aicken (2011) have both used a modified version of the Barratt Impulsiveness Scale (11th edition) in order to determine the level of self-report impulsivity in samples of offenders with ID. These studies are described in chapter one of this thesis (pages 20-21).

3.7. Measures of mental health

Kellett et al. (2003) report on the use of the Brief Symptom Inventory (BSI; Derogatis, 1993), an instrument used in the assessment and detection of mental health problems with individuals with ID. The BSI is a shorter version of the Symptom Checklist 90 – Revised (SCL-90-R; Derogatis, 1983), one of the most widely used measures in adult psychotherapy outcome studies (Lambert & Hill, 1994). The psychometric properties of the BSI, when used to assess people with ID, reported in Chapter 2, indicate that this is a reliable assessment of psychological distress for people with ID, with some evidence in support of its validity.

Kellett et al., (2003) reported that a forensic sample with ID scored significantly higher than a community sample with ID on the obsessive-compulsive, depression, hostility and psychoticism symptom scales. However, a clinical group displaying signs of psychological distress, endorsed significantly more symptoms of somatisation, interpersonal sensitivity, anxiety and phobic anxiety than the forensic group.

Lindsay and Skene (2007) adapted the Beck Anxiety Inventory (BAI) and the Beck Depression Inventory – 2nd Edition (BDI-II). The psychometric properties of these adapted instruments, reported in Chapter 2, indicate that these instruments are reliable measures of

anxiety and depression respectively, when used with people with ID. In a comparison of subsamples, the authors found that sex offenders reported significantly lower levels of anxiety and depression than those who had offended non-sexually. Lindsay and Lees (2003) administered the Beck Anxiety Inventory (BAI: Beck & Steer, 1990) and the Beck Depression Inventory (BDI: Beck et al., 1981) to 16 sex offenders with ID and a control group of 16 non-offenders with ID, as reported in Chapter 2. Slight modifications were made to the measures with simplified language and sentence structures. In addition, a visual bar graph depicting differing degrees of agreement was used to facilitate responding. The authors report strong correlations on test re-test for both measures. In addition, the mean score for sex offenders were significantly lower on both tools than the non-offending controls, similar to results from a study of sex offenders of normal intelligence (Valliant & Antonowicz, 1991).

Benson and Ivins (1992) used the Birleson Depressive Self-rating Scale (DSS: Birleson, 1981) and Howells et al. (2000) used the Hospital Anxiety and Depression Scale (HADS: Zigmond & Snaith, 1983). Benson and Ivins (1992) reported that self-report depression scores on the DSS were weakly correlated ($r=.26$) with an informant rating of depression. Howells et al. reported little evidence relating the reliability or validity of the tools used. They did report a significant reduction in mean scores on the HADS after treatment, however this was with a very small sample ($n=5$). There is little evidence provided in these studies to support the validity of these tools.

These are promising results supporting the reliability and validity tools when used with people with ID. However, this study used a small sample size ($n=32$) and included offenders in the borderline IQ range.

3.8. Discussion and conclusions

Several measures have been developed or adapted with regard to the assessment of anger and aggression in people with ID: NAS (Novaco, 2003); the PI (Novaco & Black, 2003); the AI (Benson et al., 1986); the STAXI (Spielberger, 1996), was used in two studies (Novaco & Talyor, 2004; Taylor, Novaco, et al., 2004); the DPI (Alder & Lindsay, 2007); and the IPT (Taylor, Novaco, et al., 2004). Most of these tools displayed good levels of internal consistency ($\alpha=.70-.82$) although the Anger Regulation subscale of the IPT was at the lower limit of acceptability ($\alpha=.70$) (Kline, 1999). No internal consistency has been generated for the AI. In addition, there is some evidence of the validity of some of these tools (e.g. NAS, the DPI and the IPT), provided by weak to moderate inter-correlations between scales (Alder & Lindsay, 2007; Taylor, Novaco, et al., 2004). In addition, some interesting work has been conducted in relations to threat appraisal in the expression of aggressive behaviour in samples with ID (Jahoda et al., 1998; Pert & Jahoda, 2008). However, further work is required in order to establish the psychometric properties of the SGASC. The assessment of the psychometric properties of the DPI is well advanced compared to other measures used in samples with ID. The tool demonstrates good reliability, structural qualities and promising validity (Alder & Lindsay, 2007).

In contrast to the development and adaptation of reliable and valid tools for the assessment of anger and aggression in people with ID, there have been no tools developed to assess fire-setting behaviour in this population that have acceptable levels of reliability or evidence of validity.

The psychometric assessment of sexual offenders with ID was considered in relation to three broad areas of treatment need: pro-offending attitudes; socio-affective functioning; and self management problems. In terms of pre-offending attitudes, only the QACSO has demonstrably good psychometric properties; with good internal consistency, test re-test

reliability and inter-correlation of subscales (Lindsay et al., 2007, 2008). Equally, scores on the QACSO have been shown to be significantly lower following treatment (Murphy et al., 2007; SOTSEC-ID, 2010) and at follow-up (SOTSEC-ID, 2010), indicating that it is sensitive to treatment change, providing evidence of its validity.

In relation to socio-affective functioning the RQ, MSIS and the UCLA-R have all demonstrated good levels of internal consistency (Keeling et al., 2007b), but little evidence has been presented in relation to the validity of those scores in a forensic context. A similar pattern holds for the assessment of sexual knowledge with the SSKAAT-R, ASK, GSKQ, SCEA and SKIS all demonstrating good levels of internal consistency, and for some (ASK, SCEA and SKIS) good test re-test reliability. However, little evidence is available regarding the validity of the scores on these measures. However, the large sample size used in the development of the SSKAAT-R ($n=276$) and its ability to differentiate sex-offenders from non-sex-offenders suggest that it has promising utility (Lockhard et al., 2010; Lunskey et al., 2007; Michie et al., 2006; Talbot & Langdon, 2006).

The SPSI-R has good levels on internal consistency and acceptable levels of test re-test reliability, in assessing social problem-solving skills. In addition, evidence of the validity of scores on the adapted tool was provided by improvements on three sub-scales following a social problem-solving programme (Lindsay et al., 2009), suggesting that it may have some utility in the assessment of inter-personal problem solving.

No tool has demonstrated acceptable levels of reliability in relation to the assessment of LOC or empathy in samples with ID. Two studies have used the BIS-11 to assess impulsivity in offenders with ID (Parry & Lindsay 2003; Snoyman & Aicken, 2011). Snoyman and Aicken (2011) reported good levels of internal reliability for the BIS-11. In addition, both studies identified significantly higher levels of self-report impulsivity in

violent offenders as opposed to sex offenders, the correspondence of that pattern providing some evidence of the validity of the scores.

In addition, as was outlined in Chapter 2 there are several tools assessing mental health that have good evidence of reliability and validity. Those worth noting that have been used in relation to offenders with ID include the BSI (Kellet et al., 2003), the BAI and the BAI, indeed as noted in Chapter 2 the psychometric properties of the BAI and the BDI-11 when used with samples with ID have been well developed (Lindsay & Lees, 2003; Lindsay & Skene, 2007). The evidence presented above suggests that very few psychometric tools developed for use with offenders with ID have established good levels of reliability and consistent evidence supporting the validity of the scores with samples of offenders with ID.

CHAPTER FOUR: A REVISION OF THE NOWICKI-STRICKLAND INTERNAL-EXTERNAL LOCUS OF CONTROL SCALE FOR OFFENDERS WITH AN INTELLECTUAL DISABILITY

4.0. Introduction

The aim of this chapter was to improve the readability of the Adult Nowicki-Strickland Internal-External locus of control scale (ANSIE). The revision was intended to increase the validity of responding in a sample of male offenders with ID. This was done to minimise the impact of the response biases outlined in Chapter 2 which present a major threat to the validity of self-report of people with ID. Research suggested that the validity of the original scale was likely to be impaired by response biases such as acquiescence which are prevalent in ID populations (Sigelman, Budd, et al., 1982). Acquiescence can occur when an individual does not understand the question or is uncertain about which response corresponds with their belief. The original questionnaire was developed for people without ID and consequently some questions were complex or did not match the living context of offenders with ID. It was believed that acquiescence could be countered by making the questionnaire easier to understand and answer. Consequently, attempts were made to improve the readability of the ANSIE by simplifying the language and sentence structures with slight changes in content to match the social context of detained offenders with ID. The simplification process was validated by comparing the original and adapted measures using the Flesch Reading Ease formula (Flesch, 1948) and the Flesch-Kincaid Grade Level (Kincaid, Fishburne, Rogers, & Chissom, 1975) (both of these statistics were discussed in Chapter 1 (pp.59)). Subsequently, 47 offenders with ID completed both the Modified-ANSIE (M-ANSIE) (on two occasions) and the ANSIE.

The study described in this chapter is divided into three sections. In the first part, the two versions of the questionnaire were compared using standard readability statistics, participant ratings and procedural indicators. In addition, the respondents' scores on the two completions of the M-ANSIE were compared to ascertain test retest reliability. In the second part the respondents' total scores on the two measures were compared to ascertain construct validity for the modified measure. In addition, the patterns of responding were examined to assess whether the modifications had reduced acquiescent responding. In the third part of the study an exploratory factor analysis using a principal components analysis, with an orthogonal, varimax rotation, was conducted, using a larger sample ($n=139$). The identified factor structure was compared with existing factor analyses of the ANSIE in mainstream non-ID populations and a study utilising the ANSIE with a sample of people described as having mental retardation. The findings are discussed.

LOC appears to be an important factor in the assessment of risk of offending and successful treatment change in offenders without ID (Beech et al., 1998; Johnson & Berry, 1989; Page & Scalora, 2004), as outlined in Chapter 1. The present study was driven by the necessity to psychometrically assess LOC accurately in a population of offenders with an ID. In order to reduce the threat to validity associated with response biases, outlined in Chapter 2, it is apparent that self-report inventories require adaptation for use with intellectually disabled offenders (Clare, 1993).

4.1. LOC and ID

LOC is a construct derived from social learning theory and is regarded as a fundamental personality characteristic (Rotter, 1966). It is measured on a continuum from internal to external LOC referring to the extent to which people feel that events are contingent on their own behaviour or external factors (Nowicki, 1976). This ultimately dictates the individual's

subjective evaluation of whether or not he or she is responsible for his or her own behaviour. Typically, someone functioning well has a LOC that tends to be slightly internal but can be flexible depending upon the situation (Straus, 1994). Individuals with an externally orientated LOC think they have little control over their lives, believing instead that luck, chance, fate or powerful others have more control over events than they do (Mercer & Snell, 1977).

Correlations between LOC and IQ in populations within the normal IQ range have been consistent but wide ranging in effect sizes (Brown, 1976; Crandall, Katkovsky, & Crandall, 1965; Hallahan, Gajar, Cohen, & Tarver, 1978; Milgram & Milgram, 1975; Minton, 1967; Samuel, 1980; Olendick & Olendick, 1976; Powell & Centa, 1972). For example, Brown (1976), Hallahan et al. (1978) and Olendick and Olendick (1976) all report significant moderate or large positive correlations respectively (range $r = .39-.62$), whereas Crandall et al. (1965) and Samuel (1980) report small (range $r = .16-.24$) but significant positive correlations. While these correlations are positive, it should be noted that the measures of LOC were scored in an internal direction. Consequently, those with higher IQs tended to self-report a more internal orientation than those with lower IQs.

All of these studies have methodological problems making it difficult to rely on the findings. For example, Brown (1976), Olendick and Olendick (1976) and Hallahan et al. (1978) have small sample sizes ($n = 55, 45$ and 28 , respectively). Whilst Samuel (1980) ($n = 416$) and Crandall et al. (1965) ($n = 736$) had good sample sizes their findings are limited by the measures of IQ employed. Samuel reported that IQ was estimated from administration of the performance scales only on the WISC (presumably the Wechsler Intelligence Scale for Children-Revised, Wechsler, 1974) (although this is not specified or referenced in the article). The validity of this procedure is not reported. Equally, Crandall et al. (1965) used the California Test of Mental Maturity (CTMM), an intelligence test administered by one

person to a group in order to save time and resources. Generally, such tests are not considered suitable for people in lower intelligence groups because the researcher is less aware of the needs of participants, who may under-perform due to unrecognised tiredness, confusion over instructions or attention deficits (Furr & Wilson, 1968). In addition, scores on the CTMM have been found to be uncorrelated with individually administered tests of intelligence in a sample of children with ID (Furr & Wilson, 1968), suggesting that it may not be a valid measure of intellectual functioning. Perhaps the greatest limitation of all of these studies, however, is that, except for Powell and Centa (1972), the participants were all of school age (although it should be noted that the actual age ranges are not reported in any of the studies) and therefore the extent to which these findings can be generalised to adult populations is unknown. An exception to this is the study by Powell and Centa (1972), where the relationship between mental ability and two measures of LOC was investigated with an older sample containing college students. Mental ability, as measured by the Henman-Nelson Tests of Mental Ability (No reference was given for this test and the current author has been unable to locate a reference), correlated with Rotter's Internal-External scale (Rotter, 1966) ($r=-.34$) and the Adult LOC scale (Cromwell, 1963) ($r=.34$) respectively. Again, however, this study had a small sample size ($n=43$) and the validity of the Henman-Nelson tests of Mental Ability in identifying IQ is unclear.

Studies sampling individuals with an IQ below the normal range consistently indicate that those with a lower IQs exhibit more externally orientated LOC than matched samples of average intelligence (Fincham & Barling, 1978; Gardner, Warren, & Gardner, 1977; Lawrence & Winschel, 1975; Lewis & Lawrence-Patterson, 1989; Wehmeyer, 1993; Wehmeyer, 1994; Wehmeyer, Kelchner, & Richards, 1996). Whilst all of these studies identified a significant relationship between IQ and LOC orientation they are also beset with methodological difficulties, such as missing data, small sample sizes and the validity of IQ

tests used. For example, Fincham and Barling (1978) described their low IQ group as having a mean IQ of 102 on the Peabody Picture Vocabulary Test (PPVT) (Dunn & Dunn, 1959) which is problematic because PPVT has been found to seriously underestimate IQ in people with ID (Craig & Olson, 1991). It would therefore appear that even the low IQ group in this study did not have an identifiable ID. Gardener et al. (1977) were only able to determine IQ scores for 68% of individuals in some groups in their study. In addition, they measured IQ using the CTMM, already described above in relation to Crandall et al. (1965). Lewis and Lawrence-Patterson (1989) do not report the respective IQs of their groups at all, they simply stated that one group comprised children with ID and one group comprised children without ID. The validity of IQ scores generated in these studies compared with standardised measures, such as the various versions of the Wechsler Adult Intelligence Scale (Wechsler, 1955, 1981, 1997, 2008) is unclear but clearly limits the generalisation of findings. In addition, the sample sizes are small ranging from 32 to 94 individuals spread over 2-3 groups (Fincham & Barling, 1978; Gardner, et al., 1977; Lewis & Lawrence-Patterson, 1989). Again, the samples are predominantly of school age; Fincham & Barling (9 & 10 year olds), Lewis & Lawrence-Patterson (8 & 12 year olds), although Gardener et al. extend their sample to include college students (15-21 year olds). Nevertheless, caution must be exercised in generalising these findings to adult populations.

In a review, Mamlin, Harris, and Case (2001), have challenged the conclusions that people with ID have an external LOC. Mamlin et al. reviewed 22 studies concerning the measurement of LOC in populations with ID published between 1982 and 1999. The review is critical of the studies under consideration in three main areas: firstly, participant selection and description, secondly the LOC measures used and thirdly the conclusions drawn by the authors. For example, Mamlin et al. report that the procedures for participant selection, with specific reference to the measurement and reporting of IQ, and the matching of participants

with appropriate control groups, were inadequate in all studies. They argued that the researchers should include additional procedures to ensure that each participant met the criteria for ID.

In terms of LOC measures, Mamlin et al. (2001) were critical of the lack of appropriate norms for these measures and the populations under consideration. They argue that the norming groups for the Nowicki-Strickland were seriously flawed as they were not representative of the population of the United States, lacking racial, cultural and geographical diversity. Mamlin et al. were also critical of the conclusions drawn by the researchers. They conclude that in 19 of the 22 studies the authors presented interpretations that were not supported by their data or went considerably beyond the data reported. They argue that 'externality' was arbitrarily labelled, despite the original authors simply identifying a spectrum of orientation (Nowicki & Duke, 1974a), as opposed to a demarcation of internal or external locus.

Whilst the findings of Mamlin et al. (2001) appear concerning they are also based upon extremely stringent criteria, the contravention of which do not necessarily negate the quality or indeed validity of research in this area. For example, in terms of the criticisms directed at the LOC measures these do not appear to place the finding of the studies reported in doubt. For example, they argue that tested groups do not match the groups used to norm the original measures. Whilst this is a valid criticism it does not account for why groups containing people with ID consistently score in a more external direction than matched controls. Equally, whilst the groups upon which norms were developed for the LOC scales were not necessarily based upon samples that represent the diversity of the population as a whole, this again does not account for the differences in scores between groups identified as having ID and those without. It appears that Mamlin et al. are predominantly concerned with

the use of the label 'external' as opposed to any suggestion that the scores associated with people with ID are more externally orientated than matched samples without ID.

In terms of the identification of ID, Mamlin et al. (2001) are critical that researchers have relied upon the identification of ID by the local school system without reference to how local school systems actually made those identifications. Mamlin et al. appear to suggest that only researchers can identify ID in a valid manner rather than the way such decisions are made in practice. In addition, it would appear almost impossible to match ID and non-ID groups based upon, for example, the criteria of age, level of education and academic achievement, as Mamlin et al. suggest, bearing in mind that lower levels of education and academic achievement are more likely to be indicative of an ID population (Nowicki & Strickland, 1973; Nowicki & Duke, 1974a). Consequently, strict adherence to the criteria Mamlin et al. suggest is likely to result in discarding a lot of good quality research.

In addition, Mamlin et al. (2001) excluded studies that involved an intervention. The reason for this was not given, although they state that only one study was excluded for this reason. Regardless, the review conducted by Mamlin et al. is not based upon an exhaustive literature review. For example, Wehmeyer and Palmer (1997) compared LOC orientation in people with and without ID, using a large sample ($n=431$). Consistent with Mamlin et al.'s criticism, the authors did not identify ID independently but relied upon identification by the local school district based upon federal criteria. However, the samples were drawn from four US states, which therefore should have reasonable diversity characteristics and were matched for gender and age across three groups (10.25-12.25yrs, 13.25-15.25yrs and 15.25yrs and over). LOC was determined using the CNSIE in the youngest group and the ANSIE in the other two groups. The authors found that across all age groups students with an identified ID scored significantly higher than the respective matched non-ID groups, indicating more external LOC. In conclusion, Mamlin et al. highlight a range of valid criticisms regarding the

study of LOC in people with ID. However, these criticisms are based upon extremely strict criteria, the contravention of which does not negate the overwhelming research finding, that people with ID have a more external LOC than people without ID.

4.2. The measurement of LOC

In order to determine whether LOC is more external in offenders with ID requires valid and reliable measures of LOC that can be used with samples with ID. Several scales, varying in length, type of item and dimensionality, have been developed to measure LOC. The two most common are the Internal-External scale (I-E: Rotter, 1966) and the Nowicki-Strickland Internal-External LOC scale (Nowicki-Strickland, 1973), which has several versions.

Rotter's original I-E scale (1966) probed two independent factors; perceived mastery over the course of one's life and the extent to which the individual is deemed capable of exerting an impact on political institutions (Abramowitz, 1973; Mirels, 1970) indicating the potential multi-dimensionality of the construct. Subsequent formats have identified varying content of up to six factors (Reid & Ware, 1974; Schlegel & Crawford, 1976; Dragutinovich, White, & Austin, 1983; Levenson, 1974; Walkley, 1979; Blau, 1984; Lindbloom & Faw, 1982; Klockars & Varnum, 1974).

The Nowicki-Strickland LOC Scale for Children (CNSIE) (Nowicki-Strickland, 1973) was developed to overcome deficits associated with Rotter's original scale in relation to children's vulnerability to socially desirable responding. Subsequent extension of this instrument into adult (ANSIE) (Nowicki & Duke, 1974a) and pre-school and primary school form (Nowicki & Duke, 1974b) has facilitated greater standardisation when making comparisons of scores over a wide range of ages. The development of different versions of the Nowicki-Strickland Internal/External scale has contributed to it becoming the most common LOC assessment tool in use. The present literature search revealed in excess of

1000 published articles incorporating the scale. The ANSIE is also easy to administer and score and provides a simple unitary measure of orientation. It is therefore the measure of choice for studying LOC norms in alternative adult populations.

Factor analyses suggest that LOC as measured by the ANSIE is multidimensional. Two (Dixon, McKee & McRae, 1976; Piotrowski, Dunn, Sherry, & Howell, 1983), four (Chandler & Dugovics, 1977; Piotrowski, Dunn, Sherry, & Howell, 1983), and five (Finch, Kendall, Spirito, & Mikulka, 1981; Kearney & Kearney, 1983) factor solutions have been identified.

Dixon et al. (1976) provided the first investigation of the factor structure of the ANSIE. A principal components method with varimax rotation yielded a 2 factor solution accounting for only 12.8% of the variance. The authors report that Factor I appeared to represent control of the home – interpersonal environment, and contained six items accounting for 7.2% of the variance. Factor II was concerned with the importance of luck in control, containing three items accounting for 5.6% of the variance.

Chandler and Dugovics (1977) revised the ANSIE in their study, utilising a Likert response format (strongly agree, agree with reservations, disagree with reservations, strongly disagree) rather than the typical dichotomous response format. The initial analysis for male participants identified 12 factors with eigenvalues exceeding Kaiser criterion, accounting for 45.2% of the variance. Factors were retained following varimax rotation if at least two items loaded with a magnitude of 0.40 or greater and if the content of those items was consistent. No information was provided about how judgements of consistency were made. The final factor structure comprised four factors labelled Personal Control, Powerlessness-Helplessness, Blame and Luck, which accounted for 29.1% of the variance. The study's sample size ($n=291$) was adequate for the factor analysis although the suitability of the data beyond this was not reported. In addition, decisions made about factor retention were unclear

and the resulting factor structure accounted for relatively little variance. It is unclear to what extent the atypical use of a Likert response format may have influenced the results.

Piotrowski et al. (1983) used the original ANSIE format. A principal components analysis identified four factors with eigenvalues exceeding Kaiser criterion. Only items loading greater than 0.30 were retained. Following varimax rotation all four factors were retained accounting for 50.7%, 17.5%, 15.5% and 15.3% respectively of the variance (99% in total). The factors were not interpreted or labelled. The authors also produced a two factor solution to compare with the two factor solution of Dixon et al. (1976). The authors report that 6 items from Factor I of Dixon et al. also loaded on factor one in their study, however only one item from Factor II of Dixon et al. loaded on Factor II. This study explained a remarkably high level of variance within the data.

Finch et al. (1981) used the original ANSIE response format. Principal components analysis yielded eight factors with eigenvalues above Kaiser criterion accounting for 80.6% of the variance. However, three factors were not interpreted as they contained single items. Of the remaining 5 factors, Factor I (28%) had the theme of inability to protect oneself, Factor II (11%) lack of social power, Factor III (9.2%) superstition and 'good day thinking', Factor IV (7%) futility and 'why bother?' and Factor V (5.1%) effort as useless and unfairness. This study factor analysed both male and female data together, failing to take account of potential gender differences in relation to factor structure (Dixon et al., 1976).

Kearney and Kearney (1983) again used the original version of the ANSIE. They conducted the study with males and females but conducted factor analyses separately. Principal components analysis produced 16 factors with eigenvalues greater than 1.0, accounting for 72.5% of the variance. These factors were subjected to varimax rotation. Item inclusion required a loading of 0.40 and any item that loaded on more than one factor was only retained in the factor on which it loaded highest. The authors required a minimum

of four items to establish a coherent theme for a factor. This procedure resulted in a five factor solution. Factor I had the predominant theme of powerlessness with respect to peer relationships and emphasised the inability to protect oneself from the negative feelings of others. The theme of Factor II was helplessness at home whilst Factor III emphasised hard-work vs. passivity. Factor IV appeared to refer to a generalised belief in luck, whilst Factor V was labelled as representing a sense of helplessness or futility.

The studies above appear to indicate a lack of consistency in the factor structure associated with the ANSIE. From comparing the factor structures from each study the current author was able to ascertain that on only 23 from a possible 400 occasions did a single item load on to the same factor in two or more studies. From the information provided in the studies it is difficult to draw firm conclusions about the factor structure of the ANSIE. All of the studies outlined above used samples comprising university or college students and the extent to which they were representative of the wider population is unclear. The limited sample sizes utilised in most of the studies is likely to have impacted on the lack of factor stability apparent. With reference to sample size alone, only Chandler and Dugovics (1977) with a sample size of 291, approached the 300 mark, defined by Kass and Tinsley (1979) as a cut-off for factor stability regardless of participant to variable ratio. As none of the other studies made reference to the suitability of their data for factor analysis, in terms of item communalities or use of the Kaiser-Meyer-Olkin measure of sampling adequacy (KMO) (Kaiser, 1970) it seems reasonable to assume that the resulting lack of stability has its root here.

The factor structure of the ANSIE has also been investigated in people with ID. Wehmeyer (1993) investigated the factor structure of the ANSIE using a sample of 409 participants ranging from 16 to 64 years of age and comprising 52% females and 48% males. Wehmeyer utilised a principal components analysis with a varimax rotation. Factors with an

eigenvalue greater than 1 were retained. Items were included if their loading on a factor was greater than 0.30 and a minimum of 3 items was required to establish a coherent theme for a factor. This process identified an eight-factor solution for males. Factor 1 referred to a belief in hard work and actions as a determinant of outcomes. Factor 2 was based on the theme of luck and Factor 3 focused on personal control or helplessness. Factor 4 referred to a sense of futility, with particular reference to decision making, and Factor 5 was based on perceptions of control with reference to authority figures. Factor 6 was labelled self-concept and contained items relating to physical and cognitive abilities. Factor 7 reflected control over events and choices and Factor 8 concerned control over personal actions and perceptions of others.

In a subsequent study, Wehmeyer (1994) reported that the internal consistency of the ANSIE ($\alpha=.77$) was adequate (Kline, 1999). In addition, test re-test reliability was calculated with 93 participants after a 1 month interval resulting in a moderate correlation ($r=.53$) ($p=.0001$). Whilst Wehmeyer used a fairly large sample size in this study, as with the mainstream factor analytic studies reported above, he does not appear to assess the suitability of his data for factor analysis. In addition, there is little detail presented about his participants. Wehmeyer analysed his data according to gender in order to account for potential differences in factor structures for males and females (Dixon et al., 1976; Piotrowski et al., 1983) but this obviously reduced his sample size for analysis. Whilst he describes his participants as having ‘mental retardation’ he provides no details about how this decision was reached other than the fact that participants were recruited through self-advocacy groups for people with ‘mental retardation’. Similar to the studies included in Mamlin’s et al.’s (2001) review, Wehmeyer’s studies (Wehmeyer, 1993 & Wehmeyer, 1994 (Wehmeyer and Palmer, 1997 was included)) do not appear to meet the standards outlined for adequacy in the identification of people with ID. Certainly, in the Wehmeyer (1994) study,

the level of 'disability' was determined by asking participants to evaluate themselves on seven questions about the amount of help or assistance they needed in seven 'major life areas': self-care; learning; mobility; self-direction; receptive and expressive language; capacity for independent living; and economic self-sufficiency. It is therefore unclear exactly what level of cognitive impairment these individuals had, and consequently, to what extent these findings can be generalised. Whilst Wehmeyer was clearly aware of the impact of acquiescence with regard to the ANSIE since he discussed the potential adaptation of the ANSIE in order to reduce acquiescence, he did not actually adapt the measure in an attempt to assess or reduce the impact of acquiescence.

4.3. Aims of the research

In summary, LOC appears to be an important factor in the assessment of risk of offending and successful treatment change in offenders without ID (Beech et al., 1998; Johnson & Berry, 1989; Page & Scalora, 2004), as outlined in Chapter 1. However, the relationship between LOC and risk and treatment outcome is less clear in offenders *with ID* (Langdon and Talbot, 2006; Rose et al., 2002). Consequently, there is a need to assess LOC accurately in offenders with ID. The purpose of these studies was to produce a valid measure of LOC for offenders with ID. More specifically, it aimed to improve the readability of the Adult Nowicki-Strickland Internal-External LOC scale (ANSIE), making it suitable for use with offenders with ID whilst maintaining its comparability to the original scale. In addition, the studies in this chapter aimed to examine the validity and reliability of the new scale, the Modified- Adult Nowicki-Strickland Internal-External LOC scale (M-ANSIE).

4.3.1. Hypotheses

Study 1: Creation of the M-ANSIE

1. The readability of the M-ANSIE will be significantly easier than the ANSIE when measured using standard readability statistics.
2. Participants will rate the M-ANSIE as easier to understand and answer than the ANSIE.

Study 2a: Testing the psychometric properties of the M-ANSIE

3. The scores of offenders on the M-ANSIE will be highly correlated with their scores on the ANSIE thereby demonstrating construct validity.
4. The M-ANSIE will be reliable in terms of test re-test scores.

Study 2b: Factor analysis of data for the M-ANSIE

Due to instability in the reported factor structure of the ANSIE, a hypothesis as to the factor structure of the M-ANSIE was not specified, instead its factor structure was explored using factor analysis.

4.4. METHODS

4.4.1. Sample: study 1 & 2a

Forty-seven adult male offenders with ID volunteered to participate in this study. All were residents in a regional medium secure unit for offenders with ID. For details of the sample for this study please refer to the description in Chapter 1.

4.4.2. Sample: study 2b

This study was conducted with a sample of 139 participants with ID from three types of setting: the 47 detained offenders described in Chapter 1, 46 men with ID living in

institutional type accommodation, and 46 men living in community settings. For details of these samples please refer to the description in Chapter 1.

4.4.3. Procedure: study 1

Favourable ethical opinion for this research project was gained from the Norfolk Research Ethics Committee. In addition, full approval was gained from the East Norfolk and Waveney Research Governance Committee and Research Management Team.

The first task was to facilitate the participants' understanding of the LOC assessment tool. To this end the content of the ANSIE was modified by reducing the length of all questions whilst simplifying content and vocabulary. Sentence structures were simplified and sentence content altered, where appropriate, to match the social context of participants' lives. Readability statistics were obtained for each original and revised item using *Microsoft Word 98 Grammar Check* software (Microsoft Corporation, 1993). These statistics include Flesch Reading Ease (Flesch, 1948) and Flesch-Kincaid Grade Level (Kincaid, Fishburne, Rogers, & Chissom, 1975), which measure the average number of syllables per word, average number of words per sentence and various elements of sentence composition. These statistics provide reading level estimates based on various percentages of the respective normative samples that answered comprehension questions regarding sets of standard passages. In this context American school grades correspond to specific ages for the children studying at that grade level depending upon the date of birth of the child in relation to the start of the academic year. Table 4.1 shows the ages associated with American school grades.

Finally, the original and modified items were compared to ensure that the item content remained similar. The M-ANSIE was passed to two qualified psychologists experienced in the psychometric assessment of people with intellectual disabilities to comment on its ease of reading and understanding, and the similarity of content to the

original. Minor changes to the wording of several items were made as a result of their comments. Subsequently, the original ANSIE and the M-ANSIE were passed to five qualified psychologists along with a request to judge the similarity of the content on a 4-point Likert scale (Not at all Similar, Quite Similar,

Table 4.1.

Ages Appropriate to American School Grades

American School Grade	Age
1	6-7 years
2	7-8 years
3	8-9 years
4	9-10 years
5	10-11 years
6	11-12 years
7	12-13 years
8	13-14 years
9	14-15 years
10	15-16 years
11	16-17 years
12	17-18 years

Similar, Very Similar). All judged the content as Very Similar. Examples of the original and the modified ANSIE items can be seen in Table 4.2. (See appendix 2 for the complete measures).

Table 4.2

Examples of the Original and Modified ANSIE Items

Item	Original ANSIE	Modified-ANSIE
1	Do you believe that most problems will solve themselves if you don't fool with them?	Will problems get sorted out on their own?
7	Do you feel that most of the time it doesn't pay to try hard because things never turn out right anyway?	Do you think it is worth trying hard if things never turn out right?
29	Do you believe that when bad things are going to happen they are just going to happen no matter what you do to try and stop them?	When bad things are going to happen can you stop them?

After completing each questionnaire participants were also asked to rate the ease with which they understood the questionnaire on a four point scale (very easy / easy / ok / difficult) and to highlight any particular questions they found difficult to understand. The scale was accompanied by a diagrammatic representation to aid understanding (two thumbs up, one thumb up, thumb neither up nor down, two thumbs down). Any request for repetition, clarification or further information for a particular item was interpreted as difficulty in understanding and recorded as such by the interviewer. In addition, participants were asked to rate the ease with which they answered the items. Again, they were asked to identify any particular items that they found difficult to answer. Any remarks indicating that the participant was having difficulty answering were again recorded by the interviewer. Responses that indicated an understanding of the general nature of the question but uncertainty as to whether to answer yes or no were interpreted as difficulty in answering the question and were recorded as such by the interviewer. For example, participants often framed an answer that did not incorporate a yes or no response. In these circumstances the participant was pressed for a yes or no response. If this response did not correspond to their

previously framed response it was recorded as difficult to answer. However, the answer actually recorded was the one that corresponded to the participant's original response.

4.4.4. Procedure: Study 2a

All participants completed the ANSIE and the M-ANSIE. A sub-sample of 47 detained offenders completed the M-ANSIE on two occasions in order to ascertain test-retest reliability. Half of the participants completed the original tool first and half completed the modified version first. There was a break of one to two weeks between the various completions of the ANSIE and M-ANSIE. The first completion of the M-ANSIE was used in the comparison with the ANSIE. The measures were administered by the current author and two therapeutic interventions nurses. The administration followed standardised instructions (Nowicki & Strickland, 1973) (See appendix 3). Each statement was read out and the participant asked to indicate whether he agreed with it or not. In the event of the participant being unable to answer, simple explanatory instructions were given. These were discussed and agreed between the administrators in advance. Both ANSIE and M-ANSIE were scored in an external direction using scoring keys. Participant scores were tabulated. Statistical analysis was conducted on the data using PASW Statistics 18 (SPSS Inc., 2008).

4.4.5. Procedure: study 2b

The third component of the study was an exploratory factor analysis of the data in order to examine its factor structure in relation to Wehmeyer's (1993) original factor analysis of the ANSIE with people with ID. The validity of factor analysis with dichotomous variables has been established (Bartholomew, 1980; Bartholomew, 1987; Bartholomew & Knott, 1999; Lucke & Schussler, 1986). The data analysis was conducted in line with Wehmeyer's (1993)

original study. Consequently, data were factored using a principal components analysis with a varimax rotation.

4.5. Results

The results are reported in the order that they are stated in the hypotheses section.

Throughout the results section effect sizes are reported. An effect size is an objective measure of the importance of an effect (Field, 2009). The most commonly used measures of effect size are Cohen's d and Pearson's correlation coefficient r . There is little to choose between these two statistics, particularly when samples sizes are similar, as in the studies in this thesis (Field, 2009; Dancey and Reidy, 2002). However, Pearson's r is considered easier to interpret as it is constrained to lie between 0 (no effect) and 1 (or -1) (a perfect effect), and is familiar to most researchers and clinicians (Field, 2009). The use of Pearson's r as a measure of effect size is recommended by several authors with reference to t-tests (Field, 2009; Dancey and Reidy, 2002). Consequently, the effect sizes reported here are Pearson's correlation coefficient r . Cohen (1988) provides some widely used suggestions about what constitutes small or large effect sizes which are outlined in table 4.3.

Table 4.3

Cohen's (1988) Categorisation of Effect Sizes

r value	Effect Size	Percentage of Variance Explained
.10	Small	1%
.30	Medium	9%
.50	Large	25%

4.5.1. Study 1: Creation of the M-ANSIE

Table 4.4 shows the readability statistics for the ANSIE and the M-ANSIE. The length of each item was reduced by an average of five words making the M-ANSIE items 31% shorter than the original items. The mean Flesch Reading Ease statistic increased from 86.13 for the ANSIE to 93.05 for the M-ANSIE.

Table 4.4.

Readability Statistics for the ANSIE and the M-ANSIE

Readability Statistics					
	Total Words	Words per Sentence	Letters per Word	Mean Flesch Reading Ease	Mean Flesch-Kincaid Grade Reading Level
ANSIE	643	16.08	3.8	86.13 (<i>SD</i> =9.96)	5.24 (<i>SD</i> =2.15)
M-ANSIE	440	11.00	3.9	93.05 (<i>SD</i> =7.94)	2.4 (<i>SD</i> =1.59)

Figure 4.1 illustrates the mean Flesch Reading Ease statistic for both the ANSIE and M-ANSIE with error bars. The small degree of overlap of item readability statistics, indicated by the error bars, suggests that these means are likely to be significantly different (Field, 2009). Indeed, an independent samples t-test of the Flesch Reading Ease statistics found a significant difference in reading ease ($t_{(39)} = -4.258, p < .05, r = .56$) with a large effect size.

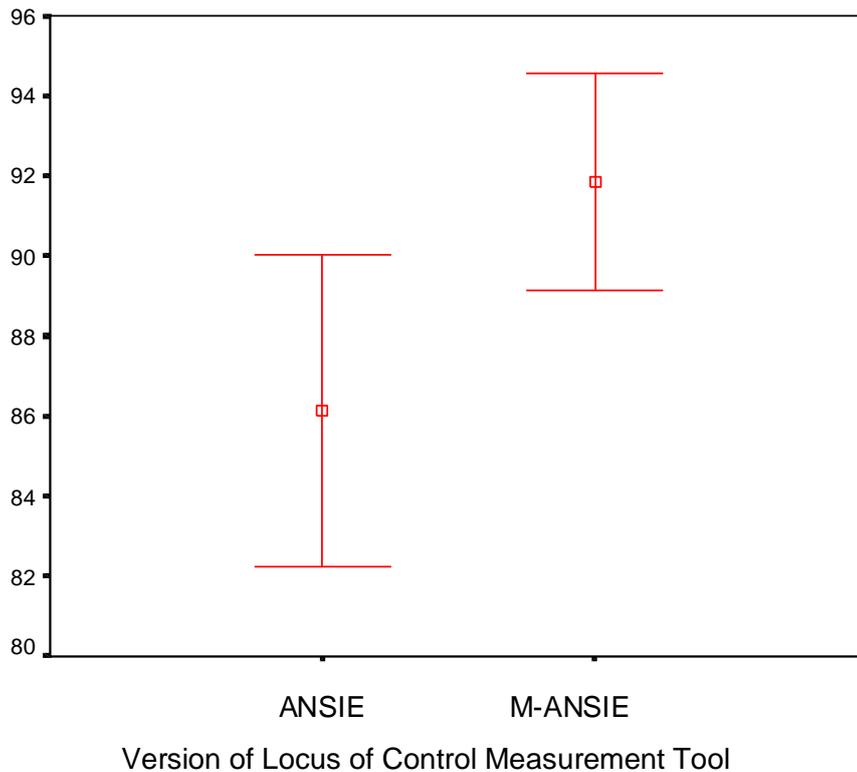


Figure 4.1. Error Bar Graph Showing Mean Flesch Reading Ease Statistics for the ANSIE and the M-ANSIE. Error Bars show 95.0% Confidence Interval (CI) of Mean

Similarly, the mean Flesch-Kincaid Grade Level statistic improved from 5.24 for the ANSIE to 2.4 for the M-ANSIE. Figure 4.2 shows the mean item Flesch-Kincaid Grade Level statistic for the ANSIE and M-ANSIE. In this instance the lack of overlap of the Grade Reading Level statistic for items of each measure indicates that these means are significantly different. A paired samples t-test of the Flesch-Kincaid Grade Level statistic showed a significant difference in grade reading level ($t_{(39)} = 7.045, p < .001, r = .75$), with a large effect size, between the ANSIE and the M-ANSIE. Also, the variability in Flesch-Kincaid Grade Level was reduced in the M-ANSIE as indicated by Levene's Test for equality of variances ($F_{(1, 78)} = 6.81, p < .05$) suggesting that the whole of the document was easier to read and not

just certain parts of it (Kolton et al, 2001). Consequently, the modifications significantly improved the readability of the M-ANSIE over the ANSIE.

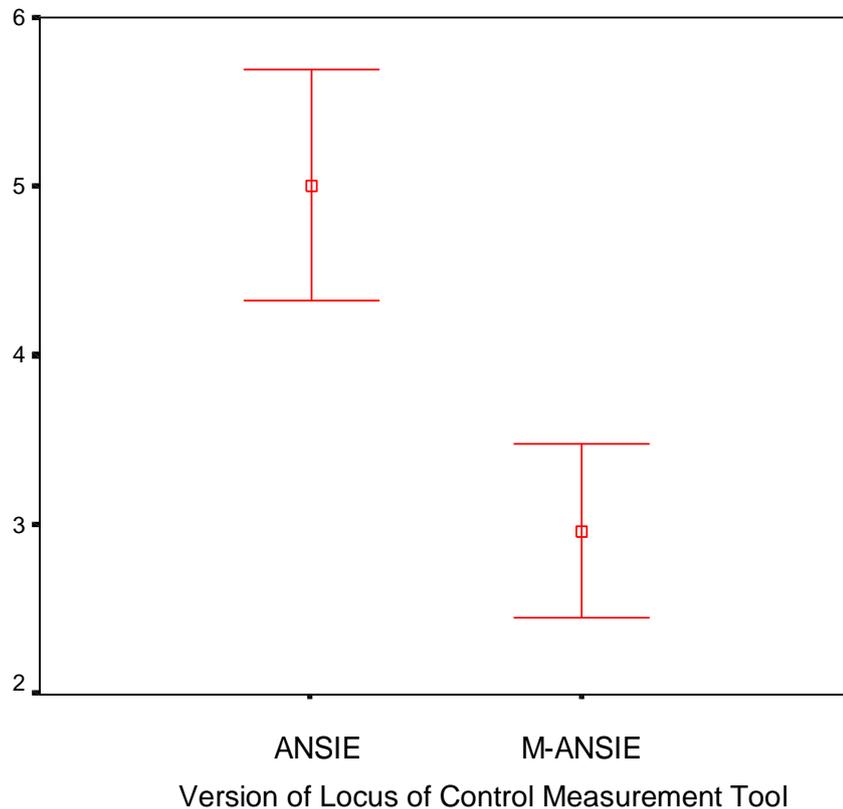


Figure 4.2. Error Bar Graph Showing Mean Item Flesch-Kincaid Grade Level Statistic for the ANSIE and the M-ANSIE. Error Bars show 95.0% CI of Mean.

The mean participant ratings for ease of understanding and ease of answering for both the ANSIE and the M-ANSIE are recorded in Table 4.5. and Figure 4.3. The absence of any overlap of participant ratings between the ANSIE and the M-ANSIE in Figure 4.3, as indicated by the error bars, suggests that these means are significantly different. Indeed, a paired samples t-test of participant's ratings indicated that participants rated the M-ANSIE as significantly easier to understand than the ANSIE ($t_{(46)} = 9.604, p < .001, r = .82$) with a large effect size.

Table 4.5 The Mean Participant Ratings for Ease of Understanding and Ease of Answering the ANSIE and M-ANSIE.

	ANSIE	SD	M-ANSIE	SD
Ease of Understanding	1.83	0.73	3.09	0.78
Ease of Answering	1.78	0.75	3.04	0.88

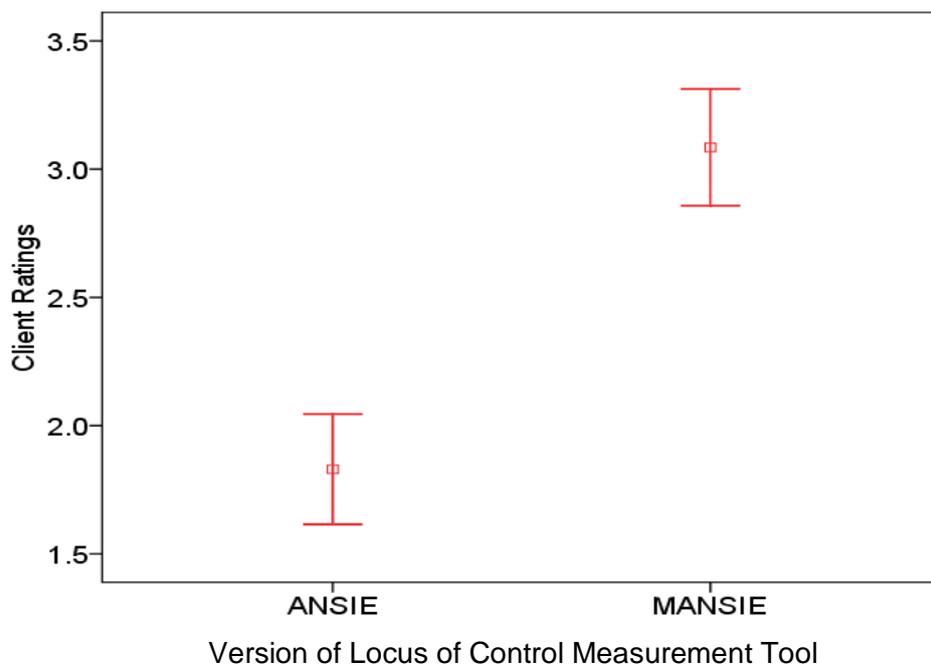


Figure 4.3. Error Bar Graph Showing Mean Participant Ratings for Ease of Understanding the ANSIE and M-ANSIE. Error Bars show 95.0% CI of Mean.

Table 4.6 and Figure 4.4 illustrate the mean client rating for ease of answering the ANSIE and the M-ANSIE. Again, the lack of overlap of the error bars in Figure 4.4. suggest that these means are significantly different. Again, a paired-samples t-test of participants' ratings of ease of answering indicates that they found the M-ANSIE significantly easier to answer than the ANSIE ($t_{(46)} = -8.022, p < .001, r = .76$) with a large effect size.

Table 4.6. Mean Number of Items Identified as Either Difficult to Understand or Difficult to Answer on the ANSIE and M-ANSIE.

	ANSIE	SD	M-ANSIE	SD
Difficult to Understand	6.60	3.52	0.90	1.05
Difficult to Answer	6.79	3.04	1.32	1.29

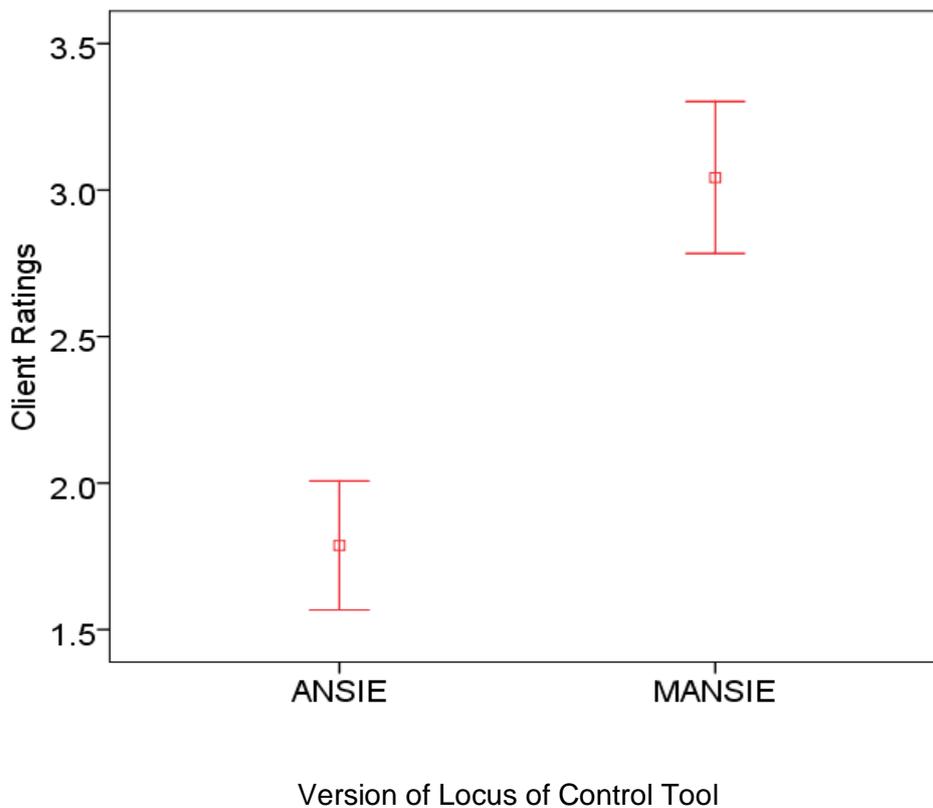


Figure 4.4. Error Bar Graph Illustrating Mean Participant Ratings for Ease of Answering the ANSIE and the M-ANSIE. Error Bars show 95.0% CI of Mean

Table 4.6 and Figure 4.5 illustrate the mean number of items recorded as being difficult to understand on the ANSIE and the M-ANSIE. Again, the lack of overlap of error bars indicates that these means are significantly different. A paired samples t-test confirmed that the number of requests for repetition, clarification or further information were

significantly reduced with the M-ANSIE ($t_{(46)} = 11.835, p < .001, r = .87$), with a large effect size, compared to the ANSIE.

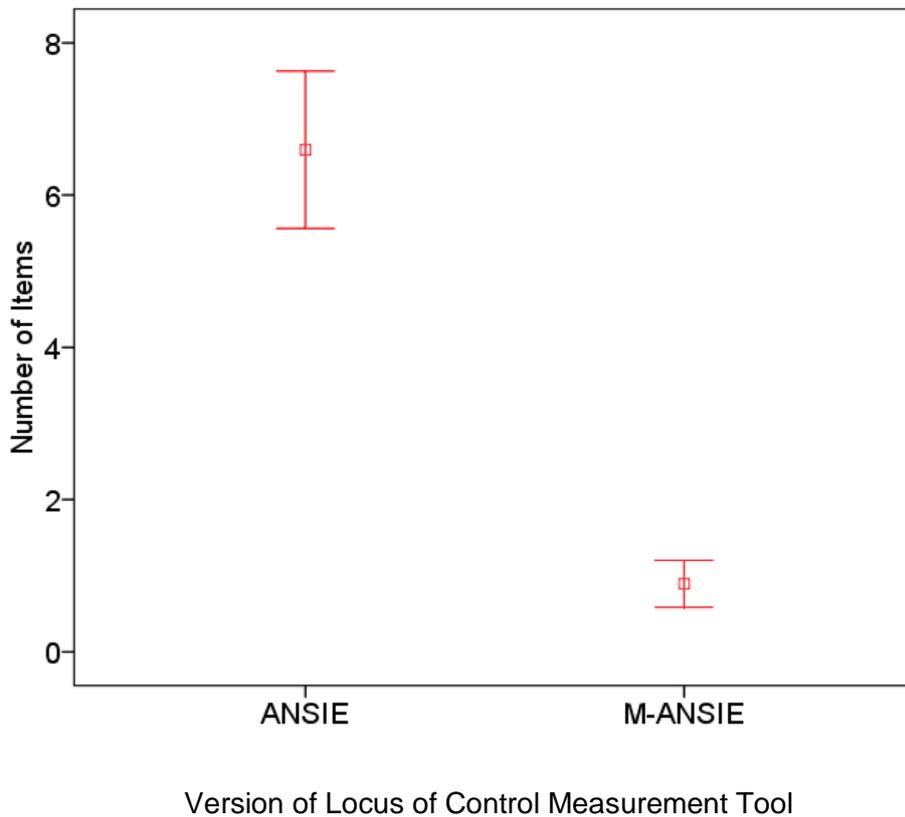


Figure 4.5. Error Bar Graph Illustrating the Mean Number of Items Identified as Difficult to Understand on the ANSIE and the M-ANSIE. Error Bars show 95.0% CI of Mean.

Figure 4.6 illustrates the mean number of items recorded as difficult to answer on the ANSIE and the M-ANSIE. The lack of overlap again indicates that these means are significantly different. A paired-samples t-test confirmed this demonstrating that the mean number of responses provided that were incongruent with associated explanations were also significantly less with the M-ANSIE than the ANSIE ($t_{(46)} = 14.175, p < .001, r = .90$) with a large effect size.

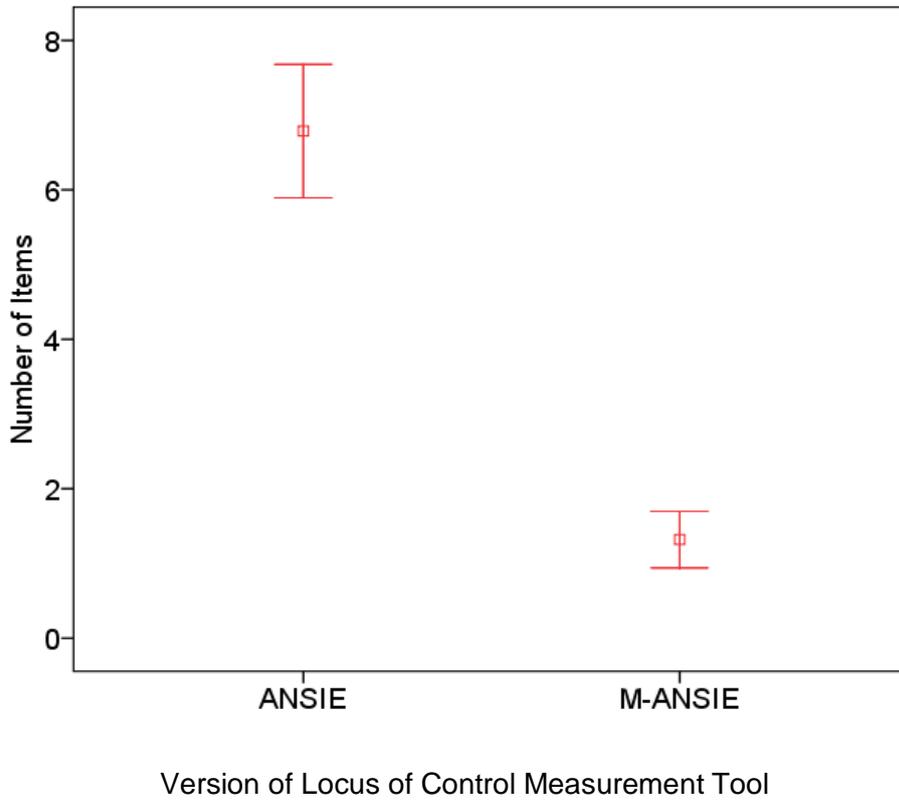


Figure 4.6. Error Bar Graph Illustrating the Mean Number of Items Identified as Difficult to Answer on the ANSIE and the M-ANSIE. Error Bars show 95.0% CI of Mean.

Each of the t-tests demonstrates that the modifications significantly improved the readability of the M-ANSIE over the ANSIE with items that were shorter and easier for participants to understand whilst maintaining their essential meaning. These changes appear to enable offenders with ID to more easily understand and respond to test items thus facilitating a more effective evaluation of the orientation of their LOC.

4.5.2. Study 2a: Testing the psychometric properties of the M-ANSIE

Kolmogorov-Smirnov tests were conducted on the ANSIE and M-ANSIE data. The results indicated that the distributions of both the ANSIE ($D_{(47)}=.12, p=.07$) and the M-ANSIE ($D_{(47)}=.13, p=.06$) data were not significantly different to a normal distribution.

Both the ANSIE and the M-ANSIE were scored in the external direction. The mean score on the ANSIE was 16.72 ($SD = 4.71$) compared with 14.45 ($SD = 4.59$) on the M-ANSIE. A paired-samples t-test indicated that there was a significant difference between the means on the two measures ($t_{(46)} = 5.16, p<.001, r=.71$) with a large effect size. The means suggest a high external orientation of LOC in this population.

The ANSIE is regarded as a valid measure of LOC (Chandler & Dugovics, 1977; Dixon et al., 1976; Finch et al., 1981; Kearney & Kearney, 1983; Piotrowski et al., 1983). Construct validity requires that instruments purporting to measure the same construct produce scores that are correlated in an expected manner (Carmines & Zeller, 1979). Consequently, construct validity for the M-ANSIE can be ascertained by correlating scores on it and scores on the ANSIE. The relationship between participants' scores on the ANSIE and M-ANSIE can be seen in Figure 4.7. The scatterplot indicates a trend in the data in the expected manner. As individuals' scores on the ANSIE increase so do their scores on the M-ANSIE. No obvious outliers are apparent although one individual did score extremely high on both measures. However, as this data followed the general trend it was not felt appropriate to exclude this individual from the analysis (Field, 2009).

