The psychopathy checklist youth version (PCL: YV): an investigation into its inter-rater reliability.

AND

Clinical Research Portfolio

Volume 1

(Volume 2 bound separately)

Sarah J Dickson, BSc Honours

Submitted in partial fulfilment of the requirements for the degree of

Doctorate in Clinical Psychology (DClinPsy)

Institute of Health and Wellbeing

College of Medical, Veterinary and Life Sciences

University of Glasgow

October 2014

©Sarah J Dickson, 2014

Declaration of Originality Form

This form **must** be completed and signed and submitted with all assignments.

Please complete the information below (using BLOCK CAPITALS).

Name:	**Sarah Jane Dickson**
Student Number:	**1103912d**
Course Name	**Doctorate in Clinical Psychology**
Assignment Number/Name	**Clinical Research Portfolio**

An extract from the University's Statement on Plagiarism is provided overleaf. Please read carefully THEN read and sign the declaration below.

I confirm that this assignment is my own work and that I have:

Read and understood the guidance on plagiarism in the Doctorate in Clinical Psychology Programme Handbook, including the University of Glasgow Statement on Plagiarism	✔
Clearly referenced, in both the text and the bibliography or references, **all sources** used in the work	✔
Fully referenced (including page numbers) and used inverted commas for **all text quoted** from books, journals, web etc. (Please check the section on referencing in the 'Guide to Writing Essays & Reports' appendix of the Graduate School Research Training Programme handbook.)	✔
Provided the sources for all tables, figures, data etc. that are not my own work	✔
Not made use of the work of any other student(s) past or present without acknowledgement. This includes any of my own work, that has been previously, or concurrently, submitted for assessment, either at this or any other educational institution, including school (see overleaf at 31.2)	✔
Not sought