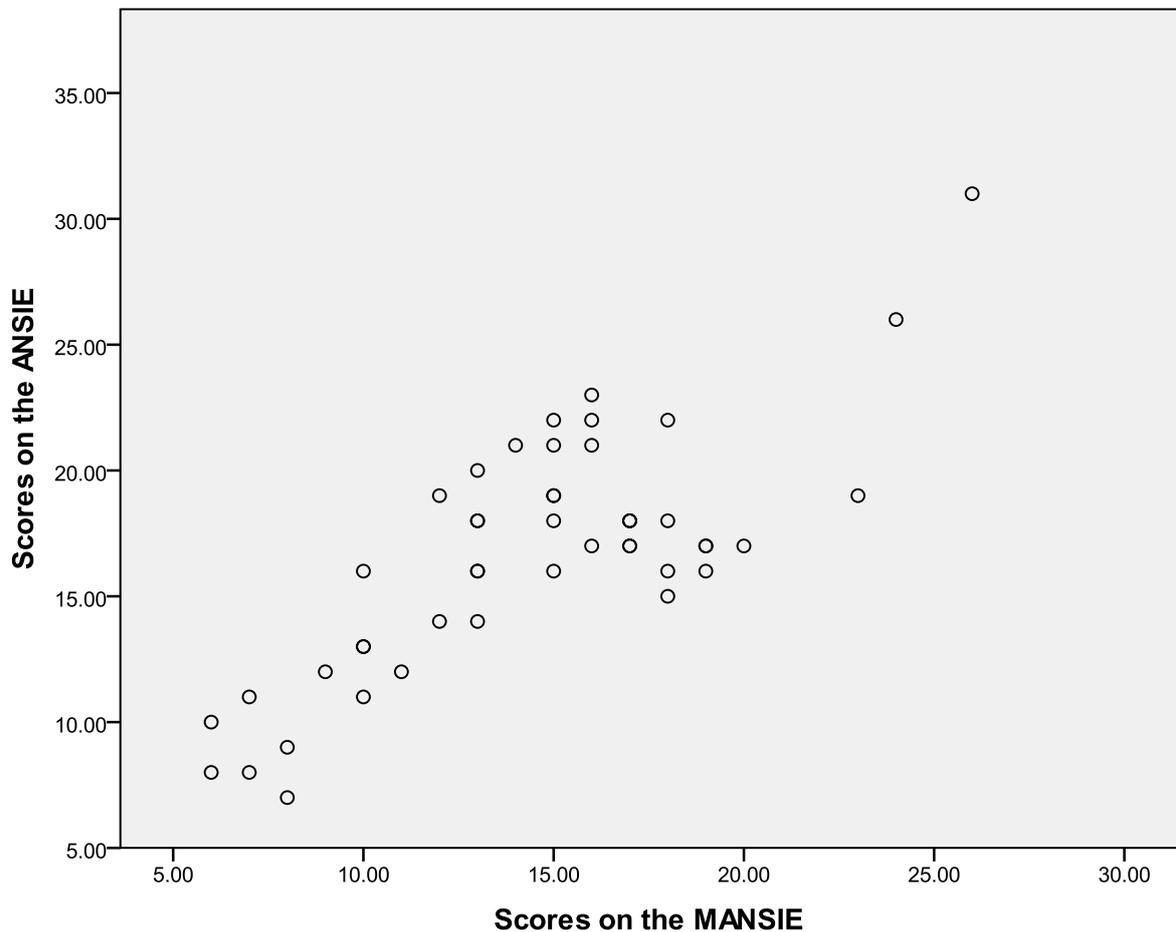


Figure 4.7. Scatterplot Illustrating the Correlation between the Participants' Scores on the ANSIE and the M-ANSIE.

Using Pearson's correlation coefficient r , a strong positive correlation was found between the total scores of the ANSIE and the M-ANSIE ($r=.788, p<.001$). Consequently, it appears that the ANSIE and the M-ANSIE are probing the same construct. Test re-test reliability was assessed by calculating a Pearson's correlation coefficient between the scores on M-ANSIE at the first administration and scores on a subsequent administration. This revealed a strong, positive correlation ($r=.797, p>.001$) between the scores, providing evidence of very good reliability (Cohen, 1988). The internal consistency of the M-ANSIE was acceptable ($\alpha=.80$) ($n=139$) (Kline, 1999).

The current author was interested in assessing to what extent the reduction in the mean score associated with the M-ANSIE was the result of a reduction in acquiescence. To that end, the number of items where an endorsed external response was indicated by a ‘yes’ was compared with the number of items where an endorsed external response was indicated by a ‘no’ for both the ANSIE and the M-ANSIE. The number of items endorsed in the external direction by a yes response or a no response are shown in Table 4.8.

Table 4.8
The Percentage of Externally Oriented Items Endorsed by Yes and No Responses

	Number of Items			Number of Responses in an External Direction		
	Total Number of Items	Number of Items with Yes External	Number of Items with No External	Total Number of External Responses	External Response ‘Yes’	External Response ‘No’
ANSIE	40	24 (60%)	16 (40%)	810	538 (66.42%)	272 (33.58%)
M-ANSIE	40	14 (35%)	26 (65%)	733	283 (38.60%)	450 (61.40%)

Whilst no statistical analysis has been conducted on the above data the findings are suggestive of a reduction in acquiescence. If responses were based solely on question content one would expect approximately 60% of external responses to the ANSIE to correspond to a ‘yes’ but only 35% on the M-ANSIE. In fact, external responses corresponding to a yes were above these percentages on both questionnaires, 66.42% on the ANSIE and 38.6% on the M-ANSIE, strongly suggesting the presence of acquiescence because participants appear to be more likely to endorse an external response if it corresponds with a yes. However, that possible over-endorsement of yes responses is more pronounced on the ANSIE by 6.42% compared with 3.6% on the M-ANSIE. This suggests a reduced bias towards acquiescence

on the M-ANSIE compared to the ANSIE, however, as already stated this has not been tested statistically.

4.5.3. Study 2b: Factor analysis of the M-ANSIE

Prior to performing the analysis, the suitability of the data for principal components analysis (PCA) was assessed. The Kaiser-Meyer-Olkin measure ($KMO = .57$), although in the barely acceptable range (Kaiser, 1974), verified the adequacy of the data for the analysis. In addition, Bartlett's test of sphericity ($\chi^2(780) = 1658.004, p < .001$) indicated that correlations between items were sufficiently large for PCA. Various authors (Field, 2009; Pallant, 2010) recommend the initial completion of both orthogonal and oblique rotations in order to aid factor interpretation and to determine the appropriate rotation method. Inspection of the Correlation Matrix produced by the Oblimin rotation indicated that there was little correlation between the identified factors, the strongest correlation, between factors 1 and 5 being just $-.166$. This indicated that the orthogonal, Varimax rotation was appropriate, which is therefore reported here.

Data for the 40 items were factored using a PCA in order to obtain eigenvalues for all items. This produced 14 factors with eigenvalues greater than Kaiser's criterion of 1, accounting for 67.4% of the variance. Consequently these 14 factors were subjected to a varimax rotation. Criterion for item inclusion was a factor loading of at least $.40$, which Stevens (2002) recommends in a sample of this size. A minimum of three items was required to establish a coherent theme for a factor in accordance with Wehmeyer (1993). This process produced 9 factors. The associated scree plot (Catell, 1966) was slightly ambiguous with two changes in gradient indicating the inclusion of 8 factors, or using the more obvious change in gradient, 3 factors (see Figure 4.8.).

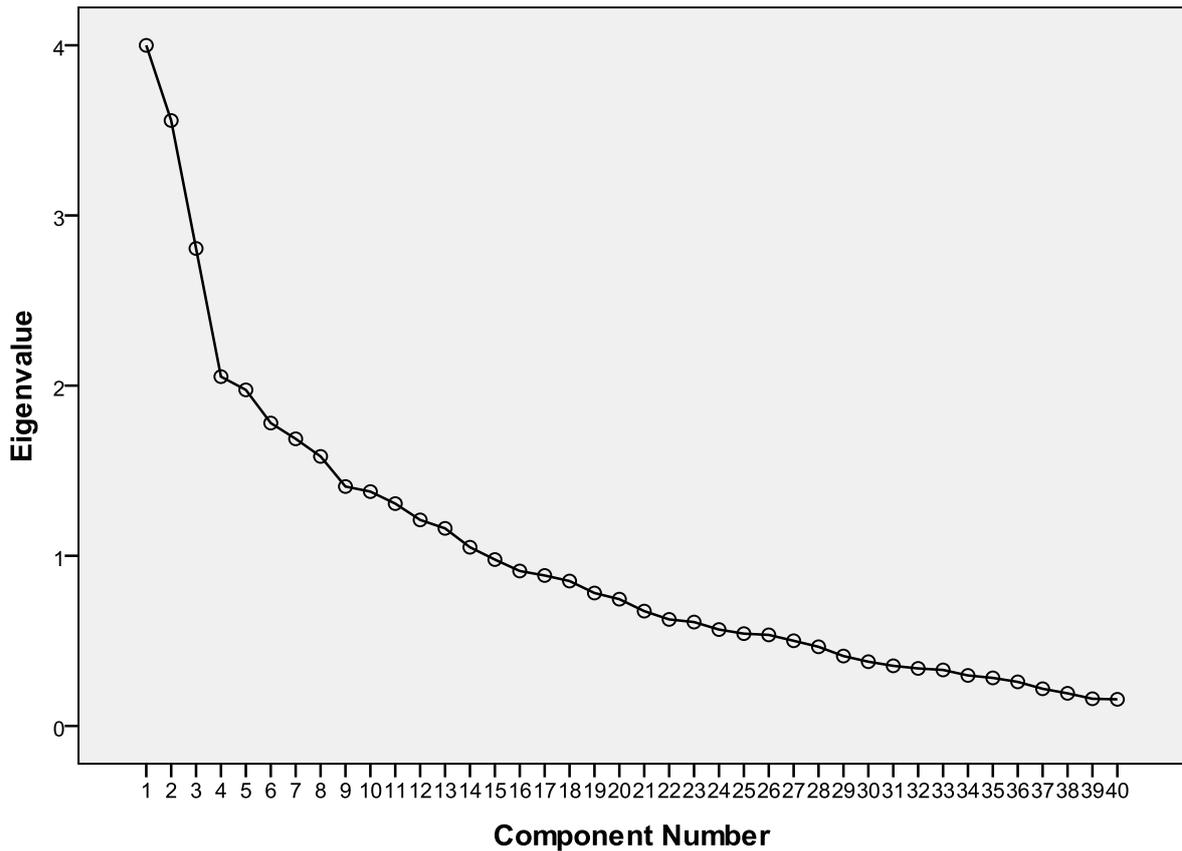


Figure 4.8. Scree Plot of Eigenvalues for the Principal Components Factor Analysis of the M-ANSIE.

Kaiser’s criterion and the scree plot methods have both been criticised as resulting in the retention of too many factors in some situations (Hubbard & Allen, 1987). In view of the difficulties identifying the number of factors to retain, a parallel analysis was conducted (Horn, 1965). Parallel analysis involves the comparison of the obtained eigenvalues with eigenvalues produced from a randomly generated data set. Those factors with eigenvalues exceeding the corresponding randomly generated eigenvalue are retained. A Monte Carlo analysis was conducted (Watkins, 2000) which indicated that 8 factors should be retained. The data analysis was therefore repeated with the specification of the extraction of 8 factors.

Table: 4.9

Eigenvalues and Contribution to the Overall Variance Explained by the 8 Factors

Factor	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	4.000	9.999	9.999
2	3.558	8.895	18.894
3	2.806	7.015	25.908
4	2.053	5.132	31.040
5	1.975	4.938	35.978
6	1.780	4.450	40.428
7	1.687	4.218	44.646
8	1.584	3.960	48.606

Table: 4.10

Summary of the Item Content of the Identified Factors

Factor 1: Inability to protect oneself	Factor 2: Superstition	Factor 3: Powerful others	Factor 4: Judgement of right and wrong	Factor 5: Success is random	Factor 6: Things just happen	Factor 7: Exerting influence	Factor 8: Futility
Doesn't like you	Wishing	Change staff's mind	Told off for no reason	Cheering or luck	Good things/hard work	Just keep trying	Hard work/get it right
Much say in what happens	Good luck charm	Get blamed/ not your fault	People say you're doing well	Clever or lucky	Will staff help	Things go well all day	Try hard/never right
Choice in friends	Four-leaf clover	Do staff listen	You do something wrong	Born good at sports	Choice in food	Change a friend's mind	Change what will happen
Someone angry	Problems sorted on own	Most people stronger than you		Don't think about problems	Does liking depend on behaviour		
Can you stop bad things	Some people born lucky	Nasty for no reason		Planning ahead			
Your enemy							
Cleverer than you							

These 8 factors accounted for 48.61% of the variance. Table 4.9 shows the individual factors' eigenvalues and the contribution to the overall variance explained. After rotation five items failed to factor. These were Question 2 ('Can you stop yourself from catching a cold?'), Question 15 ('Should staff let you make most of your own decisions'), Question 22 ('Does how hard you work make a difference to how well you do?'), Question 32 ('When good things happen is it because of hard work?') and Question 34 ('Is it easy to get friends to do what you want them to do?'). Table 4.10 indicates the items which were included and the factors they comprised. The factor loadings before and after Varimax rotation can be seen in appendix 4.

4.6. Discussion

Evidence suggests that LOC is related to offending behaviour and is an important indicator of treatment change in the mainstream non-ID population (Fisher et al., 1998). However, the relationship between LOC and offending in populations with ID is not clear (Langdon & Talbot, 2006; Rose et al., 2002). Consequently it is important that LOC orientation can be assessed accurately in this population. The validity of instruments developed to measure LOC in populations with ID has not been demonstrated due to the threat of response biases. The effects of response biases can be minimised by simplifying question and response formats and ensuring that question content is concrete and familiar. The first study in this chapter demonstrated that the M-ANSIE showed significantly improved readability in comparison to the ANSIE, thus supporting hypothesis 1. In addition, it was demonstrated that the M-ANSIE was both significantly easier to understand and significantly easier to answer than the ANSIE. This was in terms of participants' ratings, the reduced number of requests for repetition, clarification or further information made by participants, and the

number of responses provided that were incongruent with associated explanations. Thus hypothesis 2 was also supported.

The aim of study 2a was to examine the psychometric properties of the M-ANSIE. This demonstrated that participants' scores on the ANSIE and the M-ANSIE were highly correlated, providing construct validity for the M-ANSIE, as it appears to be measuring the same construct as the ANSIE, namely LOC. Thus, hypothesis 3 was also supported. However, mean scores on the ANSIE were significantly higher than mean scores on the M-ANSIE. This point is returned to below. In addition, a strong positive correlation was found between two administrations of the M-ANSIE providing evidence of test re-test reliability. Indeed, the reliability of the M-ANSIE as measured in this study was much improved compared to the data presented by Wehmeyer (1994) using the ANSIE with a similar population.

The fact that scores on the M-ANSIE were significantly lower than on the ANSIE has two possible explanations. One is that the ANSIE was over-estimating externality of LOC orientation in this population and the second is that the M-ANSIE is under-estimating LOC orientation in this population. Based upon the available evidence it would appear that the first of these options is the more likely. Evidence suggests that people with a ID are more likely to be acquiescent, responding to dichotomous questions with a 'yes' response more frequently than a 'no', regardless of question content.

In total the direction of 10 items were changed in the construction of the M-ANSIE. That is to say 10 more items on the M-ANSIE than the ANSIE had an external response that corresponded to a 'no'. If acquiescence, through affirmative responding, was prevalent in this sample one may expect to see a higher level of external responding with the ANSIE (Clare & Gudjonsson, 1993; Sigelman et al, 1981b). The fact that the mean scores on the

ANSIE were indeed significantly higher than on the M-ANSIE suggests that acquiescence through affirmative responding may have taken place during testing with the ANSIE.

In fact, the evidence presented above suggests that both the ANSIE and the M-ANSIE may both still lead to acquiescent responding in populations with an intellectual disability, however that tendency appears to be reduced in relation to the M-ANSIE. If the ANSIE displayed no response bias, one would expect the percentage of external items indicated by endorsing a 'yes' or 'no' response to match the overall percentage of items in the instrument where an external response correspond with a 'yes' or 'no' respectively. On 60% of the items on the ANSIE a 'yes' response corresponded to an external rating and consequently, on 40% of the items a 'no' response corresponded to an external rating. However, 66.42% of the total number of externally oriented responses was signified by the respondent endorsing a 'yes'. On the M-ANSIE 35% of items had an external rating that corresponded to a 'yes' response, however 38.6% of the total number of external ratings were actually indicated by endorsing a 'yes' response. Consequently, the number of 'yes' responses are 6.42% more than would be expected with the ANSIE and 3.6% higher than would be expected with the M-ANSIE. It is apparent that the percentage difference above that expected for a 'yes' endorsement representing an externally orientated response is smaller with the M-ANSIE compared to the ANSIE.

The mean scores on the ANSIE ($M=16.72$) and M-ANSIE ($M=14.45$) found in this study were lower than previous published means utilising the ANSIE, using samples with ID. Wehmeyer (1994) in a study with 282 students identified as having 'mental retardation' reported a mean ANSIE score of 18 ($SD = 3.69$, range 9-30). The eldest cohort in Wehmeyer and Palmer's (1997) study of individuals with 'mental retardation' (aged 15.25 years and above) had a mean ANSIE score of 18.2 ($SD = 4.3$). Neither of these studies gave specific figures for males only.

Langdon and Talbot (2006) in their small scale study of sex offenders with an intellectual disability reported a mean ANSIE score of 18 ($SD = 3.72$) for 12 treated sex offenders and 18.27 ($SD = 3.52$) for 11 untreated sex offenders with ID. The methodological differences between these studies and their limitations, which were outlined in Chapter 1 make it difficult to draw comparisons between the different studies.

In the literature review several studies were outlined that sought to clarify the factor structure of the ANSIE (Chandler & Dugovics, 1977; Dixon et al., 1976; Finch et al., 1981; Kearney & Kearney, 1983; Piotrowski et al., 1983). As described, these studies indicate a lack of consistency in the factor structure associated with the ANSIE. As previously explained, this may be due to the data used being unsuitable for factor analysis. Only, Wehmeyer (1993) had previously investigated the factor structure of the ANSIE with a sample of participants with ID. There was little similarity between Wehmeyer's 8 factor solution for men with an ID and the various solutions presented in other research papers. Wehmeyer argues that there is some overlap between his Factor II and Kearney and Kearney's (1983) Factor IV representing luck. However, this overlap extends only to two items, namely 21 ('If you find a four-leaf clover will it bring you luck?') and 24 ('Do you have a good-luck charm?').

The factor solution identified in this study for the M-ANSIE also overlapped most closely with Kearney and Kearney's (1983) factor structure of the ANSIE. Four items; 36 ('When someone doesn't like you is there much you can do about it?'), 20 ('Do you have a lot of choice in deciding who your friends are here?'), 23 ('If someone is angry with you is there anything you can do to stop them?') and 33 ('When someone wants to be your enemy, can you do anything to change it?') from their Factor I are also present in the Factor 1 reported here. Three items; 5 ('Do you often get blamed for things that aren't your fault?'), 9 ('Most of the time do staff listen to what clients have to say?') and 27 ('When people are nasty to

you is it usually for no reason?') from their Factor II are present in the current study's Factor 3. Two items; 32 ('When good things happen is it because of hard work?') and 35 ('Do you have much choice in what you eat?') from their Factor III are present in the current study's Factor 6. Finally, two items; 17 ('Are most people just born good at sports?') and 19 ('Is the best way to deal with problems just not to think about them?') from their Factor IV are present in the current study's Factor 5.

There is also some overlap between Wehmeyer's (1993) ANSIE factor structure and the factor structure of the M-ANSIE. Three items; 20 ('Do you have a lot of choice in deciding who your friends are here?'), 29 ('When bad things are going to happen can you stop them?') and 33 ('When someone wants to be your enemy, can you do anything to change it?') from the current study's Factor 1 are present in Wehmeyer's Factor III. Also, four of the items; 10 ('Can wishing make good things happen?'), 24 ('Do you have a good luck charm?'), 21 ('If you find a four-leafed clover will it bring you luck?') and 3 ('If you find a four-leafed clover will it bring you luck?') in the current study's Factor 2 are present in his Factor III. It is perhaps interesting to note that the factors in the present study are more similar to those of Kearney and Kearney (1983) than to those of Wehmeyer (1993). As with the present study, Kearney and Kearney utilised an adult population, whereas Wehmeyer sample ranged in age from 16-64 years. As evidence suggests that LOC orientation becomes internal with increasing age (Nowicki & Strickland, 1973) it is possible that the factor structure of the ANSIE may also alter with the age of respondent.

It is possible that the lack of clarity associated with the factor structure identified in this study is due to the fact that the factor analysis was underpowered. In an underpowered study a clear factor structure may not be identified when in fact one exists in the population. Kline (1999) suggests that 100 participants represent a minimum sample size for factor analysis, which would indicate that the sample size used in this study was adequate ($n=139$).

However, Guilford (1954) suggests a minimum sample size of 200 participants, while Cattell (1978) indicates that the figure should be at least 250. Comrey and Lee (1992) propose a scale where a minimum of 100 participants is poor, 200 is fair, 300 is good, 500 is very good and 1000 is excellent. According to these guidelines the current study was underpowered and it is possible that the lack of clarity in the factor structure resulted from this.

4.7. Conclusions

The M-ANSIE was designed to overcome some of the deficits in the existing methods used to assess individuals with ID (Clare, 1993). First, the stipulation that the M-ANSIE be administered verbally should help to counter the reading difficulties common in this population (Clare, 1993). In conjunction with this, the shorter and simpler sentence structures in the M-ANSIE, as indicated by significant improvements in readability statistics (Flesch, 1948; Kincaid, Fishburne, Rogers, & Chissom, 1975), should reduce the load on verbal memory thus facilitating non-random responding. The simplification of complex language and the substitution of familiar concepts and contexts at the expense of the unfamiliar are likely to enhance understanding, further reduce the load on cognitive mechanisms, and again facilitate non-random responding. The inclusion of items containing double negatives in the ANSIE caused people with ID significant difficulty. All questions incorporating double negatives were originally included in the ANSIE as a means of reversing externality in an attempt to control for socially desirable responding (Nowicki & Strickland, 1973). Despite the removal of double negatives from most questions there was little suggestion during testing that clients were attempting to respond in a socially desirable manner. Indeed, it may be that for this population the complex nature of the construct is sufficient to control for socially desirable responding. Alternatively, Fisher et al (1998)

suggest that this may be because offenders may not view LOC as an offence related behaviour and therefore may not feel the need to ‘fake good’ on it.

Rotter’s (1966) original formulation conceptualised LOC as a generalised expectancy or belief that defines the perceived degree of control an individual has over events in their life. However, the identification of different factors within the construct (Reid & Ware, 1974) suggests that the idea of a generalised expectancy is fallacious. For example, research with samples with ID have demonstrated a differential attribution for perceived failures and successes. Individuals with a ID are equally likely to accept responsibility for failures but significantly less likely than controls to accept responsibility for their successes (Chapman & Boersma, 1979; Pearl, Bryan, & Donahue, 1980; Tognetti, 1972). Until the exact nature of the construct can be more clearly defined it is difficult to assess whether unitary measures of LOC orientation assess meaningful factors in this population.

The establishment of LOC norms for a sample with ID is a necessity. The meaning and implications of the external orientation identified in this study are difficult to assess when there is no current means of comparing it with a matched non-offender sample. Whether an external LOC represents a criminogenic factor in this population has also yet to be established.

This study demonstrated that the LOC assessment tool of choice for many professionals, the ANSIE, caused great difficulty in understanding and answering for institutionalised offenders with ID. Consequently, this study supports the view of Clare (1993) in arguing that assessment tools must be specifically adapted for use with people with ID. The development of the M-ANSIE demonstrated that the reading ease of the scale could be significantly improved thus making it more accessible to this population. In addition, the development of the M-ANSIE resulted in an assessment that participants rated as easier to understand and answer. It also resulted in significantly fewer requests for repetition,

clarification and further information, and produced fewer responses that were incongruent with associated explanations.

A strong correlation between scores on the ANSIE and M-ANSIE provided validity for the new measure. In addition, a strong correlation between original and re-test scores on the M-ANSIE indicated that it is a reliable measure. Mean scores on the M-ANSIE were significantly lower than on the ANSIE, suggesting that the modifications to the scale may have reduced the level of acquiescent responding in this population. Support was added to this supposition from the reduction in the percentage of external responses indicated by the endorsement of a 'Yes', suggesting that the M-ANSIE is facilitating a more valid assessment of LOC orientation through a reduction in acquiescence.

The PCA analysis of the M-ANSIE indicated the retention of 8 factors accounting for 48.61% of the variance. The factor solution identified in this study shared some similarities with Kearney and Kearney's (1983) factor solution, using the ANSIE in a mainstream non-ID population. Factor 1 in both studies shared four factors and three items from their Factor 2 are also in the Factor 3 identified here. In addition, two items from Kearney and Kearney's Factor 4 were present in the Factor 5 identified here and two items from their Factor 3 were present in the Factor 5 identified here. The present study did not share many similarities with the Factor structure identified for the ANSIE in a population with ID by Wehmeyer (1993). This may have been due to methodological differences between the studies, such as the different age ranges and the potential difference in cognitive functioning of the respective participants. It should also be noted that there were few similarities between the factor structures in mainstream non-ID populations outlined above. This lack of stability may be due to the generally limited sample sizes used in these studies. Overall, the results reported in the studies in this chapter indicate that the M-ANSIE is more likely to provide an accurate assessment of LOC orientation in this population than the ANSIE.

**CHAPTER FIVE: A REVISION OF THE EYSENCK IMPULSIVITY
QUESTIONNAIRE: AN ASSESSMENT OF THE VALIDITY OF SELF-REPORT
IMPULSIVITY IN A POPULATION OF OFFENDERS WITH ID**

5.0. Introduction

The aim of this chapter was to improve the readability of the Eysenck Impulsivity Questionnaire (I7i). The revision was intended to increase the validity of responding in a sample of male offenders with ID. This was done to minimise the impact of the response biases outlined in chapter 2 which present a major threat to the validity of self-report of people with ID. Research suggested that the validity of the original scale was likely to be impaired by response biases such as acquiescence which are prevalent in populations with ID (Sigelman, Budd, et al., 1982). Acquiescence can occur when an individual does not understand the question or is uncertain about which response corresponds with their belief. The original questionnaire was developed for people without ID and consequently some questions are complex or do not match the living context of men with an ID. It was believed that acquiescence could be countered by making the questionnaire easier to understand and answer. Consequently, attempts were made to improve the readability of the I7i by simplifying the language and sentence structures with slight changes in content to match the social context of detained offenders with ID. Subsequently, 47 offenders with ID completed both the Revised-I7i (I7i-R) (on two occasions) and the I7i.

The study described in this chapter is divided into three sections. In the first part, the two versions of the questionnaire were compared using standard readability statistics, participant ratings and procedural indicators. In addition, the respondents' scores on the two completions of the I7i-R were compared to ascertain test retest reliability. In the second part

the respondents' total scores on the two measures were compared to ascertain construct validity for the revised measure. The patterns of responding were examined to assess whether the modifications had reduced acquiescent responding. In addition, the validity of both versions was assessed through comparison with scores on another self-report measure of impulsivity, the Barratt Impulsiveness Scale (Barratt, 1994; Patton et al., 1995) and a clinical assessment, which can highlight impulsive behaviour, the Behavioural Assessment of Dysexecutive Syndrome (BADS) (Alderman & Burgess, 1996). In the third part of the study an exploratory factor analysis using a principal components analysis, with an orthogonal, varimax rotation, was conducted. The findings are discussed.

5.1. Impulsivity

In Chapter 1 research was outlined, indicating that impulsivity is elevated in samples of offenders. Consequently, there is a need to assess impulsivity in offenders with ID, using a reliable and valid tool. Evidence relating to the investigation of impulsivity will be critically reviewed and summarised. Despite the broad scope of research and the variety of methodologies incorporated, there appears to be a discernable core to the definition of impulsivity. Namely, that impulsivity relates to behaviour that is socially inappropriate or maladaptive and is quickly expressed without reference to the consequences of the behaviour (Buss and Plomin, 1975; Cherek & Lane, 1999; Logue, 1995; Oas, 1985; Rachlin, 1995).

Research on impulsivity is broad-ranging, covering diverse areas such as personality (Whiteside & Lynam, 2000), personality disorder (Soloff et al., 2003), behavioural disorders such as ADHD (Barkley, 1997), pharmacology (Markovitz, 1995), a plethora of areas related to mental health (Lacey & Evans, 2006), delay of gratification (Cherek & Lane, 1999), brain damage (Frank, Samanta, Moustafa & Sherman, 2007)

and numerous physical health issues (Bickel, Odum & Madden, 1999) as well as offending behaviour (Schafer et al., 2004).

Impulsivity is associated with numerous maladaptive behaviours, for example aggression, drug use, eating disorders, high frequency alcohol consumption, gambling and suicide (Fishbein, 2000; Fishbein & Reuland, 1994; Hewitt & Flett, 1993; Johnson, Malow, Corrigan & West, 1993; Logue, 1995; Lowe & Eldredge, 1993). Consequently, it is not surprising that impulsivity is present in many psychiatric disorders: obsessive-compulsive disorders, various personality disorders, impulse control disorders, eating disorders, attention deficit hyperactivity disorder, psychoactive substance abuse disorders (APA, 2000). It is equally unsurprising therefore to find impulsivity identified as a transnosological dimension (Apter et al, 1990; van Praag, 1988). The tendency to view impulsivity as a negative behaviour is also reflected in the large body of literature detailing techniques for reducing impulsivity (Cole & Hartley, 1978; Heider, 1971; Meichenbaum & Goodman, 1971; Nelson & Birkimer, 1978; Zelniker, Jeffrey, Ault & Parsons, 1972).

5.2. Impulsivity as a personality characteristic

Eysenck and Eysenck (1969, 1976, 1977) sought to provide a hierarchical system of personality description. This hierarchical structure comprised a large number of primary traits forming the lowest level of measurement, while at the second-order level there were dimensions or super-factors based on the factor analysis of the inter-correlations between primary traits (Eysenck & Eysenck, 1985). Part of this process involved mapping the relationship between the emerging impulsiveness trait(s) and their major personality dimensions, Extraversion (E), Neuroticism (N) and Psychoticism (P). Table 5.1 illustrates the traits that define these three super-factors.

However, Eysenck, Barrett, Wilson and Jackson (1992) caution against the simplicity of the hierarchical model, with specific reference to impulsivity, which they argue breaks down into several sub-traits and as such becomes an intermediary class of trait between the primary traits and super-factors. The identification of these sub-traits is described below with reference to the development of tool to measure impulsivity. A second difficulty with the model above is that traits such as impulsivity load onto more than one super-factor (Eysenck et al, 1992) and the different sub-traits of impulsivity show different patterns of correlations with the super-factors (Eysenck & Eysenck, 1977). This eventually led to the development of an impulsivity scale drawn from items loading onto different super-factors, the development of which is described below.

Table 5.1

The Primary Traits Comprising the Super-Factors of N, P and E (Eysenck & Eysenck, 1977).

Factor 1: Neuroticism (N)	Factor 2: Psychoticism (P)	Factor 3: Extraversion (E)
Anxious	Aggressive	Sociable
Depressed	Cold	Lively
Guilt feelings	Egocentric	Active
Low self-esteem	Impersonal	Assertive
Tense	Impulsive	Sensation-seeking
Irrational	Antisocial	Carefree
Shy	Unempathic	Dominant
Moody	Creative	Surgent
Emotional	Tough-minded	Venturesome

Eysenck and Eysenck investigated the structure of impulsivity in a series of factor analytic studies (Eysenck & Eysenck, 1977; 1978; Eysenck et al., 1985; Eysenck & Zuckerman, 1978) that led to the development of the Impulsivity Questionnaire (version seven) (I7) (Eysenck et al., 1985). In these studies the authors analyse data for male and female participants separately. The present study includes only male participants so reference will only be made to male participants in the following studies. In an initial study (Eysenck & Eysenck, 1977) reported that impulsiveness in its broad sense (ImpB) broke down into four factors: narrow impulsiveness (ImpN), risk-taking, non-planning and liveliness. These are the sub-traits described by Eysenck et al. (1992). It was also reported that ImpN correlated positively with the super-factors of N and P, which they describe as “the two abnormal scales” (p.67) and consequently they conclude ImpN as pathological. However, the loading of broad impulsivity items onto E and P was problematic as it produces correlations between supposedly orthogonal factors. This precipitated the removal of such items to a separate impulsiveness scale. In addition, Eysenck and Eysenck argued that many socially abnormal groups (criminals and psychopaths etc.) lie in a plane delineated by P and E and have high impulsiveness scores. Consequently a scale comprising E and P impulsivity items is likely to be useful in assessing this factor in such groups.

In a subsequent study Eysenck and Eysenck (1978) considered the relationship between ImpB and sensation-seeking which, together with sociability, have been found to make up Extraversion (Carrigan, 1960; Eysenck & Zuckerman, 1978). Items relating to three of Eysenck and Eysenck (1977) factors of ImpN, risk-taking and non-planning (omitting liveliness which appeared to correlate more with Sociability than Impulsiveness) and four of Zucherman’s factors within the construct of Sensation Seeking namely, thrill-and-adventure-seeking, experience seeking, disinhibition and boredom susceptibility revealed significant overlap upon factor analysis, particularly between thrill and adventure seeking and risk

taking. The best solution posited two factors which were labelled Venturesomeness and Impulsiveness which subsequent study revealed to be relatively independent and represent largely different behaviours (Eysenck, 1993). This factor analysis has not been reported in detail but is mentioned in the introduction to Eysenck and Eysenck (1978).

In subsequent studies (Eysenck et al., 1985) 21 items from an empathy scale (Mehrabian & Epstein, 1972) were added for a final factor analysis. The final 54 items defined a three-factor structure: 19 items formed the impulsivity scale (I7i); 16 items formed the venturesomeness scale; and 19 formed the empathy scale. The items making up this structure can be seen in Table 5.2. Using a sample of 383 males with a mean age of 25.07 years ($SD=9.94$), the authors report the reliabilities of the scales for males are reported as .84, .85 and .69 respectively. The authors do not state what type of reliabilities these are.

The results showed an overall mean impulsiveness score of 8.76 ($SD=4.31$). However, mean scores by age groups appear to reduce as age increases for both Impulsiveness and Venturesomeness. For example, the Impulsiveness mean was 9.84 ($SD=4.13$) for 16-19 year old males but only 2.86 ($SD=3.60$) for 80-89 year old males. This difference was not tested statistically. The correlation between Impulsiveness and Venturesomeness was small ($r=.24$), suggesting that they are tapping different traits. The factor loadings onto the major personality dimensions were unclear in this study. Impulsivity correlated moderately ($r=.46$) with P and weakly ($r=.39$) with E, whereas Venturesomeness correlated weakly ($r=.22$) with P and weakly ($r=.37$) with E.

All of the studies reported above which included factor analysis and led to the development of the I7i questionnaire incorporated good sample sizes (Eysenck & Eysenck, 1977 ($n=126$, $n=235$ & $n=348$ respectively); 1978; Eysenck, Pearson, Easting & Allsopp, 1985 ($n=559$ & $n=383$ respectively). However, all of them also shared a range of methodological problems.

Table 5.2. Summary of Item Content of the Factors of the I7.

Impulsiveness	Venturesomeness	Empathy
Buy on impulse?	Enjoy water skiing?	Feel sorry for a lonely stranger?
Do and say without thinking?	Try new brands rather than reliable ones?	Emotionally involved in friends problems?
Often get into a jam?	Enjoy taking risks?	Do unhappy people irritate you?
Impulsive person?	Enjoy parachute jumping?	Nervous when others around you are nervous?
Think carefully before doing?	Is hitch-hiking dangerous?	Silly to cry out of happiness?
Do on spur of moment?	Like diving off high-board?	Do others have strong influence on your moods?
Speak before thinking?	New and exciting experiences and sensations?	Affect you when friends are upset?
Get involved when wish you hadn't?	Learn to fly a plane?	Feelings of character in film?
Carried away by new ideas?	Understand people who climb mountains?	Upset when you see someone cry?
Self-control to keep out of trouble?	Enjoy frightening things?	Laughter catching?
Everything enjoyable is illegal or immoral?	Jump straight in cold sea?	Don't worry about feelings of others?
Surprised by people's reactions to things you do and say?	Enjoy skiing very fast?	Irritated or sympathetic when someone cries?
Evening better if unplanned	Like to go scuba diving?	Enjoy watching others open presents?
Check work?	Enjoy driving fast?	Break bad news to someone?
Change interests?	Pot-holing?	Interested in friends problems?
Consider pros and cons?	Job involving danger?	Feel sorry for shy people?
Sleep on it before making decision?		Affected by feelings of those around you?
Shout back when people shout at you?		Imagine being lonely?
Make up your mind quickly?		Worry you when others are worried?

Information about the composition of the samples, indicating professional employment or student status, suggests that they are not representative of the wider population. In addition, despite the large sample size, no reference is made to the suitability of the data for factor analysis. Equally, it is unclear how new items were developed and which ones were retained and why. Items with factor loading as low as 0.07 (Eysenck & Eysenck, 1977) were retained. On no occasion do the authors report the amount of variance explained by the factor structure identified and the authors report that the number of factors was specified in advance (Eysenck et al., 1985) but do not justify this decision. It is unclear how these methodological problems may have influenced the results.

Corulla (1987) reported a study with 307 participants of which 215 were females (mean age 21.12, SD=5.38) and 32 males (mean age 21.00, SD=4.6) all of whom were described as British university students. She reports a mean impulsivity scores for male participants on the I7i of 8.76 (SD=4.79). The finding from this study must be treated with some caution as the sample is clearly not representative of the UK population, it being drawn from university students with an average age of 21 years. It is unclear how this may have affected the results.

This consideration of the personality characteristic of impulsivity appears to indicate that the concept is multifactorial. However, the variety of behaviours thought to be indicative of impulsivity varies between researchers. The number of factors identified underlying the concept are clearly likely to influence the breadth and nature of the behaviours probed in measures of impulsivity. A number of potential primary factors have been identified that are relatively independent of one another and reflect different aspects of behaviour. Researchers are clearly not agreed on the exact nature of these factors and consequently vary in their naming of such factors. However, broad commonalities may be discerned. For example, one group of factors appears to relate to the failure to analyse and reflect before engaging in a

behaviour and a second group to the consideration of the consequences of the behaviour where a risky outcome may be preferred (Eysenck & Eysenck, 1977).

Many personality inventories include items probing impulsivity (Buss and Plomin, 1975; Cattell et al, 1970; Chapman et al, 1984; Guilford et al, 1978; Haertzen et al, 1980; Jackson, 1967; Lorr & Youniss, 1973). In addition, there are several scales specifically designed to measure Impulsivity *per se* (Barratt, 1965; Eysenck and Eysenck, 1969, 1976, 1977, 1978, Eysenck et al., 1985; Lecrubier, Braconnier, Said and Payan, 1995; Plutchik and van Praag, 1995). Again all of these measures were developed for and normed on mainstream populations. Therefore, in order to assess self-report impulsivity in a sample with ID it is necessary to develop a new measure or to adapt an existing measure, and then demonstrate its reliability and validity.

While broader definitions of impulsivity probably provide a more comprehensive description of behaviours that may commonly be described as impulsive, such broad definitions do not appear to capture the pathological aspect of impulsivity described in Eysenck and Eysenck's more narrowly defined trait (Eysenck & Eysenck, 1977). In addition, it appears to be this more narrowly defined trait which is referred to in standard risk measures such as the HCR-20 (Webster et al., 1997) and Psychopathy Check List-Revised (PCL-R: Hare, 1991, 2003). Consequently, it seems appropriate to adapt a measure that assesses this narrow, pathological form of impulsivity. As described above, the Impulsivity Scale of the I7 (I7i) would appear to meet that requirement.

Several psychometric (e.g., Barratt, 1965) and behavioural (Cherek et al, 1997) measures of impulsivity have been useful in the identification of potential criminogenic impulsivity but no single measure has been as thoroughly used nor as systematically successful in this respect than the narrow definition of pathological impulsivity probed by the I7i (Eysenck, 1981; Eysenck & Eysenck, 1977; Eysenck & Eysenck, 1978; Eysenck &

McGurk, 1980). As it is specifically this pathological aspect of impulsivity, with its correlation to anti-social behaviour, that is of interest in relation to the study of offenders it appears an appropriate construct for assessment in this study.

5.3. The validity of the psychometric assessment of impulsivity

Self-report impulsivity has been found to correlate with behavioural measures of impulsivity. For example, Logan, Cowan and Davis (1984) found that poor inhibitory control in a stop task was correlated with high self-reported impulsivity. In addition, behavioural paradigms have been found to correlate with self-report psychometric measures of impulsivity in predicting violent behaviour in parolees (Cherek et al., 1997).

A drawback with such scales is that they are designed to assess a relatively stable personality characteristic. Consequently, such measures are unlikely to be effective in assessing more transitory expressions of situational impulsivity where impulsive behaviour may be limited to contexts, such as under the influence of alcohol or drugs, in particular affective states or acute mental health episodes (Lacey & Evans, 2006). However, this may be only a minor criticism if, as many have proposed, personality characteristics are indeed relatively enduring and stable (McCrae & Costa, 1994). Finally, there are serious limitations in the use of self-rating scales with populations in which impulsivity plays an important role, for example, where dissimulation or lack insight may be apparent (Lecrubier et al, 1995). Regardless, of this potential difficulty, high self-report impulsivity has been identified in various discreet populations (Fishbein, 2000; Fishbein & Reuland, 1994; Hewitt & Flett, 1993; Johnson et al., 1993; Logue, 1995; Lowe & Eldredge, 1993), suggesting that individuals are willing to disclose its presence.

5.4. Impulsivity in offenders with ID

Two studies have considered the measurement of impulsivity in offenders with ID (Parry & Lindsay, 2003; Snoyman & Aicken, 2011). Both used a modified version of the BIS-11 (as outlined in Chapter 1). The authors report that the sex offenders with ID self-reported significantly lower impulsivity ($M=33.32$ & $M=73.17$ respectively) than the non-sex offenders with ID ($M=49.32$) and violent offenders with ID (83.44) respectively. The results of these studies are limited by the small sample sizes. Barratt (1995) reported means for impulsive aggressive prisoners ($M=69.8$) and matched non-offender adults ($M=49.1$). However, it is apparent that there is some confusion regarding the most recent version and corresponding normative data for the BIS-11, with two versions of the BIS-11 in circulation, which Snoyman and Aicken (in press) labelled 'a' and 'b'. While BIS 11a is the more commonly used 30 item instrument, Parry and Lindsay (2003) used BIS 11b, a 34 item questionnaire. The scoring system used also differed from 1 to 4 and 0 to 3 respectively. Additionally, in modifying the BIS-11b for people with intellectual disabilities, Parry and Lindsay (2003) changed the items of the BIS 11b from statements to a question format, and Snoyman and Aicken (2011) expanded some of the original statements with rephrasing statements. Snoyman and Aicken's reported good internal reliability but their adaptation of the BIS-11 produced statements that were much longer than the original statements. For example, Item 4 "I have racing thoughts" became "I have racing thoughts (I have lots of things in my head at the same time)". Whether this technique promotes reliable responding requires further examination. Unfortunately the authors did not report test re-test reliability for their modified version of the BIS-11, nor did they appear to consider the presence of any response bias in their data.

5.5. Psychological assessment of people with ID

The impaired intellectual functioning associated with learning disability has implications for some assessment methods used with this population. The cognitive deficits associated with ID that are likely to impact upon psychometric assessment of this population are described in Chapter 2.

The present study was driven by the necessity to assess impulsivity accurately in a sample of offender with ID, and ultimately to provide norms for this and similar samples. However to achieve this, it is apparent that self-report inventories require adaptation for use with offenders with ID (Clare, 1993).

5.6. Aims of research

This first stage of this research was to improve the readability of the Impulsivity Scale (I7i) (Eysenck & Eysenck, 1977) whilst maintaining its comparability to the original scale and examining the validity of the new scale in relation to the original. In addition, the ease of understanding and answering the questionnaire was to be compared with the Barratt Impulsiveness Scale. The second stage of this research was to compare the validity of the original and revised scales, the Barrett Impulsiveness Scale and staff ratings of impulsivity in predicting clinical and behavioural components of impulsive behaviour.

5.6.1. Hypotheses

Study 1: Creation of the I7i-R

1. The readability of the Revised I7i-R Impulsivity Scale will be significantly easier than the original scale when measured using standard readability statistics.
2. Participants will rate the Revised I7i-R Impulsivity Scale as easier to understand and answer than the original.

Study 2a: Testing the psychometric properties of I7i-R

3. The scores of offenders on the I7i-R will be highly correlated with their scores on the I7i thereby demonstrating construct validity.
4. The I7i-R will be reliable in terms of test re-test reliability.
5. The scores on different tools measuring impulsivity will be strongly correlated.

Study 2b: Factor analysis of data for the I7i-R

No hypotheses were made regarding the factor structure of the I7i-R as no previous study has sought to analyse the impulsivity items of the I7 in isolation.

5.7.Method

5.7.1. Sample: study 1 & 2a

Forty-seven adult male offenders with learning disabilities volunteered to participate in this study. All were residents in a regional medium secure unit for offenders with ID. For details of the sample for this study please refer to the description in Chapter 1.

5.7.2. Sample: Study 2b

This study was conducted with a sample of 139 participants with ID from three types of setting (described in Chapter 1): 47 detained offenders with ID, 46 men with ID living in institutional type accommodation, and 46 men with ID living in community settings.

5.7.3. Measures

Eysenck Impulsivity Scale (I7i) (Eysenck & Eysenck, 1978) is a 19-item scale designed to measure the highly specific component of Narrow Impulsivity (ImpN). The items comprising the venturesomeness and empathy scales were not included and therefore

the scale only measures impulsivity. Completion of the scale required a simple yes or no response to each item.

Impulsivity Scale Revised (I7i-R) is a revision of this Eysenck Impulsivity Scale (I7i) modified to facilitate its use with a sample of people with ID. To this end the content of the Impulsivity Scale was modified by reducing the length of all questions and simplifying content and vocabulary. Sentence structures were simplified and sentence content altered, where appropriate, to match the social context of participants' lives. Consequently, references to home life were changed to match their current living arrangements. Finally, the original and modified items were compared to ensure that the item content remained similar. The I7i-R was passed to two qualified psychologists experienced in the psychometric assessment of people with intellectual disabilities to comment on its ease of reading and understanding, and the similarity of content to the original. Minor changes to the wording of several items were made as a result of their comments. Subsequently, the original I7i and the I7i-R were passed to five qualified psychologists along with a request to judge the similarity of the content on a 4-point Likert scale (Not at all Similar, Quite Similar, Similar, Very Similar). All judged the content as Very Similar. Examples of the original and the I7i-R items are presented in Table 5.3. (See appendix 2A for the complete measure). As with the original scale, the maximum score is 19 and the higher the score the higher the level of self-report impulsivity.

Barratt Impulsiveness Scale (BIS-11) (Patton et al., 1995) is a thirty-item questionnaire including the three subscales of attentional, motor and non-planning impulsiveness. Again, where necessary, sentence structures were simplified and sentence content altered to match the social context of participant's lives. Interviewees respond according to a four-point scale comprising never, sometimes, often or always. Each of these answers was associated with a geometric shape which was shaded to a greater or lesser

extent, so that there was no shading associated with 'Never' and complete shading associated with 'Always'. The BIS-11 and R-BIS-11 were passed to two qualified psychologists experienced in the psychometric assessment of people with intellectual disabilities to comment on the ease of reading and understanding the questionnaire, and the similarity of content to the original measure. Feedback given suggested that the modification to the BIS-11 changed the meaning of some items but they were unable to offer suggestions as to how it might be improved. Subsequently, the original BIS-11 and the R-BIS-11 were passed to five qualified psychologists along with a request to judge the similarity of the content between the two versions on a 4-point Likert scale (Not at all similar, Quite Similar, Similar, Very Similar). All judged the content as Quite Similar or Similar. Items were scored on a scale of one to four, with four representing the most impulsive option. This gives a maximum impulsivity score of 120. In addition, the R-BIS has three subscale scores; the attention subscale score being the summation of 8 specific item scores, the motor subscale score the summation of 11 specific item scores and the non-planning subscale score being the summation of the final 11 item scores.

The *Behavioural Assessment of Dysexecutive Syndrome* (BADS: Alderman & Burgess, 1996) is a battery of tests designed to assess the effects of dysexecutive syndrome, a cluster of impairments generally associated with damage to the frontal lobes of the brain. These impairments often include difficulties with impulsive behaviour such a failure to inhibit inappropriate responses or accurately assess the consequences of strategies (Burgess & Robertson, 2002). The six tests comprising the BADS are the Rule Shift Cards, Action Programme, Key Search, Temporal Judgement, Zoo Map and Modified Six Elements. The BADS was an assessment regularly used in the local ID service as part of an assessment of problem solving skills. The BADS were administered by the chief researcher.

5.7.4. Procedure: study 1

Favourable ethical opinion for this research project was gained from the Norfolk Research Ethics Committee. In addition, full approval was gained from the East Norfolk and Waveney Research Governance Committee and Research Management Team.

Readability statistics were obtained for each original and revised item using *Microsoft Word 98 Grammar Check* software (Microsoft Corporation, 1993). These statistics include Flesch Reading Ease (Flesch, 1948) and Flesch-Kincaid Grade Level (Kincaid et al., 1975), which measure the average number of syllables per word, average number of words per sentence and various elements of sentence composition. These statistics provide reading level estimates based on various percentages of the respective normative samples that answered comprehension questions regarding sets of standard passages. In this context American school grades correspond to specific ages for the children studying at that grade level depending upon the date of birth of the child in relation to the start of the academic year. The chronological ages associated with American school grades can be seen in Table 4.1 in Chapter 4.

Table 5.3

Examples of the Original and I7i-R Items

Item	Original Scale	Modified Scale
5	Do you prefer quiet parties with good conversation to 'wild' uninhibited ones?	Are quiet parties better than lively parties?
9	Do you usually think carefully before doing anything?	Do you think before you act?
15	Do you get so carried away by new and exciting ideas that you never think of possible problems?	Do you think about possible problems before you do something new?

After completing each questionnaire participants were also asked to rate the ease with which they understood the questionnaire on a 4-point scale (very easy / easy / ok / difficult) and to highlight any particular questions they found difficult to understand. Any request for repetition, clarification or further information for a particular item was interpreted as difficulty in understanding and recorded as such by the interviewer. In addition, participants were asked to rate the ease with which they answered the items. Again, they were asked to identify any particular items that they found difficult to answer. Any remarks indicating that the participant was having difficulty answering were again recorded by the interviewer. Responses that indicated an understanding of the general nature of the question but uncertainty as to whether to answer yes or no were interpreted as difficulty in answering the question and recorded as such by the interviewer. For example, participants often framed an answer that did not incorporate a yes or no response. In these circumstances the participant was pressed for a yes or no response. If this response did not correspond to their previously framed response it was recorded as difficult to answer. However, the answer actually recorded was the one that corresponded to the participant's original response.

5.7.5. Procedure: study 2a

The 47 detained offenders completed the I7i, I7i-R, R-BIS-11 and the BADS. The I7i-R was completed on two occasions in order to ascertain test-retest reliability. Half of the participants completed the original tool first and half completed the revised version first.

There was a break of one to two weeks between the various completions of the I7 and I7-R.

The first completion of the I7i-R was used in the comparison with the I7. The measures were administered by the current author and two therapeutic interventions nurses. The administration followed standardised instructions used in Chapter 4 (Nowicki & Strickland, 1973). Each statement was read out and the participant asked to indicate whether he agreed

with it or not. In the event of the participant being unable to answer, simple explanatory instructions were given. These were discussed and agreed between the administrators in advance. Both I7i and I7i-R were scored using scoring keys.

All participants completed a revised version of the BIS-11 (R-BIS-11) and the BADS.

Participant scores were tabulated. Statistical analysis was conducted on the data using PASW (Version 18) (SPSS Inc., 2010).

5.7.6. Procedure: study 2b

The third component of the study was an exploratory factor analysis of the I7i-R data in order to examine its factor structure in relation to people with intellectual disability. The validity of factor analysis with dichotomous variables has been established (Bartholomew, 1980; Bartholomew & Knott, 1999; Lucke & Schussler, 1986). The data were factored using a principal components analysis with a varimax rotation.

5.10. Results

5.10.1. Study 1: Creation of the I7i-R

The assumption of normally distributed data was assessed using histograms, and Kolmogorov-Smirnov and Shapiro-Wilk tests conducted on all of the I7i and I7i-R data. The results indicated that the distributions of all data for the I7i and the I7i-R were not significantly different to a normal distribution except for the BIS-11 Motor subscale, the BADS Test 1, Test 2, Test 3, Test 4, Test 5 and Test 6. Consequently, non-parametric statistics were conducted with these data. Table 5.4 shows the readability statistics for the I7i and the I7i-R. The length of each item was reduced by an average of 3.4 words making the I7i-R items 31% shorter than the original items. The mean Flesch Reading Ease statistic increased from 71.54 for the I7i to 97.03 for the I7i-R.

Table 5.4. Readability Statistics for the I7i and the I7i-R ($n=47$)

Readability Statistics					
	Total Words	Words per Sentence	Letters per Word	Mean Flesch Reading Ease	Mean Flesch-Kincaid Grade Reading Level
I7i	208	10.95	4.3	71.54 (SD=22.21)	5.96 (SD=3.24)
I7i-R	144	7.58	3.7	97.03 (SD=4.75)	1.25 (SD=1.05)

Figure 5.1 illustrates the mean Flesch Reading Ease statistic for both the I7i and the I7i-R with error bars. The small lack of overlap of item readability statistics, indicated by the error bars, suggests that these means are significantly different. Indeed, an independent samples t-test of the Flesch Reading Ease statistics found a significant difference in reading ease ($t_{(18)} = -4.877, p < .001, r = .75$) with a large effect size. Also, the variability in Flesch Reading Ease was reduced in the I7i-R as indicated by Levene's Test for equality of variances ($F_{(1, 36)} = 13.36, p < .001$) suggesting that the whole of the document was easier to read and not just certain parts of it (Kolton, Boer & Boer, 2001).

Similarly, the mean Flesch-Kincaid Grade Level statistic improved from 5.96 for the I7i to 1.25 for the I7i-R. Figure 5.2 shows the mean item Flesch-Kincaid Grade Level statistic for the I7i and I7i-R. In this instance the lack of overlap of Grade Reading Level statistic for items of each measure indicates that these means are significantly different. A paired samples t-test of the Flesch-Kincaid Grade Level statistic showed a significant difference in Grade Reading Level ($t_{(18)} = 5.864, p > .001, r = .81$), with a large effect size, between the I7i and the I7i-R. Also, the variability in Grade Reading Level was reduced in the I7i-R as indicated by Levene's Test for equality of variances ($F_{(1, 36)} = 11.41, p < .01$) suggesting that the whole of the document was easier to read and not just certain parts of it (Kolton et al, 2001). Consequently, the modifications significantly improved the readability of the I7i-R over the I7i.

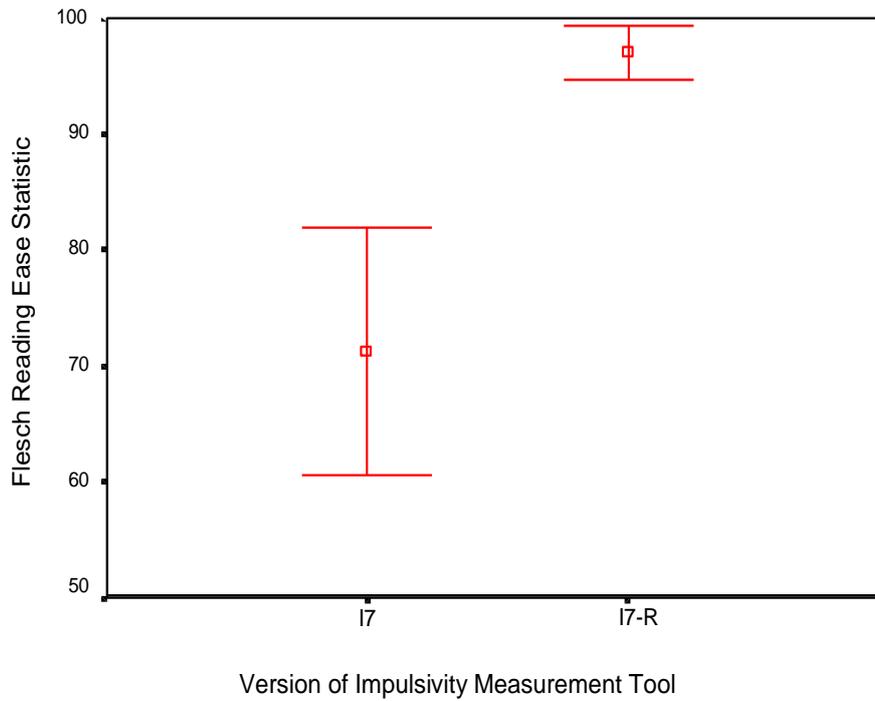


Figure 5.1. Error Bar Graph Showing Mean Flesch Reading Ease Statistic for the I7i and the I7i-R. Error Bars show 95.0% CI of Mean.

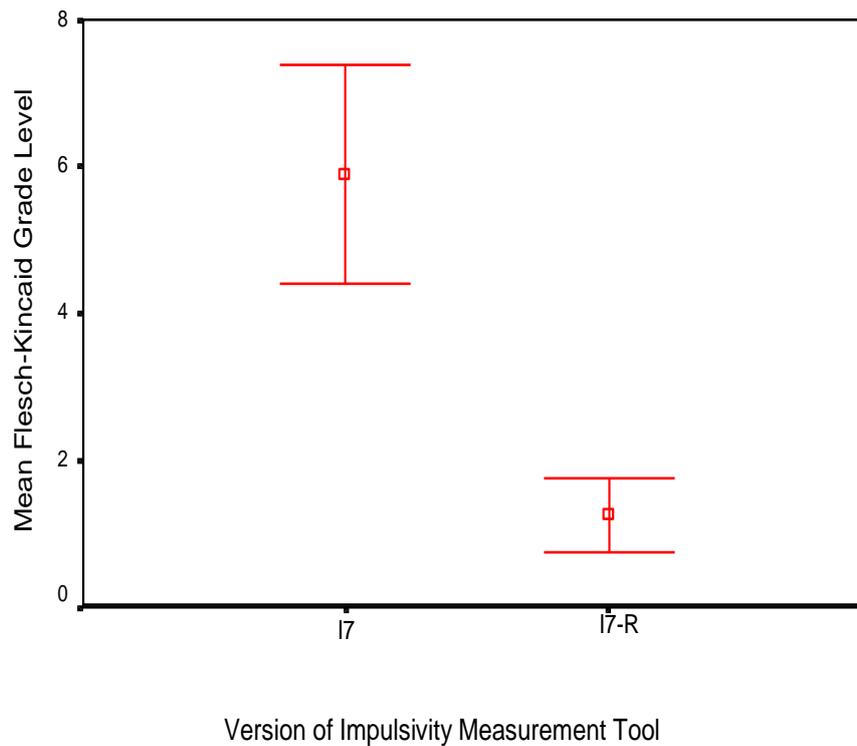


Figure 5.2. Error Bar Graph Showing Mean Item Flesch-Kincaid Grade Level Statistic for the I7i and the I7i-R. Error Bars show 95.0% CI of Mean.

The mean participant ratings for ease of understanding and ease of answering for both the I7i and the I7i-R are recorded in Table 5.5 and Figures 5.3 and 5.4

Table 5.5. The mean participant ratings for ease of understanding and ease of answering the I7i and I7i-R.

	I7I	SD	I7I-R	SD
Ease of Understanding	2.04	1.02	2.67	0.83
Ease of Answering	1.74	0.66	2.44	0.93

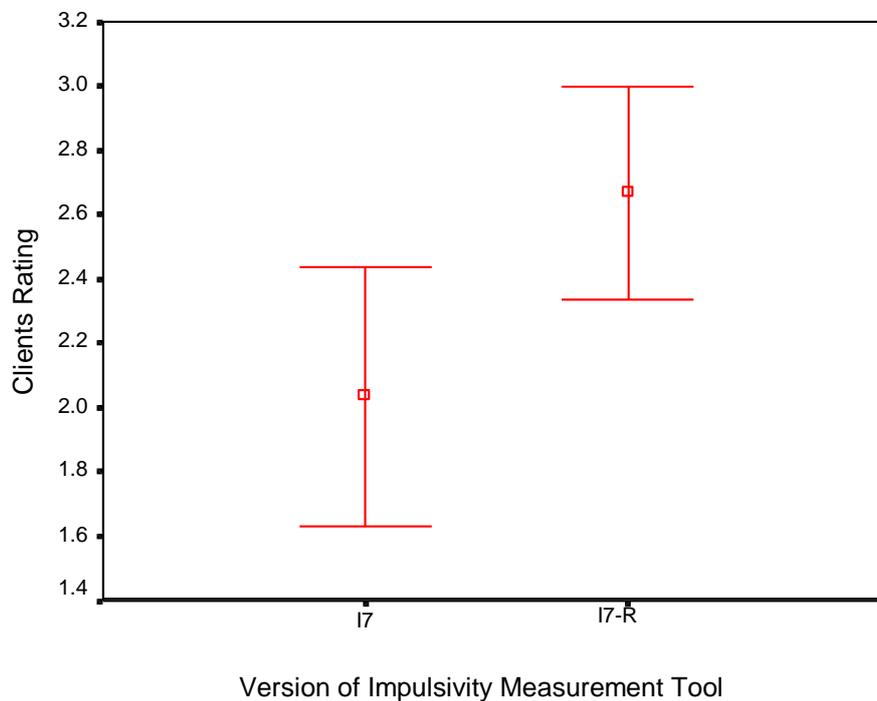


Figure 5.3. Error Bar Graph Showing Mean Participant Ratings for Ease of Understanding the I7i and I7i-R. Error Bars show 95.0% CI of Mean.

The slight overlap of participant ratings between the I7i and the I7i-R, indicated by the error bars, suggests that these means are likely to be significantly different. Indeed, a paired samples t-tests of participant’s ratings indicated that they rated the I7i-R as

significantly easier to understand than the I7i ($t_{(46)} = -2.56, p < .05, r = .35$), with a medium effect size.

Figure 5.4 illustrates the mean client rating for ease of answering the I7i and the I7i-R. The small degree of overlap suggests that these means are significantly different. A paired-samples t-test of participants' rating of ease of answering indicated that they found the I7i-R significantly easier to answer than the I7i ($t_{(46)} = -4.009, p < .001, r = .51$), with a large effect size.

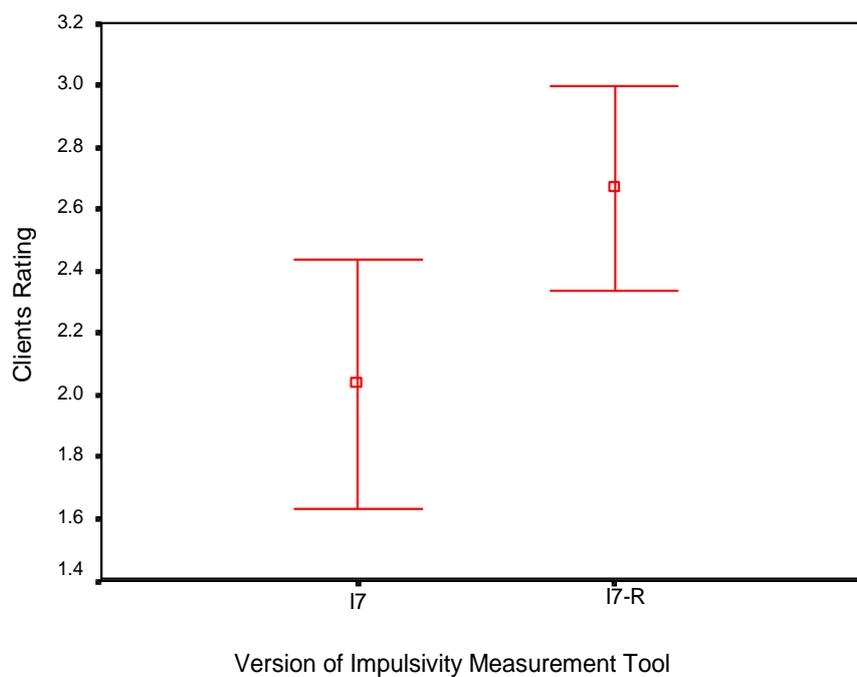


Figure 5.4. Error Bar Graph Illustrating Mean Participant Ratings for Ease of Answering the I7i and the I7i-R. Error Bars show 95.0% CI of Mean.

Table 5.6 and figures 5.5 and 5.6 illustrate the number of items recorded as difficult to understand or difficult to answer. Again, the lack of overlap of error bars in figures 5.5 and 5.6 indicates that these means are significantly different. Again, a paired samples t-test confirmed that the number of requests for repetition, clarification or further information were significantly reduced with the I7i-R compared to the I7i ($t_{(46)} = 8.503, p < .001, r = .78$), with a

large effect size. Equally, the mean number of responses provided that were incongruent with associated explanations were also significantly reduced with the I7i-R than the I7i ($t_{(46)}=4.549, p<.001, r=.56$), with a large effect size.

Table 5.6.

Mean Number of Items Identified as Either Difficult to Understand or Difficult to Answer on the I7i and I7i-R.

Mean Number of Items	I7i	SD	I7i-R	SD
Difficult to Understand	4.70	2.05	1.33	1.52
Difficult to Answer	4.26	2.64	1.81	1.66

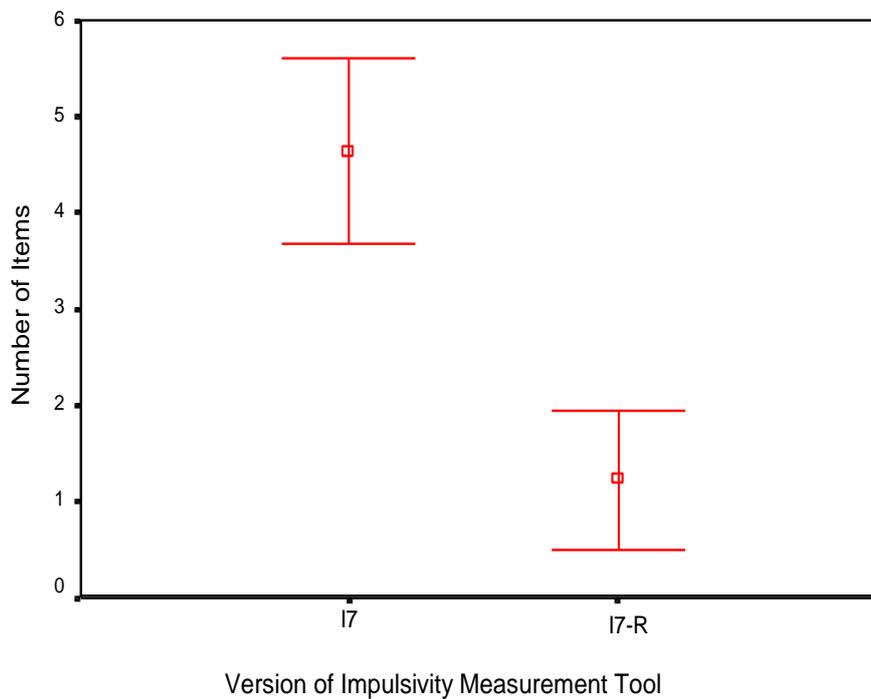


Figure 5.5. Error Bar Graph Illustrating the Mean Number of Items Identified as Difficult to Understand on the I7i and the I7i-R. Error Bars show 95.0% CI of Mean



Figure 5.6. Error Bar Graph Illustrating the Mean Number of Items Identified as Difficult to Answer on the I7i and the I7i-R. Error Bars show 95.0% CI of Mean

Each of the t-tests demonstrates that the modifications significantly improved the readability of the I7i-R over the I7i with items that were shorter and easier for participants to understand whilst maintaining their essential meaning as judged by 5 professional skilled in the assessment of people with ID. These changes appear to enable offenders with learning disabilities to more easily understand and respond to test items thus facilitating a more effective evaluation of their self-report impulsivity.

The current author was interested in assessing to what extent the reduction in the mean score associated with the I7i-R was the result of reduction in acquiescence. To that end, the number of items where an endorsed impulsive response was indicated by a ‘yes’ was compared with the number of items where an endorsed impulsive response was indicated by a ‘no’ for both the I7i and the I7i-R. The number of items endorsed in the external direction according to a ‘Yes’ or a ‘No’ response are shown in Table 4.7.

Table 5.7

The Percentage of Items Endorsed as Impulsive by Yes and No Responses

	Number of Items			Number of Impulsive Responses Endorsed		
	Total Number of Items	Number of Items with 'Yes' Impulsive	Number of Items with 'No' Impulsive	Total Number of Impulsive Responses	Impulsive Response 'Yes'	Impulsive Response 'No'
I7i	19	15 (78.95%)	4 (21.05%)	448	409 (91.26%)	39 (8.74%)
I7i-R	19	13 (68.42%)	6 (31.58%)	411	308 (74.94%)	103 (25.06%)

Whilst no statistical analysis has been conducted on the above data the findings are suggestive of a reduction in acquiescence. If responses were based solely on question content one would expect approximately 78.95% of impulsive responses to the I7i to correspond to a 'Yes' but only 68.42% on the I7i-R. In fact impulsive responses corresponding to a 'Yes' were above these percentages on both questionnaires, 91.26% on the I7i and 74.94% on the I7i-R, strongly suggesting the presence of acquiescence because participants appear to be more likely to endorse an impulsive response if it corresponds with a 'Yes'. However, that possible over-endorsement of 'Yes' responses is more pronounced on the I7i by 12.31% compared with 6.52% on the I7i-R. This suggests less of a bias towards acquiescence on the I7i-R but as already stated this has not been tested statistically.

5.8.2. Study 2a: Testing the psychometric properties of I7i-R

Cronbach's alpha for the I7i ($\alpha=.71$) and the I7i-R ($\alpha=.74$) indicated that they had acceptable levels of internal consistency for a psychological construct (Kline, 1999). In addition, the majority of items had item-to-total correlations in excess of 0.3 as recommended by Field (2009). However, several item-to-total correlations failed to reach that level for the I7i and the I7i-R. For the I7i; item 4; item 9; item 14; and item 15. For the I7i-R item 4; item 9; and

item 14. This suggests some problems with the reliability of these items. The mean scores and standard deviations obtained on the I7i, I7i-R, BIS-11 and its subscales are recorded in Table 5.8.

Table 5.8.

Means Scores and Standard Deviations Obtained on the I7i and I7i-R

Measure	Mean Score	SD
I7i	8.36	3.84
I7i-R	7.57	3.13
BIS-11	68.36	10.97
Attention	17.47	4.43
Motor	24.17	3.69
Non-planning	26.17	5.05

Participant scores on the I7i and I7i-R are illustrated in Figure 5.7. The I7i and I7i-R scores are slightly lower than those of 8.76 (4.79) and 9.84 (4.13) respectively, reported for students (Corrulla, 1987; Eysenck et al, 1985). In addition, the BIS-11 Total was lower in the present study than those reported for sex offenders (73.17%) and violent offenders (83.44%) reported by Snoyman & Aicken, 2011). The scatterplot appears to indicate a trend in the data in the expected manner with scores on the two tools mirroring each other. As individual scores on the I7i increase so their scores on the I7i-R increase. A paired-samples t-test indicated that there was a significant difference between the means on the two measures ($t_{(46)} = 2.17, p < .001, r = .30$) with a medium effect size.

The Pearson's correlation coefficient between the total scores of the I7i and the I7i-R was $r = .822$ ($p < .01$). This indicates a strong positive association between the scores on the I7i and the I7i-R. Consequently, it appears that the I7i and the I7i-R are probing the same construct. Spearman correlations between the self-report measures of impulsivity are represented in Table 5.9.

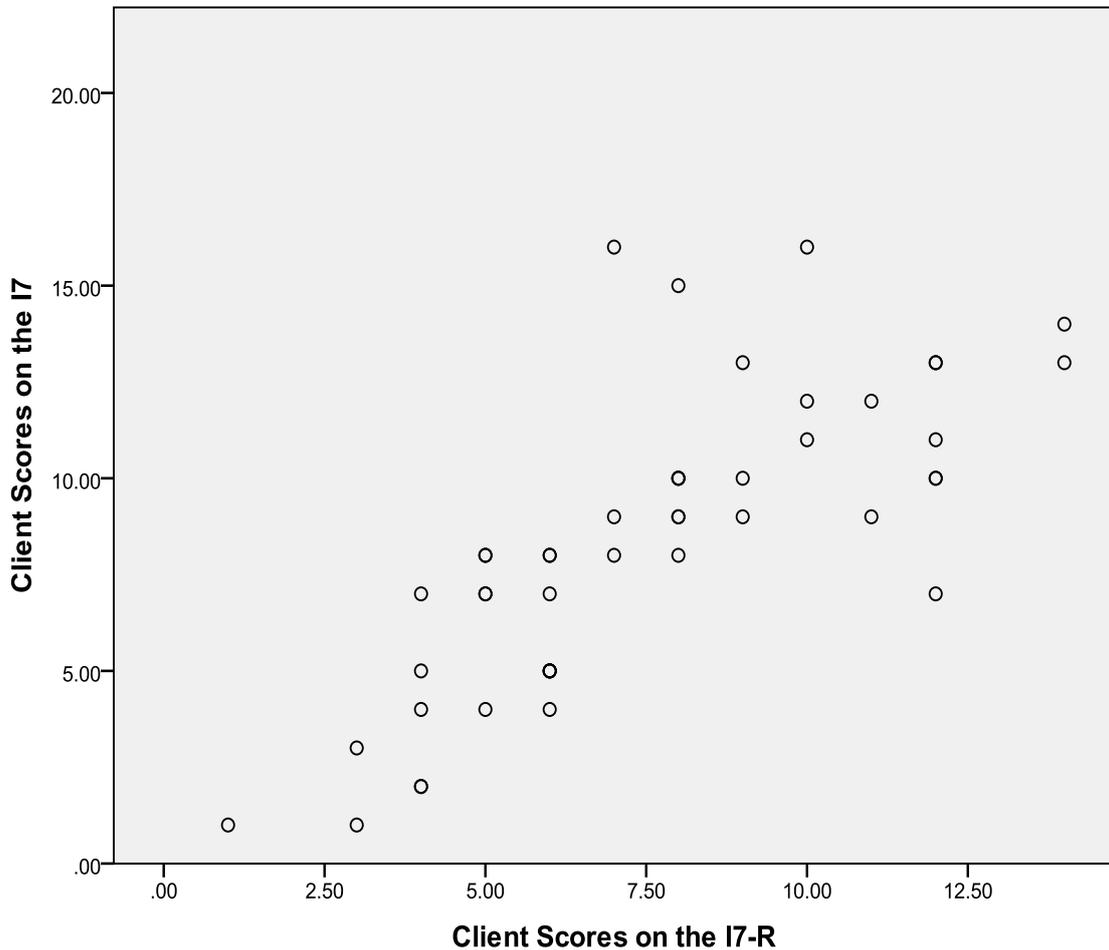


Figure 5.7. Scatterplot Illustrating the Correlation between the Participant’s Scores on the I7i and the I7i-R.

There were highly significant, moderate correlations between the total scores on these two measures and the BIS-11 Total and Factor II scores. There were also significant, moderate correlations between I7i and I7i-R scores and BIS Factor I and Factor III scores. In addition, there were highly significant, strong correlations between BIS Total and Factors I and III, as well as a highly significant, moderate correlation between BIS Total and BIS Factor II. There was also a highly significant, strong correlation between Factor I and Factor III. Correlations were also conducted between the I7i, the I7i-R and the BADS Total and sub-scale scores. The results of these correlations are recorded in Table 5.10.

Table 5.9

Spearman Correlations of Scores on Self-Report Measures of Impulsivity

	I-7	I-7R	BIS-11 Total	Factor 1	Factor 2	Factor 3
I-7	1.0					
I-7R	.822**	1.0				
BIS Total	.487**	.431**	1.0			
Factor I (Attentional)	.341*	.313*	.779**	1.0		
Factor II (Motor)	.471**	.408**	.453**	.91	1.0	
Factor III (Non-planning)	.337*	.319*	.796**	.594**	.192	1.0

*p<.05 **p<.01

Table 5.10.

Spearman correlations for the self-report measures of impulsivity and the BADS

Total and subscales.

	BADS Total	Test 1	Test 2	Test 3	Test 4	Test 5	Test 6
I7i	.342*	.410**	-.018	.438**	-.001	-.068	.190
I7i-R	.222	.373**	.007	.462**	-.038	-.126	.006
BIS Total	.323*	.147	.131	.101	.138	.187	-.378
Factor I	.185	.067	-.136	.043	.212	.220	.252
Factor II	.262	.187	.315*	-.205	.170	.210	-.049
Factor III	.458**	.261	.111	.249	.156	.149	.522**

*p<.05 **p<.01

There were few significant correlations between self-report scores and clinically impulsive behaviour. The I7i total score correlated weakly with the BADS total ($r=.342$, $p<.05$), and moderately with Test 1 (Rule Shift) ($r=.410$, $p<.01$) and Test 3 (Key Search) ($r=-.438$, $p<.01$). The I7i-R correlated weakly with performance on Test 1 (Rule Shift) ($r=-.373$, $p<.01$) and moderately with the Key Search Test ($r=.462$, $p<.01$). In addition, the BIS Factor III (Non-planning) correlated moderately with Test 6 (Modified Six Elements) ($r=.522$, $p<.01$).

5.8.3. Study 2b: Factor analysis of the I7i-R

Prior to performing the PCA analysis, the suitability of the data was assessed. The Kaiser-Meyer-Olkin measure ($KMO = .58$), although in the barely acceptable range (Kaiser, 1974), verified the adequacy of the data for the analysis. In addition, Bartlett's test of sphericity ($X^2_{(171)}=321.078$, $p<.001$) indicated that correlations between items were sufficiently large for PCA. Previous research indicates that all of the items of the I7i-R load on to a single factor when delivered as part of the whole I7 questionnaire (Eysenck et al., 1985). Consequently, any identified factors might be highly correlated and indicate the need for an oblique rotation method. Various authors (Field, 2009; Pallant, 2010) recommend the initial completion of both orthogonal and oblique rotations in order to aid factor interpretation and to determine the appropriate rotation method. Inspection of the Component Correlation Matrix produced by the Oblimin rotation indicates that there was little correlation between the identified factors, the strongest correlation, between factors 6 and 7 being just .123. This indicated that the orthogonal, Varimax rotation was appropriate, which is therefore reported here. Data for the 19 items were factored using a principal components analysis (PCA) in order to obtain eigenvalues for all items. This produced 7 factors with eigenvalues greater than Kaiser's criterion of 1, accounting for 58.25% of the variance. The identified factors were subjected

to a varimax rotation. Criterion for item inclusion was a factor loading of at least .40, which Stevens (2002) recommends in a sample of this size. A minimum of three items was required to establish a coherent theme for a factor and items loading on more than one factor were only included on the factor on which they had the highest loading. This process produced 4 factors accounting for 38.82% of the variance. The associated scree plot (Catell, 1966) was ambiguous with possible changes in gradient indicating the inclusion of 7 factors, or using the more obvious change in gradient, 2 factors (see Figure 5.8).

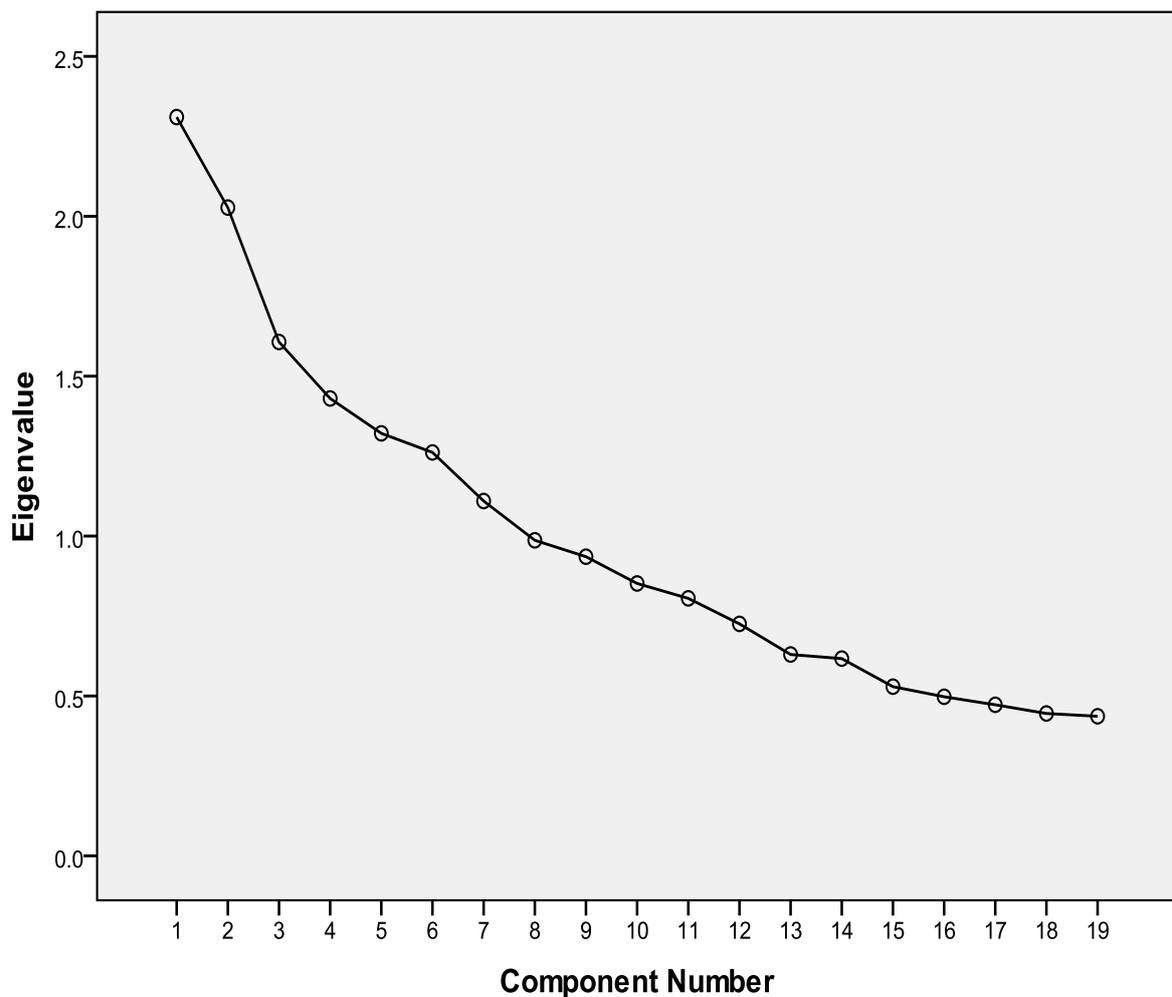


Figure 5.8. Scree Plot of Eigenvalues for the Principal Components Factor Analysis of the I7i-R.

In terms of methods for deciding on the number of factors to retain, Kaiser’s criterion and the scree plot methods have both been criticised as resulting in the retention of too many factors in some situations (Hubbard & Allen, 1987; Zwick & Velicer, 1986). In addition, the sample size used in this instance and the low communalities after extraction, with none over the 0.7 mark suggested by Field (2009), indicate that Kaiser’s criterion may not be accurate.

In view of the difficulties identifying the number of factors to retain, a parallel analysis was conducted (Horn, 1965). For details of a parallel analysis see Chapter 4. A Monte Carlo analysis was conducted (Watkins, 2000) which indicated that 6 factors should be retained. The data analysis was repeated with the specification of the extraction of 6 factors. These 6 factors accounted for 52.41% of the variance. Table 5.11 shows the individual factors’ eigenvalues and their contribution to the overall variance explained.

Table: 5.11

Eigenvalues and Contribution to the Overall Variance Explained by the 6 Factors

Factor	Eigenvalue	% of Variance	Cumulative %
1	2.310	12.159	12.159
2	2.027	10.670	22.829
3	1.607	8.458	31.287
4	1.430	7.528	38.815
5	1.321	6.954	45.769
6	1.262	6.640	52.409

The item content for these 6 factors is outlined in Table 5.8. Due to the similarity of item content it was not possible to discern any specific or particular theme for individual factors apart from Factor 3 which appeared to be about making quick decisions. Factor 3 appeared to have a theme of quick decision making. The other factors did not appear to have an identifiable theme.

Table: 5.12

Summary of the Item Content of the Identified Factors

Factor 1:	Factor 2:	Factor 3: Making Quick Decisions	Factor 4:	Factor 5:	Factor 6:
Think before making up mind	Check your work	Decide straight away	Buy things without thinking	Best things in life bad for you	Speak before you think
Think before you act	Think what might go wrong before doing it	Make up your mind quickly	Get into things you wish you could get out of	Say things before you think	Do things before you think
People shocked by things you say	Change interests		Act now and think later	Rush into things	Day out better if planned
	Say things before you think				Easy to stay out of trouble

5.9. Discussion

Evidence suggests that impulsivity is related to offending behaviour. However, the relationship between impulsivity and offending in populations with ID is not clear. Consequently it is important that impulsivity can be assessed accurately in this population, since, as was explained in Chapter 1, 2.9% of the prison population is ID with 21.7% falling in the borderline IQ range or below (Hayes et al., 2007). The validity of instruments developed to measure impulsivity in populations with ID has not been demonstrated, and we might expect measures developed with a non-ID population to have poor application due to the threat of response biases in samples with ID. The effects of response biases can be minimised by simplifying question and response formats and ensuring that question content is concrete and familiar. The first study in this chapter demonstrated that the I7i-R showed

significantly improved readability in comparison to the I7i, thus supporting hypothesis 1. In addition, it was demonstrated that the I7i-R was both significantly easier to understand and significantly easier to answer than the I7i. This was in terms of participants' ratings, the reduced number of requests for repetition, clarification or further information made by participants, and the number of responses provided that were incongruent with associated explanations. Thus hypothesis 2 was also supported.

The aim of study 2a was to examine the psychometric properties of the I7i-R. This demonstrated that participants' scores on the I7i and the I7i-R were highly correlated, providing construct validity for the I7i-R, as it appears to be measuring the same construct as the I7i, namely impulsivity. Thus, hypothesis 3 was also supported. However, mean scores on the I7 were significantly higher than mean scores on the I7i-R. This point is returned to below. In addition, a strong positive correlation was found between two administrations of the I7i-R providing evidence of test re-test reliability, thus hypothesis 4 was also supported. In addition, there were weak and moderate positive correlations between the I7i-R and the BIS-11-R and its subscales, providing construct validity for the I7i-R. However, hypothesis 5 is only partially supported because the I7i-R was not correlated to the BADS total score, unlike the I7i which has a significant medium size correlation with the BADS total. The reason for this is unclear because the patterns of intercorrelation between the I7i and the I7i-R with the BADS and its subscales are similar. Two tasks show slightly larger differences in correlation. The I7i-R had a slightly higher correlation with task 5 (Zoo Map) (-.126 to -.068) and a slightly lower correlation with the task 6 (Six Modified Elements) (.006 to .190). It is not clear why the I7i-R was not significantly correlated to the BADS total.

The fact that scores on the I7i-R were significantly lower than on the I7i has two possible explanations. One is that the I7i was over-estimating impulsivity in this population and the second is that the I7i-R is under-estimating impulsivity in this population. Based

upon the available evidence it would appear that the first of these two options is the more likely. Evidence suggests that people with a ID are more likely to be acquiescent, responding to dichotomous questions with a 'yes' response more frequently than a 'no', regardless of question content (Sigelman et al, 1981a).

In all, the direction of 3 items were changed in the construction of the I7i-R. That is to say 3 more items on the I7i-R than the I7i had an impulsive response that corresponded to a negative response. If acquiescence, through affirmative responding, was prevalent in this sample one may expect to see an increased level of self-report impulsive responding with the I7i (Clare & Gudjonsson, 1993; Sigelman et al., 1981b). Consistent with this, the mean score on the I7i was significantly higher than that on the I7i-R. This suggests that the original I7i appears to have been over representing the actual levels of self-report impulsivity in this population of offenders with ID. Consequently, acquiescence through affirmative responding may indeed have taken place during testing with the original I7i.

In fact the evidence presented above suggests that both the I7i and the I7i-R may both still lead to acquiescent responding in populations with an intellectual disability, however that tendency appears to be reduced in relation to the I7i-R. If the I7i displayed no response bias, one would expect the percentage of impulsive items indicated by endorsing a 'yes' or 'no' response to match the overall percentage of items in the instrument where an impulsive response correspond with a 'yes' or 'no' respectively. On 78.95% of the items on the I7i a 'yes' response corresponded to an impulsive rating and consequently, on 21.05% of the items a 'no' response corresponded to an impulsive rating. However, 91.26% of the total number of impulsive responses was signified by the respondent endorsing a 'yes'. On the I7i-R 68.42% of items had an impulsive rating that corresponded to a 'yes' response, however 74.94% of the total number of impulsive ratings were actually indicated by endorsing a 'yes' response. Consequently, the number of 'yes' responses are 12.31% more than would be

expected with the I7i and 6.52% higher than would be expected with the I7i-R. It is apparent that the percentage difference above that expected for a 'yes' endorsement representing an impulsive response is smaller with the I7i-R compared to the I7i.

The mean for the I7i of 8.36 ($SD=3.84$) and the I7i-R of 7.57 ($SD=3.13$) are similar to those of 8.76 ($SD=4.31$) reported by Eysenck et al (1985), for non-offending adults, and the 8.76 ($SD=4.79$) reported by Corulla (1987), with another sample of non-offending adults. This suggests that elevated impulsivity may not differentiate offenders with an ID from non-offenders with ID. A study examining whether self-reported impulsivity, as measured by the I7i-R, can indeed differentiate offenders with LD from non-offenders with LD is reported on in Chapter 7.

It is apparent from the results that the scores on the R-BIS correlated better with scores on the I7i ($r=.487$) than the I7i-R ($r=.431$). Indeed, the scores on the I7i showed higher correlation with all of the R-BIS subscales than did the I7i-R. It is not clear why this pattern of results occurred. One explanation is that the R-BIS and the I7i are providing more valid measures of impulsive behaviour. However, this explanation is not consistent with the findings suggesting that acquiescence is more pronounced in responses to the I7i than the I7i-R. In addition, this explanation is not consistent with the attempted revision of the BIS-11 outlined earlier, which was largely unsatisfactory in that response biases continued to be apparent in the scoring of the Likert scale. It is possible therefore that the increased correlation between scores on the I7i and the BIS were actually an artefact of the data and that the correlations actually represent relationships between response biases rather than question contents. However, further studies would be required to clarify the relationship between the scores on these measures.

The I7i and I7i-R were both moderately correlated with the Rule Shift Test and the Key Search Test from the BADS. The rule shift test is regarded as sensitive to impulsivity as

failure on the test may indicate a failure to inhibit an inappropriate response. Correlation with the Key Search Test was unexpected as it is more likely to indicate difficulties with planning and implementing an effective solution strategy and therefore perhaps more representative of the broader impulsivity factor.

The R-BIS mean total score obtained in this study of 68.36 ($SD=10.97$) far exceeds that reported by Parry and Lindsay (2003) of 39.07 ($SD=39.07$) in relation to the 35 offenders in their sample. It also exceeds their mean of 45.74 ($SD=16.58$) when all of their sample, including offenders and non-offenders, were considered together. It is likely that this is due to the different version of the BIS-11 used and the different scoring system employed. However, the score obtained here is lower than those reported by Snoyman and Aicken (2011) of 73.17 for sex offenders and 83.44 for violent offenders. This may have been partly due to the difficulties experienced in adapting the measure in the current study.

The success of the adaptation to the I7i was in contrast to attempted modifications of the BIS-11. The original version of the BIS-11 proved very difficult for participants to use in two respects, relating to question content and response format. Question content was such that participants requested large amounts of additional information before making a response. In addition, it was often clear from additional information provided that whilst they felt that they had understood a question and provided a response, they had actually misinterpreted the question. For example, 'Do you find it difficult to listen?' was clearly often interpreted as 'Is your hearing OK?', made apparent by references to ear appointments, syringes and complaints about others playing music too loud. On other occasions this was due to the complexity of the question, for example it proved difficult to interpret 'Do your thoughts race all over the place?', 'Do you often have irrelevant thoughts when thinking?' and 'Are you more interested in the present than dreaming about the future?' More often however it was due to the context within the question. For example, 'Do you feel awkward at talks or

presentations?', 'Do you think about job security?', 'Do you change the way you spend your days?', 'Do you change where you live a lot?' and 'Are you restless at talks and presentations?' These questions often resulted in discussions seeking to find a context for the individual that might make sense, such as education classes, church, ward meetings or one of a number of formal meetings that the individual may attend about their care. As a result of this variation in context it was felt that validity and reliability of the responses may have suffered.

The second major area of difficulty revolved around the response format of the BIS-11. There appeared to be clear evidence of invalid responding. A slight revision to the scoring was developed using geometric shapes shaded in to a greater or lesser extent. The rarely/never response was associated with an oblong with shading up to one quarter of its length. In a similar way, sometimes (one half shaded), often (three quarters shaded) and almost always/always (completely shaded). Participants' were encouraged to participate in some practice regarding their known likes and dislikes ('Do you brush your teeth in the morning?'). However, this adaptation did not prevent some respondents from responding to most questions by selecting the same option, 'sometimes'. Over 40% of participants scored 25 or more items in the same category. In addition, subsequent statements continued to indicate that invalid responses had been given. Finally, the BIS proved far more difficult to adapt than the I7i. Clarifying question content often resulted in longer questions and consequently higher Flesch Reading Ease statistics or in question content that did not obviously match the original questionnaire. For example, 'Do you get bored solving thought problems?' became 'Do you like solving maths questions?', 'Do you think about job security?' became 'Do you change the things you do at day-services?' and 'Do you like to think about difficult problems?' became 'Do you like to do sudoku?' Even after several

iterations two psychologists asked to comment on the similarity of the content between the two versions felt that there was a marked difference on a sizeable number of items.

The aim of study 2b was to examine the factor structure of the I7i-R. Work by Eysenck and colleagues (Eysenck & Eysenck, 1977; 1978; Eysenck et al., 1985; Eysenck & Zuckerman, 1978) indicated that the impulsivity items of the I7 questionnaire loaded on a single factor, when delivered and analysed in conjunction with the items of the venturesomeness and empathy subscales. This study did not identify a unitary factor for the I7i-R when administered with offenders with ID. Indeed, a Monte Carlo analysis indicated the presence of six factors; however, it was not possible to discern the theme of individual factors, beyond one relating to making decisions quickly. It is apparent that the Eysenck's constrained the identification of factors to the three expected, impulsivity, venturesomeness and empathy, without further explanation. It is possible that this prevented the identification of clusters of items within the impulsivity and other factor items. When a CPA was conducted with the specification of a single factor, that factor accounted for only 12.16% of the variance in the data. In addition, only 13 of the 19 items loaded onto the single factor, even when specifying a minimal factor loading of 0.1. Ethical issues prevented the administration of the full I7 questionnaire to all 139 participants in this study. This was partly due to the time required to complete the 63-item questionnaire with the community samples, outlined in Chapter 7. It was also felt inappropriate to assess potential empathy deficits in a community sample. In addition, there would be treatment and resource implications to any deficits identified.

The significant increase in the reading-ease, and associated reduction in grade reading level, of the I7i-R over the I7i resulted from the shorter and simpler sentence structures and the substitution of complex semantics for more straightforward language. The increase in the reading ease improved its readability by two categories from 'Fairly Easy' to 'Very Easy'

(Flesch, 1984). Flesch equates this change to the difference in language typically used in 'slick-fiction' (i.e. a category harder to read than 'pulp-fiction') to that typically used in 'comics'. However, it is likely that the Flesch-Kincaid Grade Level statistic actually underestimated the reading difficulty of the I7i for this population. This statistic was developed for use with technical language that was familiar to the reader i.e. it assumes knowledge of relatively complex words. Consequently, it underestimates language that is not familiar to the reader. This is likely to be more representative of the I7i than the I7i-R as it fails to reflect the life context of this population. This was evident in the number of requests for further information associated with administration of the I7i.

5.10. Conclusions

The development of the I7i-R was designed to overcome some of the deficits in the methods of assessing individuals with impaired intellectual functioning associated with learning disability (Clare, 1993). Firstly, the stipulation that the I7i-R be administered verbally will help to counter the reading difficulties common in this population (Clare, 1993). In conjunction with this the shorter and simpler sentence structures, as indicated by significant improvements in readability statistics (Flesch, 1948; Kincaid et al., 1975), in the I7i-R should reduce the load on verbal memory thus facilitating non-random responding.

Clearly, a dynamic study of impulsivity in this population is necessary to track its variability pre- and post- treatment and at follow-up. However, the primary aim of this study was to modify the I7i for a population with ID. However, the sample size was small and there was no comparison group so it is unwise to generalise from the results. Chapter 7 reports on a study assessing impulsivity in this sample of 47 detained offenders with ID compared with two samples of matched non-offenders.

Impulsivity has been associated with numerous antisocial behaviours. It is therefore necessary to understand whether the impulsivity observed in offenders with ID represents a vulnerability to the engagement in antisocial and criminal behaviours compared with non-offenders with ID. Such vulnerability may help to account for the possible over representation of learning disabled offenders in criminal, prison and hospital populations (Clare & Murphy, 1998).

Until the exact nature of the construct can be more clearly defined it is difficult to assess whether unitary measures of impulsivity are necessarily comparable and consequently whether they represent vulnerability or not. The construct of Narrow Impulsivity (Eysenck and Eysenck, 1978) as defined by the I7i was selected in part because it correlated with P, in particular, and N, somewhat. Eysenck speculated that this measure therefore represented something pathological as it correlated with the two ‘abnormal scales’. Specifically Eysenck suggested that it was likely to be sensitive to the impulsivity present in groups such as criminals and psychopaths. Clearly, it is important to identify whether impulsivity, as defined by the I7i-R represents a particular pathological component of the construct. Clearly, any specifically pathological component of impulsivity that represents vulnerability in a criminological sense must be identified to better inform risk assessment, treatment objectives and interventions.

A major problem for the researcher was the sparse nature of the literature base for the investigation of impulsivity in people with ID in general. Despite an extensive literature search the author was able to identify only two published works assessing impulsivity in offenders with ID, outlined above (Parry & Lindsay, 2003; Snoyman & Aicken, 2011). Perhaps of greater concern still was the lack of up to date published means for impulsivity in a representative sample of people with ID. Whilst there is some literature regarding the use of risk assessment tools such as the HCR-20 (Gray, Fitzgerald, Taylor, MacCulloch, &

Snowden, 2007; Lindsay et al., 2008) which incorporates an item regarding impulsive behaviour, these ratings of impulsive behaviour are based upon a specific behavioural description and are not self-report in nature. These studies are described in Chapter 6. The lack of research regarding ID populations is well-documented (Clare & Murphy, 1998). In addition, where suitable psychometrics exist for use with ID populations there are rarely norms for these populations (Clare & Murphy, 1998).

This study has demonstrated that the impulsivity assessment tool of choice for many professionals, the I7i, caused great difficulty in understanding and answering for an institutionalised ID offender population. Consequently, this study supports the view of Clare (1993) in arguing that assessment tools must be specifically adapted for use with learning disabled populations. The development of the I7i-R demonstrated that the reading ease of the scale could be significantly improved thus making it more accessible to this population. In addition, the development of the I7i-R resulted in an assessment that participants rated as easier to understand and answer. It also resulted in significantly fewer requests for repetition, clarification and further information and produced fewer responses that were incongruent with associated explanations. It therefore seems plausible that the I7i-R is more likely to provide an accurate assessment of impulsivity in this population. A strong correlation between scores on the I7i and I7i-R provide validity for the new measure. In addition, strong correlations between the I7i-R and the R-BIS-11 and correlations with some of the subscales of the BADS provide more construct validity. A strong correlation between original and re-test scores on the I7i-R indicate that it is a reliable measure. Mean scores on the I7i-R were significantly lower than on the I7i, suggesting that the modifications to the scale may have reduced the level of acquiescent responding in this population. Support was added to this conclusion by the reduction in the percentage of external responses indicated by the endorsement of a 'Yes', suggesting that the I7i-R is facilitating a more valid assessment of

impulsivity through a reduction in acquiescence. A Monte Carlo analysis indicated the retention of 6 factors accounting for 52.41% of the variance. The results reported in this study indicate that the I7i-R is more likely to provide an accurate assessment of impulsivity in this population than the I7i.

CHAPTER SIX: THE ROLE OF IMPULSIVITY IN THE RISK ASSESSMENT OF OFFENDERS WITH INTELLECTUAL DISABILITY IN A MEDIUM SECURE HOSPITAL SETTING

6.0. Introduction

The aim of this chapter was to assess the ability of the I7i, the I7i-R, the PCL-R and the HCR-20 in predicting institutional aggression in a sample of offenders with ID detained in a medium secure facility. Recent research (e.g., Morrissey et al., 2007a, 2007b) has suggested that assessment tools such as the HCR-20 and PCL-R do not consistently predict institutional violence or treatment progress in offenders with ID detained in a high secure hospital setting. However, this may have been due to the high levels of violence reported in these studies, which may have masked any predictive ability. The present study examined the relative ability of the I7i, I7i-R, Psychopathy Check List - Revised (PCL-R) and the HCR-20 to predict recorded institutional aggression in a sample of 47 offenders with ID living in a medium secure forensic setting. Aggressive incidents were recorded according to three definitions; Verbal/Property Aggression (Type 1), Interpersonal Physical Violence (Type 2), and High-Risk Violence (Type 3). Aggressive incidents were collated over a three month period. The study described in the chapter is divided into two sections. In the first part of the study, the mean of each measure is reported, along with the correlation of the scores on each measure with aggressive incidents. In the second part of the study a Receiver Operator Characteristic (ROC) analysis was conducted to assess whether any of the measures was able to predict any of the forms of institutional aggression.

6.1. Risk assessment in offenders with ID

The main aim in the development of risk assessment is the identification of factors associated with offending, or re-offending (Callimeri & Quinsey, 2011). As a result large scale studies in mainstream non-ID samples have led to the identification of many such risk factors (Hanson & Bussière, 1998). This in turn has led to the development of tools combining these factors as a means of increasing their predictive ability to differentiate those who are more likely to offend (Hanson, & Thornton, 2000; Harris, Rice, & Quinsey, 1993).

Many of these tools have been used in samples with ID with the aim of assessing their reliability and validity (Lindsay, Carson, et al., 2010; Lindsay et al., 2008; Morrissey et al., 2007). One means of assessing the validity of such measures has been to consider their ability to measure anti-social behaviour in institutional settings (Morrissey et al., 2007). In addition, the high levels of violence reported in such institutions (Morrissey et al., 2007) presents continuing challenges and risk to the staff in those establishments. The reliable and valid assessment of risk can inform decisions regarding risk management, with, for example, level of risk associated with level of security. This would ensure that offenders are detained at the appropriate level of security and ensure that the institutions are appropriately resources to meet those differing levels risk.

In the present study the ability of two measures commonly used for risk assessment, the HCR-20 and the PCL-R, will be considered along with measures of impulsivity outlined in chapter 5 in predicting institutional aggression. Consequently this chapter will consider the evidence regarding the use of risk assessment tools in samples with ID. In addition, the HCR-20 contains an item relating to the diagnosis of personality disorder as an identified risk factor and the PCL-R is considered an acute form of PD (Hare, 2003; Morrissey, 2010). Consequently, the applicability of mainstream approaches to personality will be considered, as will the assessment of personality disorder in samples of ID.

From the prevalence data described in chapter 2 it is not possible to conclude if a diagnosis of ID increases a person's risk of offending. What does appear clear is that offenders in the borderline range of intelligence appear to be over-represented in the offender population (Crocker et al., 2007; Hayes et al., 2007; Sheldon, 2006). However, such studies do not inform which offenders from a population of offenders with ID will offend in the future. Consequently, there is a need for reliable and valid risk assessment tools in order to identify those who pose a greater risk of reoffending. This in turn can inform decisions about the appropriate security level at which someone is detained, return to the community and risk management in that setting (Craig, Stringer, & Moss, 2006). Consequently work has been conducted on appraising risk in offenders with ID (Camilleri, & Quinsey, 2011; Harris & Tough, 2004; Lindsay & Beail, 2004; Quinsey, Book, & Skilling, 2004)

In recent years the prediction of dangerousness in intellectual disability offenders appears to be receiving greater attention from researchers and clinicians (Blacker, Beech, Wilcox, & Boer, 2011; Fitzgerald, Gray, Taylor, & Snowden, 2011; Gray et al., 2007; Lindsay, Carson, et al., 2010; Lindsay et al., 2008; Lindsay, Elliot, et al., 2004). A number of structured risk assessments developed for use in mainstream, non-ID populations have been investigated in samples with ID: Violent Risk Appraisal Guide (VRAG: Harris et al, 1993); the Historical Risk Management-20 (HCR-20: Webster et al., 1997); the Risk Matrix 2000 (RM2000: Thornton, Mann, Webster, Blud, Travers, Friendship, & Erikson, 2003); the Static-99 (Hanson, & Thornton, 2000); and The Rapid Risk Assessment for Sexual Offence Recidivism (RRASOR: Hanson, 1997). The reliability and validity of these risk assessments when used with samples with ID will now be considered.

Violent Risk Appraisal Guide (VRAG: Harris, Rice, & Quinsey, 1993)

The VRAG is currently among the most accurate predictors of which men in forensic psychiatric and prison populations will commit a new violent or sexual offence. The VRAG yields a specific probability estimate of violent or sexual recidivism. When used with a sample with ID, VRAG scores have been found to be reliable with a strong total score intra-class correlation ($r=.95$) (Gray et al., 2007), which is similar to that reported in mainstream non-ID sample (Harris, Rice, Quinsey, Lalumiere, Boer, & Lang, 2003). Inter-rater reliability, on risk categories 1-9, has been reported as 92.2% (Lindsay et al., 2008). VRAG total scores have been found to correlate: ($r=.32$) with violent or sexual misbehaviour (Quinsey et al., 2004); and ($r=.33$) with security level of detention (Lindsay et al., 2010a). In addition, using ROC analysis the VRAG has been reported to predict: clients committing another violent or sexual misbehaviour ($AUC=.69$) (Quinsey et al., 2004); violent ($AUC=.73$) and general ($AUC=.74$) reconviction after five years (Gray et al., 2007); and violent incidents across three levels of security ($AUC=.71$) (Lindsay et al., 2008).

Historical Risk Management-20 (HCR-20: Webster et al., 1997)

The HCR-20 is a 20-item structured professional judgement instrument, which is divided into three sections: Historical (10 items); Clinical (5 items); and Risk management (5 items). Each item is rated on a three-point scale, from 0, no evidence of the variable to 2, clear evidence of the variable. The total score is the sum of all items, although the authors do not recommend the use of the total as a sole indicator of risk. The tool is more a means of guiding and structuring the consideration of a comprehensive range of variables with a view to arriving at an overall understanding of an individuals' risk. A number of studies have found that the HCR-20 total score relates to institutional aggression (Belfrage, Fransson, & Strand, 2000; Dernevik, Grann, & Johansson, 2002). Boer, Frize, Pappas, Morrissey, and

Lindsay (2010) have developed suggested adaptations to the HCR-20 for offenders with ID. These guidelines generally expand upon item descriptors where necessary to increase the relevance to offenders with ID. For example, item H4: Employment Problems is extended to include the individual's ability to engage with college courses, employment placements and day services.

When used with a sample with ID, HCR-20 scores have been found to be reliable with a strong total score intra-class correlations ($r=.80-.88$) (Gray et al, 2007), which is similar to that reported in mainstream non-ID samples (Douglas, Ogloff, Nicholls, & Grant, 1999). Strong inter-rater reliability has been reported: H 89.4%; C 93.1%; and R 82.7% (Lindsay et al., 2008). HCR-20 total has also been reported to have a weak but significant negative correlation ($r=-.30$) with moves into lower levels of security (Morrissey et al., 2007b). Using ROC analysis the HCR-20 total has been reported to predict: violent ($AUC=.79$) and general ($AUC=.81$) reconviction after five years (Gray et al., 2007); and violent incidents ($AUC=.72$) across three levels of security (Lindsay et al., 2008). The HCR-20 subscales have also been reported to predict violent incidents: H ($AUC=.68$); C ($AUC=.67$); and R ($AUC=.62$), which are all approaching moderate accuracy (Lindsay et al., 2008). Equally, the HCR-20 total predicted interpersonal physical and verbal/property aggression ($AUC=.68-.77$) (Morrissey et al., 2007b).

Sexual Violence Risk Scale (SVR-20; Boer, Hart, Kropp & Webster, 1997)

The SVR-20 adopts a guided clinical assessment format to assess the risk of sexual violence recidivism. The SVR-20 contains 20 items, which are divided into three risk categories: 1) psychosocial affect; 2) sexual offences; 3) future planning. Items are coded on a 3-point scale, where "0" indicates that the item does not apply according to the available information; "1" the item probably or partially applies, and "2" indicates that the item definitely applies.

Blacker, Beech, Wilcox & Boer (2011) report that the SVR Psychosocial Affect total ($AUC=.88$) and the Overall scales ($AUC=.75$) predicted sexual recidivism in a sample of sex offenders with ID.

Risk Matrix 2000 (RM2000; Thornton et al., 2003)

RM2000 has two sub-scales: the Risk Matrix 2000/Sexual (RMS), measuring the risk of sexual recidivism; and the Risk Matrix 2000/Violent (RMV) measuring the risk of non-sexual violence. The RMV scale contains three items: age upon release, the amount of prior violence and a history of burglary. The RMV was found to significantly predict violent recidivism in sexual and sexual/violent offender groups ($AUC=.85$) (Craig, Beech, & Browne, 2006). RM2000-V significantly predicted general recidivism for sex offenders with ID (Blacker et al., 2011).

Rapid Risk Assessment for Sexual Offence Recidivism (RRASOR; Hanson, 1997)

The RRASOR is an actuarial tool containing four items: past sexual offences, age at commencement of risk, extra-familial victims, and male victims. Blacker et al. (2011) reported that the RRASOR performed at no better than chance in predicting sexual recidivism in a sample of sex offenders with ID.

Static-99 (Hanson & Thornton, 2000)

The Static-99 is a more comprehensive risk assessment tool than the RRASOR, comprising items from both the RRASOR and the Structured Assessment of Clinical Judgment (SACJ; Thornton, 1997). It consists of 10 items assessing the number and presence of prior sex

offenses, prior sentencing occasions, convictions for non-contact sexual offenses, current non-sexual violence offenses, prior non-sexual violence offenses, offenses committed against unrelated victims, offenses committed against stranger victims, offenses committed against male victims, lack of any long-term, intimate, relationships with appropriate others, and whether the offender is under the age of 25 at the time of assessment. Scores on the Static-99 correlate with sexual recidivism in mainstream non-ID populations (Hanson & Thornton, 2000). Inter-rater reliability has been reported as good for the Static-99 (97.2% for risk levels 1-4) (Lindsay et al., 2008) with a sample of offenders with ID. However, it has been reported that the Static-99 did not correlate with recidivism (Tough, 2001). In contrast, a medium significant correlation ($r=.32$) has been reported between Static-99 score and level of security for sex-offenders (Lindsay, Carson, et al., 2010). The Static-99 total has also been reported to predict sexual incidents ($AUC=.71$) (Lindsay et al., 2008).

Assessment of Risk Manageability for Intellectually Disabled Individuals who Offend (ARMIDILO: Boer, Tough, & Haaven, 2004)

The ARMIDILO incorporates 30 items including *stable* dynamic factors and *acute* dynamic factors. Stable dynamic factors are the relatively persistent characteristics of the offender; and *acute* dynamic factors, are the rapidly changing contextual factors that signal the onset of offending. *Stable* and *acute* dynamic factors were further divided into environmental and client related factors. Blacker et al., (2011) using only the *acute* and *stable dynamic client* subscales reported that the ARMIDILO predicted sexual reconviction among offenders with special needs (ARMIDILO-Stable, $AUC = .60$; ARMIDILO-Acute, $AUC = .73$). Within the special needs group, offenders with an intellectual disability were analysed separately ($n=10$). The ARMIDILO-Stable scale produced a significant predictive effect, with an AUC of $.86$. The ARMIDILO-Acute, achieved an AUC of $.75$.

Offender Group Re-conviction Scale (OGRS; Copas & Marshall, 1998).

Fitzgerald et al. (2011) assessed the predictive efficacy of the Offender Group Re-conviction Scale. The OGRS (Copas & Marshall, 1998) was developed for use in probation settings. The revised OGRS II (Taylor, 1999) estimates the probability that an offender will be re-convicted within two years of release based on 9 variables (gender, current offence category, age at first conviction, age at current conviction, number of previous custodial sentences while under 21 and total number of previous court appearances at which convicted, number of previous court appearances at which convicted of a violent offence, number of previous court appearances at which convicted of a sexual offence where these exist, any previous conviction for burglary and any previous conviction for a breach). The OGRS II produces a category risk prediction of violent/sexual crime that ranges from 'some probability' of a future violent/sexual offence, 'raised probability', 'moderate probability' or 'high probability'. If the individual has no previous violent/sexual offences the risk category is not calculated (a category of 'no history' is given).

The inter-rater reliability for the OGRS was uniformly high ($r=.96$). The authors report that there was a significant difference between the mean OGRS score for those who were reconvicted for a general offence compared to those who were not, with a large effect size. A similar significant difference was reported for those who were re-convicted of a violent offence and those not reconvicted, also with a large effect size. The AUCs for general and violent recidivism both result in exceptionally large effect sizes (Rice & Harris, 2005). This suggests that the OGRS is an excellent predictor of risk of recidivism within a population of patients with ID. However, the limitations of this study should be noted before firm conclusions are drawn. The AUCs are based on very low reconviction rates: 7 individuals were reconvicted of a violent offence and 13 individuals a general reconviction. However, the AUCs produced by the ROC analyses are non-parametric statistics and thus

may be regarded as a conservative estimate of efficacy. The authors report an average OGRS score of 39.9 and therefore the OGRS predicts that 39.9% of these offenders with ID would be re-convicted with a general offence in a two year period. In fact 13/85 (15.3%) were re-convicted with a general offence. Therefore, the OGRS is over predicting risk in this ID population. The OGRS could only be completed for 80 of the original 145 patients because some had not been convicted of the offence that led to their admission (a requirement of the OGRS) or because of missing data in the case-notes. It is possible that some non-random factors influenced these aspects of the study which may have influenced the results.

6.2. Critique of risk assessment in ID

The studies above represent a promising start to assessing the utility of risk assessments developed for offenders without ID, in populations with ID. However, there are some methodological difficulties, which suggest caution is required in interpreting the results. Some of the studies have relatively small sample sizes (Lindsay, Elliott, et al., 2004; Quinsey et al., 2004). In addition, none was truly prospective, often relying upon the retrospective coding of independent or dependent variables from case-notes (Gray et al., 2007; Hogue et al., 2006; Lindsay, Elliott, et al., 2004; Lindsay et al., 2008). In addition, this method is dependent upon the quality of the case-notes and in some instances assumes comparability of quality across settings (Hogue et al., 2006). In addition, the studies use different outcome measures which make it difficult to make comparisons between them or to generalise from them. For example, Gray et al. (2007) used post-discharge offences collated from the Home Office Offenders Index, whereas other studies rely upon a variety of different definitions that do not necessarily include convictions (Lindsay, Elliott, et al., 2004; Lindsay et al., 2008; Lindsay, Carson, et al., 2010; Quinsey et al., 2004). The relationship between these behaviours and actual offending is unclear. This is a methodologically problematic area in

research with offenders with ID (see Holland et al., 2002) in Chapter 1). In addition, it is difficult to make comparisons because the individuals are resident in different settings. Some are clearly still detained (Hogue et al., 2006; Lindsay et al., 2008), others are in community settings (Gray et al., 2007; Quinsey et al., 2004) and one contains a selection of both (Lindsay et al., 2004). It is therefore unclear to what extent all participants had the same opportunity to engage in offending or offending type behaviour.

6.3. The identification of risk factors in samples with ID

Other researchers have sought to identify factors associated with offending specifically in samples with ID (Boer, Tough, & Haaven, 2004; Fitzgerald et al, 2011; Lindsay, Elliott, & Astell, 2004a; Lindsay, Murphy, et al., 2004; McGrath, Livingston, & Falk, 2007; Quinsey et al, 2004). From this work it appears that certain common factors can be discerned: attitudes consistent with offending (Boer et al., 2004; Lindsay, Elliott, et al., 2004; Lindsay, Murphy, et al., 2004; McGrath et al., 2007); anti-social attitude/behaviour (Boer et al., 2004; Lindsay, Elliott, et al., 2004a; Lindsay, Murphy, et al., 2004; McGrath et al., 2007); family relationships (Boer et al., 2004; Holland, & Persson, 2010; Lindsay, Elliott, et al., 2004); compliance (Boer et al., 2004; Lindsay, Elliott, et al., 2004; Lindsay, Murphy, et al., 2004; McGrath et al., 2007; Quinsey et al., 2004).

Some studies have reported on the links between variables associated with criminal history and recidivism (Fitzgerald et al., 2011; Hogue et al., 2006). Hogue et al. (2006) identified a lifetime conviction for murder, having caused criminal damage and having an ICD-10 diagnosis for personality disorder were significantly related to level of security. In addition, however, they also reported that those in a high security setting were significantly younger at age of first offence than those in a community setting and were significantly more likely to have two or more convictions under the age of 18 years. Lindsay, O'Brien et al.

(2010) also reported that age at first incident varied with level of security, with the youngest in maximum secure services. Fitzgerald et al. (2011) report that those convicted of a general offence significantly differed from those who were not reconvicted on number of previous offences, number of previous acquisitive offences, number of previous drug offences, number of bail offences, history of alcohol abuse and drug abuse.

These studies suggest that anti-social characteristics, particularly at an early age impact to some extent upon future association with the criminal justice system (Boer et al., 2004; Fitzgerald et al., 2011; Hogue et al., 2006; Lindsay et al., 2004a; Lindsay et al., 2004b; McGrath et al., 2007). These factors are similar to those found in mainstream offenders and perhaps explain why risk assessment tools developed for mainstream populations have some utility in samples with ID (Gendreau et al., 2002; Keeling, Rose, & Beech, 2006b).

The limitations of the work described includes relatively small sample sizes which make it difficult to generalise from the findings; ($n=54$) Lindsay, Elliott, et al. (2004; ($n=58$) Quinsey et al. (2004). In addition, it is difficult to make comparisons between studies because of the different outcome measures used. Some rely upon formal re-conviction rates (Fitzgerald et al, 2011), one included cases where there was strong suspicion of re-offending (Lindsay et al, 2004a), and one was based upon institutional incidents (Quinsey et al., 2004). In addition, some of the studies did not report the range of IQ of participants in the study (Fitzgerald et al., 2011; McGrath et al., 2007; Quinsey et al., 2004) so it is difficult to judge how potential differences in IQ may have impacted on the results. A final problem is that studies took place in different settings which may have influenced the expression of the variables of interest. Some studies took place in institutions (Hogue et al, 2006; Lindsay, O'Brien, et al., 2010; Quinsey et al., 2004) and others were based in the community (Fitzgerald et al., 2011; McGrath et al., 2007). Consequently it is difficult to draw firm conclusions from these studies, other than to reiterate, that the variables in question appear

similar to those identified in mainstream samples (Gendreau et al., 1996; Keeling, Rose, & Beech, 2006b).

6.4. Dynamic risk factors in offenders with ID

Some of the risk assessments highlighted above, the VRAG, Static-99, and RRASOR have been criticized for their lack of clinical utility because they are comprised entirely of static items. Static risk factors are historical events or stable characteristics that are correlated with recidivism. Static risk can be used to inform decisions regarding an assessed individual, such as sentencing and level of security required. However, static scales do not predict changing imminence of recidivism during a follow-up period (Callimeri, & Quinsey, 2011) (although they do predict how quickly and what proportion of individuals will re-offend in a follow-up period) nor do they identify treatment targets. Dynamic risk factors represent components of risk that if altered, also alter a person's risk, and are capable of determining short-term risk (Quinsey, Jones, Book, & Barr, 2006). Some of the scales described above that have been used with offenders with ID, such as the RM2000 and the HCR-20 include static and dynamic factors. In addition, recent years have seen the development of tools measuring dynamic risk in ID offenders (Lindsay, Elliott, et al., 2004; McGrath et al., 2007; Steptoe et al., 2008).

The Treatment Intervention and Progress Scale for Sexual Abusers with ID (TIPS-ID: McGrath et al., 2007). The TIPS-ID is a 25-item scale assessing dynamic risks associated with sexual offending. The authors report good internal consistency for the scale ($\alpha=.91$). Intra-class correlation coefficient (ICC) of test re-test, within 6 months, for a subsample of the population ($n=40$) demonstrated a strong correlation ($r=.90$). In addition, 80% of items had moderate to strong reliabilities of at least ($r<.68$) for a single rating and

($r < .68$) for averaged ratings. The authors report some evidence of criterion related validity. For example, TIPS-ID scores were significantly higher for those under 24-hour supervision compared to those on less than 24-hour supervision, and for those individuals with a diagnosis of paraphilia. In addition, individuals who were judged to have made *significant* treatment progress scored significantly lower than individuals with *some* treatment progress. However, scores on the TIPS-ID were not significantly correlated with scores on the RRASOR and no pattern of relationship was found when a scatter-plot of these scores was examined. This suggests that the validity of scores on the TIPS-ID has yet to be demonstrated adequately.

Dynamic Risk Assessment and Management System (DRAMS: Lindsay, Murphy, et al., 2004; Steptoe, Lindsay, Murphy, & Young, 2008)

The development of the Dynamic Risk Assessment and Management System (DRAMS: Lindsay, Murphy, et al., 2004; Steptoe, Lindsay, Murphy, & Young, 2008) represents an important development in the assessment of dynamic risk in offenders with ID. The DRAMS comprises 9 items drawn from the literature on proximal/dynamic risk. These are: Mood/emotion; Antisocial behaviour; intolerant/agreeableness; Sexual self-regulation; Violence self-regulation; Agreement with routine; Substance abuse; Opportunity for victim access; Clinical items of interest. The authors reported good internal consistency ($\alpha < .75$) except clinical items ($\alpha = .17$). Inter-rater reliability ($n = 40$) produced moderate to strong correlations ($r = .46-.75$) (substance abuse and opportunity for victim access could not be scored). There were also highly significant differences in scores between a control-day, at least seven days before an incident, and the assessment one or two days before an incident. An ROC analysis indicated that mood, antisocial behaviour, intolerance/agreeableness and DRAMS total all predicted violent incidents with a medium effect size.

Short Dynamic Risk Scale (SDRS: Quinsey, 2004)

Quinsey (2004) reported on the Short Dynamic Risk Scale (SDRS), an 8-item scale that has the advantage of being quick and easy to administer. Items include: taking responsibility for behaviour; evidence of anger in the last month; coping skills displayed; and lack of consideration for others. In a field study, mean scores on the SDRS showed a significant linear trend for increasing scores, which peaked in the month of an anti-social incident. Themes common to all three instruments include mood, anti-social behaviour and agreeableness/co-operation.

Limitations of both DRAMS and the SDRS studies include the small sample sizes ($n=5$, $n=23$ & $n=22$). McGrath et al. (2007) uses a larger sample size ($n=87$) in his sample of sex-offenders based in the community. However, 42% of his sample did not have an actual conviction and therefore the extent to which they are representative of the wider population of offenders with ID is unclear. In addition, none of the studies included control groups and therefore the ability of these instruments to distinguish those who go on to re-offend has yet to be determined. A difference between the two studies was the identification of potential items. Lindsay et al. (2004b) used empirically based items, whereas McGrath et al. (2007) used clinical experience to select potential items. The outcomes of future research will determine the relative merits of both options.

6.5. Psychopathy Check List – Revised (Hare, 2003)

Psychopathy is a severe form of personality disorder, characterised by a combination of affective, interpersonal and behavioural features: shallow affect, a paucity of interpersonal life, pathological egocentricity, including the selfish, callous and remorseless use of others, and an impulsive and irresponsible lifestyle (Cleckley, 1976).

The most widely used measure of psychopathy is the Hare Psychopathy Checklist – Revised (PCL-R: Hare, 2003). Though not initially designed as a risk assessment measure (Hare, 1991), the PCL-R has been reported to correlate highly with recidivism and has been included in several actuarial assessments. Consequently, its use in samples with ID will be considered. The PCL-R is a 20-item instrument scored on the basis of a thorough file review augmented by semi-structured interview. Factor analytic studies have produced two- (Hare, 2003) and 3-factor (Cooke & Michie, 2001) solutions. According to Hare’s (2003) structure, Factor 1 represents a deceitful and manipulative interpersonal style combined with deficient affective experience (8 items) and Factor 2 represents a chronically impulsive, unstable and antisocial lifestyle (10 items). These two factors divide again into 4 facets. Facet 1 (Interpersonal) and Facet 2 (Affective) comprise Factor 1, and Facet 3 (Lifestyle) and Facet 4 (Antisocial Behaviour) comprise Factor 2 (see Table 6.1). Two items, *promiscuous sexual behaviour* and *many marital relationships* do not load onto either factor, but do load onto the super-ordinate factor of psychopathy. Cooke and Michie’s (2001) 3-factor structure is drawn from the 13 items incorporated in facets 1, 2 and 3 of Hare’s model. Their formulation removes all of the antisocial behaviour and relationship items to leave what they describe as the core of psychopathy. Cooke, Michie, Hart and Clarke (2004) using structural equation modelling reported that the most parsimonious model with the highest fit indices, suggested that the anti-social behaviour items of the PCL-R are best viewed as secondary symptoms or the consequences of psychopathy.

Psychopathy is a well-established predictor of violence in mainstream forensic populations (Webster et al., 1997). As well as its utility to predict violence in isolation, its power has also been used to augment the predictive power of other widely used risk assessment instruments, such as the Risk Matrix 2000 (Thornton et al., 2003), Historical

Clinical Risk-20 (HCR-20; Webster et al., 1997) and the Violence Risk Appraisal Guide (VRAG; Quinsey, Harris, Rice, & Cormier, 1998).

Table 6.1

PCL-R 2-Factor, 4-Facet, 20-item Model (Hare, 2003)

Psychopathy				
Factor 1 Interpersonal/Affective		Factor 2 Lifestyle/Antisocial		No factor loading
Facet 1 Interpersonal	Facet 2 Affective	Facet 3 Lifestyle	Facet 4 Antisocial	
Glib superficial charm (1)	Lack of remorse (6)	Need for stimulation (3)	Poor behavioural controls (10)	Promiscuous sexual behaviour (11)
Grandiose (2)	Shallow affect (7)	Parasitic lifestyle (9)	Early behavioural problems (12)	Many marital relationships (17)
Pathological lying (4)	Callous/lack of empathy (8)	Lack of realistic goals (13)	Juvenile delinquency (18)	
Conning/manipulative (5)	Failure to take responsibility (16)	Impulsivity (14)	Revocation of conditional release (19)	
		Irresponsibility (15)	Criminal versatility (20)	

However, there are a number of methodological and conceptual difficulties in diagnosing personality disorder generally (Alexander & Cooray, 2003) and psychopathy specifically in people with ID (Morrissey, 2003; Morrissey et al., 2005). For example, intellectual and communicative deficits are likely to result in some of the key features of psychopathy being expressed differently in this population. Consequently, the behavioural components of psychopathy specified in the item descriptions of the PCL-R manual may not

encompass the full range of indicative behaviours in this population. In addition, emotional immaturity or specific developmental disorders such as autism, the incidence of which is elevated in ID populations, are often associated with characteristics, such as inability to empathise, which are also indicative of someone with psychopathy. Potential differences in the psychological and neurobiological factors associated with these characteristics create a potential area of uncertainty for PCL-R raters. The limited psychosocial environment often experienced by people with ID is likely to limit the expression of some behaviours relevant to the PCL-R. Morrissey (2006, 2007) argues that this will affect the scoring of certain items; glibness, grandiose, conning, parasitic life-style, long-term goals and short-term marital relationships. For example, they may have experienced reduced opportunities for employment, social responsibilities, sexual and cohabiting relationships (Murphy, 1992). It may be argued therefore that ‘normal’ psychosocial functioning for people with ID should be taken into account when scoring relevant PCL-R items (e.g. Parasitic Lifestyle, Irresponsibility, Employment History, Lack of Realistic Long Term Goals etc). A further difficulty is that behaviour which might normally be understood as “offending” in mainstream populations is often interpreted as “challenging behaviour” in residential ID settings (Emerson, 1995). Consequently, violent and sexually violent behaviour is often not recorded or reported to criminal justice agencies (Clare, & Murphy, 1998). Obviously, this will affect the validity of PCL-R items that are based on recorded items and conviction.

Finally, there are difficulties associated with the standard PCL-R interview. Deficits in attention, concentration, memory and communication associated with ID are likely to affect the validity of information gained during the interview process (Finlay, & Lyons, 2001). In view of these difficulties Morrissey (2003) attempted to clarify some of these issues and improve inter-rater reliability through the development of guidelines for conducting the PCL-R with offenders with ID.

Morrissey's guidelines are based upon the series of studies incorporating the use of the PCL-R with a sample of 203 offenders with ID. The general principles of PCL-R administration used in this study adhere to those reported by Morrissey (2007). Consequently, interpretation and scoring of items was based on the 'flavour and intent' of items described in the PCL-R Second Edition manual and consideration was given to the underlying function of the behaviour, to ensure its consistency with the intent of the PCL-R item. In addition, PCL-R assessments were conducted on individuals of 21 years or over. On rare occasions this meant that individuals had been resident in the hospital for several years before completion. Equally, behaviour and presentation were only attributed to personality when other reasons for the behaviour had been ruled out. Contrary to the guidelines one person with an IQ below 55 was included in this study. It was considered that his general level of functioning was in line with others in the hospital despite his lower IQ. Morrissey indicates that a broad range of collateral information should be accessed and where possible more than one informant should be used. In addition, she states that observation of interpersonal behaviour should be included in the assessment. The guidelines also specify that extra time should be allowed for the client interview and the interview schedule revised to be more applicable to people with ID. Assessors should be appropriately clinically trained in both ID and psychopathy assessment. In addition, the characteristics identified should be 'chronic and pervasive, reflect lifetime functioning and be identifiable in a range of personal and social situations, even if it is not the full range'. Using these guidelines Morrissey and colleagues conducted a series of studies considering the use reliability and validity of the PCL-R in forensic samples with ID.

In a sample with ID, good levels of internal consistency have been reported for PCL-R total, Factor 1 and Factor 2 (Table 6.2) (Morrissey et al., 2005), although they are less good for Factor 2 than Factor 1. The levels of internal consistency for Facet 3 and 4 are below

acceptable levels, suggesting problems with the reliability of those scales with this sample. This may be explained by the interpretation of the items in the context of restricted occupational and social functioning of people with intellectual disabilities (Morrissey, 2010).

Table 6.2 Internal Consistency (Cronbach’s Alpha) for PCL-R Total, Factor and Facet Scores (Morrissey et al, 2005).

	No. items	<i>n</i>	Alpha (95 CI)
PCL-R Total	20	169	0.81 (.76-.85)
Factor 1 – Affective/Interpersonal	8	199	0.82 (.78-.86)
Factor 2 – Social Deviance	10	173	0.73 (.67-.78)
Facet 1 - Interpersonal	4	201	0.74 (.67-.79)
Facet 2 - Affective	4	200	0.79 (.74-.83)
Facet 3 - Lifestyle	5	143	0.64 (.53-.73)
Facet 4 - Antisocial	5	184	0.62 (.52-.70)

Inter-rater reliability for the total scale was found to be good. The single measure intra-class correlation coefficient was strong for the PCL-R total ($r=.89$), Factor 1 ($r=.84$) and Factor 2 ($r=.83$). These correlation coefficients are comparable to those reported for adult male psychiatric samples using the standard method of administration (Hare, 2003). In addition, evidence of the validity of PCL-R scores were provided by relationships to external criteria that were broadly similar to those identified in other populations (Hare, 2003). For example, PCL-R total ($r=.18$) and Factor 2 ($r=.26$) were weakly but significantly correlated with aggressive incidents during a six-month period (Morrissey et al., 2005). Also, correlations between the PCL-R and various risk measures were all found to be significant and in the expected direction. For example, the VRAG (with the PCL-R item removed) had a significant moderate correlation with the PCL-R total ($r=.49$) and Factor 2 ($r=.59$) and a weak but significant correlation with Factor 1 ($r=.28$). Also, there were also significant

moderate correlations between the HCR-20 (with the PCL-R item removed) and the PCL-R total score ($r=.54$), Factor 1 ($r=.33$) and Factor 2 ($r=.65$). Equally, the PCL-R total score was significantly correlated with the three subscales of the HCR-20: H scale (with the PCL-R item removed) ($r=.45$), C scale ($r=.34$) and the R scale ($r=.26$). Full Scale IQ was found not to correlate with PCL-R total, PCL-R Factor 1 or PCL-R Factor 2, indicating that the checklist was not mirroring some component of general cognitive functioning (Morrissey et al., 2005).

Confirmatory factor analysis has revealed that the three-factor structure posited by Cooke, & Michie (2001) adequately fitted the PCL-R data for participants with ID, suggesting that the same construct was being measured in ID offenders as in other UK offenders (Morrissey, et al., 2009). However, in the same study the authors found that neither the traditional 2-factor model nor the revised 4-factor model fitted the data. Regardless, it appears that there may be some equivalence between the expression of characteristics that make up the construct of psychopathy in offenders with ID and those in the prison population, particularly as no cross-group differences were identified in factor structure (Morrissey et al., 2009). However, item response theory analyses indicated that Factor 3 (impulsive and irresponsible behavioural style) of the 3-factor model underestimated psychopathy at all levels of the trait in comparison with the non-ID group. This may be because many items comprising Factor 3 (in both the 3- and 4-factor models) and Factor 4 of the 4-factor model were found to discriminate poorly. This explanation is consistent with Morrissey et al. (2005) identifying lower alpha values for Facets 3 and 4.

The PCL-R total, Factor 1, Factor 2, and PCL-R 13-Item Total predicted positive progress to lower security levels and negative progress to higher security levels (Morrissey, Hogue, et al, 2007). The point biserial correlations are duplicated in Table 6.3 and ROC analyses using dichotomous dependent variables in Table 6.4.

Significant AUCs above .70 indicate the predictive ability in relation to progress moves of PCL-R Total, Factor 1 and 13-item total. However, Morrissey, Hogue, et al. (2007) reported that the PCL-R 20-item total, the 13-item total and PCL-R Factor 1 and Factor 2 scores did not correlate significantly with measures of total aggressive incidents, interpersonal aggression, verbal and property aggression and high risk aggression displayed in a 12 month follow-up within the institution. In addition, receiver operating characteristic analyses confirmed that all of the PCL-R measures did not predict interpersonal physical or verbal/property aggression in the institutions within the 12 month follow-up.

Table 6.3

Biserial Correlations Between PCL-R 20- and 13-Item Totals, PCL-R Factors, PCL-R Facets, and Positive and Negative Progress (from Morrissey et al., 2007).

Scale	<i>n</i>	Negative Progress	Positive Progress
PCL-R Total (20)	66	.27*	-.34**
PCL-R Total (13)	60	.32**	-.22
Factor 1	64	.31**	-.21
Factor 2	58	.12	-.22
Facet 1: Interpersonal	61	.23*	-.13
Facet 2: Affective	61	.32**	-.30*
Facet 3: Lifestyle	24	.12	-.21
Facet 4: Antisocial	66	.01	-.22

* $p < .05$; ** $p < .01$

Gray et al. (2007) reported that the PCL-Screening Version (SV) predicted reconviction for violence ($AUC=.73$) and general reconviction ($AUC=.76$) with moderate accuracy, during a five year follow-up. This finding suggests that the PCL-SV, with the removal of several items in the PCL-R which are potentially problematic in samples with ID, could be a more appropriate measure to be applied when conducting risk assessment in this

population (Morrissey, 2010). However, Spiers (2007) reported no significant differences in scores on those items in a comparison of offenders with IQs of 80 and above, and those of 79 and below. These findings suggested that the items in question are not scored differently for people with ID. The significant AUC reported by Gray et al. (2007) was the result of significantly higher PCL-SV scores in the ID sample than in the comparable non-ID sample. The reasons for this were not explored; it may reflect a higher rate of psychopathy or, alternatively, it may be that the PCL-SV overestimates the latent trait in this group. It would therefore be appropriate for further analyses of PCL-SV data to be conducted in order to determine scalar equivalence of the PCL-SV across populations with and without ID (Morrissey, 2010).

Table 6.4

Predictive Accuracy of the PCL-R in Relation to Positive and Negative Progress Measures: ROC Analysis (from Morrissey, Mooney, et al., 2007).

Scale	Outcome					
	Negative Progress (n=8)			Positive Progress (n=25)		
	AUC	95% CI	<i>p</i> <	AUC	95% CI	<i>p</i> <
PCL-R Total (20)	.80	.69-.91	.01	.73	.64-.85	.01
PCL-R Total (13)	.82	.70-.94	.01	.66	.51-.80	.05
Factor 1	.84	.75-.93	.01	.65	.51-.80	.05
Factor 2	.66	.47-.85	ns	.64	.49-.78	ns
Facet 1: Interpersonal	.77	.65-.89	.05	.59	.45-.74	ns
Facet 2: Affective	.85	.73-.98	.01	.69	.55-.83	.01
Facet 3: Lifestyle	.68	.44-.93	ns	.62	.43-.82	ns
Facet 4: Antisocial	.58	.42-.74	ns	.67	.53-.81	.05

The work outlined above represents a significant step forward in risk-assessment work in samples with ID. However, there are a number of methodological difficulties with

these studies which mean that caution must be exercised in generalising from the findings and making comparisons between the various studies. On the whole the sample sizes are limited: ($n=60$) Morrissey, Hogue, et al. (2007); ($n=73$) Morrissey, Mooney, et al. (2007); and ($n=81$) Spiers (2007). Morrissey et al (2009), although adequate ($n=185$) was relatively small for item response theory and confirmatory factor analysis (Morrissey, 2010). In addition, there are discrepancies between the intellectual ability of the samples. Some samples more-or-less conformed with formal ID criteria (including two standard errors) and others extended into the borderline and low-average range: Morrissey, Hogue, et al. (2007) (IQ=43-76); Morrissey et al. (2009) (IQ<75); Spiers, 2007 (IQ<80); Morrissey, Mooney, et al (2007) (IQ=43-89); and Morrissey et al. (2005) (IQ=43-89).

None of the studies outlined above was truly predictive. Consequently, PCL-Rs were completed retrospectively from case-notes, and in some instances the same raters completed other measures used to assess the validity of the PCL-R (Morrissey et al., 2005) which may have influenced the correlations between measures reported. The assessments conducted are limited by the quality of the case-notes, the accuracy of which was not ascertained. In addition, in several of the studies PCL-Rs were completed without a patient interview (Gray et al., 2007; Morrissey et al., 2005; Morrissey, Hogue, et al., 2007; Morrissey, Mooney, et al., 2007; Morrissey et al., 2009). This is considered standard practice in research (Morrissey, 2010) but is not recommended in the PCL-R administration manual (Hare, 2003). It is unclear how this may have influenced results. Many of the studies did not incorporate matched control groups of normal intelligence (Morrissey et al., 2005; Morrissey, Hogue, et al., 2007; Morrissey, Mooney, et al., 2007; Morrissey et al., 2009). Consequently, differences between scores for samples with ID and those without ID are drawn from different studies, which may have incorporated different procedures, such as participant interviews (Hare, 2003). These differences in procedures may have influenced the results.

In addition, the individual studies have some limitations. For example, Morrissey, Hogue, et al. (2007) and Morrissey, Mooney, et al. (2007) used samples resident in high security. All were deemed to pose ‘a grave and immediate risk to the public’. There were high levels of violence, as evidenced by high base rates, which may have masked the potential predictive ability of the PCL-R. Equally, 13 individuals left the high-security setting prior to the completion of the follow-up period for positive and negative moves, and were therefore not included in the study. It is unclear whether these men had moved to more or less secure settings and consequently the influence of their exclusion is unknown. In addition, this high-security setting actually incorporated several different types of accommodation (see Morrissey, Hogue, et al, 2007). There is an assumption that the triggers and barriers to performing aggressive incidents were equally distributed across these environments but this does not appear to have been assessed. The scoring of aggressive incidents was based upon computerised records (Morrissey, Hogue, et al., 2007). It appears that no attempt was made to verify the validity of these records in relation to clinical records. A final limitation of these studies relates to the apparently high levels of missing data, which may have resulted from strict adherence to early iterations of the PCL-R guidelines (Morrissey, 2003). For example, Morrissey et al. (2005) report only 59% of the sample having all items scores. It is unclear how this may have influenced the results but it is possible that the lack of information will have limited the validity of the assessment.

Despite the limitations of these studies the work outlined above regarding the use of the PCL-R with offenders with ID represents a significant development of the risk-assessment process with this population. However, it is worth noting that a number of discrepancies are apparent in relation to the PCL-R (PCL-SV) in samples with ID. Morrissey et al (2005) reported weak but significant correlations between PCL-R Total, PCL-R Factor 2 and aggressive incidents during a six-month follow-up period. Equally, Gray et al. (2007)

reported that the PCL-SV predicted general and violent reconviction during a five-year follow-up period. However, Morrissey et al. (2007a) reported that the PCL-R 20-item total, 13-item total, Factor 1 and Factor 2 scores did not predict any form of institutional aggression in a high secure setting. This latter finding appears to be at odds with research in mainstream non-ID offenders where the PCL-R appears to be particularly successful in predicting violent recidivism (Harris et al., 1993; Hemphill, Hare, & Wong, 1998). One of the aims of the present study was to investigate the ability of the PCL-R to predict institutional aggression in a medium secure environment.

The purpose of the present study was to compare the predictive validity of simple unitary measures of impulsivity with two standardised risk assessment measures commonly used in forensic settings which incorporate items relating to impulsivity; the Psychopathy Check List-Revised (PCL-R: Hare, 1991, 2003) and the Historical Clinical Risk-20 (HCR-20; Webster et al., 1997). Whilst these risk assessment tools are common in forensic settings there is little published evidence of their utility in populations of offenders with ID. Hare (2003) describes the PCL-R as being primarily a measure of inferred personality traits. In addition, the HCR-20 includes items relating to PCL-R score and the presence or absence of a personality disorder. Consequently, it is necessary to consider the applicability of personality theory to people with ID. First, I will outline the work on personality development in people with ID before considering the application of mainstream personality approaches to people with ID, conducted by Lindsay and colleagues.

6.6. Personality development in people with ID

Over a period of 40 years Zigler and colleagues (Luthar & Zigler, 1988; Yando & Zigler, 1991; Zigler, 1961; Zigler, Bennett-Gates, Hodapp, & Henrich, 2002; Zigler & Burack, 1989) have investigated the development of personality in people with ID. Zigler and colleagues

(Zigler & Burack, 1989) argue that personality development in individuals with ID is affected by the same factors as those of normal intellectual functioning. The life experiences of people with ID, however have been shown to lead to the development of personalities characterised by five features; positive reaction tendency, negative reaction tendency, expectancy of success, outer-directedness, and effectance motivation. *Positive reaction tendency* describes a heightened motivation to interact with, and become dependent upon a supportive adult. *Negative reaction tendency* describes an initial wariness when interacting with unfamiliar adults. *Expectancy of success* indicates the degree to which one expects to succeed or fail when faced with a novel task. *Outer-directedness* depicts the tendency of someone with ID to look to others for cues to the solution of difficult or ambiguous problems. *Effectance motivation* refers to the pleasure derived from tackling and solving difficult problems.

Zigler and colleagues (Zigler, & Burack, 1989; Zigler, & Bennett-Gates, 1999; Zigler et al., 2002) provide a developmental perspective of the personality structure in people with ID. The daily experiences and adaptive efforts of people with ID are seen to influence their motivational and personality structures. However, there are some limitations to this research. The sample used in the Zigler et al. (2002) study was not representative of the wider population of people with ID. For example, only those with cultural-familial ID were included, those with an organic aetiology were excluded, as were those with secondary diagnoses, such as hearing impairment or psychopathology. In addition, the emphasis on personality development has resulted in the selection of children as participants (Luthar, & Zigler, 1988; Yando, & Zigler, 1991; Zigler, 1961; Zigler, & Burack, 1989; Zigler et al., 2002). The application of this theoretical perspective to adults with ID is less well developed. Research has yet to be conducted to assess the compatibility of Zigler's work with the more widely researched perspective of trait theory (Eysenck & Eysenck, 1978). In

addition, research has yet to be published considering the relationship between the factors of the EZPQ and offending behaviour.

This work on the development on personality development in ID is of considerable clinical interest but has not yet been integrated with mainstream personality theory or with conceptualisations of personality disorder (Reid, Lindsay, Law, & Sturmey, 2004; Moreland, Hendy & Brown, 2008). This means that there are difficulties using this approach to define what actually constitutes disorder of personality in this population, or determining it equivalence to those without ID. Consequently, consideration shall be given to the application of mainstream personality theory to people with ID.

6.7. The application of mainstream personality theory to people with ID

The review of risk assessment in relation to offenders with ID highlights the role of personality disorder generally and the more acute form of personality disorder, described as psychopathy, specifically. For example, the HCR-20 contains items relating to both personality disorder and psychopathy, as defined by the PCL-R. However, little work has been conducted concerning the application of personality theory developed for intellectually normal populations in populations with ID. Lindsay, Rzepecka, and Law (2007) have adapted the Neuroticism Extraversion Openness Personality Inventory – Revised (NEO-PI-R) (Costa, & McCrae, 1995) for use with individuals with ID. The NEO-PI-R is a widely used questionnaire which measures personality as defined by the Five Factor Model (FFM) (McCrae, & Costa, 1999). The FFM is now generally agreed to comprehensively represent the fundamental dimensions of normal personality (Wiggins, & Trapnell, 1997). The FFM comprises five major dimensions, each containing six facets. The domains and their facets are shown in table 6.5.

The NEO-PI-R is a widely researched instrument, with well established reliability and validity (Costa & McCrae 1995a, 1995b). The Lindsay, Rzepecka, et al. (2007) study represented the first attempt to consider the utility of the NEO-PI-R to assess the personality of people with ID and consequently to determine the utility of the FFM in understanding the personality of people with ID. This study demonstrated the successful adaptation of a widely accepted questionnaire measuring the five factor model of personality for use with people with ID and provided some evidence of its validity in that scores on the original and the adapted version when completed by people of normal intellectual ability were highly correlated (r range 0.91-1.0). However, in a subsequent study with 40 participants with IQs in the mild ID range the authors reported discrepancies between self and observer-rating. Specifically, statistical analysis revealed that participants rated themselves as higher on E, A and C than observers.

The development of adapted NEO-PI-R is promising, but this work needs extending to a larger population. In addition, it might be interesting to examine self-report and rater-report on the adapted questionnaire for an intellectually normal group to assess the normal correspondence of scores. Another limitation of this study is the length of the questionnaire, with completion time at 3-4hours for a person with ID. This is likely to impact upon reliability and validity of the adapted measure (Finlay & Lyons, 2001).

Research has demonstrated that circumplex measures of personality and personality disorder developed for mainstream populations have good reliability when used with samples with ID (Lindsay, Steptoe, Hogue, Mooney, Taylor, and Morrissey, 2009; Lindsay, Taylor, Hogue, Mooney, Steptoe, and Morrissey, 2010). In addition, evidence supporting the concurrent validity of these measures was reported, with correlations with other psychometrics, in a manner expected (Lindsay, Taylor, et al., 2010).

Table 6.5 The Factors and Facets of the Five Factor Model (McCrae, & Costa, 1999)

Neuroticism (N)	Extraversion (E)	Openness (O)	Agreeableness (A)	Conscientiousness (C)
anxiety	warmth	fantasy	trust	competence
anger hostility	gregariousness	aesthetics	straightforwardness	order
depression	assertiveness	feeling	altruism	dutifulness
self-consciousness	activity	actions	compliance	achievement-striving
impulsiveness	excitement seeking	ideas	modesty	self-discipline
vulnerability	positive emotions	values	tender-mindedness	deliberation

These studies indicate that the factor structure of personality in people with ID can be mapped according to the same co-ordinates used for people with normal intellectual ability. Having determined that approaches to personality assessment developed for mainstream, non-ID populations, have utility in assessing personality in people with ID will consider the literature regarding the identification of personality disorder in this population

6.8. Personality disorder in ID

A substantial number of studies have been published concerning the prevalence of personality disorder (PD) in people with ID. Reference to the participants in these studies indicates that they can be divided into four broad categories; community, institution, community and institution, and offenders.

Several studies have considered the prevalence of PD in institutions (Ballinger & Reid, 1987; Craft, 1959; Day, 1985; Flynn, Matthews, & Hollins, 2002). The authors report mean PD prevalence ranging from 1% (Day, 1985) to 92% (Flynn et al., 2002). Cooper, Smiley, Morrison, Williamson, and Allan (2007), Eaton, and Menolascino (1982), Jacobson

(1990), Khan, Cowan, and Roy (1995); Lidher et al. (2005) and Reiss (1990) all report on studies with community-based samples. The authors report the prevalence of PD ranging from 0.7% (Cooper et al., 2007) to 39% (Reiss, 1990). Several studies have samples which included participants living in the community *and* institutional accommodation (Bouras & Drummond, 1992; Deb & Hunter, 1991; Goldberg et al., 1995; Gostason, 1987). The authors report the prevalence of PD ranging from 3% (Gostason, 1987) to 91% (Goldberg et al., 1995). Two of the studies reported community and institutional prevalence separately. Deb and Hunter (1991) reported lower prevalence (18.6%) in the community compared with institutions (31.65%), by contrast, Goldberg et al., (1995) reported 57% prevalence in the institutional sample and 91% in the community sample. The final category of studies related to PD in samples comprising offenders with ID (Alexander, Crouch, Halstead, & Piachaud, 2006; Alexander, Hiremath, Chester, Green, & Hoare, 2011; Day, 1984). The authors report prevalence of PD ranging from 26% (Alexander et al., 2006) to 47% (Alexander et al., 2011; Day, 1984). The most striking aspect of these studies is the range of prevalence of PD; spanning 0.7% to (Cooper et al., 2007) to 92% (Flynn et al., 2002). It seems reasonable to conclude that these differences are too great to be explained by real differences (Alexander & Cooray, 2003). Consideration of the studies reported indicates that a variety of methodological problems may have contributed to the disparate estimates.

Investigation of the presence of PD in populations with ID has led to large number of studies being conducted when compared to other areas of study for this population.

However, despite this volume of work a number of consistent methodological problems limit much of the findings (Alexander & Cooray, 2003; Torr, 2008). Several of the studies are limited by small sample sizes (Alexander et al., 2006, 2011; Ballinger & Reid, 1987; Deb & Hunter, 1991; Day, 1984, 1985; Flynn et al., 2002; Khan et al., 1995; Lidher et al., 2005). In addition, there are differences in the samples that make it difficult to compare between

studies. For example, some studies include children (Eaton & Menolascino, 1982; Jacobson, 1990) or adults below the age of 21 years (Day, 1984). This is problematic because developmental delay associated with ID means that personality features may not have stabilised until later in life and therefore identification below the age of 21 years may not be valid (Alexander & Cooray, 2003; Borthwick-Duffy, 1994; Morrissey, 2006 & 2007). In response to these difficulties current guidelines for both main diagnostic systems (ICD & DSM) specify the minimum age for the diagnosis of PD as 21 years of age (Royal College of Psychiatrists, 2001; Fletcher, Loschen, Stavrakaki & First, 2007). One sample only included adults over the age of 40 years (Day, 1985).

In addition, the definition of PD can be problematic in a population with ID where confounding factors, such as behavioural problems associated with ID, can be interpreted as aspects of personality, leading to elevated rates of immature, unstable and dependent PD in some older studies (Corbett, 1979). However, more recent studies have not found this as problematic (Lindsay, Steptoe, et al, 2007). There are also differences in the range of IQs included that make it difficult to generalise results. Some studies include participants in the borderline range (Day, 1984, 1995), and some extend into the average range (Alexander et al., 2006).

The reliability and validity of the identification of ID is questionable in a number of studies because either, the instrument used is not specified, or it is not the 'gold-standard' WAIS (BPS, 2001); (Alexander et al., 2011; Ballinger & Reid, 1987; Bouras & Drummond, 1992; Cooper et al., 2007; Craft, 1959; Day, 1984, 1985; Deb & Hunter, 1991; Eaton & Menolascino, 1982; Flynn et al., 2002; Goldberg., et al, 1995; Gostason, 1987; Jacobson, 1990; Khan et al., 1995).

There are also difficulties associated with differences in the classification systems used, whether ICD (Alexander et al., 2002; Day, 1984, 1995; Ballinger & Reid, 1987) or

DSM (Craft, 1959; Eaton & Menolascino, 1982; Reiss, 1990), particularly as the diagnostic method has been found to influence prevalence rates (Cooper, Smiley, Morrison, Williamson, & Allan, 2007). Equally, over time the changes within these formats are likely to impact upon the identification of PD making comparisons between these studies problematic due to the time over which they have been completed; DSM-I (Craft, 1959), DSM-R (Goldberg et al., 1995; Reiss, 1990), DSM-III (Deb & Hunter, 1991; Eaton & Menolascino, 1982), DSM-IV (Cooper et al., 2007).

Communication difficulties may hinder the gathering of appropriate subjective information necessary for the identification of particular traits. Indeed, few of the above studies included a clinical interview with the participant as part of the process of identifying PD, although Goldberg et al. (1995) is an exception. Equally, several personality disorders, such as paranoid PD, assume a level of cognitive ability, which may be absent in those with ID. Alexander and Cooray (2003) argue that this is compounded by a lack of valid and reliable instruments for the identification of PD in this population.

It is also apparent that some samples are not representative of the wider ID population, for a variety of reasons. Some studies only included participants with a mental disorder as well as ID (Cooper et al., 2007; Day, 1995; Eaton & Menolascino, 1982; Flynn et al., 2002). Others only assessed those released (Lidher et al., 2005), or only those referred to a service (Bouras & Drummond, 1992; Goldberg., et al, 1995). One study only included participants who had not been released from an institution (Alexander et al., 2006) and another only included a subsample of those identified as suitable, without further explanation (Flynn et al., 2002). Others used non-random sampling methods (Jacobson, 1990; Reiss, 1990). All of these factors may have led to slight differences in the participants included, which may have influenced the results.

In addition, the validity of the identification of PD is questionable for several reasons. In several studies there is a lack of clarity about the process of identifying PD, with no reference to the reliability or validity of the assessment (Craft, 1959; Eaton & Menolascino, 1982; Day, 1984, 1995). Some authors reported difficulties with the instrument used (PIMRA - Bouras and Drummond, 1992; SAP – Flynn et al., 2002). In one study only a selection of PDs were assessed without explanation (Deb & Hunter, 1991). In addition, one study presented difficulty interpreting the results (Goldberg., et al., 1995). A final methodological problem for all of these studies was that none included a matched control group of normal ability.

Two further studies are worthy of more attention: Lindsay et al. (2006); and Lindsay, Steptoe, et al. (2007). These studies attempted to address some of the criticism of previous research on personality disorder in people with ID. Using a sample of 164 participants drawn from three levels of security, Lindsay and colleagues conducted a thorough assessment of the presence or absence of the 93 traits comprising the 10 personality disorder categories of the DSM-IV. Significant attempts were made to ensure the integrity of the classification, with assessments completed from four sources; a file review, a clinician familiar with the participant (psychologist or psychiatrist), by observer rating from nursing or care staff who knew the participant well, and using the Structured Assessment of Personality (SAP: Mann, Jenkins, Cutting, & Cowan, 1981) an observer based assessment procedure. Final diagnoses were made on the 10 types of PD according to the DSM-IV guidelines. A definite categorical diagnosis was made if at least three data sources were consistent in placing the participant above the threshold for that PD. Inter-rater reliability was assessed by comparing decisions made based on file information and decisions made by clinicians independent of the research team as these were considered to be used most often in clinical settings. The percentage agreement for all PDs was over 85%, except anti-social, for which it was 74.3%. In all cases

kappa (k) was .47 or above, which although low was significant the 1% level. Agreement on specific PD categories was calculated at 84.7%, $k=.69$, which is a good effect size (Cicchetti & Sparrow, 1981). The authors reported that 64 participants (39.3%) satisfied the diagnostic criteria for a PD and 30 (17.1%) satisfied the criteria for Severe Personality Disorder: a PCL-R score over 30; or, a PCL-R score over 25 plus one form of personality disorder, except for ASPD; or, co-morbidity of two or more forms of PD. The number of patients meeting the criteria for SPD was significantly higher in the high secure setting than the other two settings. The authors report that there were significant differences in scores on the VRAG, RM2000/V, RM2000/C and the RM2000/S for those with and without the diagnosis of SPD. There were no significant differences on the Static-99. The results indicate that measures of personality disorder vary more consistently with actuarial risk of violence than actuarial risk of sexual offending in this population of offenders. Equally, Lindsay, O'Brien, et al. (2010) Using a case-note review of 477 participants the authors report that offenders detained in a high-security setting were significantly more likely to have a PD than those accessing generic community services and community specialist forensic services but not those in low and medium secure services ($n=477$).

A two-factor solution was generated using maximum likelihood confirmatory factor analysis with oblique rotation. This solution accounted for 37.2% of the variance in the data and the factor structure was largely consistent with mainstream PD research. In addition, the diagnosis of PD was found to be unrelated to IQ suggesting that features of ID are not being mistaken for features of PD. Equally, the fact that no diagnosis of dependent PD was made in the whole study suggests that factors associated with immaturity and mental delay are not being mistaken for features of PD.

These carefully conducted studies indicate that the diagnosis of personality disorder in people with ID can be conducted using the same classification system as in mainstream

population. However, some caution is required in generalising from these results as this sample comprised male forensic participants and therefore the applicability of the results to the wider population of people with ID is unclear. In addition, the authors caution that the factor analysis was second order, being at the PD level, and argue that a primary analysis is required at the trait level to elucidate personality disorder in this population. In addition, the authors included a category of 'possible PD' where the numbers of traits expressed in a category was one below the threshold. Whilst this was not included in the factor analysis it will have affected the reported kappas. In addition, it is unfortunate that the authors did not report specifically on the relative validity of the sources of data. For example, because of its use in many previous studies, it would have been interesting to know to what extent the data gained on the SAP was consistent with information from other sources.

6.9. Aims of the Research

Research reported in Chapter 1 indicated that violent offenders with ID self-reported high impulsivity than sex-offenders with ID (Parry & Lindsay, 2003; Snoyman & Aicken, 2011). Consequently, it seems reasonable to determine whether self-report impulsivity predicts institutional violence in a population of offenders with ID. Evidence presented in this chapter was more equivocal in relation to the use of the PCL-R in samples with ID. A significant correlation was been reported with institutional aggression (Morrissey et al., 2005) and the PCL-SV has been found to predict violent reconviction (Gray et al., 2007). However, Morrissey, Hogue, et al. (2007) reported that the PCL-R did not predict institutional violence in a high secure setting. In addition, the HCR-20 was reported to predict violent incidents across three levels of security (Lindsay et al., 2008) and to predict interpersonal/physical and verbal/property aggression in a high secure setting (Morrissey, Hogue, et al., 2007). The aim of the study was to investigate the predictive ability of the I7i, I7i-R, PCL-R and the HCR-20

with a population of offenders with ID, detained in a medium security forensic psychiatric setting. As with all populations in detention, particularly long-term detention, it is not appropriate to assess reconviction rates as the outcome measure. Consequently, this study used institutional aggression as the outcome variable. Recent research suggested that the impulsivity measures would be better at predicting institutional aggression than the PCL-R total and factor scores but not as good as the HCR-20.

6.9.1. Hypotheses

1. The PCL-R 20-item total, PCL-R 13-item total, PCL-R factors 1 and 2, HCR-20 total and HCR-20 H, C, and R subscales and the I7i and I7i-R will all be significantly correlated with institutional aggression.
2. The HCR-20 total score will predict verbal/property aggression.
3. The HCR-20 total score will predict interpersonal physical aggression.
4. The I7i-R total score predict verbal/property aggression.
5. The I7i-R total score will predict interpersonal physical aggression.

6.10. Method

6.10.1 Sample

Forty-seven adult male offenders with learning disabilities volunteered to participate in this study. All were residents admitted to a regional medium secure unit for offenders with learning disabilities. See Chapter 1 for the details of the sample.

6.10.2. Measures

I7i and the **I7i-R** were used. Details of these questionnaires and the administration procedure can be seen in Chapter 5.

Psychopathy Checklist-Revised (PCL-R: Hare, 2003) comprises 20 items scored between 0 (does not apply), 1 (may apply) and 2 (definitely applies). The most recent structural model of the PCL-R, described in the second edition manual (Hare, 2003) was used for the purposes of the study. The sum of scores on all twenty items furnishes the total score; Factor 1 (interpersonal and affective items) comprises 8 items and Factor 2 (antisocial behaviour and lifestyle items) comprises 10 items. In addition, Cooke and Michie's 13-item total was calculated which excluded the five antisocial behaviour items and two other items (Items 11 *Promiscuous sexual behaviour*, and 17 *Many marital relationships*). The PCL-R was completed for all 47 participants using data gathered from files and interviews with at least one appropriate clinical informant, largely in line with guidelines developed for using the PCL-R with offenders with ID (Morrissey, 2006). All assessments incorporated an interview with a member of direct care staff and where an individual had moved from another facility for people with ID within 6 months of assessment, an interview was also conducted with a member of direct care staff in that establishment. On 42 occasions at least one member of the client's family were also interviewed. On the five occasions when family members were not interviewed, three were because there were no surviving family members or they could not be contacted and two were because surviving family members refused to be interviewed. In both of these instances that index offence was the killing of another family member. Clinical observations were included in the assessment and the standard procedure was broken down into several interviews depending upon the needs of the individual being interviewed. The raters were members of the multi-disciplinary team (psychologist and psychiatrists) appropriately trained in administration of the PCL-R and experienced in working with and assessing people with ID. However, the scoring of items was more consistent with Morrissey's revised guidelines in that greater emphasis was placed on scoring items despite the limited contexts in which that behaviour may have been expressed, as long

as its expression was consistent in those context that were available. For example, *parasitic lifestyle* was assessed, where little no other information was available, upon continued examples of exploitation, theft or coercion of others in the hospital. Equally, lack of realistic long-term goals was assessed according to the extent that they differed from those goals expressed generally, such that wanting to be a racing driver ‘cos I’d be better at it than them on the TV’ appeared to be more consistent with the flavour and intent of the item criteria than someone who wanted to work in a charity shop ‘where my friends work’.

Historical Clinical Risk-Management-20 (HCR-20: Webster, Douglas, Eaves, & Hart, 1997) is an instrument designed to predict future violence. It has three sub-scales comprising 10 historical items (including one that relates to the PCL-R score), 5 clinical items and 5 risk management items. Again, each item is scored on a three-point scale (0, 1 or 2) the exact contents of which vary from item to item. The HCR-20 total score is the sum of all items. The HCR-20 was completed for all 47 participants using data gathered from files and interviews with at least one appropriate clinical informant. Again, raters were members of the multi-disciplinary team (psychologist and psychiatrists) appropriately trained in administration of the HCR-20. Reliability was checked by using a blind-rater to re-score 19 cases. Internal consistency for the total scale and three sub-scale scores was good ($\alpha = .82-.87$). Inter-rater reliability and was found to be very good ($r = .90$) (Dancey & Reidy, 2002).

Aggressive Incidents: Follow-up institutional data were collated over a three-month period and coded in line with Morrissey, Hogue, et al’s (2007) procedure. The three month follow-up period was used in order to maximise the number of participants included. Using this time period individuals who were resident on the unit for a three month assessment period could be included. Each incident was coded according to an agreed description of verbal/property aggression, interpersonal physical aggression and high-risk aggression. Incidents were summed for each individual participant to create three continuous variables:

all aggressive incidents, verbal aggression or aggression to property (labelled Type 1 aggression) and interpersonal physical aggression (labelled Type 2 aggression). In addition, high-risk aggression (labelled Type 3 aggression) was also recorded which was intended to capture the degree of actual or potential harm to others. For example, an incident of verbal aggression may have not escalated due to staff interventions but may still have been interpreted as high risk. Consequently, all incidents were rated as low, medium or high and all incidents rated as high were also recorded as Type 3 aggression.

Behavioural descriptions were agreed with staff in advance so that incidents could be coded appropriately and reliably. The definitions of each behaviour were as straight-forward as possible in order to conform every-day understanding and interpretations.

Physical aggression was described as physical aggression directed at an individual. This included hitting, scratching, biting, spitting (at the person), throwing objects at a person, taking items (from the person) without consent, blocking passage and refusing to perform necessary tasks, which led to or were likely to lead to direct confrontation. These final two items are more difficult to describe but were physical acts often associated with wider bullying (by certain individuals) of others.

Verbal aggression was described as verbal communication that was likely to cause harm to another person. This included swearing and insults directed at others. Also, verbal communication that was likely to cause distress to an average person. For example, jokes were not routinely included but if a joke was particularly unpleasant or the jokes appeared to be part of a wider pattern of bullying behaviour it was included. Generally, these were not included for isolated incidents but for patterns of behaviour.

Property aggression included aggression directed at objects. This included damaging or destroying objects. Pushing or lifting objects during an incident, where there is no other explanation for the behaviour (i.e. moving a chair to get out of a room). Throwing objects

(not directed at a person). Theft of objects was included as property aggression (not from the person).

High-risk aggression generally related to the possibility of harm to others. This included incidents when one or both individuals involved in an incident are displaying high-levels of violence, which might include the use of a weapon. In addition, it included incidents involving certain individuals who were unable to manage their levels violence, without external support of staff. Also included were incidents involving several clients at a time and incidents that appeared to have involved some planning. This also included all incidents where police action was required. In addition, in some instances this included objects thrown at others, such a hot drinks.

6.10.3. Procedure

Favourable ethical opinion for this research project was gained from the Norfolk Research Ethics Committee. In addition, full approval was gained from the East Norfolk and Waveney Research Governance Committee and Research Management Team.

PCL-R and HCR-20 assessments were completed on all participants. In addition, all aggressive incidents were recorded and coded according to criteria described below.

6.11. Results

In order to examine the normality of the data, Kolmogorov-Smirnov and Shapiro-Wilk test were completed along with production of normality plots. The results indicated that PCL-R Factor 1, PCL-R Factor 2, Type 1 aggression, Type 2 aggression and Type 3 aggression were all non-normally distributed. PCL-R Factors 1 and 2 are both positively skewed and all of the aggression data sets are negatively skewed. Consequently, non-parametric tests are reported where necessary. In total 78.7% of the sample were involved in

Table 6.6

Means and Standard Deviations for PCL-R 20- and 13-Item Totals, PCL-R Factors, HCR-20 Total and H-, C-, and R-Subscales, and Aggression Types 1, 2 and 3.

	Mean	<i>SD</i>	Range	Median
PCL-R	19.6	8.2	9-32	24
Factor 1	7.7	5.8	2-12	11
Factor 2	8.6	5.6	3-13	12
13-Item Total	12.9	5.7	3-21	18
HCR-20 Total	26.5	6.1	15-37	27
HCR-20-H	13.2	3.2	6-19	14
HCR-20-C	6.4	1.6	3-10	6
HCR-20-R	6.7	2.2	2-10	7
Overall Aggression	12.6	21.6	0-105	6
Type 1 Aggression	9.9	16.6	0-75	5
Type 2 Aggression	2.7	6.2	0-30	0
Type 3 Aggression	3.5	10.1	0-57	0

at least one aggressive incident during the three months of data collection. Seventy-eight point seven percent had engaged in at least one Type 1 aggressive incident (verbal/property aggression, range 0-75), 44.7% had engaged in at least one Type 2 aggressive incident (physical aggression against a person, range 0-30) and 25.5% had engaged in an incident which was judged to reflect a significant risk of harm to others (Type 3, range 0-57). The means, standard deviations and ranges for PCL-R, Factor 1, Factor 2, 13-Item Total, HCR-20 Total, the HCR-20 subscales (H, C & R) and the different forms of aggression are shown in

Table 6.6. The relatively high means for the PCL-R items and the HCR-20 and subscales are noted and will be returned to in the discussion.

Table 6.7. Spearman's rho correlations for the independent variables with institutional behaviour exhibited over a 3 month period.

Measure	Total Violent Incidents	Type 1 Aggression	Type 2 Aggression	Type 3 Aggression
I7i	.140	.130	-.036	-.012
I7i-R	.016	-.009	-.112	-.171
PCL-R Total	.688***	.647***	.603***	.379***
Factor 1	.693***	.663***	.578***	.431***
Factor 2	.441***	.411***	.398***	.210
13-item	.665***	.626***	.607***	.420***
Total				
HCR-20	.592***	.539***	.599***	.379***
Total				
H	.552***	.481***	.658***	.429***
C	.411***	.366*	.365*	.199
R	.464***	.422***	.485***	.274

*p<.05; **p<.01; ***p<.001

The continuous dependent variables were correlated with the various independent measures and the outcomes are shown in Table 6.7. As the dependent variables were not normally distributed, non-parametric correlations (Spearman's rho) were employed using ranked data. None of the self-report impulsivity measures correlated significantly with the observations of aggressive behaviour. The PCL-R total score, PCL-R Factors 1 and 2 and the PCL-R 13-item total all showed moderate correlations with all types of aggression (Dancey & Reidy, 2002). The HCR-20 total and sub-scale scores all showed significant moderate correlations with all types of aggression except the Clinical and Risk Management Items

failing to reach significance with Type 3 Aggression (High-Risk Aggression). All of the significant correlations are positive in direction meaning that as the score on the assessment tool increases so does the number of aggressive incidents.

6.11.1. Multiple Regression

Consideration was given to conducting a multiple regression to determine the extent to which the continuous dependent variables of verbal violence, physical violence and high-risk violence could be predicted by those variables with potential predictive capability identified from the correlations. During the process of predictor selection consideration was given to the inter-correlation of variables. Field (2009) indicates that correlations between predictors in excess of 0.8 should be avoided as this is likely to lead to problems with multicollinearity. Therefore, the best predictors were selected ensuring that no variables correlating with each other more than .08 were included. When such instances occurred the less predictive variable was excluded and the next best predictor included. Regardless, initial investigation indicated that the data were not suitable for multiple regression due to problems relating to homoscedasticity and multicollinearity. As a consequence of this, consideration was given to analysing the data using logistic regression utilising a categorical dependent variable. However, the data were found to be unsuitable for logistic regression due to contravention of the assumptions of linearity of logit and multicollinearity, however ROC analysis was proceeded with for each predictor individually. This analysis was conducted rather than individual logistic regression analyses because ROC has a number of advantages over logistic regression (Bennell, 2005) and because it was similarly used by Morrissey et al. (2007).

6.11.2. Receiver Operating Characteristic Analysis

ROC analyses were conducted to ascertain the predictive accuracy of the various measures. These analyses focused on the I7, I7R and those risk assessment tools used to assess the likelihood of future aggression, namely the PCL-R (including factors 1 and 2), PCL 13-item total, HCR-20 total and subscale scores. These analyses take account of the differing base rates of the dependent variables and the fact that the instruments used produce a range of scores, many of which could potentially be used as the cut-off point for making predictions (Bennell, 2005). This procedure entails the dichotomisation of the target variable i.e. type of violence, as present or absent. Then the proportion of the sample where there was correct prediction of the target variable in question (proportion of Hits) is plotted against the proportion of the sample where behaviour was incorrectly predicted (proportion of False Alarm) for each level of the measure. The 'area under the curve' (AUC) was calculated for each of the specified scales. The AUC can range from 0 to 1. An AUC of .5 represents a predictive capability no greater than chance, whereas an AUC of 1 would represent perfect prediction. According to criteria proposed by Swets (1988), AUCs between 0.50 and 0.70 indicate low accuracy, AUCs between 0.70 and 0.90 indicate moderate accuracy and AUCs between 0.90 and 1.00 indicate high accuracy. The AUCs were calculated for the selected measures in relation to verbal/property aggression, physical aggression and high-risk aggression only (see table 6.8). Overall aggression (yes/no) was not included as the figure were identical to verbal/property aggression.

Table 6.8 displays the results of these analyses, including the 95% confidence intervals for the AUC statistic and the probability that it differs from .5 (chance). It can be seen that for verbal/property aggression whilst the PCL-R Total, Factor 1 and HCR-20 Total AUC all reach the cut-off for moderate accuracy only the PCL-R Factor 1 reached significance. By contrast, for physical aggression the HCR-20-H reached the level of high

accuracy and was highly significant (Swets, 1988). In addition, the PCL-R Total, Factor 1, 13-item total, HCR-20 total, and HCR-20-R all reached the level of moderate accuracy (Swets, 1988) and were highly significant (Swets, 1988). In addition, the AUCs for PCL-R Factor 2 and HCR-20-C were approaching 0.7 but remained in the low accuracy range (Swets, 1988) and were significant.

For high-risk aggression the PCL-R Factor 1 reached the level of moderate accuracy and was highly significant. The PCL 13-Item total, the HCR-20-H and IQ all reached the level of moderate accuracy and were significant. The PCL-R Total and HCR-20 Total, whilst approaching *AUC* of 0.7, remained in the low accuracy range but were significant. Figures 6.1 to 6.9 show the ROC curves for all of the independent and dependent variables included in the ROC analyses.

Table 6.8. ROC Analyses: ROCs and 95% confidence intervals for three types of aggression occurring within 3 months of assessment.

Measure	Verbal/Property Aggression			Physical Aggression			High Risk Aggression		
	AUC	95% CI	<i>p</i> <	AUC	95% CI	<i>P</i> <	AUC	95% CI	<i>p</i> <
I7	.63	.46-.81	.203	.53	.36-.70	.748	.54	.36-.72	.639
I7R	.57	.39-.75	.499	.48	.31-.65	.856	.45	.28-.61	.533
PCL-R	.70	.51-.89	.054	.83	.72-.95	.000**	.69	.53-.84	.034*
Factor 1	.71	.53-.90	.043*	.82	.70-.95	.000**	.73	.58-.88	.009**
Factor 2	.64	.47-.82	.157	.69	.52-.86	.028*	.57	.39-.76	.399
13-Item	.69	.51-.87	.071	.83	.71-.95	.000**	.71	.56-.87	.016*
HCR20	.70	.52-.87	.061	.83	.71-.95	.000**	.69	.53-.85	.030*
Total									
H	.66	.50-.82	.116	.90	.81-.98	.000**	.73	.58-.86	.010*
C	.65	.44-.86	.149	.67	.52-.83	.043*	.59	.42-.75	.319
R	.60	.40-.80	.343	.76	.63-.90	.002**	.63	.46-.80	.134

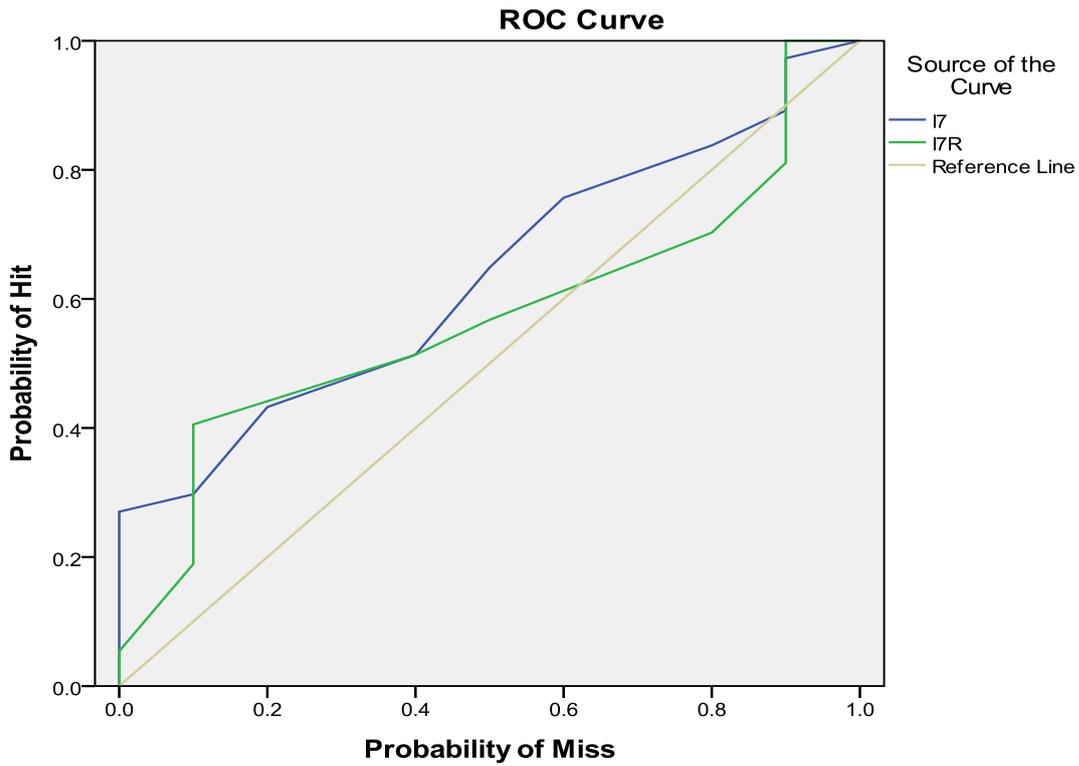


Figure 6.1 ROC Curves for Verbal/Property Aggression as Predicted by the I7i, I7i-R and IQ.

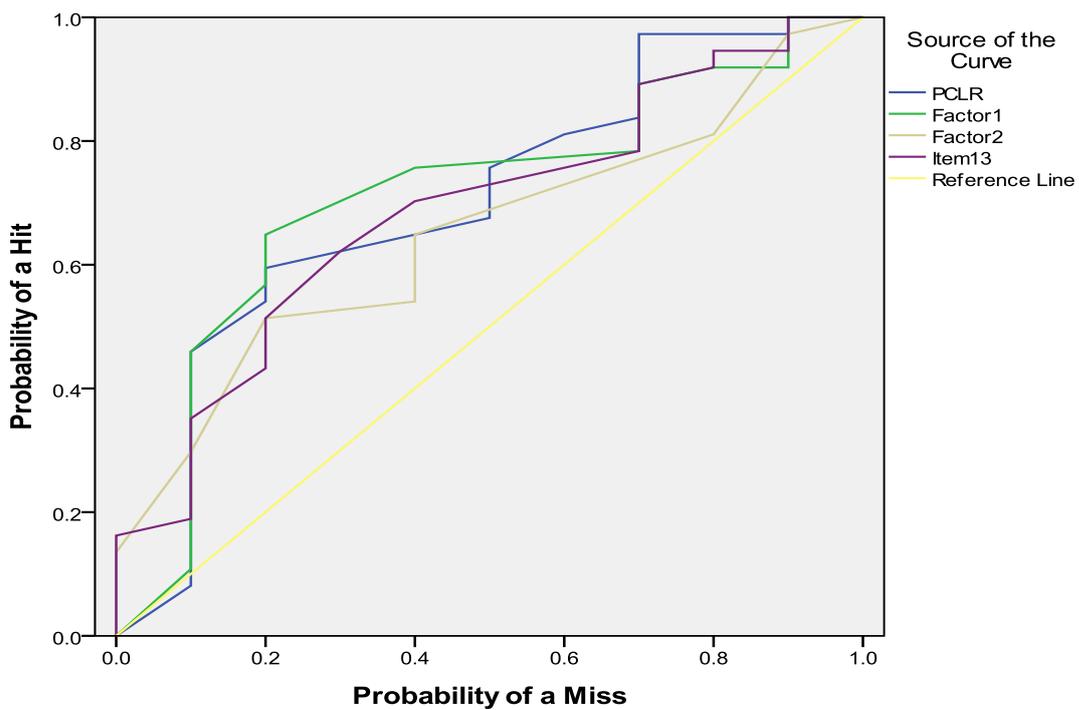


Figure 6.2 ROC Curves for Verbal/Property Aggression as Predicted by the PCL-R, Factor 1, Factor 2 and 13-Item Total.

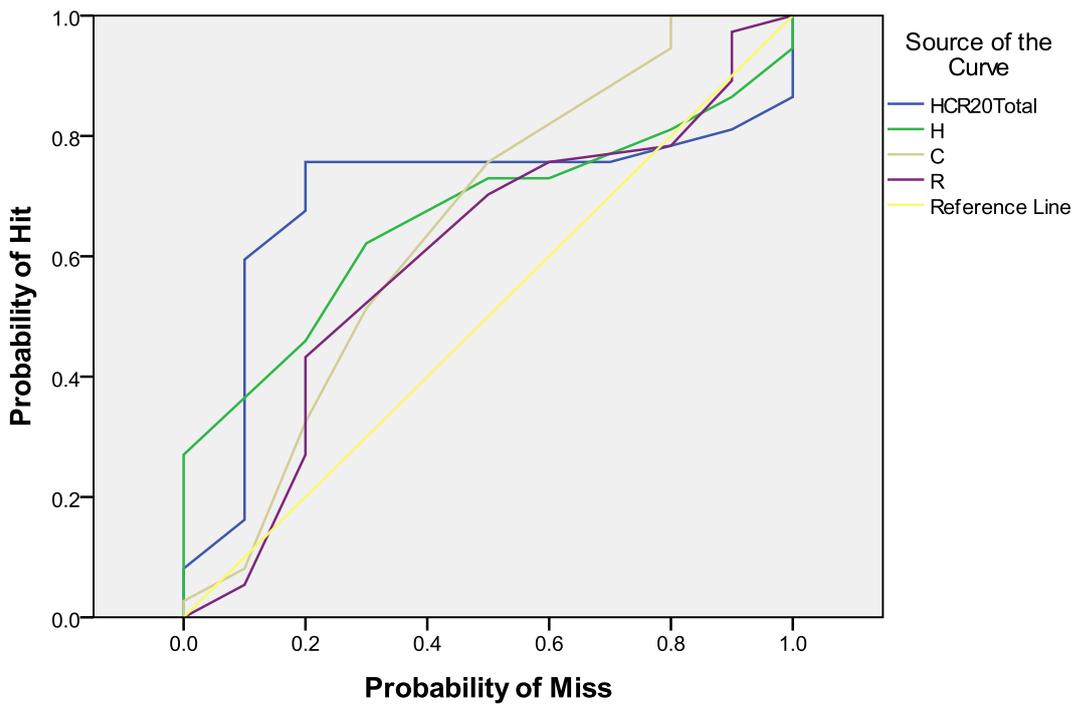


Figure 6.3 ROC Curves for Verbal/Property Violence as Predicted by the HCR-20, H-subscale, C-subscale and R-subscale.

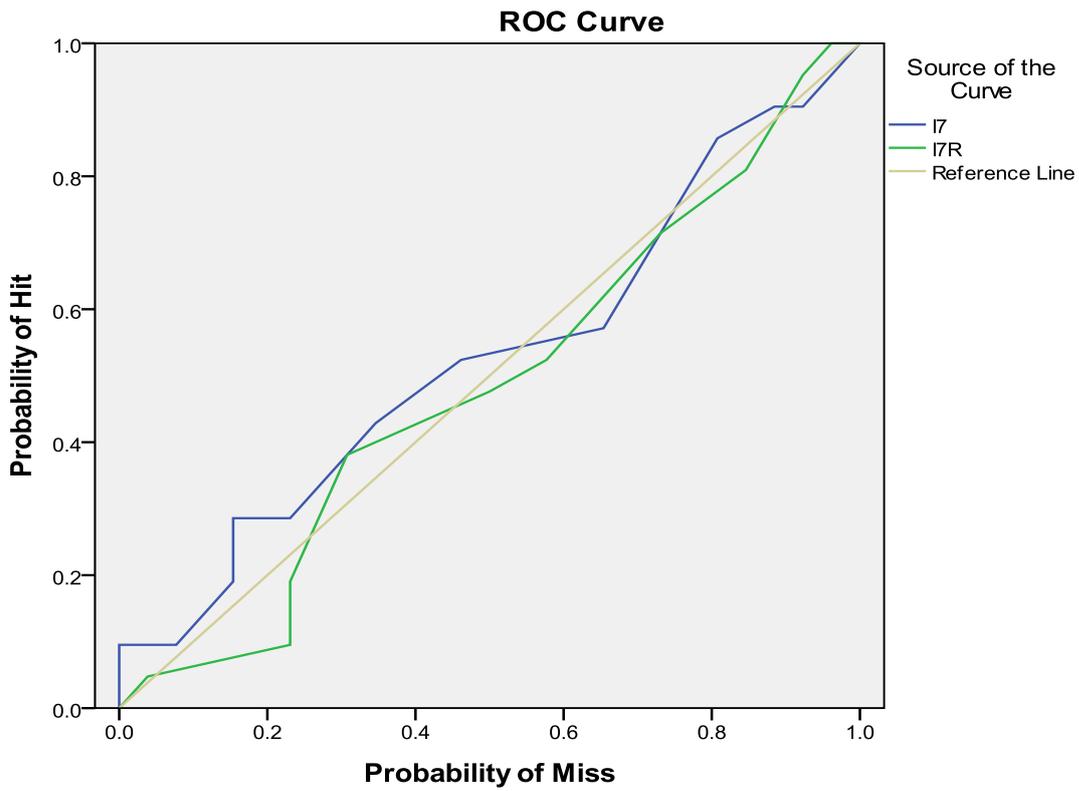


Figure 6.4 ROC Curves for Physical Violence as Predicted by the I7i, I7i-R and IQ.

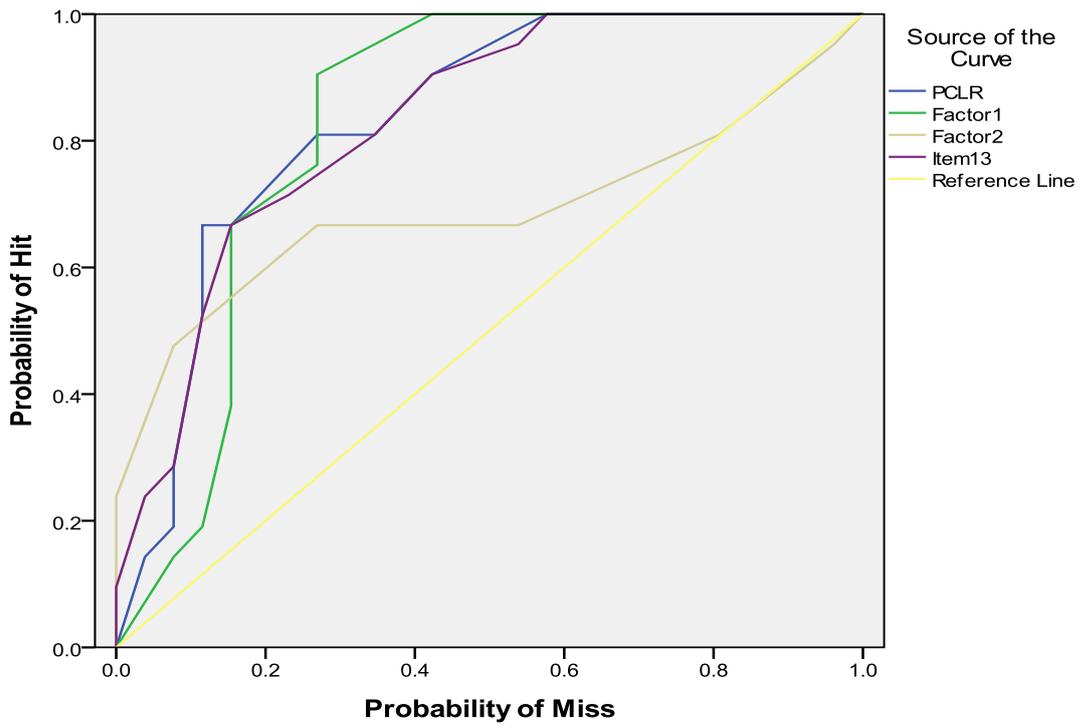


Figure 6.5 ROC Curves for Physical Violence as Predicted by the PCL-R, Factor 1, Factor 2 and 13-Item Total.

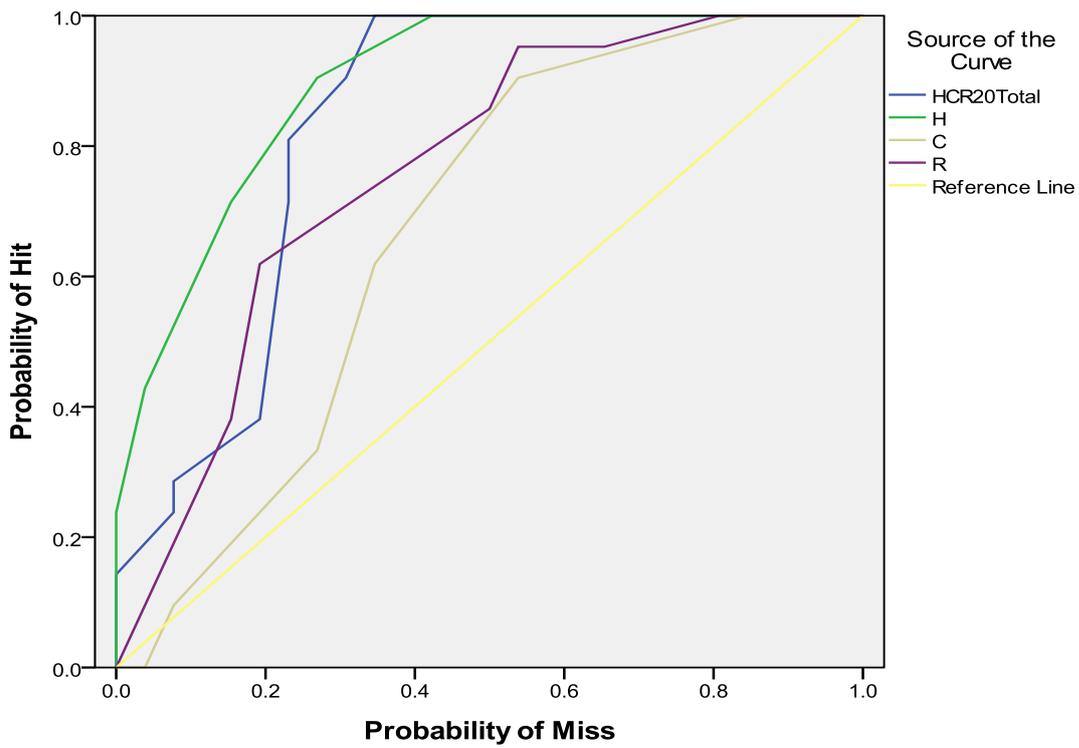


Figure 6.6 ROC Curves for Physical Violence as Predicted by the HCR-20, H-subscale, C-subscale and R-subscale.

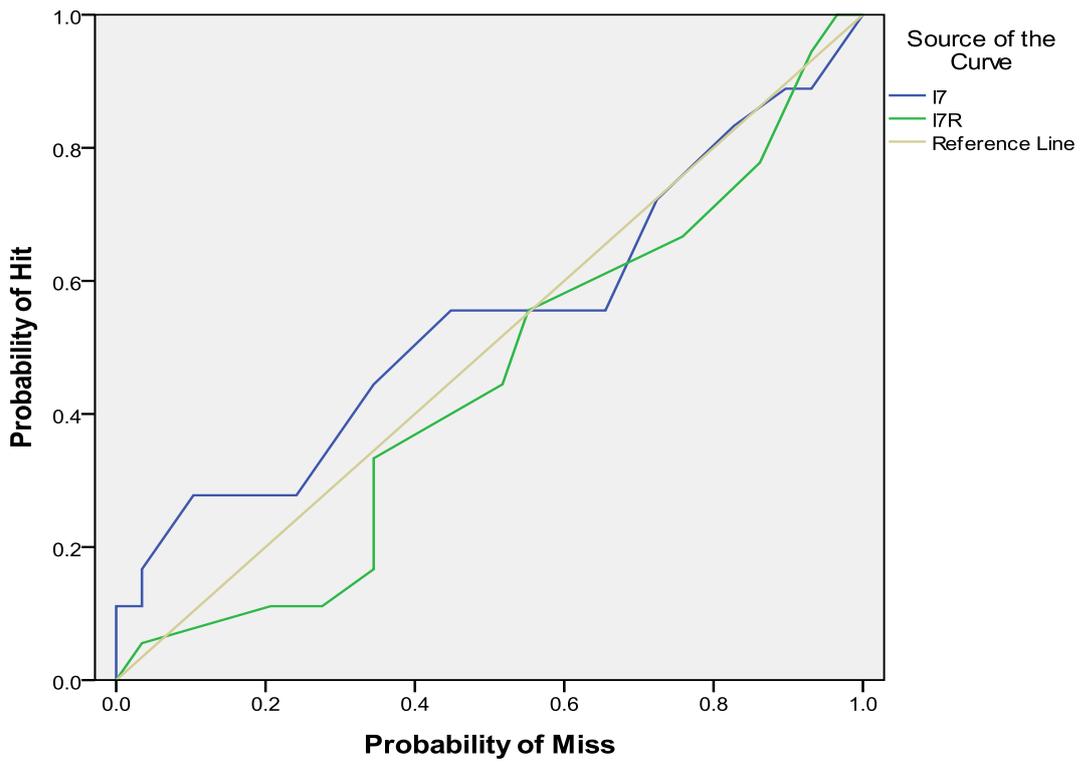


Figure 6.7 ROC Curves for High-Risk Violence as Predicted by the I7i, I7i-R and IQ.

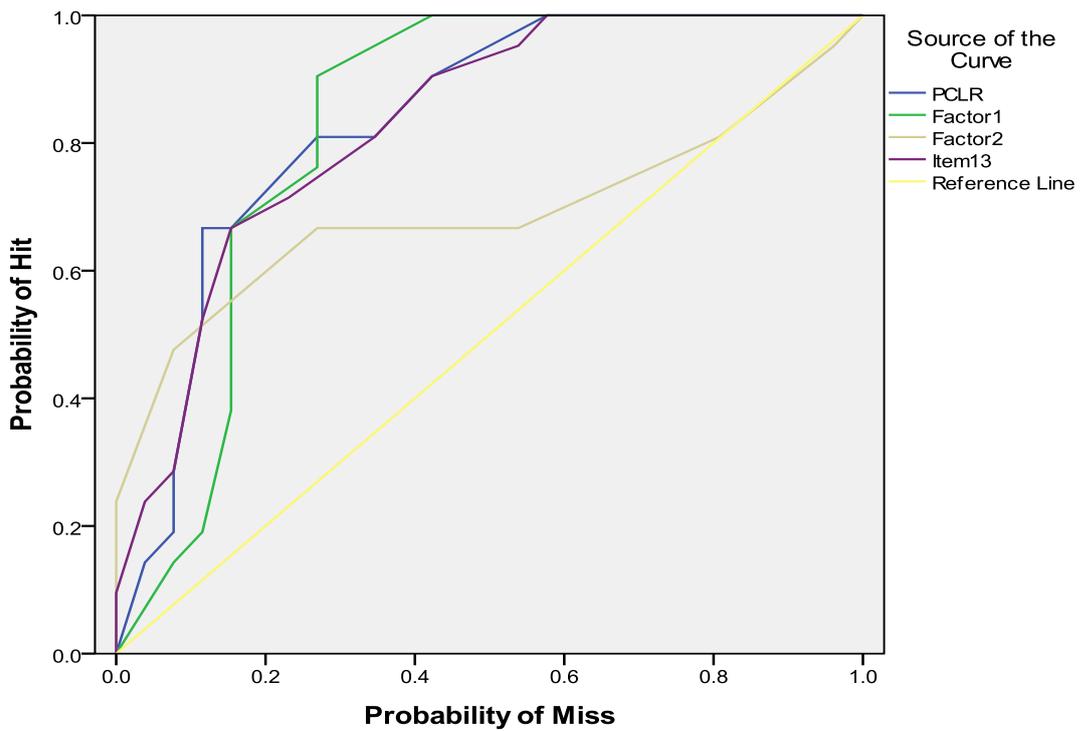


Figure 6.8 ROC Curves for High-Risk Violence as Predicted by the PCL-R, Factor 1, Factor 2 and 13-Item Total.

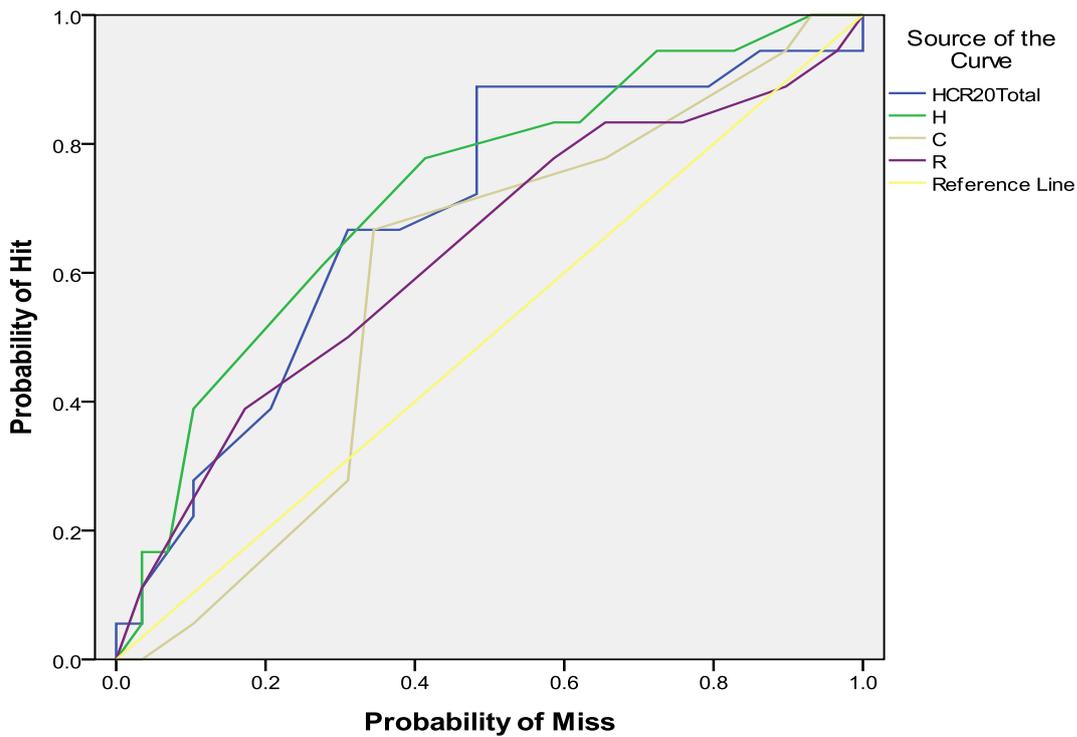


Figure 6.9 ROC Curves for High-Risk Violence as Predicted by the HCR-20, H-subscale, C-subscale and R-subscale.

6.12. Discussion

Findings from research studies conducted in a *high* secure setting are equivocal about the success of risk assessment tools, designed for mainstream non-ID offenders, when used with offenders with ID. Morrissey, Hogue, et al. (2007) found that the PCL-R 20-item total, PCL-R 13-item total, PCL-R Factor 1 and HCR-20 Total score all had utility in determining treatment progress in a high secure setting. By contrast, Morrissey, Mooney, et al. (2007) found, that of the assessments tested in the present study, only the HCR-20 Total score significantly predicted interpersonal physical aggression and verbal/property aggression ($AUC=.68-.77$). The aim of the present study was to investigate the predictive ability of the I7i, I7i-R, PCL-R and the HCR-20 with a population of offenders with ID, detained in a *medium* security forensic psychiatric setting.

The findings from the present study suggest that the PCL-R total, the PCL-R 13-item total, the HCR-20 total, HCR-20-H and the HCR-20-R all have utility in predicting interpersonal physical aggression in a medium secure psychiatric setting. The HCR-20-H reached the level of high predictive accuracy and PCL-R total, Factor 1, PCL-R 13-item total, HCR-20 total and HCR-20-R all reached the level of moderate predictive accuracy. These findings contradict those from a previous study in a high secure setting (Morrissey, Hogue, et al., 2007). Of the measures used in the present study, Morrissey, Hogue, et al. (2007) only found the HCR-20 total reached the level of low predictive accuracy for interpersonal physical aggression and was significant. Morrissey, Hogue, et al (2007) reported none of the PCL-R related independent variables were significantly predictive of any form of aggression. In addition, the PCL-R total and Factor 1 scores obtained in the present study, in a medium secure setting, are higher than those reported by Morrissey, Hogue, et al. (2007) and Morrissey (2005) in a high secure setting. In relation to the PCL-R, the results obtained by Morrissey, Hogue, et al. (2007) may have been influenced by the sample used. In 2003, 73 individuals had been assessed using a variety of file-based and informant based measures as part of a wider three-site study conducted in 2003-2004 (described in Hogue et al., 2006). Morrissey, Hogue, et al. (2007) used information for 60 of those participants still remaining in a high-secure setting at 12-month follow-up. Morrissey et al suggest that the individuals who had left high-security were likely to be of lower risk and therefore presumably lower PCL-R score. However, they do not provide any evidence to support this view. Regardless, this would have had the effect of increasing mean PCL-R scores in the high secure environment and therefore does not explain why scores obtained in a medium secure setting would be higher. Another explanation for this finding is poor reliability at one or both sites so that PCL-R scores are elevated or depressed erroneously. This seems unlikely as all staff are appropriately trained on PCL-R administration. Alternatively, scores at the high secure

setting may simply have been lower and associated with lower levels of violence, which may have influenced the predictive accuracy of the PCL-R. Morrissey, Hogue, et al. (2007) also relied upon computerised records in order to rate the levels of aggression. However, they do not attempt to demonstrate the validity of this source in relation to clinical records. In addition, there were different procedures used in the completion of the PCL-R. Morrissey, Hogue, et al. (2007) relied upon file review and an interview with a clinical informant. In the present study, this procedure was augmented by an interview with the participant, with at least one family member and a second clinical informant if the participant had been transferred from another institution; prison, hospital or community home. Hare (2003) does not encourage file only reviews but accepts this as a necessity in certain circumstances. He notes, however, that in assessments based solely on file reviews the PCL-R scores tend to be considerably lower than those obtained with the standard PCL-R procedure (p.57).

In addition, it is clear that there were different emphases placed on the data collection procedure and the rating process. For example, in the present study the risk assessment tools were routinely used as a means of structuring the information gathering process. Consequently, information not available during initial reviews was actively pursued in order to increase the validity of the risk assessment process. This explanation is also supported by the significantly lower amount of missing data in this study. Morrissey, Hogue, et al. (2007) recorded that three items in their study had a notable level of missing data: Item 9 (Parasitic Lifestyle) 24% omitted, Item 15 (Irresponsibility) 28% omitted and Item 17 (Many Short-Term Marital Relationships) 58% omitted. The exact number of assessments with complete data is not stated but based on the figures provided can be no more than 42%. Hare (2003) reports on a range of offender samples containing 8017 male offenders and states that 94.7% contained complete data sets (p. 50). The present study contained no missing data for all those included in the study. Again, this may be due to the fact that raters knew clients and

often client's families well and so were able to gain more relevant information over an extended period of time. In two instances where a family member was not available for interview, raters felt that sufficient information was available from clinical informants to make a valid rating. Finally and linked to the pursuit of relevant information in the risk assessment process, there was a much greater emphasis on scoring all items on the information available in the present study. Consequently, the procedure used coincides to a greater extent with Morrissey's amended guidelines (Morrissey, 2007, 2010).

Contrary to expectation, none of the self-report measures of impulsivity predicted future institutional violence. It appears therefore that self-report measures of impulsivity reflect the tendency to engage in impulsive behaviour *per se* rather than purely impulsive aggressive behaviour. It was also apparent during the completion of self-report measures of impulsivity that participants were considering the expression of the trait in relation to everyday life rather than solely in relation to the occurrence of specific aggressive incidents. In particular several participants who engaged in high levels of apparently impulsive violent behaviour, self-reported low-levels of impulsive behaviour. This may seem to question the validity of self-report measures of impulsivity. However, the items contained within the measures did not specifically relate to institutional violence and aggression. In addition, the fact that someone engages in apparently impulsive behaviour does not necessarily mean that the behaviour is indeed impulsive as determined by the individual. For example, when discussing apparently impulsive aggressive incidents, participants often indicated that the incident in question was actually part of a longer running difficulty and therefore subjectively they felt that they had thought about these behaviours in advance and therefore did not consider them impulsive. Within such discussions there was however clear evidence of questionable validity of self-report impulsive behaviours as participants frequently sought to

justify their behaviours, viewing them as rational behaviours in specific situations and therefore not impulsive.

Limitations in the present study are likely to have influenced the results. For example, the present study was not truly predictive in nature in the sense the Morrissey, Mooney, et al. (2007) study was. In most cases the PCL ratings preceded the 3-month data collection period. However, those clients who came to the medium secure unit for assessment had data relating to violent incidents collected before the assessment was completed. In addition, as the assessments and data collection were completed as part of the normal clinical processes on the unit it is possible that the recording and reporting of dependant variable data influenced the rating of risk assessments, as the raters were not blind to this process. In addition, many of these clients in the present study were extremely well known to raters. Indeed, most clients ($n=33$) were known to raters in excess of one year prior to the completion of ratings, with the majority of those being known for over two years ($n=26$). Only 14 participants had PCL ratings finalised within the first year of admission. There were various reasons for this, foremost of which was that historically PCL ratings were not completed on a routine basis because of difficulties administering the assessment, concerns over labelling individuals, difficulties interpreting scores and the uncertain validity of results. The fact that participants were well known to raters, including familiarity with client history, may have led to more accurate completion of PCL items or may have placed undue emphasis on contemporaneous presentation in the completion of the risk assessment tools. In addition, this familiarity extended to the individual's behaviour on the unit. Again, as the raters were also present on the unit during the working week, this may have led to a more accurate completion of PCL items. Alternatively, there may be contamination in the data collection process whereby those individuals who demonstrated more difficult behaviours on the unit were more likely to be perceived as having traits consistent with

psychopathy. However, test re-test reliability was good suggesting this was unlikely. However, re-tests were also completed by individuals who knew the clients well and who had also been involved intimately in routine multi-disciplinary discussions which included consideration of personality traits. It is possible that PCL ratings may have been pulled towards some team understanding, rather than representing the raters' individual opinions. It would be interesting perhaps to compare reliability data for ratings completed by individuals who were not familiar with clients on a day-to-day basis with those who were. In addition, it may be illuminating to trace PCL-R item ratings from admission to discharge to assess the stability of such ratings in this population.

Finally, the participants in this study had a higher mean IQ than the participants in the Morrissey, Hogue, et al. (2007) study. The mean IQ in the Morrissey, Hogue, et al. (2007) study was 66.2 ($SD=8.9$, range 43-76), whereas in the present study the mean IQ was 70.6 ($SD=8.0$, range 51-84). If, as Morrissey et al. suggest, lower levels of psychosocial functioning reduces the links between the lifestyle indicators present in the PCL-R and psychopathy, then it is possible that those with higher levels of cognitive functioning and therefore potentially higher levels of psychosocial functioning will maintain the relationship present in mainstream non-ID populations. This argument is consistent with findings from prevalence studies concerning the prevalence of personality disorder in people with ID. For example, Jacobson (1990) reported increased prevalence of PD as IQ increased within a population of people with ID.

As already mentioned Morrissey, Mooney, et al. (2007) only found the HCR-20 total reached the level of low predictive accuracy and was significant in their sample. In contrast, in the present study, the HCR-20-H subscale reached the level of high predictive accuracy for physical aggression and was highly significant. In addition, HCR-20 total and HCR-20-R both reached the level of moderate predictive accuracy for physical aggression and were

highly significant. The HCR-20-C also reached the level of low predictive accuracy for physical aggression but was significant nonetheless. Finally, HCR-20-H and HCR-20 total were both significantly accurate predictors of high-risk aggression, at moderate and low levels of accuracy respectively. In addition, the HCR-20 scores obtained in this study were higher than those found by Gray et al. (2007) and Lindsay et al. (2008). In the present study the mean HCR-20 total is 26.5 ($SD=6.1$) compared with 22.32 ($SD=5.98$) in Gray et al. and 21.22 ($SD=6.48$) in Lindsay et al (2008). Interestingly, however is the fact that it appears that this difference is primarily due to differences in the HCR-20-C and HCR-20-R subscales. The mean HCR-20-H subscale in the present study was 13.2 ($SD=3.2$), which is compatible with 12.43 ($SD=3.17$) in Gray et al. and 13.75 ($SD=2.72$) in the medium/low secure population in the Lindsay et al. study. By contrast the mean HCR-20-C in the present study was 6.4 ($SD=1.6$) compared with 5.86 ($SD=5.86$) in Gray et al and 4.84 ($SD=2.34$) in Lindsay et al. Finally, the mean HCR-20-R in the present study was 6.7 ($SD=2.2$) compared with 4.09 ($SD=2.45$) in Gray et al and 2.63 ($SD=1.42$) in Lindsay et al. There may be several reasons for these findings. Firstly, it is possible that different figures reported represent different risks within the different populations but it is difficult to understand why these differences become more pronounced, with particular reference to the Risk Management items. However, it is possible that the differences represent different aspects of the data collection process and different uses made of the HCR-20 in the various studies. In the present study the HCR-20 was very much a *live* process which was used to structure decisions about an individual's progress and the risks presented by current treatment strategies, including community access. Consequently, the HCR-20-R subscale was used in a dynamic sense to structure decisions about risk at a particular point in a treatment pathway and was consequently updated on a regular basis. This appears to be different to the more cross-sectional use of the HCR-20 in these other two studies which perhaps represent a snap-shot of

Risk-Management issues at a specific time. It is also not clear in these two studies to what extent the clinical team was involved in the completion of the HCR-20 and therefore to what extent it accurately reflects current risk management issues.

A limitation of the present study relates to the limited data collection period of three months utilised compared with the one year follow-up period used by Morrissey, Hogue, et al. (2007). The three-month period was used in order to maximise the number of participants whose data could be used in the analysis. Five participants (10.6 %) remained on the unit for an assessment period of three-months only and a further 11 (23.4%) had left the unit within one-year of the completion of the assessments used in this study. Again, it may be that assessment completion was influenced by current behaviour and that the relationship over a longer period of time may have been reduced. However, the single best predictor of aggression was the HCR-20-H which, of the three HCR-20 sub-scales one would expect to have the most stability over time.

6.13. Conclusion

The study reported in this chapter indicated that the I7i and the I7i-R scores did not correlate with any form of institutional aggression. In addition, the I7i and the I7i-R did not predict any form of institutional aggression. In contrast the PCL-R Total, Factor 1, Factor 2 and 13-Item Total had numerous moderate correlations with all forms of institutional aggression (although PCL-R Total and Type 3 High Risk Aggression and Factor 2 and Type 2 Interpersonal/Physical Aggression were only weak correlations). Only Factor 2 and Type 3 High Risk Aggression were not significantly correlated. In addition, the PCL-R and the HCR-20 appear to have some utility in the prediction of aggression in a medium secure setting. In the present study this was particularly true of Type 2 Interpersonal/Physical aggression. The HCR-20-H reached the level of high predictive accuracy and PCL-R total,

Factor 1, PCL-R 13-item total, HCR-20 total and HCR-20-R all reached the level of moderate predictive accuracy. In addition, PCL-R Factor 1 predicted Type 1 Verbal/property Aggression and Type 3 High Risk Aggression with moderate accuracy and the 13-Item Total and HCR-20-H predicted Type 3 High Risk Aggression with moderate accuracy.

**CHAPTER SEVEN: A COMPARISON OF SELF-REPORT IMPULSIVITY AND
LOCUS OF CONTROL BETWEEN INSTITUTIONALISED OFFENDERS WITH ID,
INSTITUTIONALISED NON-OFFENDERS WITH ID AND NON-
INSTITUTIONALISED NON-OFFENDERS WITH ID.**

7.0. Introduction

The preceding chapters have outlined the adaptation of self-report measures of impulsivity and locus of control in order to facilitate valid responding in populations with an ID.

Research in mainstream, non-ID populations, indicates that high impulsivity and an external locus of control (LOC) differentiates offenders from non-offenders. An external LOC has also been associated with ID *per se*. In addition, an external loc has been associated with institutionalised living. This means that the presence of an external loc in detained ID offenders is difficult to interpret. The purpose of this study was to examine the levels of self-report LOC, as measured by the M-ANSIE, and self-report impulsivity, as measured by the I7i-R, in three samples with ID: Group 1 comprised 47 male offenders with ID detained in a Medium Secure Unit; Group 2 comprised 46 male non-offenders with ID living in similar institutional type accommodation; and Group 3 comprised 46 male non-offenders with ID living in non-institutional type accommodation.

7.1. Locus of Control

Locus of Control (LOC) is a construct derived from social learning theory and is regarded as a fundamental personality characteristic (Rotter, 1966). It runs on a continuum from internal to external referring to the extent to which a person feels that events are contingent on his or her own behaviour or external factors (Nowicki, 1976). This ultimately dictates the individual's subjective evaluation of whether or not he or she is responsible for his or her

own behaviour (Lefcourt, 1976). Typically, someone functioning well has a locus of control that tends to be slightly internal but can be flexible depending upon the situation. Individuals with an externally orientated locus of control do not think they have much control in their lives, believing instead that luck, chance, fate or powerful others have more control over events than they do (Mercer & Snell, 1977).

Research outlined in Chapter 4 suggested that individuals with ID have a more external orientation of locus of control than those without ID (Brown, 1980; Dudley-Mailing et al., 1982; Hallahan, Gajar, Cohen & Tarver, 1978; Wehmeyer, 1993, 1994; Wehmeyer et al., 1996).

7.2. LOC and criminal behaviour

Evidence presented in Chapter 1 indicated that an external LOC is associated with the commissioning of numerous forms of criminal behaviour. For example, an externally orientated locus of control has been identified in: delinquents (Beck & Ollendick, 1976; Ducette & Wolk, 1972; Duke & Fenhagen, 1975; Eitzen, 1974; Elenewski, 1975; Martin, 1975; Parrott & Strongman, 1984); sexual offenders (Beck-Sander, 1995; Gudjonsson, 1990); violent offenders (Beck-Sander, 1995; Fisher et al, 1998; Wiehe, 1987); other non-violent offenders (Fisher et al., 1998); and fire setters (Hall, 1995).

7.3. LOC and other potential criminogenic factors

LOC orientation is also associated with other factors implicated in offending and antisocial behaviours. Wiehe (1987) found an inverse relationship between locus of control and empathy as measured by the Hogan Empathy Scale (Hogan, 1969). Abusing mothers demonstrated significantly less empathic ability and a more external locus of control orientation compared with non-abusing mothers.

Numerous studies have identified a relationship between a poor self-image and an external locus of control orientation in delinquent populations (Duke & Fenhagen, 1975; Martin, 1975; Fitts & Hamner, 1969; Friedberg, 1982; Martinez, Hays & Solway, 1977). This suggests that the combination of a poor self-image combined with a view of one's behaviour as ineffectual in achieving goals or gaining rewards may act as strong motivators for anti-social behaviours. An external orientation of locus of control has been found to correlate with numerous other factors associated with offending. For example, risk taking (Ducette & Wolk, 1972), alcohol use (Distefano, Pryer & Garrison, 1972; Goss & Morosko, 1970) and the presence of a diagnosed conduct disorder in adolescents (Powell & Rosén, 1999). In contrast, internal locus of control has been correlated with resistance to coercion, an important factor in the prevention of delinquency (Gore, 1962; Strickland, 1965).

Locus of control has also been related to other criminogenic factors such as impulsiveness (Clark, 1994; Wiehe, 1987) and moral judgements (Bachrach, Huesmann, & Peterson, 1976). Cognitive distortions in paedophiles have been found to be highly significantly correlated with external attributions (Gudjonsson, 1990). However, the relationship between cognitive distortions and LOC in samples with ID is not straightforward. For example, Langdon and Talbot (2006) reported a significant reduction in cognitive distortions expressed post-intervention but with no corresponding change to the orientation of LOC. LOC has been linked with Suggestibility in terms of a correlation between LOC and Total Suggestibility (Gudjonsson & Lister, 1984) and Yield 1 (Gudjonsson & Lister, 1984; Liebman et al., 2002), with those with an external LOC tending to score higher on the GSS.

7.4. Locus of control and institutionalisation

Only a small amount of evidence is available regarding locus of control orientation and institutionalised living. This is largely confined to studies with samples of elderly adults moving into institutions. There is evidence that living in an institution appears to promote external locus of control (Beck & Ollendick, 1976).

Newly admitted nursing home residents with internally orientated locus of control were reported to be behaviourally more competent than those externally orientated (Erber & Dye, 1982). Equally, encouraging decision making, personal responsibility, and freedom among nursing home residents has been found to result in more favourable ratings of alertness, sociability and activity by carers (Langer & Rodin, 1976). However, more established residents rated by carers as well adjusted to institutionalised life have been found to hold external orientations (Felton & Kahana, 1974; Lawton & Nahemow, 1973). It may be that adjustment to an institutional environment is compatible with perceived external control. Internal orientation may be non-adaptive in an institutional environment particularly when the establishment allows for little flexibility. This corresponds with the 'congruency hypothesis' (Watson & Bauml, 1967) which states that individuals will prefer situations that are congruent with their locus of control orientation. Hence, externals will thrive in contexts where they perceive an absence of control. Rotter (1975) argues that greater differentiation in locus of control orientation is likely in less structured, less familiar and more ambiguous environments. Consequently, it has been argued that an external orientation is a desirable condition for long-term institutional residents, particularly as perceived locus of control has been found to be independent of anxiety, morale and cognitive competency (Erber & Dye, 1982).

The evidence suggests that there is a tendency for locus of control orientation to become more external unless measures are in place to encourage maintenance of an internal

orientation (Lawton & Nahemow, 1973). However, caution is required in generalising from the results of these studies because of several methodological problems. There is little evidence available that specifically considers LOC. In addition, all of the studies reported include relatively small sample sizes and take place in a single institution (Beck & Ollendick, 1976; Erber & Dye, 1982; Felton & Kahana, 1974; Langer & Rodin, 1976). In addition, there is no assessment of the particular needs of the individuals in the different studies and therefore it is difficult to judge how representative each sample is of all institutionalised older adults. In addition, the samples in some of the studies included older adults who were in the early stages of dementia. It is possible that more external orientation of locus of control is a result of the progression of dementia rather than the effect of institutionalisation *per se*.

An external LOC has also been reported in other forms of institutional care. For example, Beck and Ollendick (1976) reported high external LOC in a sample of institutionalised delinquents. Some authors have attributed external LOC to institutional living and not as a criminogenic factor (Little & Kendall, 1978; Wehmeyer, 1998; Wehmeyer et al, 1995). Whilst none of these studies specifically assessed LOC orientation in institutionalised populations with ID it seems likely that similar factors will affect that population as well. However, it is insufficient to assume this similarity between elderly mainstream adults in residential and care homes and intellectually disabled offenders in forensic environments. Research must empirically demonstrate this correspondence.

Wehmeyer and colleagues have reported on links between accommodation and self-determination in populations with ID (Wehmeyer, 1994, 1994b, Wehmeyer et al., 1995; Tossebro, 1995). Locus of control has been identified as a key component of self-determination, along with self-advocacy, decision-making skills and independent living skills (Powers, Sowers, Turner, Nesbitt, Knowles & Ellison, 1996; Wehmeyer, 1994). Wehmeyer (1994,1997) reported significantly higher ANSIE scores for participants with ID than those

without ID. However, his research does not go on to demonstrate an empirical link with place of residence, although Wehmeyer argues that factors such as residence in an institution play a part in the development of an institutional personality. However, Wehmeyer does not provide research findings to demonstrate this link between accommodation and LOC orientation.

Research suggests that self-determination is improved in smaller living units (Tossebro, 1995). Consequently, it is possible that incarcerated offenders living in institutions are likely to experience less self-determination and consequently a more external LOC than those living more independently. In addition, it could be speculated that controlled environments, such as medium secure units, are likely to have a negative impact upon self-determination and hence LOC orientation. Consequently, individuals living in less restrictive environments may experience greater self-determination and hence have a more internally orientated LOC. Tossebro (1995) argues that there were differences in the components of care within the smaller living units that may have influenced the results, rather than the size of unit per-se. For example, the smaller living units tended to have more emphasis on engaging individuals' on tasks selected by them, rather than group activities. Consequently, Tossebro (1995) argues that the size of the unit is a proxy for the quality of staff to patient inter-actions.

Clearly, at present when a measurement of LOC is taken with detained, ID offenders the interpretation of that measure is problematic because the external orientation may denote the ID, institutionalisation or a forensic risk factor. However, any elevation in external LOC identified in a forensic ID population, over mainstream means, tends to be viewed as a treatment objective and hence as being informative regarding that individual's risk. It is possible that this criterion is too harsh and that a detained non-offender ID population will also show a more external orientation than mainstream means. It is possible that any

elevation in self-report impulsivity and external LOC currently identified in people with ID in forensic settings should not necessarily be regarded as a forensic need and consequently should not inform forensic risk assessment. As has already been mentioned, moderate correlations have been reported between LOC and impulsivity in samples of offenders (Clark, 1994; Wiehe, 1987).

7.5. Locus of control and offenders in prison

Several studies have reported that LOC is elevated in samples of offenders detained in prison (Pugh, 1993; Levenson, 1975). However, most studies have used LOC measures specifically developed for prison populations (Pugh, 1994). Pugh (1992) reports on the use of the ANSIE in a sample of offenders in prison. The mean score for the ANSIE is not reported. However, Pugh (1992) reports that there is a moderate correlation between scores on the ANSIE and scores on the Prison Locus of Control Scale ($r=.63$). This suggests that elevated scores on such scales indicate elevated externality as measured on the ANSIE.

Research with offenders in prison suggests that offenders with an internally orientated LOC are better adjusted to prison-life (Pugh, 1998). For example, Pugh (1998) reported that offenders detained in segregation cells, for frequent and continual rule violations, scored significantly higher on LOC scales, signifying a more external LOC, than a group of trustee offenders, who were rated as well adjusted to prison life. More specifically, internally orientated offenders in prison, have been reported to experience significantly fewer problems associated with incarceration (Goodstein, 1979; Pugh, 1994, 1998; Zamble & Goodstein, 1986), demonstrated better problem-solving skills (Pugh, 1992, 1993, 1998), are disciplined less frequently and less severely (Levenson, 1975), report significantly fewer problems related to stress (Pugh 1994, 1998), and report lower levels of non-psychotic depression (Pugh, 1998). In addition, internally orientated offenders in prison are more likely to

participate in treatment, education and occupational programmes (Groh & Goldenberg, 1976). After release offenders with a more internal orientation of LOC have continued to report better adjustment in terms of the variables reported above (Goodstein, 1979).

The findings from these studies are consistent in suggesting that those with more internal LOC are better adjusted than those with more external LOC. There are methodological difficulties which mean that the results must be treated with some caution. For example, most of the studies employ reasonable sample sizes ($n \geq 100$) (Pugh, 1992, 1993, 1998, Zamble & Goodstein, 1986). However, many of the comparisons in the studies only utilise a portion of the overall sample, so that in effect the sample size is much reduced ($n=28$) (Pugh, 1998). In addition, many of the outcomes of these studies are based upon pencil and paper assessments rather than behaviour. Consequently, problems with incarceration (Prison Problem Scale, Zamble & Porporino, 1988), problem-solving (Problem-solving Inventory: Heppner & Petersen, 1982), stress (Perceived Stress Scale: Cohen, 1983), and depression (Generalised Contentment Scale: Hudson, 1982), are all based upon self-report. No assessment was made to the validity of these measures in relation to actual behaviour in prison. In addition, it is also unclear whether elevated scores on the Prison Control Scale (Zamble & Porporino, 1988) and the Prison Locus of Control Scale (Pugh, 1992) relate to external LOC on mainstream measures of LOC such as the ANSIE. Equally, all of the studies reported above did not include a control group of non-offenders as a means of determining whether scores were external for the offenders in prison. Regardless, of the methodological problems associated with these studies the results are consistent in suggesting that those with an internal orientation of LOC are better adjusted to prison life, and continue to report better outcomes after release (Goodstein, 1997). Such findings may have implications for detained offenders with ID and suggest that developing an internal LOC may represent a treatment goal for those in institutions.

7.6. Impulsivity and Criminal Behaviour

Links between impulsivity and criminal behaviour are well established. For example, in juvenile and adolescent populations impulsivity has been found to correlate with re-offending, both sexual and non-sexual (Miner, 2002; Prentky, Harris, Frizzell & Righthand, 2000) and the development of delinquent behaviour (White et al., 1994). White et al. (2002) found that aggressive offences committed under the influence of alcohol were more likely to be committed by more impulsive adolescent males.

Similar links between impulsivity and criminal behaviour have been identified in adult populations. For example, impulsivity has been found to correlate with three forms of offending in adult child molesters (Prentky et al, 1997), alcohol use and violence (Hamberger & Hastings, 1991) and intimate partner violence (Cunradi, Caetano, Clark & Schafer, 1999; Schafer et al., 2004). In addition, impulsivity has been found to differentiate male and female violent and non-violent parolees from controls (Cherek & Lane, 1999; Cherek et al., 1997) and to predict re-offence risk across domains of criminal behaviour (Prentky & Knight, 1991).

Junger, West and Timman (2001) argue that evidence supports the presence of an underlying trait which may represent a general disregard for the long term adverse consequences of one's actions and suggested that this trait may be understood as risk-taking, impulsiveness or lack of self-control. Indeed, impulsivity is identified as a central tenant in Gottfredson and Hirschi's (1990) General Theory of Crime in which 'low self-control' describes a relatively stable predisposition to commit crime over an individual's life course.

The evidence presented above suggests that LOC and impulsivity are associated with offending behaviour. Equally, LOC has been reported to correlate with impulsivity (Clark, Waller & Fisher, 1994; Wiehe, 1987). For example, Clark et al. (1994) reported a moderate correlation ($r=.50$) between scores on the ANSIE and the I7i. Clark et al. reported a mean M-

ANSIE score of 15.11 and a mean I7i score of 10.67 for a sample of sex offenders. However, this study used a small sample ($n=18$), none of whom were reported as having ID. This I7i mean score appears high when compared with Corulla (1987) 8.76 (4.79) and Eysenck et al. (1985) 7.93 (4.12). The present study provides an opportunity to assess the relationship between LOC and impulsivity in a sample of offenders with ID.

The work of Wehmeyer and others (Tossebro, 1995; Wehmeyer, 1994a, 1994b, Wehmeyer et al., 1995) suggests that LOC forms a crucial component of self-determination, which is in turn, influenced by factors such as living environment. There are a number of methodological problems associated with the other evidence presented above. For example, many of the studies used small sample sizes and therefore caution must be exercised in generalising from the results (Beck & Ollendick, 1976; Cherek & Lane, 1999; Cherek, Moeller, Dougherty & Rhoades, 1997; Eitzen, 1974). Those that include larger sample sizes (Wehmeyer, 1994, 1997; Wehmeyer et al, 1995) do not specify how ID was determined nor do they report the mean and range of IQs of either the experimental or the control group. It is therefore difficult to make informed comparisons between the groups in individual studies, and between results from different studies. In addition, several of the studies linking LOC and impulsivity with crime use broad definitions of crime and do not specify whether an individual has actually been convicted (Beck & Ollendick, 1976; Eitzen, 1974).

In addition, most of the research outlined above relates to main-stream populations, without ID. Parry and Lindsay (2003) and Snoyman and Aicken (2011), reported higher rates of self-report impulsivity, as measured by the BIS-11, in violent offenders with ID in comparison with sexual offenders with ID. However, neither study included a non-offender control group and so were unable to demonstrate whether the rate of impulsivity was elevated or not. In addition, both studies used the BIS-11. There was some confusion regarding different versions of the tool in circulation and consequently making comparisons between

studies is problematic. It is also important to develop and/or adapt a range of measures to assess areas of interest which may help to define the domain of phenomenon in question (Carmines & Zeller, 1979). In addition, the inter-correlation of different tools can be used as a means of validating the data that they produce (Carmines & Zeller, 1979).

Research findings from mainstream populations and their implications for risk assessment and treatment initiatives are frequently extrapolated into populations in ID. However, research has yet to demonstrate a link between impulsivity and criminality in an ID population and consequently its application in such populations is potentially erroneous and clearly not evidence based.

7.7. Aims of research

Given these findings, the present study was devised to explore LOC and impulsivity in three groups: (1) institutionalised offenders with ID, (2) institutionalised non-offenders with ID, and (3) non-institutionalised non-offenders with ID. If LOC and impulsivity are criminogenic factors in an ID population, significant differences between group 1 and the other two groups should be observed. If LOC and impulsivity are influenced by living in institutional type accommodation then significant differences between group 3 and the other two groups would be expected.

7.7.1. Hypotheses

1. All groups will display a high external locus of control.
2. M-ANSIE total score will be significantly higher in offenders (group 1) than in non-offender groups (groups 2 and 3).

3. M-ANSIE total score will be significantly higher in institutional groups (groups 1 and 2) than those living in the community (group 3).
4. All groups will self-report high levels of impulsivity.
5. I7i-R scores will be significantly higher in offenders than in control groups.
6. Scores on the I7i-R will be significantly correlated with scores on the M-ANSIE.

7.8 Method

7.8.1. Participants

Details of the samples can be found in Chapter 1.

7.8.2. Design

A between-subjects design was employed to allow comparisons across the three groups outlined above.

7.8.3. Measures

The M-ANSIE is described in Chapter 4 and the I7i-R is described in Chapter 5.

7.8.4. Procedure

The same procedure was followed to gain informed consent as described in Chapter 1.

Where consent was gained, the participants completed the M-ANSIE and the I7i-R, a process that took 15–20 minutes. Questionnaire items were read aloud to each participant and the response recorded appropriately. Participants were reminded that they could withdraw consent at any time, even during the data collection process.

7.9. Results

Kolmogorov-Smirnov tests were conducted on the M-ANSIE data and the I7i-R data to assess the normality of the data. The results indicated that the distributions of both the M-ANSIE and the I7i-R data were not significantly different to a normal distribution across all groups. Parametric tests could therefore be conducted. Levene's tests for the M-ANSIE and the I7i-R were not significant indicating that the assumption of homogeneity of variance across the three groups had not been violated. Means scores and standard deviations for total scores on the M-ANSIE and the I7i-R in institutionalised offenders (Experimental Group: Group 1), institutionalised non-offenders (Institutional-type Sample: Group 2) and non-institutionalised non-offenders (Community Sample: Group 3) are reported in Table 7.1. The means for both measures were lower than expected in all groups. In addition, the mean scores for the different groups appear remarkably similar, contrary to expectations.

In order to determine if the institutionalised offender group, institutionalised non-offender group and non-institutionalised non-offender groups significantly differed with respect to scores on the M-ANSIE and the I7i-R, two one-way ANOVAs were conducted. The analysis indicated that there was not a significant difference between the groups on the M-ANSIE ($F_{(2,136)}=0.58, p=.56$) or the I7i-R ($F_{(2, 136)}=.11, p=.90$).

A Pearson's correlation coefficient was conducted to assess whether there was a relationship between the M-ANSIE scores and the I7-R scores. There was no significant relationship between these two variables and the size of the correlation is very small ($r=.06, p=.26$).

Table 7.1. Mean scores and standard deviations for total scores on the M-ANSIE and I7i-R and in Institutionalised Offenders, Institutionalised Non-Offenders and Non-Institutionalised Non-Offenders.

	Group	Mean	Standard Deviation
M-ANSIE	Institutionalised Offenders	14.02	4.08
	Institutionalised Non-offenders	14.96	5.15
	Non-institutionalised Non-offenders	14.89	4.74
I7-R	Institutionalised Offenders	8.51	4.08
	Institutionalised Non-offenders	8.83	2.81
	Non-institutionalised Non-offenders	8.61	2.81

7.10. Discussion

The purpose of the present study was to examine self-reported orientation of LOC and levels of impulsivity in three groups of people with ID. Group 1 comprised institutionalised offenders with ID, Group 2 institutionalised non-offenders with ID and Group 3 non-institutionalised non-offenders with ID. As expected all three groups reported high external orientation of LOC, using the Fisher et al. (1998) cut-off of 12, and thus hypothesis 1 was supported. However, locus of control scores did not differ significantly between offenders and non-offenders, and consequently hypothesis 2 was not supported. Total score on the M-ANSIE did not differ significantly between those in institutional accommodation and those in the community. Consequently hypothesis 3 was not supported. The mean M-ANSIE scores for the three samples in the present study were: Group 1 ($M=14.02$, $SD=4.08$); Group 2 ($M=14.96$, $SD=5.15$); and Group 3 ($M=14.89$, $SD=4.74$). Wehmeyer and Palmer (1997) reported a mean ANSIE score of 18.2 ($SD=4.3$) in a sample of non-offenders with ID and Langdon and Talbot (2006) reported a mean of 18.0 ($SD=3.72$) for a group of sex offenders

with ID. It is difficult to determine why the mean scores in the present study were lower than those previously found in samples with ID. However, in the study reported in chapter 4, participants scored significantly lower on the M-ANSIE ($M=14.45$, $SD=4.59$) than on the ANSIE ($M=16.72$, $SD=4.74$). Therefore, it is possible that the lower mean score reported in the present study is the result of using the M-ANSIE rather than the ANSIE. This would imply that the other studies over-estimated the externality of LOC orientation. Alternatively, the sample used in the present study had a considerably higher IQ ($M=70.55$, $SD=8.03$) than, for example, the one reported by Talbot and Langdon (2006) ($M=64.9$, $SD=6.79$). Langdon and Talbot (2006) constrained their sample to those in the mild and borderline IQ range, whereas the present study included several participants in the low average range. It is possible that this may have affected the results, particularly as LOC orientation has been reported to be more external in samples with lower IQs (Samuel, 1980).

Self-report impulsivity did not appear to be elevated when compared with the means reported by Corulla (1987) and Eysenck et al. (1985). The means for the three samples in the present study were: Group 1 ($M=8.51$, $SD=4.08$); Group 2 ($M=8.83$, $SD=2.81$); and Group 3 ($M=8.61$, $SD=2.81$). Corulla (1987) reported a mean scores of 8.76 ($SD=4.79$), for a sample of male university students, and Eysenck et al. (1985) reported a mean of 7.93 ($SD=4.12$) for a diverse sample from the mainstream non-ID population, which again included students. Consequently, hypothesis 4 was not supported. The sample in the present study contained a group of older men ($n=11$) who were over 50 years of age. Eysenck et al. (1985) reported that impulsivity reduced uniformly as age increased, although they did not report whether this difference was statistically significant. It is possible that the lower impulsivity was due to the older mean age of participants in the present study ($M=31.5$ years, $SD=12.0$) compared to the other studies, for example, Corulla (1987) ($M=21.0$ years, $SD=4.6$). I7i-R scores did not differ significantly between offenders and non-offenders, and consequently hypothesis 5 was

not supported. In addition, total scores on the I7i-R did not correlate significantly with total scores on the M-ANSIE and consequently hypothesis 6 was not supported.

Locus of control orientation has long been regarded as a criminogenic factor in mainstream populations. However, evidence is beginning to accumulate to suggest that this relationship is somewhat different in offender populations with ID. Langdon and Talbot (2006) concluded that the shift in locus of control orientation associated with the successful treatment of mainstream non-ID offenders is not apparent in the successful treatment of offenders with ID. The results from this study support findings from earlier studies in suggesting that ID populations generally possess high external orientation of locus of control (Wehmeyer, 1993). Consequently, within these populations a high external orientation of locus of control does not appear to differentiate offenders from non-offenders. One may speculate whether this propensity for an externally orientated locus of control represents a criminogenic vulnerability in ID populations or whether orientation of locus of control has a different meaning in such populations.

There was no significant difference between scores for those living in institutional type accommodation and those living in the community. This was unexpected as previous research suggested that LOC tends to be elevated in people living institutional accommodation (Beck & Ollendick, 1976). It is possible that care practices have changed since the research on institutionalisation was conducted. Promoting an internal LOC appears to be an established treatment goal for those working with people with ID (Bernie-Smith et al., 2006).

The present study did not identify high levels of self-reported impulsivity in any group. Consequently, self-reported impulsivity failed to differentiate offenders from non-offenders. It appears that the criminogenic utility of self-reported impulsivity is not apparent in samples with ID. This is interesting because an earlier study (Logan et al, 1984) found that

self-report impulsivity correlated with observations of actual impulsive behaviour. However, this relationship failed to extend to forensic impulsive behaviours such as institutional aggression (reported in Chapter 6). This suggests that offenders with ID are no more likely to self-report impulsivity than those in the mainstream non-ID, non-offender population. Again, an examination of the link between self-report impulsivity and actual impulsive behaviour in a non-offender population would help to clarify this issue.

There are limitations to the present study. The first, related to the relationship between self-report and actual behaviour in people with ID. For example, research has previously demonstrated links between self-report impulsivity and actual impulsive behaviour in a mainstream non-ID sample (Logan et al, 1984). Research has yet to demonstrate a similar relationship between self-report impulsivity and impulsive behaviour in a sample with ID.

There were also potential difficulties with the control samples of non-offenders. No information was available about whether these men had criminal records. The NHS case-notes for all participants were examined and no report of criminal activity was recorded for any of the non-offender samples. However, a high proportion of the men (78%) had instances of aggressive behaviour recorded in their case-notes. In some instances (19%) this has led to specific measures to address this behaviour, such as staff meetings or suspensions from day-services. On a few occasions (4%) the police were involved in the follow-up of incidents. However, there was no record of the police taking action against one of the men in the community samples. However, the quality of case-notes varied and the validity and reliability of the case-notes from different establishments is unknown.

A further limitation of the present study was the lack of a control group of non-offenders from the mainstream non-ID population. This would have helped to determine whether the scores obtained on the adapted measures, the M-ANSIE and the I7i-R, were

actually elevated. These tools have been demonstrated to be reliable when used with people with ID, however it is uncertain what the scores actually mean when compared with the ANSIE and I7i scores reported in other studies. The scores were lower than expected in all groups but this may be a feature of the tools themselves. Consequently, comparison with a sample without ID would determine a baseline in that population against which the samples with ID could be compared.

7.11. Conclusion

This study found that self-report impulsivity and LOC did not differentiate offenders with ID from non-offenders with ID. In addition, self-report Impulsivity and LOC did not differentiate those living in institutional-type accommodation from those living in non-institutional accommodation. The results from this study support those of Talbot and Langdon (2006) in suggesting that the relationship between LOC and offending does not follow the same pattern as that identified in mainstream non-ID offenders. The mean scores on the M-ANSIE reported in this study suggest that there is an external orientation of LOC in people with ID, in that mean scores, were above the cut-off determined by Fisher et al. (1998). It is therefore possible that the elevated LOC scores on the M-ANSIE generally mask any specific criminogenic utility. In addition, there are no established means for these adapted measures in the mainstream non-ID population. Consequently, the comparisons between scores on the adapted measures and the original measures may imply lower scores in the present sample but this may simply result from the different measures used. Certainly, the evidence presented in Chapter 4 and Chapter 5 reported significantly lower scores on the adapted measures than the original measures. Consequently, it would be useful to extend this research to include those without an ID to further clarify the relationship between self-report impulsivity and LOC and offending behaviour.

CHAPTER EIGHT: CONCLUSIONS

This thesis has attempted to advance our understanding of offending and the assessment of offending in people with ID. In addition, attempts were made to generate findings that would be of practical use by those employed in the criminal justice system. This has been done by demonstrating the adaptation of two tools measuring self-report LOC and impulsivity respectively. Data relating to the reliability of these tools was also reported. This thesis has attempted to add to our empirical knowledge of response biases and the adaptation of assessment tools to minimise the threat these response biases pose to the reliability and validity of self-report of people with ID. The evidence presented suggests that acquiescence is present in response to questions using yes/no formats, despite adaptation of questions designed to overcome those response biases. Empirical knowledge has also been gained in relation to the predictive accuracy of two risk assessment tools, the PCL-R and the HCR-20, developed for mainstream non-ID offenders, when used in relation to offenders with ID. The evidence presented suggests that both tools have some utility in predicting aggressive incidents in offenders with ID. The thesis has also sought to clarify whether two factors associated with offending in mainstream non-ID samples, LOC and impulsivity, are also elevated in a sample of offenders with ID.

Chapter 1 critically reviewed and summarised the literature on the prevalence of people with ID in the CJS. A formal definition of ID was provided but this was contrasted with the samples used in studies of people with ID, which often included those in the borderline intelligence range. The literature review suggests that people with ID, when including those in the borderline intelligence range, are significantly over-represented in the CJS. However, the evidence also indicated that the prevalence figures varied depending upon the criteria used to define ID. Hayes et al. (2007) highlighted this point clearly. If ID were

determined by IQ alone, as it is in many studies (Birmingham et al, 1996; Vanny et al, 2009), the rate varies according to the IQ cut-off used. For example, 7% of the sample of offenders had IQs below 70, 16% had IQs below 75 (allowing for two standard errors), and 31% fell within the borderline and below (<80). However, if the sample were identified using the VABS score as well, only 3% fell below IQ 70 on both instruments, 9% below 75, and 22% below 80. There are two main points to draw from this research. First, there are a significant number of offenders in prison with considerable cognitive deficits. Secondly, the samples that researcher identify are likely to differ, depending upon the criteria they use to identify ID. Consequently, it is important that researchers report the cognitive ability of their sample, with specific reference to the range of IQ included. Evidence suggests that the range of IQ in a sample affects the response rates and reliability of self-report assessments in samples with low IQ (Alder & Lindsay, 2007; Glenn et al., 2003; Lindsay & Lees, 2003; Payne & Jahoda, 2004; Rojahn et al., 1994; Sigelman et al., 1981a, 1981b), as reported in Chapter 2.

The Prison Reform Trust's *No One Knows* project (Loucks, 2007; Talbot & Riley, 2008) have also highlighted the plight of the large number of people, 20-30% (Loucks, 2007) with significant problems with thinking and understanding who are in prison. They argue that these intellectual difficulties interfere with individuals' ability to cope with many aspects of the CJS. *No One Knows* has deliberately avoided clear definitions of learning difficulties/intellectual disabilities, in an attempt not to be overly inclusive or exclusive, but rather to consider the needs of those who find some activities that involve thinking or understanding difficult, and who need additional support to meet their everyday needs. In view of the large number of people with IQs in the borderline and below range (IQ<85 allowing for two standard errors) in the CJS and the largely unidentified and unmet needs of these individuals, the present study included a sample of people in a secure forensic service

with IQs below 85, and with significant impairments of social /adaptive functioning as determined using the VABS.

The evidence presented also suggested that people with ID commit a wide variety of offences (Barron, Hassiotis & Banes, 2004; Hayes 1993; Hayes & Craddock, 1992; Holland, 2004; Lindsay, 2002; Lindsay, O'Brien, et al., 2010; Simpson & Hogg, 2001; Walker & McCabe, 1973). It is difficult to determine clear patterns in the offending behaviour of people with ID for several reasons. Different researcher have sampled at different stages of the CJS. For example, adolescents with ID self-report higher levels of bullying/threatening others; stealing valuable items from houses/ shops/school; using weapons against others; starting fires; deliberately destroying property; and stealing from someone in the street, compared to those without ID (Dickson et al., 2005). In contrast, Lindsay, O'Brien, et al. (2010) assessed the offence histories of those in offender services for people with ID and reported that fire setting, theft and road traffic offences do not feature prominently. However, one common theme that arises from these prevalence studies is the conclusion that people with ID engage in the commissioning of a wide range of offences (Holland, 2004). Consequently, there is a need for a range of assessment tools to assess risk and need in this population.

Several authors have noted the lack of psychometric tools specifically developed for, or adapted from existing mainstream non-ID assessments, for use with offenders with ID (Lindsay, 2002). This has prompted the development of psychometric tools specifically for this population (Lindsay, Whitefield, et al., 2007) and the adaptation of existing ones (Keeling et al., 2007b; Williams et al, 2007). Previous research in mainstream non-ID samples suggests that self-report external orientation of LOC (Beck & Ollendick, 1976; Beck-Sander, 1995; Duke & Fenhagen, 1975; Eitzen, 1974; Elenewski, 1975; Martin, 1975; Parrott & Strongman, 1984) and high self-report impulsivity (Cherek & Lane, 1999; Cherek

et al., 1997; Cunradi et al., 1999; Hamberger & Hastings, 1991; Prentky et al., 1997; Schafer et al., 2004) are associated with offending behaviour. However, the present author was unable to find any study that demonstrated that this was the case in offenders with ID when compared with a matched control group of non offenders. Consequently, one of the aims of the studies reported in this thesis was the adaptation of reliable self-report measures of LOC and impulsivity to ensure their accurate measurement.

Due to the high numbers of people in the borderline range in the CJS (Crocker et al., 2007; Hayes et al., 2007), and the largely unmet needs of the people within that range (Loucks, 2006, 2007), the inclusion of people in the borderline range in many previous studies (Alexander et al., 2006; Alexander et al., 2002; Crossland et al., 2005; Day 1984, 1995; Eaton & Menolascino, 1982; Hogue et al., 2006), and the inclusion of people with IQs in the borderline range in the forensic service where the research was conducted, it was considered appropriate to include them in the sample for this thesis.

Chapter 2 critically reviewed and summarised the literature on response biases in samples with ID. These biases represent a threat to the reliability and validity of self-report for people with ID. The evidence presented demonstrated that people with ID are vulnerable to several forms of response bias, such as acquiescence, nay-saying, recency and suggestibility. In addition, evidence was presented to suggest that within the population with ID, susceptibility to response biases are related to level of intellectual functioning, with those with lower IQs being more acquiescent (Budd, et al., 1981; Sigelman, et al., 1981b; Sigelman, et al., 1980). This suggests that the need for adapted measures becomes more acute as IQ reduces. In addition, it implies that the use of adapted measures is likely to have a differential impact on results, depending upon the ability of the sample used. Consequently, it is important that researchers specify the range of participant IQs in a sample to facilitate comparison and to facilitate meta-analyses.

It is also apparent that there is a similarity between acquiescence and suggestibility, in that it seems likely that the expression of acquiescence will be apparent in the measurement of suggestibility. This seems likely because ten of the leading questions on the GSS use a yes/no response format. The evidence for this relationship is equivocal (Gudjonsson, 1986; Gudjonsson, 1990; Gudjonsson & Clare, 1995). However, all of these studies used Winkler's acquiescence scale (Winkler et al., 1982). It would appear that a more direct test might be enlightening, by incorporating a series of irrelevant, ambiguous questions into the GSS format. The implications of this might be that it is not simply leading questions that provide a threat to the validity of police interviews of people with ID.

Perhaps in response to the threat of acquiescence, the literature review highlighted a large number of studies utilising Likert-scales with this population. Evidence was presented indicating that Likert-scales can be used reliably by people with ID and that psychometric tools are generating evidence to support their validity with samples with ID, for example using the BAI and the BDI (Alder & Lindsay, 2007; Glenn et al, 2003; Lindsay & Lees, 2003; Lindsay & Skene, 2007).

Chapter 3 critically reviewed the self-report psychometric tools that have been used in the assessment of offenders with ID. The review indicates that there are very few tools that have been developed or adapted for use with offenders with ID, that have good levels of reliability and validity. However, the review also suggests that the use of simple language and sentence structures and the use of specific procedures, including verbal administration of assessments, can lead to the development and adaptation of tools that are reliable and valid for use with offenders with ID.

Chapter 4 reported on the modification of the ANSIE (Nowicki & Duke, 1974a), a measure of LOC, commonly used in mainstream non-ID offenders. The M-ANSIE was demonstrated to be a reliable tool measuring LOC. Initial evidence of its validity was

provided through its strong correlation with the ANSIE. In addition, CPA produced an eight factor solution accounting for 48.61% of the variance. The themes of the factors were; inability to protect oneself, superstition, powerful others, judgement of right and wrong, success is random, things just happen, exerting influence and futility. Some evidence of the validity of the factor structure was provided by similarities to the factor structure generated in a mainstream non-ID sample. In addition, some tentative evidence was presented to suggest that the M-ANSIE has resulted in a reduction in acquiescent responding, however this was not statistically tested.

Chapter 5 reported on the adaptation of the I7i Questionnaire (Eysenck et al., 1985). The revised I7i (I7i-R) was demonstrated to have good reliability. Some evidence of its construct validity was provided by its strong correlation with the original I7i-R and weak and moderate correlations with the BIS-11 and its subscales and weak and moderate correlation with Test 1 and Test 3 of the BADS. In addition, CPA produced a six factor solution accounting for 52.41% of the variance. However, the themes of these factors were difficult to interpret. The results from Chapter 4 and Chapter 5 suggest that measures adapted for use with people with ID can produce reliable measures with some evidence to support their validity.

Chapter 6 assessed the ability of the I7i, I7i-R, PCL-R Total, Factor 1, Factor 2, 13-Item Total and the HCR-20 and its subscales to predict institutional violence. Previous research has suggested that the PCL-R and PCL-SV is reliable when used with people with ID and is able to predict various forms of violence and movement to detention at higher levels of security (Gray et al., 2007; Morrissey, Mooney, et al., 2005; Morrissey et al., 2009). In addition, previous research indicated some utility in predicting institutional aggression (Lindsay et al, 2008; Morrissey, Mooney, et al, 2007). However, Morrissey, Hogue, et al (2007) reported that the PCL-R did not predict any form of aggression in a high secure

setting. Consequently, the PCL-R was selected to determine whether it could predict aggression in a medium secure setting.

An ROC analysis indicated that the I7i and the I7i-R did not predict any form of institutional violence. However, the PCL-R Total, Factor 1, Factor 2, 13-Item Total and the HCR-20 and its subscales all predicted physical aggression with moderate to high accuracy. In relation to the HCR-20 the results support those of Gray et al. (2007) in demonstrating good levels of reliability. In addition, HCR-20-H was a strong predictor of physical interpersonal aggression, the HCR-20 Total and HCR-20-R were moderate predictors and the HCR-20-C a weak predictor of institutional aggression and thus supporting previous findings (Lindsay et al., 2008; Morrissey, Mooney, et al., 2007b).

In relation to the PCL-R the results support those of Morrissey et al. (2005) in demonstrating good levels of reliability. In relation to the PCL-R, the PCL-R Total, Factor 1 and 13-Item total all predicted interpersonal/physical aggression with moderate accuracy and Factor 2 predicted interpersonal/physical aggression with low accuracy. The findings generally are supportive of previous research in identifying links between PCL-R scores and violence (Gray et al., 2007; Morrissey et al., 2005, 2009). They suggest that the PCL-R has utility in predicting institutional violence in offenders with ID in a medium secure setting. These results also support previous findings that suggest that factors associated with offending in mainstream non-ID samples are similar to those found in samples with ID (Boer et al, 2004; Hogue et al, 2006; Lindsay, Elliott, et al, 2004, Lindsay, Murphy, et al., 2004; Quinsey et al, 2004). Consequently, risk assessment tools used with samples with ID are likely to have some utility in samples with ID. In addition, these finding implied that LOC and impulsivity might also be implicated in offending in samples with ID.

Chapter 7 assessed whether self-report LOC, as measured by the M-ANSIE, and self-report impulsivity, as measured by the I7i-R, were significantly higher in offenders with ID

when compared with non-offenders with ID. In addition, the influence of being institutionalised was also examined by including institutionalised non-offenders and non-institutionalised non-offenders in the study. The results indicated that M-ANSIE scores and I7i-R did not differ between any groups. The mean scores for all groups for LOC were external in that they were in excess of 12 (Fisher et al., 1998). The results suggest that relationships between LOC and impulsivity, and offending are different in samples with ID, in that they do not appear to be elevated in offender samples.

The overall findings suggest some similarities between offenders of normal intelligence and offenders with ID. Researchers have identified a range of factors, such as general anti-social attitudes and behaviour in ID offender samples that appear to match those identified in offenders with-ID (Boer et al., 2004; Hogue et al., 2006; Lindsay, Elliott, et al, 2004; Lindsay, Murphy, et al., 2004; Quinsey et al, 2004). Consequently, it is not surprising to find that risk assessment tools developed for mainstream non-ID offenders have some utility in offenders with ID. Research studies have generally reported that risk assessment tools developed in mainstream non-offender populations have some utility in samples with ID (Gray et al., 2007; Lindsay et al, 2008; Morrissey et al, 2005, 2009; Morrissey, Mooney, et al., 2007). The findings from the research presented in this thesis indicate that both the PCL-R and the HCR-20 are able to predict institutional aggression in offenders with ID in a medium secure setting. However, the findings reported in this thesis suggest that the relationships between factors associated with offending in mainstream, non-ID samples are not the same in offenders with ID. Self-report impulsivity has been identified as a criminogenic factor in mainstream offender populations. However, as reported in Chapter 7, self-report impulsivity as measured by the I7i-R did not differentiate offenders with ID from non-Offenders with ID.

In addition, in mainstream non-ID samples, a self-reported high external LOC has been found to differentiate offenders from non-offenders. However, as reported in Chapter 7, self-report high external LOC, as measured by the M-ANSIE, did not differentiate offenders with ID from non-Offenders with ID. This suggests that the relationship between these variables and offending is different in people with ID compared to those without ID.

One difference between the sample of offenders with ID used in this thesis is that they are detained in a hospital, rather than prison, or detention centre of delinquents. It is possible that differences in these environments has played a role in masking potential links between LOC, impulsivity and offending, and impacted upon the predictive accuracy of the PCL-R and the HCR-20. For example, it is possible that the environment in a specialist ID Forensic Service may be structured in a manner that promotes greater decision making opportunities and facilitates choice for the individual in many aspects of their life. This may in turn promote a more internal LOC and lower impulsivity, with improved decision making opportunities potentially leading to a better understanding of the consequences of decisions and a reduction in impulsivity (Zigler & Burack, 1989). In addition, all those entering such a facility will be subject to interventions developed to facilitate changes in thinking. This may be through formal one-to-one or group interventions, structured sessions with their named nurse as well as care-plans conducted by all members of the care team. It is possible that these interventions have influenced the orientation of LOC and the expression of impulsive behaviour. In addition, it is possible that those entering medium secure units as opposed to prison, share some factor not common to all offenders with ID. For example, they may be those who represent the greatest risk. As such this may reduce the variability of scores on psychometrics and mask the full utility of such measures.

References

- Abbeduto, L., Furman, L., & Davies, B. (1989). Relation between receptive language and mental age of persons with mental retardation. *American Journal on Mental Retardation*, 93, 535-543
- Abbeduto, L., & Nuccio, J. B. (1991). Relation between receptive language and cognitive maturity in persons with mental retardation. *American Journal on Mental Retardation*, 96 (2), 143-149
- Abel, G. G., Becker, J. V., & Cunningham-Rathner, J. (1984). Complications, consent and cognitions in sex between children and adults. *International Journal of Law and Psychiatry*, 7, 89-103.
- Abramowitz, S. I. (1973). Internal–external control and social–political activism, a test of the dimensionality of Rotter’s internal–external scale. *A Journal of Consulting and Clinical Psychology*, 40, 196-201.
- Achenbach, T. M., & Edelbrock, C. (1983). *Manual for the Child Behavior Checklist*. Burlington VA: University of Vermont.
- Ager, A. (1990). *The Life Experience Checklist*. Kidderminster: British Institute of Learning Disabilities
- Ainslie, G. (1975). Specious reward: A behavioural theory of impulsiveness and impulse control. *Psychological Bulletin*, 82, 463-496.
- Alder, L., & Lindsay, W.R. (2007). Exploratory factor analysis and convergent validity of the Dundee Provocation Inventory. *Journal of Intellectual and Developmental Disability*, 32 (3), 190-199.
- Alexander, R., & Cooray, S. (2003). Diagnosis of personality disorders in learning disability. *British Journal of Psychiatry*, 44, 528-531.

Alexander, R.T., Crouch, K., Halstead, S., & Piachaud, J. (2006). Long-term outcome from a medium secure service for people with intellectual disability. *Journal of Intellectual Disability Research*, 50 (4), 305-315.

Alexander, R.T., Hiremath, A., Chester, V., Green, F., Gunaratna, I., & Hoare, S. (2011). Evaluation of treatment outcomes for a medium secure unit for people with intellectual disability. *Advances in Mental Health and Intellectual Disabilities*, 5 (1), 22-32.

Alexander, R.T, Piachaud, J., Odebivi, L., & Gangadharan, S.K. (2002). Referrals to a forensic service in the psychiatry of learning disability. *British Journal Forensic Practice*, 4, 29-33.

American Association on Mental Retardation (AAMR) (1992). *Mental Retardation. Definition, classification and systems of supports*, 9th Ed. Washington D. C.: American Association on Mental Retardation.

American Psychiatric Association (APA) (1994). *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV)*. American Psychiatric Association: Washington DC.

Al Nhar, T. (1992). Measurement of locus of control and its relationship with academic achievement – a critical review. *Dirasat*, 19, 72-89.

Apter, A., van Praag, H. M., Plutchik, R., Sevy, S., Korn, M., & Brown, S. L. (1990). Interrelationships among anxiety, aggression, impulsivity and mood: a serotonergically linked cluster? *Psychiatry Research*, 32, 191-199.

Atkinson, D. (1988). Research interviews with people with mental handicaps. *Mental Handicap Research*, 1, 75-90.

Babinski, L. M., Hartsough, C. S., & Lambert, N. M. (1999). Childhood conduct problems, hyperactivity-impulsivity, and inattention as predictors of adult criminal activity. *Journal of Child Psychology and Psychiatry*, 40 (3), 347-355.

Bachrach, R., Huesmann, L. R., & Peterson, R. A. (1976). Relation between locus of control and development of moral judgement. *Child Development*, 48, 1340-1352.

Bain, S. A., & Baxter, J. S. (2000). Interrogative suggestibility: The role of interviewer behaviour. *Legal and Criminological Psychology*, 5, 123-133.

Bain, S. A., Baxter, J. S., & Fellowes (2004). Interacting influences on interrogative suggestibility. *Legal and Criminological Psychology*, 9 (2), 239-252.

Baker, W., & Bramston, P. (1997). Attributional and emotional determinants of aggression in people with mild intellectual disabilities. *Journal of Intellectual and Developmental Disability*, 22 (3), 169-175.

Balla, D., & Zigler, E. (1979). Personality development in retarded persons. In N. R. Ellis (Ed) *Handbook of Mental Deficiency: Psychological Theory and Research* (2nd Ed). Hillsdale, NY: Lawrence Erlbaum.

Ballinger, B.R., & Reid, A.H. (1987). A standardised assessment of personality disorder in mental handicap. *British Journal of Psychiatry*, 150, 108-109.

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

Barkley, R.A. (1997). Behavioural inhibition, sustained attention, and executive functions: constructing a unified theory of ADHD. *Psychological Bulletin*, 121, 65-94.

Barratt, E. S. (1965). Factor analysis of some psychometric measures of impulsiveness and anxiety. *Psychological Reports*, 16, 547-554.

Barratt, E. S. (1994). Impulsiveness and aggression. In Monahan. J and Steadman, H. J. (Eds.), *Violence and mental disorder: developments in risk assessment* (pp. 61-79). Chicago: University of Chicago Press.

Barron, P., Hassiotis, A., & Banes, J. (2004). The effectiveness of treatment for male sex offenders with learning disabilities: a review of the literature. *Journal of Sexual Aggression*, 10, 215-236.

Bartel, N.R., Bryen, D., & Keehn, S. (1973). Language comprehension in the mentally retarded child. *Exceptional Children*, 39, 375-382.

Bartholomew, D.J. (1980). Factor analysis for categorical data. *Journal of Royal Statistical Society*, 42 (3), 293-321.

Bartholomew, K., & Horowitz, L.M. (1991). Attachment styles among young adults: a test of a four-category model. *Journal of Personality and Social Psychology*, 61, 226-244.

Basquill, M., Nezu, C.M., Nezu, A.M., & Klein, T.L. (2004). Aggression-related hostility bias and social problem-solving deficits in adult males with mental retardation. *American Journal on Mental Retardation*, 109, 255-263.

Battle, J. (1992). *Culture-Free Self-Esteem Inventory*. Austin, TX: Pro-Ed.

Baxter, J. S., & Boon, J. C. W. (2000). Interrogative suggestibility: the importance of being earnest. *Personality and Individual Differences*, 28, 753-762.

Baxter, J. S., Jackson, M., & Bain, S. A. (2003). Interrogative suggestibility: interactions between interviewees' self-esteem and interviewer style. *Personality and Individual Differences*, 35, 1285-1292.

Beail, N. (2002). Interrogative suggestibility, memory and intellectual disability. *Journal of Applied Research in Intellectual Disabilities*, 15, 129-137.

Beck, A.T., Brown, G., Steer, R.A., Eidelson, J.I., & Riskind, J.H. (1987). Differentiating anxiety and depression: A test of the cognitive content-specificity hypothesis. *Journal of Abnormal Psychology*, 96, 179-183.

Beck, D.C., Carlson, G.A., Russell, A.T., & Brownfield, F.E. (1987). Use of depression rating instruments in developmentally and educationally delayed adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry*, 26, 97-100.

Beck, S., & Ollendick, T. (1976). Personal space, sex of experimenter, and locus of control in normal and delinquent adolescents. *Psychological Reports*, 38, 383-387.

Beck, A.T., & Steer, R.A. (1990). *Beck Anxiety Inventory manual*. San Antonio, TX: The Psychological Corporation.

Beck, A.T., Steer, R.A., & Carbin, M.G. (1988). Psychometric properties of the Beck Depression Inventory: twenty-five years of evaluation. *Clinical Psychology Review*, 8, (1), 77-100.

Beck, A.T., Ward, C.H., Mendolism, M., Mock J., & Erbaugh, J. (1981). An inventory for measuring depression. *Archives of General Psychiatry*, 4, 561-571.

Beck-Sander, A. (1995). Childhood abuse in adult offenders: The role of control in perpetuating cycles of abuse. *Journal of Forensic Psychiatry*, 6, 486-498.

Beech, A., Fisher, D., & Becket, R. (1998). *STEP 3: An evaluation of the prison sex offender treatment programme*. London: Home Office publication.

Beirne-Smith, M., Patton, J.R., & Kim, S.H. (2003). *Mental Retardation: an introduction to intellectual disabilities*. New Jersey: Pearson Education.

Belfrage, H., Fransson, G., & Strand, S. (2000). Prediction of violence using the HCR-20: a prospective study in two maximum security correctional institutions. *The Journal of Forensic Psychiatry*, 11, 167-175.

- Belson, W., & Duncan, J. A. (1962). A comparison of the checklist and the open response questioning systems. *Applied Statistics*, 11, 120-132.
- Belter, R.W., & Brinkmann, E. (1981). The construct validity of the Nowicki-Strickland locus of control scale for children. *Psychological Reports*, 48, 427-432.
- Bender, M.P., Aitman, J.B., Biggs, S.J., & Haug, U. (1983). Initial findings covering a sexual knowledge questionnaire. *Mental Handicap*, 11, 168-169.
- Bennell, C. (2005). Improving police decision making: general principles and practical applications of receiver operating characteristic analysis. *Applied Cognitive Psychology*, 19, 1157-1175.
- Benson, B. A. (1992). Teaching anger management to persons with mental retardation. *Unpublished PhD Thesis*. University of Illinois at Chicago.
- Benson, B. A. (1994). Anger management training: A self-control programme for people with mild mental retardation. In N. Bouras (Ed), *Mental Health on Mental Retardation* (pp. 224-232). Cambridge: Cambridge University Press.
- Benson, B. A., & Ivins, J. (1992). Anger, depression and self-concept in adults with mental retardation. *Journal of Intellectual Disability Research*, 36, 169-175.
- Benson, B. A., Johnson Rice, C., & Miranti, S. V. (1986). Effects of anger management training with mentally retarded adults in group treatment. *Journal of Consulting and Clinical Psychology*, 54, 728-729.
- Bickel, W.K., Odum, A.L., & Madden, G.J. (1999). Impulsivity and cigarette smoking: delay discounting in current, never, and ex-smokers. *Psychopharmacology*, 146 (4), 447-454.
- Birleson, P. (1981). The validity of depressive disorder in childhood and the development of a self-rating scale: a research report. *Journal of Child Psychology and Psychiatry*, 22, 73-88.
- Birmingham, L., Mason, D., & Grubin, D. (1996). Prevalence of mental disorder in remand prisoners: consecutive case study. *British Medical Journal*, 313 (7071), 1521-1524.

Black, L., & Novaco, R. W. (1993). Treatment of anger with a developmentally handicapped man. In R. A. Wells & V. J. Giannetti (Eds.), *Casebook of the Brief Psychotherapies*. New York: Plenum Press.

Black, L. (1994). Helping people with learning difficulties express anger in socially acceptable ways: the development of a treatment intervention and outcome measures. *Unpublished Doctoral Dissertation*. University of St Andrews, Fife.

Blacker, J., Beech, A.R., Wilcox, D.T., & Boer, D.P. (2011). The assessment of dynamic risk and recidivism in a sample of special needs sexual offenders. *Psychology, Crime and Law*, 17 (1), 75-92.

Blackwell, S.L., McIntyre, C.W., & Murray, M.E. (1983). Information processed from brief visual displays by learning-disabled boys. *Child Development*, 54, 927-940.

Blair, E., Sudman, S., Bradburn, N. M., & Stocking, C. (1977). How to ask questions about drinking and sex: response effects in measuring consumer behaviour. *Journal of Marketing Research*, 14, 316-321.

Blanchard, R., Kolla, N.J., Cantor, J.M., Klassen, P.E., Dickey, R., Kuban, M.E., & Blak, T. (2007). IQ, handedness and pedophilia in adult male patients stratified by referral source. *Sex Abuse*, 19, 285-309.

Blau, G. J. (1984). Brief note comparing the Rotter and Levenson measures of locus of control. *Perceptual and Motor Skills*, 58, 173-174.

Boon, J.C. W., & Baxter, J. S. (2000). Minimising interrogative suggestibility. *Legal and Criminological Psychology*, 5, 273-284.

Booth, T., Simons, K., & Booth, W. (1989). Perspectives of people with learning difficulties on relocation and community care. *International Journal of Rehabilitation Research*, 12, 167-174.

Borthwick-Duffy, S.A. (1994). Epidemiology and prevalence of psychopathology in people with mental retardation. *Journal of Consulting and Clinical Psychology*, 62 (1) 17-27.

Bouras, N., & Drummond, C. (1992). Behaviour and psychiatric disorders of people with mental handicaps living in the community. *Journal of Intellectual Disability Research*, 36, 349-357.

Bowen, E., & Gilchrist, E. (2004). Do Court- and self-referred domestic violence offenders share the same characteristics? A preliminary comparison of motivation to change, locus of control and anger. *Legal and Criminological Psychology*, 9, 279-294.

Bower, D.P., Frize, M., Pappas, R., Morrissey, C., & Lindsay, W.R. (2010). Suggested adaptations to the SVR-20 for offenders with intellectual disabilities. In L.A. Craig, W.R. Lindsay, & K.D. Browne (Eds.), *Assessment and Treatment of Sexual Offenders with Intellectual Disabilities: a Handbook* (pp 193-209). Chichester: Wiley.

Bower, D.P., Tough, S., & Haaven, J. (2004). Assessment of risk manageability of intellectually disabled sex offenders. *Journal of Applied Research in Intellectual Disabilities*, 17, 275-183.

Bramston, P., & Fogarty, G. (2000). The assessment of emotional distress experienced by people with an intellectual disability: a study of different methodologies. *Research in Developmental Disabilities*, 21, 487-500.

Bramston, P., Fogarty G., & Cummins, R.A. (1999). The nature of stressors reported by people with an intellectual disability. *Journal of Applied Research in Intellectual Disabilities*, 12, 1-10.

Bramston, P., & Miochie, C. (2001). Disability and stress: a study in perspectives. *Journal of Intellectual and Developmental Disability*, 26 (3), 233-242.

Bray, D., & Foreshaw, N. (1996). *Sex Offender's self appraisal scale. Version 1.1*. Preston: Lancashire Care NHS Trust.

British Psychological Society (BPS) (2001) *Learning Disability: definitions and contexts*. Professional Affairs Board of the British Psychological Society, Leicester.

Bromley, J., & Emerson, E. (1995). Beliefs and emotional reactions of care staff working with people with challenging behaviour. *Journal of Intellectual Disability Research*, 39, 341-352.

Bromley, J., Emerson, E., & Caine, A. (1998). The development of a self-report measure to assess the location and intensity of pain in people with intellectual disabilities. *Journal of Intellectual Disability Research*, 42 (1), 72-80.

Brown, A.L. (1972). A rehearsal deficit in retardates continuous short-term memory: keeping track of variables that have few of many states. *Psychonomic Science*, 29, 373-376.

Brown, A.L., Campion, J.C., & Murphy, M.D. (1974). Keeping track of changing variables: long-term retention of a trained rehearsal strategy with retarded adolescents. *American Journal of Mental Deficiency*, 78, 446-453.

Brown, A.L., & Scott, M.S. (1971). Recognition memory for pictures in pre-school children. *Journal of Experimental Child Psychology*, 11, 401-412.

Brown, G., Johnson, P., & Peedie, P. (2002). *Working with learning disabled firesetters: a cognitive behavioural approach in care and treatment of offenders with a learning disability*. Bournemouth: Nursing Praxis International.

Brown, R. (1980). Locus of control and its relationship to intelligence and achievement. *Psychological Reports*, 46, 1249-1250.

Brownmiller, S. (1975). *Against Our Will: Men, Women and Rape*. New York: Bantam.

Broxholme, S. L., & Lindsay, W. R. (2003). Development and preliminary evaluation of a questionnaire on cognitions related to sex offending for use with individuals who have mild intellectual disabilities. *Journal of Intellectual Disability Research*, 47 (6), 472-482.

- Budd, E. C., Sigelman, C. K., & Sigelman, L. (1981). Exploring the outer limits of response bias. *Sociological Focus*, 14, 297-307.
- Bull, R. (1995). Interviewing witnesses with communicative disability. In R. Bull & D. Carson (eds) *Handbook of Psychology in Legal Contexts*. Chichester: Wiley.
- Bull, R. (2010). The investigative interviewing of children and other vulnerable witnesses: psychological research and working/professional practice. *Legal and Criminological Psychology*, 15, 5-23.
- Burgess, P. W., & Robertson, I. H. (2002). Principles of rehabilitation of frontal lobe function. In D. T. Stuss & R. T. Knight (Eds.), *Principles of Frontal Lobe Function* (pp. 557-572). New York: Oxford University Press.
- Buss, A. H., & Plomin, R. (1975). *A temperament theory of personality development*. Wiley: New York.
- Camilleri, J.A., & Quinsey, V.L. (2011). Appraising the risk of sexual and violent recidivism among intellectually disabled offenders. *Psychology, Crime and Law*, 17 (1), 59-74.
- Cant, R., & Standen, P. (2007). What professionals think about offenders with learning disabilities in the criminal justice system. *British Journal of Learning Disabilities*, 35, 174-180.
- Cantor, J.M., Blanchard, R., Christensen, B.K., Dicky, R., Klassen, P., & Beckstead, P.E., ... Kuban, M.E. (2004). Intelligence, memory, and handedness in paedophilia. *Neuropsychology*, 18 (1), 3-14.
- Cantor, J.M., Blanchard, R., Robichaud, L.K., & Christensen, B.K. (2005). Quantitative reanalysis of aggregate data on IQ in sexual offenders. *Psychological Bulletin*, 131, 555-568.
- Cantor, G.N., & Ryan, T.J. (1962). Retention of verbal paired-associates in normals and retardates. *American Journal of Mental Deficiency*, 66, 861-865.

Cardone, D., & Dent, H. (1996). Memory and interrogative suggestibility: The effects of modality of information presentation and retrieval conditions upon the suggestibility scores of people with learning disabilities. *Legal and Criminological Psychology*, 1, 165-177.

Carmines, E.G., & Zeller, R.A. (1979). *Reliability and Validity Assessment*. London: Sage.

Carrigan, P. M. (1960). Extraversion-introversion as a dimension of personality: a reappraisal. *Psychological Bulletin*, 57, 329-360.

Castles, E. E., & Glass, C. R. (1986). Training in social and interpersonal problem solving skills for mildly and moderately retarded adults. *American Journal of Mental Deficiency*, 91, 1, 35-42.

Cattell, R. B. (1978). *The Scientific Use of Factor Analysis*. New York: Plenum.

Cattell, R. B., Eber, H. W., & Tatsuoka, M. M. (1970). *Handbook for the Sixteen Personality Factor Questionnaire*. Champaign, IL: IPAT.

Chandler, T.A., & Dugovics, D.A. (1977). Sex differences in research on locus of control. *Psychological Reports*, 41, 47-53.

Chapman, J.W., & Boersma, F.J. (1979). Learning disabilities, locus of control, and mother attitudes. *Journal of Educational Psychology*, 71 (2), 250-258.

Chapman, K., & Oakes, P. (1995). Asking people with learning disabilities their view on direct psychological interventions. *Clinical Psychology Forum*, 81, 28-33.

Chapman, L. J., Chapman, J. P., Numbers, J. S., Edell, W. S., Carpenter, B. N., & Beckfield, D. (1984). Impulsive nonconformity as a trait contributing to the prediction of psychotic like and schizotypal symptoms. *Journal of Nervous and Mental Disease*, 11, 681-91.

Charman, T., & Clare, I. C. H. (1992). Education about the laws and social rules relating to sexual behaviour, *Mental Handicap*, 20, 74-81.

Cherek, D. R., & Lane, S. D. (1999). Laboratory and psychometric measurements of impulsivity among violent and nonviolent female parolees. *Biological Psychiatry*, 46, 273-280.

Cherek, D. R., Moeller, F. G., Dougherty, M., & Rhodes, H. (1997). Studies of violent and nonviolent male parolees: II laboratory and psychometric measurement of impulsivity. *Biological Psychiatry*, 41, 523-529.

Chui, L. (1990). Self-esteem of gifted, normal and mildly mentally handicapped children. *Psychology in the Schools*, 27, 263-268.

Cicchetti, D.V., & Sparrow, S.S. (1981). Developing criteria for establishing interrater reliability of specific items: applications to assessment of adaptive behaviour. *American Journal of Mental Deficiency*, 86, 127-137.

Clare, I. C. H., (1993). Issues in the assessment and treatment of male sex offenders with mild learning disabilities. *Sexual and Marital Therapy*, 8, 167-180.

Clare, I. C. H., & Gudjonsson, G. H. (1992). Devising and Piloting an Experimental Version of the 'Notice to Detained Persons'. Royal Commission on Criminal Justice Research Study No. &. London: HMSO.

Clare, I. C. H., & Gudjonsson, G. H. (1993). Interrogative suggestibility, confabulation and acquiescence in people with mild learning difficulties (mental handicap): implications for reliability during Police interrogation. *British Journal of Clinical Psychology*, 32, 295-301.

Clare, I. G. H., & Gudjonsson, G. H. (1995). The vulnerability of suspects with intellectual disabilities during police interviews: a review and experimental study of decision-making. *Mental Health Research*, 8 (2), 110-128.

Clare, I. C. H., & Murphy, G. H. (1998). Working with offenders or alleged offenders with intellectual disabilities. In E. Emerson, C. Hatton, J. Bromley & E. Caine (Eds) *Clinical Psychology in People with Intellectual Disabilities* (pp 154-177). Chichester: Wiley.

Clare, I. C. H., Murphy, G. H., Cox, D. & Chaplin, E. H. (1992). Assessment and treatment of fire-setting: a single-case investigation using a cognitive-behavioural model. *Criminal Behaviour and Mental Health*, 2, 253-268.

Clark, A., Waller, G. & Fisher, D. (1994). Sex offenders: associations between impulsiveness and locus of control. *Unpublished Doctoral Thesis*: University of Birmingham.

Clark, L., & Lewis, D. (1977). *Rape: the price of coercive sexuality*. Toronto: Woman's Educational Press.

Cleckley, H. (1976). *The mask of sanity* (5th Ed.), St Louis, MO: Mosby.

Cline, R. E. (1975). A description of self-esteem measures among educable mentally retarded children and their non-retarded peers. *Dissertation Abstracts International*, 36, 2133.

Cohen, S. (1983). A global measure of perceived stress. *Journal of Health and Social Behaviour*, 24, 385-396.

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd Ed). Hillsdale, NJ: Earlbaum.

Cole, P. & Hartley, D. (1978). The effects of reinforcement and strategy training on impulsive responding. *Child Development*, 49, 381-384.

Comrey, A. L. & Lee, H. B. (1992). *A First Course in Analysis*. New York: Plenum.

Cone, J.D. (1971) Locus of control and social desirability. *Journal of Consulting and Clinical Psychology*, 36, 449.

Conroy, J. & Bradley, V. (1985). *The Pennhurst Longitudinal Study: Report of Five Years of Research and Analysis*. Philadelphia: Temple University Developmental Disabilities Centre, Boston Human Services Institute.

Cooke, D. J. & Michie, C. (2001). Refining the construct of psychopathy: towards a hierarchical model. *Psychological Assessment*, 13, 171-188.

Cooke, D.J., Michie, C., Hart, S.D., & Clarke, D. (2004). Searching for the pan-cultural core of psychopathic personality disorder. *Personality and Individual Differences*, 39, 283-295.

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

Cooper, S., Smiley, E., Jackson, A., Finlayson, J., Allan, L., Mantry, D., & Morrison, J. (2009). Adults with intellectual disabilities: prevalence, incidence and remission of aggressive behaviour and related factors. *Journal of Intellectual Disability Research*, 33 (3), 217-232.

Coopersmith S. (1981). *Coopersmith Self-esteem Inventory*. Palo Alto: Consulting Psychologists Press.

Copsa, J., & Marshall, P. (1998). The offender group reconviction scale: a statistical reconviction score for use by Probation Officers. *Applied Statistics*, 47, 159-171.

Corbett, J.A. (1979). *Psychiatric Illness in Mental Handicap*. London: Gaskell Press.

Corulla, W.J. (1987). A psychometric investigation of the Eysenck Personality Questionnaire (revised) and its relationship to the I.7 impulsiveness questionnaire. *Personality and Individual Differences*, 8 (5), 651-658.

Costa, P. T. & McCrae, R. R. (1995) Primary traits of Eysenck's P-E-N system: three- and five-factor solutions. *Journal of Personality and Social Psychology*, 69 (2), 308-317.

Craft, M. (1959). Mental disorder in the defective: a psychiatric survey among in-patients. *American Journal of Mental Deficiency*, 63, 829-834.

Craig, L.A., Beech, A.R., & Browne, K.D. (2006). A cross validation of the Risk Matrix 2000 Sexual and Violent scales. *Journal of Interpersonal Violence*, 21, 1-22.

Craig, A.R., Franklin, J.A., & Andrews, G. (1984). A scale to measure locus of control behaviour. *British Journal of Medical Psychology*, 57, 173-180.

Craig, L. & Hutchinson, R. B. (2005). Sexual offenders with learning disabilities: risk, recidivism and treatment. *Journal of Sexual Aggression*, 11 (3), 1-16.

Craig, L.A., Stringer, I., & Moss, T. (2006). Treating sexual offenders with learning disabilities in the community: a critical review. *International Journal of Offender Therapy and Comparative Criminology*, 50 (4) 369-390.

Crandall, V.C., Katkovsky, W. & Crandall, V.J. (1965). Children's beliefs in their own control reinforcements in intellectual – academic achievement situations. *Child Development*, 36, 91-109.

Crocker, A.G., Cote, G., Toupin, J., & St-Onge, B. (2007). Rate and characteristics of men with an intellectual disability in pre-trial detention. *Journal of Intellectual and Developmental Disability*, 32 (2), 143-152.

Crocker, T. M. (1989). Assessing consumer satisfaction with mental handicap services: a comparison between different approaches. *British Journal of Mental Subnormality*, 35, 94-100.

Cromer, R.F. (1972). The learning of linguistic surface structure cues to deep structure by educationally subnormal children. *American Journal of Mental Deficiency*, 77, 346-353.

Cromer, R.F. (1974). Receptive language in the mentally retarded: processes and diagnostic distinctions. In R.L. Schiefelbusch & L.L. Lloyd (Eds.), *Language Perspectives – acquisition, retardation and intervention* (pp. 237-267). Baltimore: University Park Press.

Cronbach, L.J. (1971). Test validation. In R.L. Thorndike (Ed.), *Educational Measurement*. Washington, D.C.: American Council on Education.

Crossland, S., Burns, M., Leach, C., & Quinn, P. (2005). Needs assessment in forensic learning disability. *Medicine, Science and the Law*, 45 (2) 147-153.

Cummins, J.P., & Das, J.P. (1980). Cognitive processing, academic achievement and WISC-R performance in EMR children. *Journal of Consulting and Clinical Psychology*, 46, 777-779.

Cummins, R.A. (1993). *Comprehensive Quality of Life Scale-Intellectual Disability (COMQol-ID)* (4th edn). Victoria, Australia: Deakin University.

Cunradi, C. B., Caetano, R. C., Clark, C. L. & Schafer, J. (1999). Alcohol-related problems and intimate partner violence among White, Black and Hispanic couples in the U.S. *Alcoholism: Clinical and Experimental Research*, 23, 1492-1501.

Dagnan, D., Dennis, S. & Wood, H. (1994). A pilot study in the satisfaction of people with learning disabilities with the services they receive from a community psychology service. *British Journal of Developmental Disabilities*, 40, 38-44.

Dagnan, D., & Sandhu, S. (1999). Social comparison, self-esteem and depression in people with intellectual disabilities. *Journal of Intellectual Disability Research*, 43 (5), 372-379.

Damon, W. & Hart, D. (1982). The development of self-understanding from infancy through adolescence. *Child Development*, 53, 841-864.

Dancey, C.P., & Reidy J. (2010). *Statistics without maths for psychology*. 4th Edition. Harlow: Pearson Prentis Hall

Darke, J. L. (1990). 'Sexual aggression: achieving power through humiliation'. In W.L. Marshall, D.R. Laws, & H.E. Barbaree (Eds.), *Handbook of Sexual Assault: Issues, Theories and Treatment of the Offender*. New York: Plenum.

Daruna, J. H. & Barnes, P. A. (1993). A neurodevelopmental view of impulsivity. In W. G. McCown, J. L. Johnson, & M. B. Shure (Eds.), *The impulsive client: theory, research and treatment*. Washington DC: American Psychological Association.

Davis, M. (1983). Measuring individual differences in empathy: evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, 44, 113-126.

Day, K. (1985). Psychiatric disorder in the middle-aged and elderly mentally handicapped. *The British Journal of Psychiatry*, 147, 660-667.

Day, K. (1994). Male mentally handicapped sex offenders. *British Journal of Psychiatry*, 165, 630-639.

Deb, S., & Hunter, D (1991). Psychopathology of people with mental handicap and epilepsy. III: Personality Disorder. *The British Journal of Psychiatry*, 159, 830-834.

Dent, H. R. (1986). An experimental study of the effectiveness of different techniques of questioning mentally handicapped witnesses. *British Journal of Clinical Psychology*, 25, 13-17.

Department of Health and Home Office (1992). *Review of health social services for mentally disordered offenders and others requiring similar services* (Chair: Dr. J. Reed). London, HMSO.

Department of Health, (1990). *The care programme approach for people with a mental illness referred to the specialist psychiatric service*. HC(90)23/LASSL(90)11.DOH:London.

Dernevik, M., Grann, M., & Johansson, S. (2002). Violent behaviour in forensic psychiatric patients: risk assessment and different risk management levels using the HCR-20. *Psychology, Crime and the Law*, 8, 93-111.

- Derogatis, L.R. (1983). *SCL-90-R; administration, scoring and procedures: manual II*. Townson, MD: Clinical Psychometric Research.
- Derogatis, L.R. (1993). *Brief symptoms inventory: administration scoring and procedures manual, 3rd edition*. Minneapolis, MN: National Computer Systems.
- Dickman, S.J. (1990). Functional and dysfunctional impulsivity: personality and cognitive correlates. *Journal of Personality and Social Psychology*, 58 (1), 95-102.
- Dickson, K., Emerson, E., & Hatton, C. (2005). Self-reported anti-social behaviour: prevalence and risk factors amongst adolescents with and without intellectual disability. *Journal of Intellectual Disability Research*, 49 (11), 820-826.
- Diener, E. & Fujita, F. (1995) Resources, personal strivings and subjective well-being: a nomothetic and idiographic approach. *Journal of Personality and Social Psychology*, 68, 926-935.
- Digman, J. M. (1990). Personality structure: emergence of the five-factor model. *Annual Review of Psychology*, 41, 417-440.
- Distefano, M.K., Jr., Pryer, M.W. & Garrison, J.L. (1972). Internal–external control among alcoholics. *Journal of Clinical Psychology*, 28, 36-37.
- Dixon, D.N., McKee, C.S., & McRae, B.C. (1976). Dimensionality of three adult, objective locus of control scales. *Journal of Personality Assessment*, 40 (3), 310-319.
- Douglas, K.S., Ogloff, J.R.P., Nicholls, T.L., & Grant, I. (1999). Assessing risk for violence among psychiatric patients: the HCR-20 violence risk assessment scheme and the Psychopathy Checklist – Screening Version. *Journal of Consulting and Clinical Psychology*, 67, 917-930.
- Dragutinovich, S., White, K.D. & Austin, D.A. (1983). Factor analysis and norms for an 18-item version of Reed and Ware’s internal-external scale. *Canadian Journal of Behavioural Science*, 15, 259-265.

Ducette, J., & Wolk, S. (1972). Locus of control and extreme behaviour. *Journal of Consulting and Clinical Psychology*, 39, 253-258.

Dudek, J. A., Nezu, C. M., Nezu, A. M., Wasserman, A., Good, W., Chung, S. & McGuffin, P. (1997). *Psychological, Verbal and Emotional Abuse Questionnaire*. Unpublished questionnaire, Allegheny University of the Health Sciences, Philadelphia.

Dudley-Mailing, C. L., Snider, V., & Tarrier, S. G. (1982). Locus of control and learning disability: A review and discussion. *Perceptual and Motor Skills*, 54, 503-514.

Duke, M. & Fenhagen, E. (1975). Self-parental alienation and locus of control in delinquent girls. *Journal of Genetic Psychology*, 127, 103-107.

Dukes, E., & McGuire, B.E. (2009). Enhancing capacity to make sexually-related decisions in people with an intellectual disability. *Journal of Intellectual Disability Research*, 53 (8), 727-734.

Dunn, L.M., & Dunn, L.M. (1981). *Peabody picture vocabulary test (Rev. Ed.)*. Circle Pines, MN: American Guidance Service.

Duttweiler, P. C. (1984) The Internal Control Index: A newly developed measure of locus of control. *Educational and Psychological Measurement*, 44, 209-221.

D’Zurilla, T.J., & Nezu, A.M. (1990). Development and preliminary evaluation of the social problem-solving inventory. *Journal of Consulting and Clinical Psychology*, 2 (2), 156-163.

D’Zurilla, T.J., Nezu, A.M., & Maydeu-Olivares, A. (2002). *The Social Problem-Solving Inventory – Revised (SPSI-R): Technical Manual*. North Tonawanda, NY: MHS.

Eaton, L.F., & Menolascino, F.J. (1982). Psychiatric disorders in the mentally disordered: types, problems, and challenges. *American Journal of Psychiatry*, 139, 10, 1297-1303.

Einfeld, S. L. & Tonge, B. J. (1991). The Developmental Behaviour Checklist: the development and validation of an instrument for the assessment of behavioural and emotional disturbance in children and adolescents with mental retardation. *Journal of Autism and Developmental Disorders*, 25, 81-104.

Eitzen, D.S. (1974). Impact of behavior modification techniques on locus of control of delinquent boys. *Psychological Reports*, 35, 1317.

Elenewski, R. (1975). Runaway youth: a study of personality factors and the locus of control dimension in adolescent runaways. *Dissertations Abstracts International*, 35, 1975A-1976A (University Microfilms No 74-23, 125, 74-456, 88).

Emerson, E. (1995). *Challenging behaviour: Analysis and intervention in people with learning disabilities*. Cambridge, UK: Cambridge University Press.

Emerson, E. (1998). Assessment. In E. Emerson., C. Hatton, J, Bromley & A. Caine (eds), *Clinical Psychology and People with Intellectual Disabilities*. Chichester: Wiley.

Emerson, E. (2005). Use of the strengths and difficulties questionnaire to assess the mental health needs of children and adolescents with intellectual disabilities. *Journal of Intellectual and Development Disability*, 30, 1, 14-23.

Erber, J.T., & Dye, C.J. (1982). A profile of the newly admitted nursing home resident. *International Journal of Aging and Human Development*, 15, 307-313.

Evenden, J. L. (1999). Impulsivity: a discussion of clinical and experimental findings. *Journal of Psychopharmacology*, 13, 2, 180-192.

Everington, C. & Fulero, S. M. (1999). Competence to confess: measuring, understanding and suggestibility of defendants with mental retardation. *Mental Retardation*, 37, 212-220.

Eysenck, S. G. B. (1993). The I7: development of a measure of impulsivity and its relationship to the superfactors of personality. In McCown, W. G., Johnson, J. L. & Shure, M. B (Eds.), *The impulsive client: theory, research and treatment*. Washington DC: American Psychological Association.

Eysenck, H.J., Barrett, P., Wilson, G., & Jackson, C. (1992). Primary trait measurement of the 21 components of the P-E-N system. *European Journal of Psychological Assessment*, 8 (2), 109-117.

Eysenck, H. & Eysenck, S. (1969). *Personality Structure and Measurement*. London: Routledge.

Eysenck, H. & Eysenck, S. (1976). *Psychoticism as a dimension of personality*. London: Hodder & Stoughton.

Eysenck, S. G. B. & Eysenck, H. J. (1977). The place of impulsiveness in a dimensional system of personality description. *British Journal of Social and Clinical Psychology*, 16, 57-68.

Eysenck, S.B.G., & Eysenck, H.J. (1978). Impulsiveness and venturesomeness: their position in a dimensional system of personality description. *Psychological Reports*, 43, 1247-1255.

Eysenck, H.J., & Eysenck, M.W. (1985). *Personality and Individual Differences: a Natural Science Approach*. New York: Plenum Press.

Eysenck, H. J. & Gudjonsson, G. H. (1989). *The causes and cures of criminality*. Plenum: New York.

Eysenck, S.B.G., & McGurk, B.J. (1980). Impulsiveness and venturesomeness in a detention centre population. *Psychological Reports*, 47, 1299-1306.

Eysenck, S. B. G., Pearson, P. R., Eastings, G. & Allsopp, J. F. (1985). Age norms for impulsiveness, venturesomeness and empathy in adults. *Personality and Individual Differences*, 6, 613-619.

Eysenck, S. B. G. & Zuckerman, M. (1978). The relationship between sensation seeking and Eysenck's dimensions of personality. *British Journal of Psychology*, 69, 483-487.

Fazel, S., Xenitidis, K., & Powell, J. (2008). The prevalence of intellectual disabilities among 12,000 prisoners – a system review. *International Journal of Law and Psychiatry*, 31, 369-373.

Felton, B. & Kahana, E. (1974). Adjustment and situationally-bound locus of control among institutionalized aged. *Journal of Gerontology*, 29, 295-301.

Field, A. (2009). *Discovering statistics using SPSS (3rd Edition)*. London: Sage.

Finch, A.J., Kendall, P.C., Spirito, A., & Mikulka, P.J. (1981). Multidimensionality of the Nowicki-Strickland locus of control scale for adults. *Journal of Personality Assessment*, 45 (2), 151-154.

Finch, A.J., Saylor, C.F., & Nelson, W. (1983). The Children's Inventory of Anger: A self report inventory. Paper presented at the American Psychological Association meeting, Anaheim, CA.

Fincham, F. & Barling, J. (1978). Locus of control and generosity in learning disabled, normal achieving and gifted children. *Child Development*, 49, 530-533.

Finlay, W. M. L. & Lyons, E. (2001). Methodological issues in interviewing and using self report questionnaires with people with mental retardation. *Psychological Assessment*, 3, 319-335.

Fischer, H. L., Krajicek, M. J. & Borthick, W. A. (1973). *Sex Education for the Developmentally Disabled: A Guide for Parents, Teachers and Professionals*. Baltimore, MD: University Park Press.

Fishbein, D. (2000). Neuropsychological function, drug abuse and violence. *Criminal Justice and Behaviour*, 27 (2), 139-159.

Fishbein, D. H. & Reuland, M. (1994). Psychological correlates of frequency and type of drug use among jail inmates. *Addictive Behaviours*, 19 (6), 583-598.

Fisher, D., Beech, A. & Brown, K. (1998). Locus of control and its relationship to treatment change and abuse history in child sexual abusers. *Legal and Criminological Psychology*, 3, 1-12.

Fitts, W.M. & Hamner, W.T. (1969). *The self-concept and delinquency*. Nashville, TN: Nashville Mental Health Centre.

Fitzgerald, S., Grey, N.S., Taylor, J., & Snowden, R.J. (2011). Risk factors for recidivism in offenders with intellectual disabilities. *Psychology, Crime and Law*, 17 (1), *Special issue: intellectual disability and criminal behaviour*, 43-58.

Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology*, 32, 221- 233.

Fletcher, R., Loschen, E., Stavrakaki, C., & First, M. (Eds.). (2007). *Diagnostic Manual – Intellectual Disability (DM-ID): A Textbook of Diagnoses of Mental Disorders in Persons with Intellectual Disability*. Kingston, NY: NADD Press

Flynn, M. C. & Saleem, J. K. (1986). Adults who are mentally handicapped and living with their parents: satisfaction and perceptions regarding their lives and circumstances. *Journal of Mental Deficiency Research*, 30, 379-387.

Flynn, A., Matthews, H., & Hollins, S. (2002). Validity of the diagnosis of personality disorder in adults with learning disability and severe behavioural problems. *British Journal of Psychiatry*, 180, 543-546.

Fogarty, G.J., Bramston, P., & Cummins, R.A. (1997). Validation of the lifestress inventory for people with a mild intellectual disability. *Research in Developmental Disabilities*, 18 (6), 435-456.

Forchuk, C., Martin, M.L., & Griffiths, M. (1995). Sexual knowledge interview schedule: reliability. *Journal of Intellectual Disability Research*, 13 (1), 35-39.

Freedman, M., Leach, L., Kaplan, E., Winocur, G., Shulman, K., & Denis, D.C. (1994). *Clock drawing: a neurophysiological analysis*. New York: Oxford University Press.

Friedberg, R.D. (1982). Locus of control and self-concept in a status offender population. *Psychological Reports*, 50, 289-290.

Fryson, R. (2005). *Young people with learning disabilities who show sexually inappropriate or abusive behaviours*. Nottingham: Anne Frank Trust.

Furr, K.D., & Wilson, F.R. (1968). The California short-form test of mental maturity 1963 Edition as a screening device for educable mentally retarded programs. *Journal of School Psychology*, 7 (4), 47-49.

Gardener, D.C., Warren, S.A., & Gardener, P.L. (1977). Locus of control and law knowledge: a comparison of normal, retarded and learning disabled adolescents. *Adolescence*, 12, 103-109.

Gendreau, P., Goggin, C., & Smith, P. (2002). Is the PCL-R really the “unparalleled” measure of offender risk? A lesson in knowledge accumulation. *Criminal Justice and Behaviour*, 29, 397-426.

Gerjuoy, I. R. & Winters, J. J. (1966). Lateral preferences for identical geometric forms: retardates. *Perception and Psychophysics*, 1, 104-106.

Gilbert, P., & Allen, S. (1994). Assertiveness, submissive behaviour and social comparison. *British Journal of Clinical Psychology*, 33, 295-306.

Glenn, E., Bihm, E.M., & Lammers, W.J. (2003). Depression, anxiety, and relevant cognitions in persons with mental retardation. *Journal of Autism and Developmental Disorders*, 33 (1), 69-76.

Goldberg, L. R. (1990) An alternative “description of personality”: the big five factor structure. *Journal of Personality and Social Psychology*, 59 (6), 1216-1229.

Goldberg, L. R. (1993) The structure of phenotypic personality traits. *American Psychologist*, 48 (1), 26-34.

Goldberg, B., Gitta, M.Z. & Puddephatt, A. (1995). Personality and trait disturbances in an adult mental retardation population: significance for psychiatric management. *Journal of Intellectual Disability Research*, 39 (4), 284-294.

Gollay, E., Freedman, R., Wyngaarden, M. & Kurtz, N. (1978). *Coming back: the community experiences of deinstitutionalized mentally retarded people*. Cambridge, M: ABT Books.

Goodman, S.H., & Walters, L.K. (1987). Convergent validity of five locus of control scales. *Educational and Psychological Measurement*, 47, 743-747.

Goodman, R. (1997). The strengths and difficulties questionnaire: a research note. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 38, 581-586.

Goodman, W., Leggett, J., & Garrett, T. (2007). Locus of control in offenders and alleged offenders with learning disabilities. *British Journal of Learning Disabilities*, 35, 192-197.

Goodstein, L.D. (1979). Inmate adjustment to prison and the transition to community life. *Journal of Research in Crime and Delinquency*, 16, 246-272.

Gopnick, A., & Astington, J.W. (1988). Children’s understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child Development*, 59, 26-37.

Gore, P. S. (1962). *Individual differences in the prediction of subject compliance to experimenter bias*. Unpublished doctoral dissertation, Ohio State University.

Goss, A., & Morosko, I. E. (1970). Relations between a dimension of internal-external control and the MMPI with an alcoholic population. *Journal of Consulting and Clinical Psychology*, 34, 189-192.

Gostasson, R. (1987). Psychiatric illness among the mildly mentally retarded. *Upsala Journal of Medical Science*, 44, (Suppl.), 115-124.

Gottfredson, M. R. & Hirschi, T. (1990). *A General Theory of Crime*. Stanford CA: Stanford University Press.

Gray, N.S., Fitzgerald, S., Taylor, J., MacCulloch, M.J., & Snowden, J.R. (2007). Predicting future reconviction in offenders with intellectual disabilities: the predictive efficacy of VRAG, PCL-SV, and the HCR-20. *Psychological Assessment*, 19 (4), 474-479.

Greenwald, H. J. & Satow, Y. (1970). A short social desirability scale. *Psychological Reports*, 27, 131-135.

Griffin, D.W., & Bartholomew, K. (1994). The metaphysics of measurement: the case of adult attachment. In K. Bartholomew and D. Perlman (Eds.), *Advances in Personal Relationships, Vol.5: Attachment Processes in Adulthood* (pp. 17-52). London: Kingsley.

Groh, T., & Goldenberg, E. (1976). Locus of control within subgroups in a correctional institution. *Criminal Justice and Behaviour*, 3, 169-178.

Groth, N.A., Hobson, W.F. & Gary, T.S. (1982) 'The Child Molester: Clinical Observations'. In J. Conte & D. Shore (Eds.), *Social Work and Child Sexual Abuse*. New York: Haworth.

Gudjonsson, G.H. (1983). Suggestibility, intelligence, memory recall and personality: an experimental study. *British Journal of Psychiatry*, 142, 35-37.

Gudjonsson, G.H. (1986). The relationship between interrogative suggestibility and acquiescence: empirical findings and theoretical implications. *Personality and Individual Differences*, 7, 195-199.

Gudjonsson, G. H. & Clare, I. C. H. (1995). The relationship between confabulation and intellectual ability, memory, interrogative suggestibility and acquiescence. *Personality and Individual Differences*, 19, 333-338.

Gudjonsson, G. H. (1990). Cognitive distortions and blame attribution among paedophiles. *Sexual and Marital Therapy*, 5, 183-185.

Gudjonsson, G. H. (1992). *The psychology of interrogations, confessions and testimony*. Chichester: Wiley.

Gudjonsson, G. H. (2003). *The psychology of interrogations and confessions: A handbook*. Chichester: Wiley.

Gudjonsson, G. H. & Henry, L. (2003). Child and adult witnesses with intellectual disability: The importance of suggestibility. *Legal and Criminological Psychology*, 8, 2, 241-252.

Gudjonsson, G. H. (1988). The relationship of intelligence and memory to interrogative suggestibility: The importance of range effects. *British Journal of Clinical Psychology*, 27, 85-187.

Gudjonsson, G. H., Clare, I., Rutter, S. & Pearse, J. (1993) *Persons at risk during interviews in police custody: The identification of vulnerabilities*. Research study No 12. The Royal Commission on Criminal Justice. London: HMSO.

Gudjonsson, G.H, & Lister, S. (1984). Interrogative suggestibility and its relationship with perceptions of self-concept and control. *Journal of the Forensic Science Society*, 24, 99-110.

Gudjonsson, G.H., & Singh, K.K. (1989). The revised Gudjonsson Blame Attribution Inventory. *Personality and Individual Differences*, 10 (1), 67-70.

Guilford, J. (1954). *Psychometric Methods (2nd Edition)*. New York: McGraw-Hill.

Guilford, J. (1975). Factors and factors of personality. *Psychological Bulletin*, 82, 802-814.

Guilford, J. P., Guilford, J. S. & Zimmerman, W. S. (1978). *The Guilford-Zimmerman Temperament Survey: manual for instructions and interpretations*. San Diego, CA: Edits.

Haertzen, C. A., Martin, W. R., Ross, F. E. & Neider, M. S. (1980). Psychopathic state inventory (PSI): development of a short test for measuring psychopathic states. *International Journal of Addiction*, 15, 137-146.

Hagiliassis, N., Gulbenkoglou, H., Di Marco, M., Young., & Hudson, A. (2005). The anger management project: a group intervention for anger in people with physical and multiple disabilities. *Journal of Intellectual and Developmental Disability*, 30 (2), 86-96.

Hall, A. (1995). Using groupwork to understand arsonists. *Nursing Standard*, 9, 25-28.

Hallahan, D. P., Gajar, A. H., Cohen, S. B., & Tarver, S.G. (1978). Selective attention and locus of control in learning disabled and normal children. *Journal of Learning Disabilities*, 11, 231-236.

Hanson, R.K. (1997). The development of a brief actuarial risk scale for sexual offence recidivism. User Reports 1997-04, Department of the Solicitor General of Canada. Retrieved from www.sgc.gc.ca/publications/corrections/199704_e.pdf

Hanson, R.K., & Bussiere, M.T. (1998). Predicting relapse: a meta-analysis of sexual offender recidivism studies. *Journal of Consulting and Clinical Psychology*, 66, 348-362.

Hanson, R.K., & Thornton, D. (2000). Improving risk assessments for sex offenders: a comparison of three actuarial scales. *Law and Human Behaviour*, 24, 119-136.

Hamberger, L. K. & Hastings, J. E. (1991). Personality correlates of men who batter and non violent men: some continuities and discontinuities. *Journal of Family Violence*, 6, 131-147.

Hare, R. D. (2003). *The Hare Psychopathy Checklist – Revised (2nd ed)*. Multi-Health Systems: Toronto, ON.

Harris, G.T., Rice, M.E., & Quinsey, V.L. (1993). Violent recidivism of mentally disordered offenders: the development of a statistical prediction instrument. *Criminal Justice and Behaviour*, 20, 315-355.

Harris, G.T., Rice, M.E., Quinsey, V.L., Lalumiere, M.L., Bower, D., & Lang, C (2003). A multisite comparison of actuarial risk assessments for sex offenders. *Psychological Assessment*, 15, 413-425.

Harris, A.J.R., & Tough, S. (2004). Should actuarial risk assessments be used with sex offenders who are intellectually disabled? *Journal of Applied Research in Intellectual Disabilities*, 17, 235-241.

Harper, D. C. & Wadsworth, J. S. (1993). Grief with adults with mental retardation: preliminary findings. *Research in Developmental Disabilities*, 14, 313-330.

Hartley, S.L., & MacLean, Jr. W.E. (2006). A review of the reliability and validity of Likert-type scales for people with intellectual disability. *Journal of Intellectual Disability Research*, 50, 11, 813-827.

Hayes, S. (1993). *People with an intellectual disability and the Criminal Justice System: appearances before local courts*. Research Report 4. Sydney, NSW: Law Reform Commission.

Hayes, S. (1996). Recent research on offenders with learning disabilities. *Tizard Learning Disabilities Review*, 1, 7-15.

Hayes, S. (1997). Prevalence of intellectual disability in local courts. *Journal of Intellectual and Developmental Disability*, 22, 2, 71-85.

Hayes, S. & Craddock, G. (1992). *Simply criminal* (2nd Edition). Sydney: Federation Press.

Hayes, S., Shackell, P., Mottram, P., & Lancaster, R. (2007). The prevalence of intellectual disability in a major UK prison. *British Journal of Learning Disabilities*, 35, 162-167.

Heal, L. W. & Sigelman, C. K. (1995). Response biases in interviews of individuals with limited mental ability. *Journal of Intellectual Disability Research*, 39, 331-340.

Heider, E. (1971) Information processing and the modification of an “impulsive conceptual tempo.” *Child Development*, 42, 1276-1281.

Hemphill, J.F., Hare, R.D., & Wong, S. (1998). Psychopathy and recidivism: a review. *Legal and Criminological Psychology*, 3, 139-170.

Henry, L. A. & Gudjonsson, G. H. (1999). Eyewitness memory and suggestibility in children with learning disabilities. *American Journal on Mental Retardation*, 104, 491-508.

Henry, L. A. & Gudjonsson, G. H. (2003). Eyewitness memory, suggestibility, and repeated recall sessions in children with mild and moderate intellectual disabilities. *Law and Human Behaviour*, 27 (5), 481-505.

Henry, L.A., & Gudjonsson, G.H. (2007). Individual and developmental differences in eyewitness recall and suggestibility in children and intellectual disabilities. *Applied Cognitive Psychology*, 21, 361-381.

Heppner, P.P., & Petersen, C.H. (1982). The development and implications of a personal problem-solving inventory. *Journal of Counselling Psychology*, 29, 66-75.

Hewitt, P. L. & Flett, G. L. (1993). Perfectionism and goal orientation in impulsive and suicidal behaviour. In W. G. McCown, J. L. Johnson, & M. B. Shure (Eds.), *The impulsive client: theory, research and treatment*. Washington DC: American Psychological Association.

Heyman, W. B. (1990). The self-perception of a learning disability and its relationship to academic self-concept and self-esteem. *Journal of Learning Disabilities*, 23, 472-475.

- Hingsburger, D., Griffiths, D., & Quinsey, V. (1991). Detecting counterfeit deviance: differentiating sexual deviants from sexual inappropriateness. *Habilitation Mental Health Care Newsletter*, 10, 51-54.
- Hirschi, T. (1969) *Causes of Delinquency*. Berkeley, CA: University of California Press.
- Hobson, R. P., Ouston, J. & Lee, A. (1989). Recognition of emotion by mentally retarded adolescents and young adults. *American Journal on Mental Retardation*, 93, 434-443.
- Hodgins, S. (1992): Mental disorder, intellectual deficiency and crime: evidence from a birth cohort. *Archives of General Psychiatry*, 49, 476-483.
- Hogue, T., Lindsay, W. R., Morrissey, C. & Cooke, D. C. (2005). *Can we use the PCL-R to assess psychopathy in people with ID?* Paper presented at the British and Irish Group for the Study of Personality Disorders conference, Glasgow, UK.
- Hogue, T., Steptoe, L., Taylor, J.L., Lindsay, W.R., Mooney, P., Pinkney, L., ... O'Brien, G. (2006). A comparison of offenders with intellectual disability across three levels of security. *Criminal Behaviour and Mental Health*, 16, 13-28.
- Holder, H. B. & Kirkpatrick, S. W. (1991). Interpretation of emotion from facial expressions in children with and without learning disabilities. *Journal of Learning Disabilities*, 24, 170-177.
- Hollon, S.D., & Kendall, P.C. (1980). Cognitive self-statements in depression: development of an automatic thoughts questionnaire. *Cognitive Therapy and Research*, 4, 383-395.
- Holland, A. J. (2004). Criminal behaviour and developmental disability: an epidemiological perspective. In W. R. Lindsay, J. L. Turner & P. Sturmey (Eds), *Offenders with Developmental Disabilities*. Chichester: Wiley.

Holland, T., Clare, I. C. H. & Mukhopadhyay, T. (2002). Prevalence of criminal offending by men and women with intellectual disability and characteristics of offenders: implications for research and service development. *Journal of Intellectual Disability Research*, 46 (Suppl.1), 6-20.

Holland, S., & Persson, P. (2010). Intellectual disability in the Victorian prison system: characteristics of prisoners with an intellectual disability released from prison in 2003-2006. *Psychology, Crime and Law*, 17 (1), 25-41.

Holmes, N., Shah, A., & Wing, L. (1982). The disability assessment schedule: a brief screening device for use with the mentally retarded. *Psychological Medicine*, 12, 879-890.

Home Office Offenders Index (2000) London: Home Office: Research, Development, Statistics.

Horn, J.L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30, 179-185.

Howells, P. M., Rogers, C. & Wilcock, S. (2000). Evaluating a cognitive/behavioural approach to anger management skills to adults with learning disability. *British Journal of Learning Disabilities*, 28, 137-142.

Hubbard, R., & Allen, S.J. (1987). An empirical comparison of alternative methods for principal component extraction. *Journal of Business Research*, 15 (2), 173-190.

Hudson, W.W. (1982). *The Clinical Measurement Package: A Field Manual*. Chicago, IL: Dorsey.

Ireland, J. L. (2004). Compiling forensic risk assessment reports. *Forensic Update*, 77, 15-22.

Irving, B. & McKenzie, I. K. (1989). *Police interrogation: the effects of the Police and Criminal Evidence Act 1984*. London: The Police Foundation of Great Britain.

Irving, B. (1980). *Police interrogation. A case study of current practice*. Royal Commission on Criminal Procedure Research Study No. 2, London: HMSO.

Jacobson, J.W. (1990). Do some mental disorders occur less frequently among persons with mental retardation? *American Journal on Mental Retardation*, 94 (6), 596-602.

Jackson, D. N. (1967). *Personality Research Form Manual*. Port Huron, MI: Research Psychologist Press.

Jackson, H.F., Glass, C., & Hope, S. (1987). A functional analysis of recidivistic arson. *British Journal of Clinical Psychology*, 26, 175-185.

Jackson, T. L. & Altman, R. (1996). Self-management of aggression in an adult male with mental retardation and severe behaviour disorders. *Education and Training in Mental Retardation and Developmental Disabilities*, 31, 55-65.

Jacoub, E., & Hall, I. (2008). The sexual lives of men with mild learning disability: a qualitative study. *British Journal of Learning Disabilities*, 37, 5-11.

Jahoda, A., Pert, C., Squire, J., & Trower, P. (1998). Facing stress and conflict: a comparison of the predicted responses and self-concepts of aggressive and non-aggressive people with intellectual disability. *Journal of Intellectual Disability Research*, 42 (5), 360-369.

Jahoda, A., Markova, I. & Cattermole, M. (1988). Stigma and the self-concept of people with mild mental handicap. *Journal of Mental Deficiency Research*, 32, 10115.

Johnson, S. J. (2002). Risk assessment in people with intellectual disability: the evidence base. *Journal of Intellectual Disability Research*, 46(suppl. I), 47- 56.

Johnson, T.C. & Berry, C. (1989). Children who molest: A treatment program. *Journal of Interpersonal Violence*, 4, 185-203.

Johnson, W. L., Marlow, R. M., Corrigan, S. A. & West, J. A. (1993) Impulsive behaviour and substance abuse. In W. G. McCown, J. L. Johnson, & M. B. Shure (Eds), *The impulsive client: theory, research and treatment*. Washington DC: American Psychological Association.

Jones, R. S. P., Miller, B., Williams, H. & Goldthorp, J. (1997). Theoretical and practical issues in cognitive-behavioural approaches for people with learning disabilities: a radical behavioural perspective. In B. Stenfert Kroese, D. Dagnan. & K. Loumidis (Eds), *Cognitive Behaviour Therapy for People with Learning Disabilities*. London: Routledge.

Junger, M., West, R. & Timman, R. (2001) Crime and risky behaviour in traffic: an example of cross-situational consistency. *Journal of Research in Crime and Delinquency*, 38, 4, 439-459.

Kabzems, V. (1985) The use of self-report measures with mentally retarded individuals. *Mental Retardation and Learning Disability Bulletin*, 14, 248-252.

Kaiser, H.F. (1970). A second-generation little jiffy. *Psychometrika*, 35, 401-415.

Kandel, E., Mednick, S. A., Kirkegaard-Sorensen, L., Hutchings, B., Knop, J., Rosenberg, R. & Schulsinger, F. (1988). IQ as a protective factor for subjects at high risk for antisocial behaviour. *Journal of Consulting and Clinical Psychology*, 56, 224-226.

Kass, R.A., & Tinsley, H.E.A. (1979). Factor analysis. *Journal of Leisure Research*, 11, 120-138.

Kaufman, A.S., & Kaufman, N.L. (1990). *Kaufman Brief Intelligence Test manual*. Minnesota: American Guidance Service.

Kazdine, A.E., Matson, J.L., & Senatore, V (1983). Assessment of depression in mentally retarded adults. *American Journal of Psychiatry*, 140 (8), 1040-1043.

Kearney, M.J., & Kearney, J.F. (1983). Investigation of sex differences and consistency of factor patterns on the adult Nowicki-Strickland internal-external control scale. *Journal of Clinical Psychology*, 39, 416-421.

Keeling, J. A. & Rose, J. L. (2006). The adaptation of a cognitive-behavioural treatment programme for special needs sexual offenders. *British Journal of Learning Disabilities*, 34, 110-116.

Keeling, J.A., Rose, J.L., & Beech, A.R. (2006). An investigation into the effectiveness of a custody-based cognitive-behavioural treatment for special needs sexual offenders. *The Journal of Forensic Psychiatry and Psychology*, 17 (3), 372-392.

Keeling, J.A., Beech, A.R., & Rose, J.L. (2007a). Assessment of intellectually disabled sex offenders: the current position. *Aggression and Violent Behaviour*, 12, 229-241.

Keeling, J.A., Beech, A.R., & Rose, J.L. (2007b). A preliminary evaluation of the adaptation of four assessments for offenders with special needs. *Journal of Intellectual & Developmental Disability*, 32 (2), 62-73.

Keeling, J.A., Beech, A.R., & Rose, J.L. (2007c). Comparing sexual offender treatment efficacy: mainstream sexual offenders and sexual offenders with special needs. *Journal of Intellectual and Developmental Disability*, 32 (2), 117-124.

Keilty, J., & Connelly, G. (2001). Making a statement: an exploratory study facing women with an intellectual disability when making a statement about sexual assault to the Police. *Disability and Society*, 16 (2), 273-291.

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

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

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

Kennedy, C.H. (1993). Sexual consent and education assessment. Philadelphia, PA: Drexel University.

Kennedy, C.H. (1999). Assessing competency to consent to sexual activity in a cognitively impaired population. *Journal of Forensic and Neuropsychology*, 1, 17-33.

Khan, A., Cowan, C., & Roy, A. (1997). Personality disorders in people with learning disabilities: a community survey. *Journal of Intellectual Disability Research*, 41, 4, 324-330.

Kifer, E. (1975). Relationships between academic achievement and personality characteristics: a quasi-longitudinal study. *American Educational Research Journal*, 12, 191-210.

Kincaid, J. P., Fishburne, R. P., Rogers, R. L. & Chissom, B. S. (1975) *Derivation of new readability formula for navy enlisted personnel*. Millington, Tennessee: Navy Research Branch.

King, N., Lancaster, N., Wynne, G., Nettleton, N. & Davis, R. (1999). Cognitive behavioural anger management training for adults with mild intellectual disability. *Scandinavian Journal of Behaviour Therapy*, 28, 19-22.

Kiresuk, T. & Sherman, R. (1968) Goal attainment scaling: a general method of evaluating comprehensive mental health programmes. *Community Mental Health Journal*, 4, 443-53.

Kline, P. (1999). *The handbook of psychological testing (Second Edition)*. London: Routledge.

Klockars, A.J. & Varnum, S.W. (1974) A test of the dimensionality assumptions of Rotter's internal-external scale. *Journal of Personality Assessment*, 38, 377-383.

Knopp, F. H. & Lackey, L. B. (1987). *Sexual offenders identified as intellectually disabled: A summary of data from 40 research providers*. Orwell, VT: The Safer Society Press.

Kolton, D. J. C., Boer, A., & Boer, D. P. (2001). A revision of the Abel and Becker Cognition Scale for intellectually disabled sex offenders. *Sexual Abuse: A Journal of Research and Treatment*, 13, 217-219.

Kovacs, M. (1985). The children's depression inventory (CDI). *Psychopharmacology Bulletin*, 21, 995-998.

Krupski, A. (1980). Attention processes: research, theory, and implications for special education. In B.K. Keogh (Ed.), *Advances in Special Education* (Vol.1, pp. 101-140). Greenwich, CT: JAI Press.

Krupski, A. (1985). Variations in attention as a function of classroom task demands in learning handicapped and CA-matched non handicapped children. *Exceptional Children*, 52 (1), 52-56.

Krupski, A. (1987). Attention: the verbal phantom strikes again – a response to Samuels. *Exceptional Children*, 54 (1), 62-65.

Lally, J. (1993). Staff issues: training, support and management. In I. Fleming & B. Stenfert Kroese (Eds.), *People with Learning Disabilities and Severe Challenging Behaviour* (pp. 141-163). Manchester: Manchester University Press.

Lambrick, F. & Glaser, W. (2004). Sex offenders with an intellectual Disability. *Sexual Abuse*, 16, 4, 381-392.

Langdon, P.E., Maxted, H., Murphy, G.H., & SOTSEC-ID Group, (2007). An exploratory evaluation of the Ward and Hudson offending pathways model with sex offenders who have intellectual disability. *Journal of Intellectual and Developmental Disability*, 32 (2), 94-105.

Langdon, P. E. & Talbot, T. J. (2006). Locus of control and sex offenders with an intellectual disability. *International Journal of Offender Therapy and Comparative Criminology*, 50, 4, 391-401.

Langer, E.J. & Rodin, J. (1976). The effects of choice and enhanced personal responsibility for the aged: a field experiment in an institutional setting. *Journal of Personality and Social Psychology*, 34, 191-198.

Lawrence, E. & Winchell, J.E. (1975). Loss of control: implications for special education. *Exceptional Children*, 42, 483-490.

Lawton, M.P. & Nahemow, L. (1973). Ecology and the aging process. In Eisdorfer, C. And Lawton, M.P. (Eds) *The Psychology of Adult Development and Aging*. Washington: American Psychological Association.

Lecrubier, Y., Braconnier, A., Said, S & Payan, C. (1995). The impulsivity rating scale (IRS): preliminary results. *European Psychiatry*, 10, 331-338.

Lefcourt, H.M. (1972). Recent developments in the study of locus of control. In B.A. Maher (Ed.), *Progress in Experimental Personality Research*. New York: Academic Press.

Leggett, J., Goodman, W., & Dinani, S. (2007). People with learning disabilities' experiences of being interviewed by the Police. *British Journal of Learning Disabilities*, 35, 168-173.

Leone. C., & Burns. J. (2000). The measurement of locus of control: assessing more than meets the eye? *Journal of Psychology*, 134, 63-76.

Leung, P.W., Salili, F., & Baber, F.M. (1986). Common adolescent problems in Hong Kong: Their relationship with self-esteem, locus of control, intelligence and family environment. *Psychologia: An International Journal of Psychology in the Orient*, 29, 91-101.

Levenson, H. (1974). Activism and powerful others: distinctions within the concept of internal-external control. *Journal of Personality Assessment*, 38, 377-383.

Levenson, H. (1975). Multi-dimensional locus of control and prison inmates. *Journal of Applied Social Psychology*, 5, 342-347.

Lewis, S.K., & Lawrence-Patterson, E. (1989). Locus of control of children with learning disabilities and perceived locus of control by significant others. *Journal of Learning Disabilities*, 22 (4), 255-257.

Lewis, C., & Morrissey, C. (2010). The association between self-report and informant reports of emotional problems in a high secure intellectual disability sample. *Advances in Mental Health and Intellectual Disabilities*, 4 (2), 44-49.

Lidher, J., Martin, D.M., Jayaprakash, M.S. & Roy, A. (2005). Personality disorders in people with learning disabilities: follow-up of a community survey. *Journal of Intellectual Disability Research*, 49 (11), 845-851.

Lindbloom, G., & Faw, T. (1982). Three measures of locus of control: What do they measure? *Journal of Personality Assessment*, 46, 70-71.

Lindsay, W. R. (Ed) (2002). Offenders with intellectual disability (special issue.) *Journal of Applied Research in Intellectual Disabilities*, 15.

Lindsay, W. R., Allan, R., MacLeod, F., Smart, N. & Smith, A. H. (2003). Long-term treatment and management of violent tendencies of men with learning disabilities convicted of assault. *Mental Retardation*, 41, 47-56.

Lindsay, W.R., Allan, R., Parry, C., MacLeod, F., Cottrell, J., Overend, H., & Smith, A.H.W. (2004). Anger and aggression in people with intellectual disabilities: treatment and follow-up of consecutive referrals and a waiting list comparison. *Clinical Psychology and Psychotherapy*, 11, 255-264.

Lindsay, W. R. & Beail, N. (2004). Risk assessment: actuarial prediction and clinical judgement of offending incidents and behaviour for intellectual disability services. *Journal of Applied Research in Intellectual Disabilities*, 1, 229-234.

Lindsay, W. R., Bellshaw, E., Culross, G., Staines, C. & Michie, A. (1992). Increases in knowledge following a course of sex education for people learning disabilities. *Journal of Learning Disability Research*, 36, 531-539.

Lindsay, W.R., Carson, D., O'Brien, G., Holland, A.J., Johnston, S., Taylor, J.L., ... Price, K. (2010). The relationship between assessed risk and service security level for offenders with intellectual disability. *The Journal of Forensic Psychiatry and Psychology*, 21 (4), 537-548.

Lindsay, W.R., Elliot, S.F., & Astell, A. (2004). Predictors of sexual offence recidivism in offenders with intellectual disabilities. *Journal of Applied Research in Intellectual Disabilities*, 17, 299-305.

Lindsay, W.R., Hamilton, C., Moulton, S., Scott, S., Doyle, M., & McMurrin, M. (2009). Assessment and treatment of social problem solving in offenders with intellectual disability. *Psychology, Crime and Law*, 17 (2), 181-197.

Lindsay, W.R., Hastings, R.P., Griffith, D.M., & Hayes, S.C. (2007). Trends and challenges in forensic research on offenders with intellectual disability. *Journal of Intellectual and Developmental Disability*, 32, 55-61.

Lindsay, W.R., Hogue, T., Taylor, J.L., Mooney, P., Steptoe, L., Johnston, S., ... Smith, A.H.W. (2006). Two studies on the prevalence and validity of personality disorder in three forensic intellectual disability samples. *Journal of Forensic Psychiatry and Psychology*, 17 (3), 485-506.

Lindsay, W.R., Hogue, T., Taylor, J.L., Steptoe, L., Mooney, P., O'Brien, G. ... Smith, A.H.W. (2008). Risk assessment in offenders with intellectual disability: a comparison across three levels of security. *International Journal of Offender Therapy and Comparative Criminology*, 52 (1), 90-111.

Lindsay, W.R., & Lees, M.S. (2003). A comparison of anxiety and depression in sex offenders with intellectual disability and a control group with intellectual disability. *Sex Abuse: A Journal of Research and Treatment*, 15 (4), 339-345.

Lindsay, W.R., Marshall, I., Neilson, C., Quinn, K., & Smith, A.H.W. (1998). The treatment of men with a learning disability convicted of exhibitionism. *Research in Developmental Disabilities*, 19 (4), 295-316.

Lindsay, W. R., Michie, A., Baty, F., Smith, A. & Miller, S. (1994). The consistency of reports about feelings and emotions from people with intellectual disability. *Journal of Intellectual Disability Research*, 38, 61-66.

Lindsay, W.R., & Murphy, L., Smith, G., Murphy, D., Edwards, Z., Chittock, C., ... Young, S.J. (2004). The dynamic risk assessment and management system: an assessment for immediate risk of violence for individuals with offending and challenging behaviour. *Journal of Applied Research in Intellectual Disabilities*, 17 (4), 267-274.

Lindsay, W. R., Neilson, C. Q., Morrison, F. & Smith, A. (1998) The treatment of six men with a learning disability convicted of sex offences with children. *British Journal of Clinical Psychology*, 37, 83-98.

Lindsay, W.R., O'Brien, G., Carson, D., Holland, A.J., Taylor, J.L., Wheeler, J.R., ... Johnston, S. (2010). Pathways into services for offenders with intellectual disabilities: childhood experiences, diagnostic information, and offence variables. *Criminal Justice and Behaviour*, 37 (6), 678-694.

Lindsay, W. R., Overend, H., Allan, R., Williams, C. & Black, L. (1998). Using specific approaches for individual problems in the management of anger and aggression. *British Journal of Learning Disabilities*, 26, 44-50.

Lindsay, W.R., Rzepecka, H., & Law, J. (2007). An exploratory study into the use of the five factor model of personality with individuals with intellectual disabilities. *Clinical Psychology and Psychotherapy*, 14, 428-437.

Lindsay, W.R., & Skeine, D.D. (2007). The Beck Depression Inventory II and the Beck Anxiety Inventory in people with intellectual disabilities: factor analyses and group data. *Journal of Applied Research in Intellectual Disabilities*, 20, 401-408.

Lindsay, W. R. & Smith, A. H. W. (1998). Responses to treatment for sex offenders with intellectual disability: a comparison of men with 1- and 2- year probation sentences. *Journal of Intellectual Disability Research*, 42, 346-353.

Lindsay, W. R., Smith, A. H. W., Law, J., Quinn, K., Anderson, A., Smith, A. & Allen, R. (2004) Sexual and non-sexual offenders with intellectual and learning disabilities: a comparison of characteristics, referral patterns and outcome. *Journal of Interpersonal Violence*, 19, 875-890.

Lindsay, W.R., Steptoe, L., & Beech, A.T. (2008). The Ward and Hudson pathways model of the sexual offence process applied to offenders with intellectual disability. *Sexual Abuse: A Journal of Research and Treatment*, 20 (4), 379-392.

Lindsay, W.R., Steptoe, L., Hogue, T.E., Mooney, P., Taylor, J.L., & Morrissey, C. (2009). Structure, fit and coherence of two circumplex assessments of personality in a population with intellectual disabilities. *Journal of Intellectual Disability Research*, 53, 6, 529-537.

Lindsay, W.R., Steptoe, L., Hogue, T.E., Taylor, J.L., Mooney, P., Haut, F., ... O'Brien, G. (2007). Internal consistency and factor structure of personality disorders in a forensic intellectual disability sample. *Journal of Intellectual and Developmental Disability*, 32 (2), 134-142.

Lindsay, W.R., Taylor, J.L., Hogue, T.E., Mooney, C., Steptoe, L., & Morrissey, C. (2010). Relationship between assessed emotion, personality, personality disorder and risk in offenders with intellectual disability. *Psychiatry, Psychology and Law*, 17 (3), 385-397.

Lindsay, W.R., Whitefield, E., & Carson, D. (2007). An assessment for attitudes consistent with sexual offending for use with offenders with intellectual disabilities. *Legal and Criminological Psychology*, 12, 55-68.

Little, V. L. & Kendall, P. C. (1978). Note on locus of control and academic achievement in institutionalized juvenile delinquents. *Journal of Abnormal Child Psychology*, 6, 281-283.

Lockhart, K., Guerin, S., Shanahan, S., & Koyle, K., (2010). Expanding the test of counterfeit deviance: are sexual knowledge, experience and needs a factor in the sexualised challenging behaviour of adults with intellectual disability? *Research in Developmental Disabilities*, 31, 117-130.

Logan, G. D., Cowan, W. B. & Davis, K. A. (1984). On the ability to inhibit responses in simple and choice reaction time tasks: A model and a method. *Journal of Experimental Psychology: Human Perception and Performance*, 10, 276-291.

Logue, A. W. (1995). *Self-control*. Englewood Cliffs, NJ: Prentice-Hall.

Loper, A.B., & Reeve, R.E. (1983). Response bias on a locus of control measure by learning-disabled children. *Journal of Abnormal Child Psychology*, 11 (4), 537-548.

Lorr, M. & Youniss, R. P. (1973). An inventory of interpersonal style. *Personality Assessment*, 30, 165-173.

Loucks, N. (2007). *No One Knows: Offenders with Learning Difficulties and Learning Disabilities. Review of prevalence and associated needs*. Prison Reform Trust, March 2007.

Lovett, D. L. & Harris, M. B. (1987). Important skills for adults with mental retardation: the client's point of view. *Mental Retardation*, 25, 351-356.

Lowe, M. R. & Eldredge, K. L. (1993). The role of impulsiveness in normal and disordered eating. In W. G. McCown, J. L. Johnson, & M. B. Shure (Eds.), *The impulsive client: theory, research and treatment*. Washington DC: American Psychological Association.

Lunsky, Y., & Benson, B.A. (1997). Social support in supported living: a comparison of staff and consumer ratings. *American Journal on Mental Retardation*, 102, 280-284.

Lunsky, Y., & Benson, B.A. (2001). Association between perceived social support and strain, and positive and negative outcome for adults with mild intellectual disability. *Journal of Intellectual Disability Research*, 45, 2, 106-114.

Lunsky, Y., Frijters, J., Griffiths, D.M., Watson, S.L., & Williston, S. (2007). Sexual knowledge and attitudes of men with intellectual disability who sexually offend. *Journal of Intellectual and Developmental Disability*, 32 (2), 74-81.

Luthar, S., & Zigler, E. (1988). Motivational factors, school atmosphere, and SES: determinants of children's probability task performance. *Journal of Applied Developmental Psychology*, 9, 477-494.

Lyall, I., Holland, A. J. & Collins, S. (1995). Offending by adults with learning disabilities: identifying need in one health district. *Mental Handicap Research*, 8, 99-109.

Lyall, I., Holland, A. J., Collins, S. & Styles, P. (1995). Incidence of persons with a learning disability detained in Police custody: a needs assessment for service development. *Medicine, Science and the Law*, 35, 61-71.

Mackinnon, S., Bailey, B. & Pink, L. (2004). *Understanding learning disabilities: a video based training resource for trainers and managers to use with staff*. Brighton: Pavilion Publishing Ltd.

MacMahon, K.M.A., Jahoda, A., Espie, C.A., & Broomfield, N.M. (2006). The influence of anger-arousal level on attribution of hostile intent and problem solving capacity in an individual with a mild intellectual disability and a history of difficulties with aggression. *Journal of Applied Research in Intellectual Disabilities*, 19, 99-107.

Malloy, C., Nezu, C. M. & Nezu, A. M. (1995). *The Denial Questionnaire*. Unpublished manuscript, Allegheny University of the Health Sciences, Philadelphia.

Mamlin, N., Harris, K. R. & Case, L. P. (2001). A methodological analysis of research on locus of control and learning disabilities: rethinking a common assumption. *The Journal of Special Education*, 34, 214-225.

Mann, A.H., Jenkins, R., Cutting, J.C., & Cowen, P.J. (1991). The development and use of a standardised assessment of abnormal personality. *Psychological Medicine*, 11, 839-847.

March, P. (1992). Do photographs help adults with severe mental handicaps to make choices? *British Journal of Mental Subnormality*, 38, 122-129.

Markovitz, P. (1995). Pharmacotherapy of impulsivity, aggression and related disorders. In E. Hollander & D.J. Stein (Eds.), *Impulsivity and Aggression* (pp. 263-287). Chichester: Wiley.

Martin, F. (1975). Internal-external control among predelinquent and delinquent middle school pupils. *Psychology*, 12, 2426.

Martinez, M., Hays, J.R. & Solway, K.S. (1977). Comparative study of delinquent and non delinquent Mexican-American youths. *Psychological Reports*, 44, 215-221.

Masi, G., Brovedani, P., Mucci, M., & Favilla, L. (2002). Assessment of anxiety and depression in adolescents with mental retardation. *Child Psychiatry and Human Development*, 32, 3, 227-237

Mason, J. (1997). People on probation with a learning disability: identification, prevalence and characteristics. [Cited in Murphy, G. H. & Mason, J. (1999) People with development disabilities who offend. In N. Bouras (Ed.), *Psychiatric and Behaviour Disorders in Developmental Disabilities and Mental Retardation*, pp. 226-245. Cambridge: University Press.]

Mason, J., & Murphy, G.H. (2002a). People with intellectual disabilities on Probation: an initial study. *Journal of Community and Applied Social Psychology*, 12, 44-55.

Mason, J., & Murphy, G.H. (2002b). Intellectual disability among people on Probation: prevalence and outcome. *Journal of Intellectual Disability Research*, 46 (3), 230-238.

Matson, J. L., Helsel, W. J., Bellack, A. S. & Senatore, V. (1983). Development of a rating scale to assess social skill deficits in mentally retarded adults. *Applied Research in Mental Retardation*, 4, 399-407.

Matson, J.L., Senatore, V., & Kazdine, A.E. (1984). Psychometric properties of the psychopathology instrument for mentally retarded adults. *Applied Research in Mental Retardation*, 5, 81-89.

McAfee, J.K., Cockeram, J., & Wolfe, P.S. (2001). Police reactions to crimes involving people with mental retardation: a cross-cultural experimental study. *Education and Training in Mental Retardation and Development Disabilities*, 36 (2), 160-171.

McBrien, J., Hodgetts, A., & Gregory, J. (2003). Offending and risk behaviours in community services for people with intellectual disabilities in one local authority. *Journal of Forensic Psychiatry and Psychology*, 14, 280-297.

McBrien, J., & Murphy, G. (2006). Police and carers' views on reporting alleged offences by people with intellectual disabilities. *Psychology, Crime and Law*, 12 (2), 127-144.

McCrae, R. R. & Costa, P. T., Jr. (1994). The stability of personality: observations and evaluations. *Current Directions in Psychological Science*, 3, 173-175.

McCrae, R. R. & Costa, P. T., Jr. (1999). A five-factor theory of personality. In L. A. Pervin, & O. P. John (Eds.), *Handbook of Personality: Theory and Research*. New York: Guilford.

McEvoy, J. (1989). Investigating the concept of death in adults who are mentally handicapped. *The British Journal of Mental Subnormality*, 35, 115-121.

McGarry, M.S. (1979) An exploration of personality change as a function of skill acquisition in adult retardates. *Rehabilitation Psychology*, 26, 57-60.

McGrath, R.J., Livingstone, J.A., & Faulk, G. (2007). A structured method of assessing dynamic risk factors among sexual abusers with intellectual disabilities. *American Journal on Mental Retardation*, 112 (3), 221-229.

McNulty, C., Kissi-Deborah, R. & Newsom-Davies, I. (1995). Police involvement with clients having intellectual disabilities: a pilot study in South London. *Mental Handicap Research*, 8, 129-136.

McVilly, K. R. (1995). Interviewing people with learning disability about their residential service. *British Journal of Learning Disabilities*, 23, 138-142.

Mehrabian, A., & Epstein, N. (1972). A measure of emotional empathy. *Journal of Personality*, 40, 525-543.

Meichenbaum, D. & Goodman, J. (1971). Training impulsive children to talk to themselves: A means of developing self-control. *Journal of Abnormal Psychology*, 77, 115-126.

Mencap (2010). Communication and people with learning disabilities. Available from www.mencap.org.uk

Mercer, D.C. & Snell, M.E. (1977). *Learning theory research in mental retardation: implications for teaching*. Columbus: Merrill.

Microsoft Corporation. (1993). Microsoft Word (Version 6.0) [Computer software]. Redmond, WA: Author.

Michie, A.M., Lindsay, W.R., Martin, V., & Grieve, A. (2006). A test of counterfeit deviance: a comparison of sexual knowledge in groups of sex offenders with intellectual disability and controls. *Sexual Abuse: A Journal of Research and Treatment*, 18 (3), 271-278.

Milgram, N.A., & Milgram, R.M. (1975). Dimensions of locus of control in children. *Psychological Reports*, 37, 523-538.

Miller, D.T. (1980). Locus of control and ability to tolerate gratification delay – when it is better to be an external. *Journal of Clinical Psychology*, 36, 137-141.

Miller, R.S., & Lefcourt, H.M. (1982). The assessment of social intimacy. *Journal of Personality Assessment*, 46, 514-518.

Milne, R., Clare, I. C. H. & Bull, R. (2002). Interrogative suggestibility among witnesses with mild intellectual disabilities: the use of an adaptation of the GSS. *Journal of Applied Research in Intellectual Disabilities*, 15, 8-17.

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

Miner, M. H. (2002). Factors associated with recidivism in juveniles: an analysis of serious juvenile sex offenders. *Journal of Research in Crime and Delinquency*, 39 (4), 421-436.

Ministry of Justice, (2010). Prison population (monthly). Available from <http://www.justice.gov.uk/publications/statistics-and-data/prisons-and-probation/population>

Minskoff, E. H., Sautter, S. W., Sheldon, K. L., Steidle, E. F. & Baker, P. D. (1988). A comparison of learning disabled adults and high school students. *Learning Disabilities Research*, 3 (2), 115-123.

Minton, H.L. (1967). Power as a personality construct. In B.A. Maher (Ed.), *Progress in experimental personality research (Vol 4)*. New York: Academic Press.

Mirels, H.L. (1970). Dimensions of internal versus external control. *Journal of Consulting and Clinical Psychology*, 34, 226-228.

Mischel, W., Shoda, Y. & Rodriguez, M. L. (1989). Delay of gratification in children. *Science*, 244, 933-938.

Moffatt, C.W., Hanley-Maxwell, C., & Donnellan, A.M. (1995). Discrimination of emotion, affective perspective-taking and empathy in individuals with intellectual disability. *Education and Training in Intellectual Disability and Developmental Disabilities*, 30, 76-85.

Moffitt, T. E., Caspi, A., Dickson, N., Silva, P. & Stanton, W. (1996). Childhood onset versus adolescent onset antisocial conduct problems in males: natural history from ages 3-18 years. *Development and Psychopathology*, 8, 399-424.

Moffitt, T. E., Gabrielli, W. F., Mednick, S. A. & Schulsinger, F. (1991). Socio-economic status, IQ and delinquency. *Journal of Abnormal psychology*, 90, 152-157.

Moore, E., Adams, R., Elsworth, J. & Lewis, J. (1997). An anger management group for people with learning disability. *British Journal of Learning Disabilities*, 25, 53-57.

Moreland, J., Hendy, S., & Brown, F. (2008). The validity of a personality disorder diagnosis for people with intellectual disability. *Journal of Applied Research in Intellectual Disabilities*, 21, 219-226.

Morgenstern, M. & Klass, E. (1991). Standard intelligence tests and related assessment techniques. In J. L. Matson & J. A. Mulick (Eds), *Handbook of Mental Retardation (2nd edition)* (pp. 195-210). New York: Pergamon.

Morrissey, C. (2003). The use of the PCL-R in forensic populations with intellectual disabilities. *British Journal of Forensic Practice*, 5, 20-24.

Morrissey, C. (2006). Guidelines for using the PCL-R with offenders with intellectual disabilities. Unpublished manuscript.

Morrissey, C. (2010). Psychopathy and other personality disorders in sexual offenders with intellectual disabilities. In L.A. Craig, W.R. Lindsay & K.D. Browne (Eds.), *Assessment and Treatment of Sexual Offenders with Intellectual Disabilities: a Handbook*. (pp 153-175). Chichester: Wiley.

Morrissey, C., Cooke, D., Michie, C., Hollin, C., Hogue, T., Lindsay, W.R., & Taylor, J.L. (2009). Structural, item, and test generalizability of the Psychopathy Checklist – Revised to offenders with intellectual disabilities. *Assessment*, 17 (1), 16-29.

Morrissey, C., Hogue, T. E., Mooney, P., Allen, C., Johnson, S., Hollin, C., Lindsay, W. R. & Taylor, J. L. (2007) Predictive validity of the PCL-R in offenders with intellectual disability in a high secure hospital setting: institutional aggression. *The Journal of Forensic Psychiatry and Psychology*, 18 (1), 1-15.

Morrissey, C., Hogue, T., Mooney, P., Lindsay, W. R., Steptoe, L., Taylor, J. & Johnston, S. (2005). Applicability, reliability and validity of the Psychopathy Checklist – Revised in offenders with intellectual disabilities: Some initial findings. *International Journal of Forensic Mental Health*, 4, 207-220.

Morrissey, C., Hogue, T.E., Mooney, P., Allen, C., Johnston, S., Hollin, C., Lindsay, W.R., & Taylor, J.L. (2007a). Predictive validity of the PCL-R in offenders with intellectual disability in a high security hospital setting: institutional aggression. *Journal of Forensic Psychiatry & Psychology*, 18, 1, 1-15.

Morrissey, C., Mooney, P., Hogue, T.E., Lindsay, W.R., & Taylor, J.L. (2007b). Predictive validity of the PCL-R for offenders with intellectual disability in a high security hospital: treatment progress. *Journal of Intellectual and Developmental Disability*, 32 (2), 125-133.

Muckley, A. (1997). *Firesetting: addressing offending behaviour*. Unpublished Manual. Redcar and Cleveland Psychological Service.

Murphy, G. H. (1992). Community adjusted social integration, work and social competence for people with intellectual disabilities or mental retardation. *Current Opinion in Psychiatry*, 5, 831-835.

Murphy, G. H. & Clare, I. C. H. (1996). Analysis of motivation in people with mild learning disabilities (mental handicap) who set fires. *Psychology, Crime and Law*, 2, 153-164.

Murphy, G. H. & Clare, I. C. H. (1998). People with learning disabilities as offenders or alleged offenders in the UK criminal justice system. *Journal of the Royal Society of Medicine*, 91, 178-182.

Murphy, G. H. & Clare, I. C. H. (1998). Working with offenders or alleged offenders with intellectual disabilities. In E. Emerson, C. Hatton, J. Bromley & A. Caine (Eds) *Clinical Psychology and People with Intellectual Disabilities* (pp 154-176). Chichester: Wiley.

Murphy, G. H. & Clare, I. C. H. (1991). MEITS: a service option for people with mild mental handicap and challenging behaviour or psychiatric problems. 2. Assessment, treatment and outcome for service users and service effectiveness. *Mental Handicap Research*, 4, 180-206.

Murphy, W. D., Coleman, E. M. & Haynes, M. R. (1983). Treatment and evaluation issues with the mentally retarded sex offender. In J. G. Greer & I. R. Stuart (Eds.), *The Sexual Aggressor: Current perspective on treatment* (pp. 22-41). New York: Reinhold.

Murphy, G. H., Harnett, H. & Holland, A. J. (1995). A survey of intellectual disabilities among men on remand in prison. *Mental Handicap Research*, 8, 81-98.

Murphy, G. H. & Mason, J. (1999) People with development disabilities who offend. In N. Bouras (Ed.), *Psychiatric and Behaviour Disorders in Development Disabilities and Mental Retardation* (pp. 226-245). Cambridge: University Press.

Murphy, G., Powell, S., Guzman, A-M., & Hays, S-J. (2007). Cognitive-behavioural treatment for men with intellectual disabilities and sexually abusive behaviour: a pilot study. *Journal of Intellectual Disability Research*, 51 (11), 902-912.

Nelson, W. & Birkimer, J. (1978). Role of self-instruction and self-reinforcement in the modification of impulsivity. *Journal of Consulting and Clinical Psychology*, 46, 183.

Nezu, C.M., Nezu, A.M., & Arean, P. (1991). Assertiveness and problem-solving training for mildly mentally retarded persons with dual diagnoses. *Research in Developmental Disabilities*, 12 (4), 371-386.

Nezu, C. M., Nezu, A. M. & Dudek, J. A. (1998). A cognitive behavioural model of assessment and treatment for intellectually disabled sexual offenders. *Cognitive and Behavioural Practice*, 5, 25-64.

Nezu, C. M., Nezu, A. M., Good, W. & Saad, R. (1997). *Social problem solving tasks for persons with intellectual disabilities – adapted for sex offenders*. (PST-Sex offenders). Unpublished test and manual. Allegheny University of the Health Sciences, Philadelphia.

Novaco, R. W. (1975) *Anger control: the development and evaluation of an experimental treatment*. Lexington, MA: Heath.

Novaco, R. W. (1986). Anger as a clinical and social problem. In R. Blanchard & C. Blanchard (Eds.), *Advances in the study of aggression (vol. 2)* (pp. 1-86). New York: Academic Press.

Novaco, R. W. (1988). Novaco Provocation Inventory. In M. Hersen & A. S. Bellack (Eds.) *Dictionary of Behavioral Assessment Techniques*. New York: Pergamon Press.

Novaco, R. W. (1994). Anger as a risk factor for violence among the mentally disordered. In J. Monahan & H. J. Steadman (Eds), *Violence and Mental Disorder: developments in risk Assessment* (pp. 21-59). Chicago, IL: University of Chicago Press.

Novaco, R. W. (2003). *The Novaco Anger Scale and Provocation Inventory (NAS-PI)*. Los Angeles: Western Psychological Services.

Novaco, R. W. & Taylor, J. L. (2004). Assessment of anger and aggression in male offenders with developmental disabilities. *Psychological Assessment*, 16, 1, 42-50.

Nowicki, S. (1976). *Adult Nowicki-Strickland internal-external locus of control scale*. Test manual available from S Nowicki Jr, Department of Psychology, Emory University, Atlanta, GA 30322, USA.

- Nowicki, S., & Duke, M. P. (1974a). A locus of control scale for college as well as non-college adults. *Journal of Personality Assessment*, 38, 136-137.
- Nowicki, S., & Duke, M. P. (1974b). A preschool and primary internal-external control scale. *Developmental Psychology*, 10, 874-880.
- Nowicki, S., & Strickland, B. R. (1973). A locus of control scale for children. *Journal of Consulting and Clinical Psychology*, 40, 148-155.
- Nunn, G.C., Montgomery, J.D., & Nunn, S.J. (1986). Criterion-related validity of the Nowicki-Strickland Locus of Control Scale with academic achievement. *Psychology: A Quarterly Journal of Human Behavior*, 23, 9-11.
- O'Malley, P. M. & Bachman, J. G. (1979). Self-esteem and education: sex and cohort comparisons among high school seniors. *Journal of Personality and Social Psychology*, 37, 1153-1159.
- Oas, P. (1985). The psychological assessment of impulsivity: A review. *Journal of Psychoeducational Assessment*, 3, 141-156.
- Ollendick, D.G. & Ollendick, T.H. (1976). The interrelationship measures of locus of control, intelligence, and achievement in juvenile delinquents. *Educational & Psychological Measurement*, 36, 1111-1113.
- Page, G. L. & Scalora, M. J. (2004). The utility of locus of control for assessing juvenile amenability to treatment. *Aggression and Violent Behaviour*, 9, 523-534.
- Pallant, J. (2007). *SPSS survival manual*. Milton Keynes: Open University Press.
- Parrot, C.A. & Strongman, K.T. (1984). Locus of control and delinquency. *Adolescence*, 19, 459-471.

Parry, C.J., & Lindsay, W.R. (2003). Impulsiveness as a factor in sexual offending by people with mild intellectual disability. *Journal of Intellectual Disability Research*, 47 (6), 483-487.

Patton, J. H., Stanford, M. S. & Battatt, E. S. (1995) Factor structure of the Barratt Impulsiveness Scale. *Journal of Clinical Psychology*, 51, 768-774.

Pauhus, D.L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology*, 46 (3), 598-609.

Payne, R., & Jahoda, A. (2004). The Glasgow Social Self-Efficacy Scale – a new scale for measuring social self-efficacy in people with intellectual disability. *Clinical Psychology and Psychotherapy*, 11, 265-274.

Pearl, R., Bryan, T., & Donahue, M. (1980). Learning disabled children's attributions for success and failure. *Learning Disability Quarterly*, 3 (1), 3-9.

Perner, J., & Wimmer, H. (1985). "John thinks that Mary thinks that": attribution of second-order beliefs by five-to ten-year-old children. *Journal of Experimental Child Psychology*, 39, 437-471.

Pert, C., & Jahoda, A. (2008). Social goals and conflict strategies of individuals with mild to moderate intellectual disabilities who present problems of aggression. *Journal of Intellectual Disability Research*, 52 (5), 393-403.

Piotrowski, C., Dunn, B.R., Sherry, D., & Howell, W.L. (1983). Factor structure on the adult Nowicki-Strickland I-E scale in a college population. *Educational and Psychological Measurement*, 43, 1211-1215.

Plutchik, R. & vanPraag, H. M. (1995). The nature of impulsivity: definitions, ontology, genetics and relations to aggression. In E. Hollander, & D. J. Stein (Eds), *Impulsivity and Aggression*. New York: Wiley.

Powell, R. (2003). Psychometric properties of the Beck Depression Inventory and Zung Self Rating Depression Scale in adults with mental retardation. *Mental Retardation*, 41, 88-95.

Powell, A., & Centa, D. (1972). Adult locus of control and mental ability. *Psychological Reports*, 30, 829-830.

Powell, K. M. & Rosén, L. A. (1999). Avoidance of responsibility in conduct disordered adolescents. *Personality and Individual Difference*, 27, 327-340.

Power, M. J., Champion, L. A., & Aris, S. J. (1988). Development of a measure of social support: the significant others (SOS) scale. *British Journal of Clinical Psychology*, 27 (4), 349-358.

Powers, L., Wilson, R., Matuszewskit, J., Phillips, A., Resin, C., Schumacher, D., & Gensert, J. (1996). Facilitating adolescent self-determination: What does it take? In D. J. Sands & M. L. Wehmeyer (Eds.), *Self-determination across the life span: Independence and choice for people with disabilities* (pp. 257-284). Baltimore: Paul H Brookes.

Prentky, R. A. & Knight, R. A. (1991). Identifying critical dimensions for discriminating among rapists. *Journal of Consulting and Clinical Psychology*, 59, 643-661.

Prentky, R. A., Lee, A. S., Knight, R. A & Cerce, D. D. (1997). Recidivism rates among child molesters and rapists: A methodological analysis. *Law and Human Behaviour*, 6, 635-659.

Prentky, R., Harris, B., Fizzell, K. & Righthand, S. (2000). An actuarial procedure for assessing risk in juvenile offenders. *Sexual Abuse: A Journal of Research and Treatment*, 12, 2, 71-93.

Proctor, T., & Beail, N. (2007). Empathy and theory of mind in offenders with intellectual disability. *Journal of Intellectual and Developmental Disability*, 32 (2), 82-93.

Prosser, H. & Bromley, J. (1998). Interviewing people with intellectual disabilities. In E. Emerson, C. Hatton, J. Bromley & A. Caine (Eds), *Clinical Psychology and people with intellectual disabilities* (pp. 99-113). Chichester: Wiley.

Pugh, D.N. (1992). Prisoners and locus of control: initial studies of a specific scale. *Psychological Reports*, 70, 523-530.

Pugh, D.N. (1993). The effects of problem-solving ability and locus of control on prisoner adjustment. *International Journal of Offender Therapy and Comparative Criminology*, 37, 163-176.

Pugh, D.N. (1994). Revision and further assessments of the Prison Locus of Control Scale. *Psychological Reports*, 74, 979-986.

Pugh, D.N. (1998). Studies of validity for the Prison Control Scale. *Psychological Reports*, 82, 739-744.

Quinsey, V.L., Book, A., & Skilling, T.A. (2004). A follow-up of deinstitutionalised men with intellectual disabilities and histories of antisocial behaviour. *Journal of Applied Research in Intellectual Disabilities*, 17, 243-253.

Quinsey, V. L., Harris, G. T., Rice, M. E. & Cormier, C. (1998). *Violent offenders: Appraising and managing risk*. Washington DC: American Psychological Association.

Quinsey, V.L., Jones, G.B., Book, A.S., & Barr, K.M. (2006). The dynamic prediction of anti-social behaviour among forensic psychiatric patients: a prospective field study. *Journal of Interpersonal Violence*, 21 (12), 1539-1565.

Rachlin, H. (1995). Self-control: beyond commitment. *Behavioural and Brain Sciences*, 18, 109-159.

Rapley, M. (1995). Black swans: conversation analysis of interviews with people with learning disabilities. *Clinical Psychology Forum*, 84, 17-23.

Reed, J. (1990). Identification and description of adults with mental handicaps showing physically aggressive behaviour. *Mental Handicap Research*, 3, 126-136.

Reed, J. (1997). Understanding and assessing depression in people with learning disabilities: a cognitive behavioural approach. In B. Stenfert Kroese, D. Dagnan & K. Loumidis (Eds.), *Cognitive Behaviour Therapy for People with Learning Disabilities*. London: Routledge.

- Register, P. A. & Kihlstrom, J. F. (1998). Hypnosis and interrogative suggestibility. *Personality and Individual Difference*, 9, 549-558.
- Reich, W., Shayka, J. J. & Taibleson, C. (1992). *Diagnostic Interview for Children and Adolescents, DICA-R-C, DSM-III-R Version, Revised Version of DICA for children ages 6-12*. Unpublished manuscript, Washington University, St Louis, MO.
- Reid, A.H., Lindsay, W.R., Law, J., & Sturmey, P. (2004). The relationship of offending behaviour and personality disorder in people with developmental disabilities. In W.R. Lindsay, J.L. Taylor & P. Sturmey (Eds.), *Offenders with Developmental Disabilities* (pp 289-303). Chichester: Wiley.
- Reid, D. & Ware, E. E. (1974). Multidimensionality on internal versus external control: Addition of a third dimension and nondistinction of self versus others. *Canadian Journal of Behavioral Science*, 6, 131-142.
- Reiss, A. J. & Rhodes, A. L. (1961). The distribution of juvenile delinquency in the social class structure. *American Sociological Review*, 26, 720-732.
- Reiss, S. (1988). *Test manual for the Reiss Screen for Maladaptive Behaviour*. Columbus, OH: International Diagnostic Systems.
- Reiss, S. (1990). Prevalence of dual diagnosis in community-based day programs in the Chicago Metropolitan area. *American Journal of Mental Retardation*, 94, 578-585.
- Renwick, S. J., Black, L., Ramm, M. & Novaco, R. W. (1997). Anger treatment with forensic hospital patients. *Legal and Criminological Psychology*, 2, 103-116.
- Reynolds, W.M. (1989). *Reynolds Child Depression Scale: Professional manual*. Odessa, FL: Psychological Assessment Resources, Inc.
- Reynolds, W.M., & Baker, J.A. (1988). Assessment of depression in persons with mental retardation. *American Journal on Mental Retardation*, 93, 93-103.

Rice, M. E. & Chaplin, T. C. (1979) Social skills training for hospitalised male arsonists. *Journal of Behaviour Therapy and Experimental Psychiatry*, 10, 105-8.

Rice, M.E., & Harris, G.T. (2005). Comparing effect sizes in follow-up studies: ROC area, Cohen's D, and R. *Law and Human Behaviour*, 29, 615-620.

Rodgers, M. (1978). *Empathic ability as a function of internal-external locus of control and impulsivity-reflection*. Unpublished doctoral dissertation, University of Nebraska.

Rojahn, J., Rabold, D. E. & Schneider, F. (1995). Emotion specificity in mental retardation. *American Journal on Mental Retardation*, 99, 477-486.

Rojahn, J. & Warren, V. J. (1997). Emotion recognition as a function of social competence and depressed mood in individuals with intellectual disability. *Journal of Intellectual Disability Research*, 41, 469-475.

Rojahn, J., Warren, V.J. & Ohringer, S. (1994). A comparison of assessment methods for depression in mental retardation. *Journal of Autism and Developmental Disorders*, 24, 3, 305-313.

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

Rose, J., Jenkins, R., O'Conner, C., Jones, C. & Felce, D. (2002). A group treatment for men with learning disabilities who sexually offend or abuse. *Journal of Applied Research in Learning Disabilities*, 15, 138-150.

Rose, J., Loftus, M., Flint, B., & Carey, L. (2005). Factors associated with the efficacy of a group intervention for anger in people with intellectual disabilities. *British Journal of Clinical Psychology*, 44, 305-317.

Rose, J., O'Brien, A., & Rose, D. (2009). Group and individual cognitive behavioural interventions for anger. *Advances in Mental Health and Learning Disabilities*, 3 (4), 45-50.

Rose, J. & West, C. (1999) Assessment of anger in people with intellectual disabilities. *Journal of Applied Research in Intellectual Disabilities*, 12, 211-224.

Rose, J., West, C. & Clifford, D. (2000). Group interventions for anger in people with intellectual disabilities. *Research in developmental disabilities*, 21, 171-181.

Rosen, M., Clark, G. R. & Kivitz,, M. S. (1977) *Habilitation of the handicapped. New dimensions in programmes for the developmentally disabled*. Baltimore: University Park Press.

Rosen, M., Floor, L. & Zisfein, L. (1974). Investigating the phenomenon of acquiescence in the mentally handicapped: theoretical model, test development and normative data. *British Journal of Mental Subnormality*, 20, 58-68.

Rosenberg, M. (1979). *Conceiving the Self*. Basic Books: New York.

Rosenberg, M., Schooler, & Schoenbach, C. (1989). Self esteem and adolescent problems; modelling reciprocal effects. *American Sociological Review*, 54, 1004-1016.

Rossiter, R., Hunnisett, E. & Pulsford, M. (1998). Anger management training and people with moderate to severe learning disabilities. *British Journal of Learning Disabilities*, 26, 67-74.

Rothenberg, B. (1969). Conservation of number among four and five-year-old children: some methodological considerations. *Child Development*, 40, 382-406.

Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs: General and Applied*, 80 (1, Whole No.609).

Rotter, J. B. (1975). Some problems and misconceptions related to the construct of internal versus external control of reinforcement. *Journal of Consulting and Clinical Psychology*, 43, 56-67.

Royal College of Psychiatrists (2001). *Diagnostic criteria for learning disability (DC-LD)*. London: Gaskell.

Russell, D.W. (1996). The UCLA Loneliness Scale (Version 3): reliability, validity, and factor structure. *Journal of Personality Assessment*, 66, 20-40.

Ryan, G. (1989). 'Victim to Victimiser'. *Journal of Interpersonal Violence*, 4, 325-41.

Sex Offender Treatment Services Collaborative – Intellectual Disabilities (SOTSEC-ID), (2010). Effectiveness of group cognitive-behavioural treatment for men with intellectual disabilities at risk of sexual offending. *Journal of Applied Research in Intellectual Disabilities*, 23, 537-551.

Salter, A. C. (1988). *Treating Child Sex Offenders and Victims: a practical guide*. Newbury Park, CA: Sage.

Samuel, W, (1980). Mood and personality correlates of IQ by race and sex of subjects. *Journal of Personal and Social Psychology*, 38, 993-1004.

Saunders, D. G. (1991). Procedures for adjusting self-reports of violence for social desirability bias. *Journal of Interpersonal Violence*, 6, 336-344.

Saylor, C. F., Benson, B. A. & Einhaus, L. (1985). Evaluation of an anger management program for aggressive boys in residential treatment. *Journal of Child and Adolescent Psychotherapy*, 2, 5-15.

Schafer, J., Caetano, R. & Cunradi, C. B. (2004). A path model of risk factors for intimate partner violence among couples in the United States. *Journal of Interpersonal Violence*, 19, 2, 127-142.

Schlegel, R. P., & Crawford, C. A. (1976). Multidimensionality of internal-external locus of control: Some additional data bearing on the validity of self-control as a third dimension. *Canadian Journal of Behavioral Science*, 8, 375-387.

Schloss, P.J., Alper, S., & Jayne, D. (1994). Self-determination for persons with disabilities: choice, risk, and dignity. *Exceptional Children*, 60, 215-225.

Schuman, H. & Presser, S. (1977). Question wording as an independent variable in survey analysis. *Sociological Methods and Research*, 6, 151-170.

Schuman, H. & Presser, S. (1979). The open and closed question. *American Sociological Review*, 44, 692-712.

Scullin, M. H. & Ceci, S. J. (2001). A suggestibility scale for children. *Personality and Individual Differences*, 30, 843-856.

Seaward, S., & Rees, C. (2001). Responding to people with a learning disability who offend. *Nursing Standard*, 15 (37), 36-39.

Shelton, D. (2006). A study of young offenders with learning disabilities. *Journal of Correctional Healthcare*, 12 (1), 36-44.

Sherer, M., Maddux, J.E., Mercandate, B., Prentice-Dunn, S., Jacobs, B., & Rogers, R.W. (1982). The Self-Efficacy Scale: construction and validation. *Psychological Reports*, 51, 663-671.

Shields, I.W., & Whitehall, G.C. (1991). *The Pride in Delinquency Scale*. Paper presented at the Eastern Ontario Correctional Psychologists' Winter Conference. Burrits Rapids, Canada.

Sigelman, C. K. & Budd, E. C. (1986). Pictures as an aid in questioning mentally retarded persons. *Rehabilitation Counseling Bulletin*, 29, 173-181.

Sigelman, C. K., Budd, E.C., Spanhel, C. L. & Schoenrock, C. J. (1981a) . Asking questions of retarded persons: a comparison of yes-no and either-or formats. *Applied Research in Mental Retardation*, 2, 347-357.

Sigelman, C. K., Budd, E.C., Spanhel, C. L. & Schoenrock, C. J. (1981b). When in doubt, say yes: acquiescence in interviews with mentally retarded persons. *Mental Retardation*, 19, 53-58.

Sigelman, C. K., Budd, E. C., Winer, J. L., Schoenrock, C. J. & Martin, P. W (1982). Evaluating alternative techniques on questioning mentally retarded persons. *American Journal of Mental Deficiency*, 86, 511-518.

Sigelman, C. K. Schoenrock, C. J., Budd, E. C., Winer, J. L., Spanhel, C. L., Martin, P.W., ... Bensberg, G. J. (1981). Communicating with mentally retarded persons: asking questions and getting answers. *Research and Training Centre in Mental Retardation*. Lubbock, T.X: Texas Tech University.

Sigelman, C.K., Schroeck, C.J., Spanhel, C.L., Hromas, S.G., Winer, J.L., Budd, E.C., & Martin, P.W. (1980). Surveying mentally retarded persons: responsiveness and response validity in three samples. *American Journal of Mental Deficiency*, 84 (5), 479-486.

Sigelman, C. K., Winer, J. L. & Schoenrock, C. J. (1982). The responsiveness of mentally retarded persons to questions. *Education and Training for the Mentally Retarded*, 17, 120-124.

Simpson, M. K. & Hogg, J. (2001). Patterns of offending among people with intellectual disability: a systematic review. Part I. Methodology and prevalence data. *Journal of Intellectual Disability Research*, 45, 384-396.

Singh, N.N., Lancioni, G.E., Winton, A.S.W., Singh, A.N., Adkins, A.D., & Singh, J. (2008). Clinical and benefit-cost outcomes of teaching a mindfulness-based procedure to adult offenders with intellectual disabilities. *Behaviour modification*, 32 (5), 622-637.

Snoyman, P., & Aicken, B. (2011). Self-reported impulsivity in male offenders with low cognitive ability in New South Wales prisons. *Psychology, Crime and Law*, 17 (2), 151-164.

Soloff, P.H. (2003). Is there any drug treatment of choice for the borderline patient? *Acta Psychiatrica Scandinavica*, 89, 50-55.

Sondennaa, E., Rasmussen, K., Palmstierna, T., & Nottestad, J. (2008). The prevalence and nature of intellectual disability in Norwegian prisons. *Journal of Intellectual Disability Research*, 52 (12), 1129-1137.

Sparrow, S.S., Balla, D.A., & Cicchetti, D.V. (1984). *Vineland Adaptive Behaviour Scales Interview Edition Survey Form manual*. Circle Pines, MN: American Guidance Service.

Spielberger, C. D. (1996). *State-Trait Anger Expression Inventory professional manual*. Odessa, FL: Psychological Assessment Resources.

Spiers, L. (2007). An evaluation of the PCL-R in assessing prisoners with low intellectual functioning. *The British Journal of Forensic Practice*, 9 (1), 10-15.

Stalker, K., & Harris, P. (1998). The exercise of choice by adults with intellectual disabilities: A literature review. *Journal of Applied Research in Intellectual Disabilities*, 11, 60-76.

Stenfert Kroese, B., Cushway, D. & Hubbard, C. (1998). The conceptualization of dreams by adults with learning disabilities. *Journal of Applied Research in Intellectual Disabilities*, 11, 146-155.

Stenfert Kroese, B., Gillott, A. & Atkinson, V. (1998). Consumers with intellectual disabilities as service evaluators. *Journal of Applied Research in Intellectual Disabilities*, 11, 116-128.

Stephoe, L., Lindsay, W.R., Forrest, D., & Power, M. (2006). Quality of life and relationships in sex offenders with intellectual disability. *Journal of Intellectual and Developmental Disability*, 31 (1), 13-19.

Stephoe, L.R., Lindsay, W.R., Murphy, L., & Young, S.J. (2008). Construct validity, reliability and predictive validity of the dynamic risk assessment and management system (DRAMS) in offenders with intellectual disability. *Legal and Criminological Psychology*, 13 (2), 309-321.

- Straub, W.F. (1979). Effect of a simulation game on locus of control of institutionalised delinquent youth. *Psychological Reports*, 45, 659-662.
- Strickland, B. R. (1965). The prediction of social action from a dimension of internal external control. *Journal of Social Psychology*, 66, 353-358.
- Sudman, S. & Bradburn, N. M. (1974). *Response Effects in Surveys: A Review and Synthesis*. Chicago, IL: Aldine.
- Sugg, B. (1987). Community care: the consumer's point of view. *Community Care*, 86, 6.
- Swallow, S.W., & Kuiper, N.A. (1988). Social comparison and negative self evaluations: an application to depression. *Clinical Psychology Review*, 8, 55-76.
- Swanson, C. K. & Garwick, G. B. (1990). Treatment of low functioning sex offenders: group therapy and interagency co-ordination. *Mental Retardation*, 28, 155-161.
- Swets, J.A. (1988). Measuring the accuracy of diagnostic systems. *Science*, 240, 1285-1293.
- Talbot, J., & Riley, C. (2007). No One Knows: Offenders with Learning Difficulties and Learning Disabilities. *British Journal of Learning Disabilities*, 35, 154-161.
- Tata, P. R. & Gudjonsson, G. H. (1990). The effects of mood and verbal feedback on interrogative suggestibility. *Personality and Individual Differences*, 11, 1079-1085.
- Taylor, R. (1999). Predicting reconvictions for sexual and violent offences using the revised Offender Group Reconviction Scale. Available at <http://www.ncjrs.gov/App/Publications/abstract.aspx?ID=182569>
- Taylor, J.L. & Halstead, S. (2001). Clinical risk assessment for people with learning disabilities who offend. *British Journal of Forensic Practice*, 3, 1, 22-32.

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

Taylor, J.L., Novaco, R.W., Guinan, C. & Street, N. (2004). Development of an imaginal provocation test to evaluate treatment for anger problems in people with intellectual disabilities. *Clinical Psychology and Psychotherapy*, 11, 233-246.

Taylor, J.L., Robertson, A., Thorne, I., Belshaw, T., & Watson, A. (2006). Responses of female fire-setters with mild and borderline intellectual disabilities to a group intervention. *Journal of Applied Research in Intellectual Disabilities*, 19 (2), 179-190.

Taylor, J.L., Thorne, I., Robertson, A., & Avery, G. (2002). Evaluation of a group intervention for convicted arsonists with mild and borderline intellectual disabilities. *Criminal Behaviour and Mental Health*, 12 (4), 282-293.

Taylor, J.L., Thorne, I. & Slavkin, M.L. (2004). Treatment of fire-setting behaviour. In W. R. Lindsey, J. L. Taylor & P. Sturmey (Eds.), *Offenders with Development Disabilities*. Wiley: Chichester.

Thornton, D., Mann, R., Webster, S., Blud, L., Travers, R., Friendship, C., & Erikson, M. (2003). Distinguishing and combining risk for sexual and violent recidivism. *Annals of the New York Academy of Science*, 989, 225-235.

Torr, J. (2008). Personality disorder and offending in people with learning disabilities. *Advances in Mental Health and Learning Disabilities*, 2 (1), 4-10.

Tossebro, J. (1995). Impact of size revisited: relation of number of residents to self determination and deprivatization. *American Journal on Mental Retardation*, 100, 1, 59-67.

Tough, S.E. (2001). *Validation of two standard risk assessments (RRASOR 1997; static-99, 1999) on a sample of adult males who are developmentally disabled with significant cognitive deficits*. Unpublished Masters thesis, University of Toronto, Toronto, ON, Canada.

- Tully, B. & Cahill, D. (1984). Police interviewing of mentally handicapped persons: an experimental study. *Journal of Forensic Psychiatry*, 3, 167-170.
- Tyrer, F., McGrother, C.W., Thorp, C.F., Donaldson, M., Bhaumik, S., Watson, J.M., & Hollin, C. (2006). Physical aggression towards others in adults with learning disabilities: prevalence and associated factors. *Journal of Intellectual Disability Research*, 50, 4, 295-304.
- Valliant, P.M., & Antonowicz, D.H. (1991). Cognitive behavioural therapy and social skills training improves personality and cognition in incarcerated offenders. *Psychological Reports*, 68 (1), 27-33.
- Vanny, K.A., Levy, M.H., Greenberg, D.M., & Hayes, S.C. (2009). Mental illness and intellectual disability in Magistrates' Courts in New South Wales, Australia. *Journal of Intellectual Disability Research*, 53 (3), 289-297.
- Van Praag, H. M. (1988). Serotonergic mechanisms and suicidal behaviour. *Psychiatr Psychobiol*, 5, 1-11.
- Verri, A., Cummins, R.A., Petito, F., Vallero, E., Monteath, S., Gerosa, E., & Nappi, G. (1999). An Italian-Australian comparison of quality of life among people with intellectual disability living in the community. *Journal of Intellectual Disability Research*, 43 (6) 513-522.
- Voelker, S. L., Shore, D. L., Brown-Moore, C., Hill, L. C., Miller, L. T. & Perry, J. (1990). Validity of self-report of adaptive behaviour skills by adults with mental retardation. *Mental Retardation*, 28, 305-309.
- Walker, N. & McCabe, S. (1973). *Crime and insanity in England*. Edinburgh, UK: University Press.
- Walker, T. & Cheseldine, S. (1997). Towards outcome measurements: monitoring effectiveness of anger management and assertiveness training in a group setting. *British Journal of Learning Disabilities*, 25, 134-137.

Walkley, F. H. (1979). Internal control, powerful others and chance: A confirmation of Levenson's factor structure. *Journal of Personality Assessment*, 43, 532-535.

Watkins, M.W. (2000). Monte Carlo PCA for parallel analysis. Available at <http://www.softpedia.com/get/Others/Home-Education/Monte-Carlo-PCA-for-parallel-analysis>

Watson, D. & Bauml, E. (1967). Effects of locus of control and expectations of future control upon present performance. *Journal of Personality and Social Psychology*, 6, 333-340.

Webster, C. D., Douglas, K. S., Eaves, D. & Hart, S. D. (1997). *The HCR-20 scheme (Version 2): the assessment of risk for violence*. Burnaby, BC: Mental Health, Law and Policy Institute, Simon Fraser University.

Wechsler Adult Intelligence Scale-III (UK) (1998). The Psychological Corporation: London.

Wechsler Adult Intelligence Scale-IV (UK) (2010). The Psychological Corporation: London.

WAIS-R, Wechsler, D. (1981). The Psychological Corporation: London.

Wehmeyer, M. L. (1994b). Reliability and acquiescence in the measurement of locus of control with adolescents and adults with mental retardation. *Psychological Reports*, 75, 527-537.

Wehmeyer, M. L. (1993). Factor structure and construct validity of a locus of control scale with individuals with mental retardation. *Educational and Psychological Measurement*, 53, 1055-1066.

Wehmeyer, M. L. (1994). Perceptions of self-determination and psychological empowerment of adolescents with mental retardation. *Education and Training in Mental Retardation and Developmental Disabilities*, 29, 9-21.

Wehmeyer, M. L. (1994a). Employment status and perceptions of control of adults with cognitive and developmental disabilities. *Research in Developmental Disabilities*, 15, 119-131.

Wehmeyer, M. L. (1994b). Perceptions of self-determination and psychological empowerment of adolescents with mental retardation. *Education and Training in Mental Retardation and Developmental Disabilities*, 29, 9-21.

Wehmeyer, M. L. (1994c). Reliability and acquiescence in the measurement of locus of control with adolescents and adults with mental retardation. *Psychological Reports*, 75, 527-537.

Wehmeyer, M. L., Kelchner, K. & Richards, S. (1996). Individual and environmental factors related to the self-determination of adults with mental retardation. *Journal of Vocational Rehabilitation*, 5, 291-305.

Wehmeyer, M.L., & Palmer, S.B. (1997). Perceptions of control of students with and without cognitive disabilities. *Psychological Reports*, 81, 195-206.

Weidman, A. (1983). The compulsive adolescent substance abuse-psychological differentiation and family process. *Journal of Drug Education*, 13, 161-172.

West, D. J. & Farrington, D. P. (1973). *Who becomes delinquent?* London: Heinemann.

White, H. R., Tice, P. C., Loeber, R. & Stouthamer-Loeber, M. (2002). Illegal acts committed by adolescents under the influence of alcohol and drugs. *Journal of Research in Crime and Delinquency*, 39 (2), 131-152.

White, J. L., Moffitt, T. E., Caspi, A., Bartusch, D. J., Needles, D. J. & Stouthamer-Loeber (1994). Measuring impulsivity and examining its relationship to delinquency. *Journal of Abnormal Psychology*, 103, 192-205.

White, R., & Willner, P. (2005). Suggestibility and salience in people with intellectual disabilities: an experimental critique of the Gudjonsson Suggestibility Scale. *The Journal of Forensic Psychiatry and Psychology*, 16 (4), 638-650.

Wiehe, V. R. (1987) Empathy and locus of control in child abusers. *Journal of Social Science Research*, 9, 17-29.

Wiehe, V.R. (1987) Impulsivity, locus of control and education. *Psychological Reports*, 60, 1273-1274.

Willner, P. (2008). Clarification of the memory artefact in the assessment of suggestibility. *Journal of Intellectual Disability Research*, 52, 318-326.

Willner, P. (2011). Assessment of capacity to participate in court proceedings: a selective critique and some recommendations. *Psychology, Crime & Law*, 17, 2, 115-129.

Wimmer, H., & Perner, J. (1983). Beliefs about belief: representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13, 103-128.

Winkler, J. D., Kanouse, D. E. & Ware, J. E. Jr. (1982). Controlling for acquiescence response set in scale development. *Journal of Applied Psychology*, 67, 5, 555-561.

Winter, N., Holland, A. J. & Collins, S. (1997). Factors predisposing to suspected offending by adults with self-report learning disabilities. *Psychological Medicine*, 27, 599-607.

Wish, J., McCombs, K. F. & Edmonson, B. (1979). *Manual for the Socio-sexual Knowledge and Attitudes Test*. Stoelting Corp: Chicago.

Wolk, S., & Brandon, J. (1977). Runaway adolescents' perceptions of parents and self. *Adolescence*, 12, 175-186.

Woodcock, K.A., & Rose, J. (2007). The relationship between the recognition of facial expressions and self-reported anger in people with intellectual disabilities. *Journal of Applied Research in Intellectual Disabilities*, 20, 279-284.

World Health Organization (WHO) (1992). *The ICD-10 Classification of Mental and Behavioural Disorders*. Geneva: World Health Organisation.

Wylie, R. C. (1979). *The self-concept: A review of methodological considerations and measuring instruments*. Nebraska: University of Nebraska Press.

Wyngaarden, M. (1981). Interviewing mentally retarded persons: issues and strategies. In R. H. Bruininks, C. E. Meyers, B. B. Sigford & K. C. Lakin (Eds.), *Deinstitutionalisation and Community Adjustment of Mentally Retarded People. Monologue 4*. Washington D.C.: AAMD.

Yando, R., & Zigler, E. (1971). Outerdirectedness in the problem-solving of institutionalised and noninstitutionalised normal and retarded children. *Developmental Psychology*, 4, 277-282.

Zamble, E., & Porporino, F.J. (1988). *Coping, behaviour, and adaptation in prison inmates*. New York: Springer-Verlag

Zelniker, T., Jeffrey, W., Ault, R., & Parsons, J. (1972). Analysis and modification of search strategies of impulsive and reflective children on the Matching Familiar Figures Test. *Child Development*, 43, 321-325.

Zigler, E. (1961). Social deprivation and rigidity in the performance of feeble minded children. *Journal of Abnormal and Social Psychology*, 62, 413-421.

Zigler, E. & Balla, D. (1977). Impact of institutional experience on the behaviour and development of retarded persons. *American Journal of Mental Deficiency*, 82, 1-11.

Zigler, E., Bennett-Gates, D., Hodapp, R., & Henrick, C.C. (2002). Assessing personality traits of individuals with mental retardation. *American Journal on Mental Retardation*, 107(3), 181-193.

Zigler, E., & Burack, J.A. (1989). Personality development and dually diagnosed person. *Research in Developmental Disabilities*, 10, 225-240.

Zigmond, A. S. & Snaith, R. P. (1983). The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica*, 67, 361-370.

Zung, W. (1965). Self-rating depression scale. *Archives of General Psychiatry*, 12, 63-70.

APPENDIX 1: Research Invitation

My name is Mark Kells. You might know me or have seen me on the unit. I am a Psychologist.

Picture of the Researcher here

You are being invited to take part in a research project. It is important that you understand why the research is being done and what you will have to do. Ask me or your named nurse if there is anything that you want to know. Take your time to think about whether or not you want to take part or not. You can discuss it with me or your named nurse if you want. Even if you agree to take part you can drop out at any time. If you do drop out no one will be annoyed with you and it will not affect your care here.

What is the project about?

The study is about two things. One is the amount of control that you think you have over the things that you do and what happens to you. The other is about whether you do things without thinking about them before hand. Everyone feels differently about these things. There are no right or wrong answers. I am just interested in what you think. We will compare the answers we get from people who live here with people who live in the community.

Why have I been chosen?

You have been chosen because I am interested in what people with an intellectual disability think about these things.

What will happen to me if I take part?

I will read you some questions and you have to say if you agree with them or not. Remember there are no right or wrong answers. I am just interested in what you think. It takes 15 to 20 minutes to complete all of the questions.

Will anyone know what I have said?

No. No-one will know what you have said. Your name will not be written on the answer sheet so no-one will know what you said.

What will happen to the results of the project?

The results might be written in an academic journal, which is like a magazine for other researchers. If this happens your name will not be mentioned. You will also be able to go to a meeting on the unit to hear what I found out.

For more information please contact

Mark Kells

You can speak to me when you see me on the unit and book a meeting with me.

You can ring me on extension 2417.

Or you can ask your named nurse to book a meeting with me.

APPENDIX 2: ANSIE – Form NC

YES NO

- ___ 1. Do you believe that most problems will solve themselves if you don't fool with them?
- ___ ___ 2. Do you believe that you can stop yourself from catching a cold?
- ___ ___ 3. Are some people just born lucky?
- ___ ___ 4. Most of the time, do you feel that getting good grades means a great deal to you?
- ___ ___ 5. Are you often blamed for things that just aren't your fault?
- ___ ___ 6. Do you believe that if somebody studies hard enough, he or she can pass any subject?
- ___ ___ 7. Do you feel that most of the time it doesn't pay to try hard because things never turn out right anyway?
- ___ ___ 8. Do you feel that if things start out well in the morning that it's going to be a great day, no matter what you do?
- ___ ___ 9. Do you feel that most of the time parents listen to what their children have to say?
- ___ ___ 10. Do you believe that wishing can make good things happen?
- ___ ___ 11. When you get rejected, does it usually seem it's for no good reason at all?
- ___ ___ 12. Most of the time do you find it hard to change a friend's opinion?
- ___ ___ 13. Do you think that cheering, more than luck, helps a team to win?
- ___ ___ 14. Do you feel that it is nearly impossible to change your parents' mind about anything?
- ___ ___ 15. Do you believe that your parents should allow you to make most of your own decisions?
- ___ ___ 16. Do you feel that when you do something wrong there's very little you can do to make it right?
- ___ ___ 17. Do you believe that most people are just born good at sports?
- ___ ___ 18. Are most of the other people your age and sex stronger than you are?
- ___ ___ 19. Do you feel that one of the best ways to handle most problems is just not to think about them?
- ___ ___ 20. Do you feel that you have a lot of choice in deciding who your friends are?
- ___ ___ 21. If you find a four leaf clover, do you believe that it might bring good luck?
- ___ ___ 22. Do you often feel that whether or not you do your homework has much to do with what kinds of grades you get?
- ___ ___ 23. Do you feel that when a person your age decides to angry with you, there's little you can do to stop him or her?
- ___ ___ 24. Have you ever had a good luck charm?
- ___ ___ 25. Do you believe that whether or not people like you depends on how you act?

- ___ ___ 26. Did your parents usually help you if you ask them to?
- ___ ___ 27. Have you ever felt that when people were angry with you, it was usually for no reason at all?
- ___ ___ 28. Most of the time, do you feel that you can change what might happen tomorrow by what you do today?
- ___ ___ 29. Do you believe that when bad things are going to happen they just are going to happen no matter what you do to try to stop them?
- ___ ___ 30. Do you think that people can get their own way if they just keep trying?
- ___ ___ 31. Most of the time, do you find it useless to try to get your own way at home?
- ___ ___ 32. Do you feel that when good things happen, they happen because of hard work?
- ___ ___ 33. Do you feel that when somebody your age wants to be your enemy, there's little you can do to change matters?
- ___ ___ 34. Do you feel that it's easy to get friends to do what you want them to do?
- ___ ___ 35. Did you usually feel that you had little to say about what you got to eat at home?
- ___ ___ 36. Do you feel that when someone doesn't like you there's little you can do about it?
- ___ ___ 37. Do you usually feel that it is almost useless to try in school because most other students are just plain smarter than you are?
- ___ ___ 38. Are you the kind of person that believes that planning ahead makes things turn out better?
- ___ ___ 39. Most of the time, did you feel that you had little to say about what your family decided to do?
- ___ ___ 40. Do you think it's better to be smart than to be lucky?

APPENDIX 2: Modified ANSIE

Yes No

- 1. Will problems get sorted out on their own?
- 2. Can you stop yourself from catching a cold?
- 3. Are some people are just born lucky?
- 4. Does people saying that you're doing well mean a lot to you?
- 5. Do you often get blamed for things that aren't your fault?
- 6. If you work hard enough at something can you get it right?
- 7. Do you think it is worth trying hard if things never turn out right?
- 8. If things go well in the morning will things go well all day no matter what you do?
- 9. Most of the time do staff listen to what clients have to say?
- 10. Can wishing make good things happen?
- 11. When you get told off it is usually for no reason?
- 12. Do you find it hard to change a friend's mind?
- 13. Does cheering help a team to win more than luck?
- 14. Is it almost impossible to change staff's mind about things?
- 15. Should staff let you make most of your own decisions?
- 16. When you do something wrong is there much you can do to put it right?
- 17. Are most people just born good at sports?
- 18. Are most people stronger than you?
- 19. Is the best way to deal with problems just not to think about them?
- 20. Do you have a lot of choice in deciding who your friends are here?
- 21. If you find a four-leafed clover will it bring you luck?
- 22. Does how hard you work make a difference to how well you do?
- 23. If someone is angry with you is there anything you can do to stop them?
- 24. Do you have a good luck charm?
- 25. Does people liking you depend on how you behave?

- ___ ___ 26. Will staff usually help if you ask?
- ___ ___ 27. When people are nasty to you is it usually for no reason?
- ___ ___ 28. Can you change what will happen tomorrow, by what you do today?
- ___ ___ 29. When bad things are going to happen can you stop them?
- ___ ___ 30. Can people get what they want if they just keep trying?
- ___ ___ 31. Is there any point in trying to get your own way on the unit?
- ___ ___ 32. When good things happen is it because of hard work?
- ___ ___ 33. When someone wants to be your enemy, can you do anything to
change it?
- ___ ___ 34. Is it easy to get friends to do what you want them to do?
- ___ ___ 35. Do you have much choice in what you eat?
- ___ ___ 36. When someone doesn't like you is there much you can do about it?
- ___ ___ 37. Is there much point in trying hard if most other people are cleverer
than you?
- ___ ___ 38. Does planning ahead make things turn out better?
- ___ ___ 39. Do you have much say in what happens to you?
- ___ ___ 40. Is it better to be clever than lucky?

APPENDIX 2A: Impulsiveness (I7i)

- | | | |
|---|-----|----|
| 1. Do you often buy things on impulse? | YES | NO |
| 2. Do you generally do and say things without stopping to think? | YES | NO |
| 3. Do you often get into a jam because you do things without thinking? | YES | NO |
| 4. Are you an impulsive person? | YES | NO |
| 5. Do you usually think carefully before doing anything? | YES | NO |
| 6. Do you often do things on the spur of the moment? | YES | NO |
| 7. Do you mostly speak before thinking things out? | YES | NO |
| 8. Do you often get involved in things you later wish you could get out of? | YES | NO |
| 9. Do you get so 'carried away' by new and exciting ideas, that you never think of possible snags? | YES | NO |
| 10. Do you need to use a lot of self-control to stay out of trouble? | YES | NO |
| 11. Would you agree that almost everything enjoyable is illegal or immoral? | YES | NO |
| 12. Are you often surprised by people's reactions to what you do or say? | YES | NO |
| 13. Do you think an evening out is more successful if it is unplanned or arranged at the last moment? | YES | NO |
| 14. Do you usually work quickly, without bothering to check? | YES | NO |
| 15. Do you often change your interests? | YES | NO |
| 16. Before making up your mind, do you consider all the advantages and disadvantages? | YES | NO |
| 17. Do you prefer to 'sleep on it' before making decisions? | YES | NO |
| 18. When people shout at you, do you shout back? | YES | NO |
| 19. Do you usually make up your mind quickly? | YES | NO |

APPENDIX 2A: Impulsiveness Revised (I7i-R)

20. Do you buy things without thinking?
21. Do you say things without stopping to think?
22. Do you act now and think later?
23. Do you do things without thinking about them?
24. Do you think before you act?
25. Do you do things without stopping to think?
26. Do you speak now and think later?
27. Do you get into things you wish you could get out of?
28. Do you think about possible problems before you do something new?
29. Do you find it easy to stay out of trouble?
30. Are all the best things in life bad for you?
31. Are people shocked by things you say?
32. Is a day out better if it's planned?
33. Do you check your work?
34. Do you often change your interests?
35. Do you think before making up your mind?
36. Do you like to decide straight away?
37. When people shout at you, do you shout back?
38. Do you make up your mind quickly?

APPENDIX 3: Nowicki-Strickland Internal-External Locus of Control Directions for Administration and Scoring.

Administration.

The administration and scoring of the Nowicki-Strickland LOC life-span scales requires no special preparation other than knowing the test materials well enough to read them. However, this is especially true with children and people with an intellectual disability. The test administrator should pronounce the words clearly and slowly when he/she is reading the scale items to the participants. It is suggested that with these client groups the examiner read the items aloud to make sure that all participants understand and to keep them working at the same pace. When reading the items aloud, the examiner ought to repeat each item twice.

The scales can be administered to groups of any size or to an individual depending on the test situation. With younger children or in cases where examinee handicaps may make personal attention more important, the scales should be administered in smaller groups or individually.

The exception to the general instruction for administration of the different scales will be covered in each section describing the specified scale. However, some general comments are appropriate here. The instructions for each scale are generally the same and go as follows:

“We are trying to find out what people like you think about certain things. We want to you answer the following questions the way you feel. There are no right or wrong answers. Don’t take too much time answering any one question, but do try to answer them all”.

For those who might have difficulty understanding the task it is suggested that the examiner have a practice session on the identification and meaning of yes and no. The usual procedure is to present two questions to see if the participant understands what they are being asked to do. Generally, the participants are asked (1) Are you a man? (2) Do you have four noses? After each question the examiner to correct and explain as necessary. In most cases these directions and the additional help are sufficient for the successful completion of the scale.

It might be proper at this point to comment on a frequently asked question: “What should I do if I can answer both yes and no to a question?” The usual response to this question has been to assure the participant that this is not an unusual happening and to tell him that if it is a little more yes than no then answer yes; if it is a little more no than yes then answer no. They are urged to pick one or other of the responses and to try to answer that and all items.

Scoring.

For all the scales, the score is the total number of items answered in an externally controlled direction. The externally keyed responses are presented in tables to the end of the sections relating to each of the tests.

APPENDIX 4: Pattern/Structure Coefficients Before Rotation

Item	Factor							
	1	2	3	4	5	6	7	8
Question39	.591	.082	-.128	.181	.053	-.007	.152	-.228
Question31	.524	-.194	-.161	.130	-.058	.089	.340	-.236
Question1	.520	-.350	.161	.023	-.109	-.162	-.100	.091
Question18	.519	-.254	-.099	-.019	-.147	.297	.144	.077
Question15	.510	.054	-.288	-.021	-.117	.016	-.164	.089
Question23	.508	.325	-.039	.049	.258	.012	-.210	-.013
Question36	.481	.452	.015	.255	.024	-.209	.001	-.258
Question37	.475	.152	.142	-.345	.129	-.235	.199	-.014
Question5	.468	.031	-.008	.102	-.316	.099	-.377	-.085
Question3	.465	-.366	.037	-.247	.058	.058	-.126	.077
Question14	.308	-.219	-.149	.133	-.303	.292	-.219	.199
Question24	.153	-.690	.275	.032	.092	-.022	-.189	.077
Question10	.289	-.609	.053	.220	.398	-.305	.003	-.037
Question33	.160	.576	-.280	.039	.131	-.186	-.249	.001
Question21	.132	-.574	-.064	.110	.246	-.234	-.036	.166
Question16	.164	.437	.340	-.041	-.172	-.249	-.034	.108
Question29	.358	.420	-.151	.086	.119	-.034	-.136	.048
Question17	.332	-.374	.284	-.232	.090	.325	.235	.042
Question34	.200	.360	-.138	.003	.129	.357	.306	.249
Question11	-.046	.015	.597	.083	-.200	-.409	-.213	-.160
Question19	.423	-.012	.466	-.133	.271	-.141	.035	.203
Question25	-.127	.207	.464	.132	.198	.090	.060	-.276
Question35	.169	-.167	.449	.225	-.109	.069	.038	-.412
Question22	.275	.294	.382	-.242	.045	-.115	.064	.174
Question27	.149	.047	.381	.162	-.299	.006	-.196	.277
Question4	-.105	.212	.375	-.169	-.060	-.192	-.239	-.120
Question2	.233	.128	-.154	-.506	-.121	-.021	.143	-.233
Question32	-.096	.058	.400	.453	.062	.203	.202	-.064
Question28	.342	.140	-.122	-.443	-.175	-.039	-.093	-.370
Question20	.352	.170	-.249	.425	.168	-.130	.318	-.157
Question26	-.001	.207	.335	.400	-.129	.240	.211	-.178
Question13	.013	.267	.248	-.394	.320	.341	.029	.006
Question30	.095	.064	.255	-.272	-.501	-.030	.379	.139
Question40	.022	-.023	.213	-.277	.485	.152	-.020	-.109
Question9	.279	.086	.111	.130	-.372	.231	.013	.112
Question38	-.148	.038	.367	-.074	.181	.386	-.078	-.138
Question12	.331	.069	.135	.125	.238	.359	-.446	.069
Question8	.079	-.278	.128	-.094	-.224	-.277	.366	.026
Question6	.042	.362	.231	.151	-.063	.022	.076	.554
Question7	.106	.221	.010	.147	.323	-.254	.277	.402

APPENDIX 4: Pattern/Structure Coefficients After Rotation

Item	Factor							
	1	2	3	4	5	6	7	8
Question36	.735	-.088	.053	.171	-.069	.163	.006	-.001
Question39	.603	.133	.168	-.189	.006	.113	.115	-.109
Question20	.596	.070	-.111	-.305	-.221	.204	.044	.090
Question23	.578	.021	.160	.055	.230	-.081	-.223	.085
Question29	.515	-.152	.114	.028	.052	-.118	-.171	.149
Question33	.512	-.312	-.071	.140	-.064	-.260	-.284	.113
Question37	.416	.118	-.036	.137	.338	-.178	.379	.054
Question31	.383	.225	.198	-.381	-.031	.148	.303	-.202
Question10	.124	.842	-.130	-.098	-.048	.041	-.043	-.029
Question24	-.273	.699	.200	.060	.086	.051	-.028	-.090
Question21	-.079	.667	-.052	-.133	-.106	-.137	-.041	.062
Question1	.165	.510	.371	.121	.000	-.060	.183	-.011
Question3	.067	.426	.321	-.053	.271	-.229	.100	-.119
Question14	-.037	.103	.614	-.177	-.117	-.078	-.086	-.017
Question5	.276	.034	.578	.160	-.075	-.030	-.100	-.189
Question9	.076	-.124	.496	.001	-.038	.158	.139	.081
Question18	.141	.200	.482	-.350	.119	-.013	.219	-.059
Question27	-.063	.042	.421	.349	-.046	.152	.027	.263
Question15	.376	.049	.391	-.105	-.043	-.291	-.016	-.026
Question11	-.030	.164	-.019	.725	-.096	.259	.108	-.042
Question4	-.024	-.106	-.072	.545	.131	.035	-.011	-.061
Question16	.249	-.194	.082	.483	.050	.041	.180	.252
Question34	.227	-.311	.107	-.383	.253	.029	.050	.338
Question13	-.011	-.220	-.075	.015	.670	.011	-.043	.049
Question40	.000	.131	-.222	-.004	.555	.021	-.121	-.066
Question17	-.108	.329	.245	-.210	.476	.129	.293	-.050
Question19	.221	.367	.069	.254	.431	.037	.162	.288
Question38	-.202	-.088	.000	.065	.426	.315	-.173	-.085
Question22	.210	-.041	.063	.327	.360	-.027	.239	.268
Question32	-.035	.003	-.014	-.006	.015	.656	-.048	.184
Question26	.085	-.186	.102	.018	-.022	.637	.047	.064
Question35	.071	.212	.134	.189	.053	.546	.115	-.280
Question25	.043	-.083	-.207	.227	.248	.498	-.076	-.021
Question30	-.122	-.194	.216	.135	.036	.026	.661	.101
Question8	-.083	.235	-.041	.029	-.134	.014	.536	.007
Question12	.182	.110	.381	.035	.333	.079	-.477	.046
Question2	.203	-.191	.004	-.025	.224	-.297	.352	-.316
Question6	.026	-.193	.174	.136	.038	.066	.034	.658
Question7	.254	.089	-.241	-.066	.017	-.044	.065	.590
Question28	.309	-.151	.134	.125	.187	-.276	.216	-.465

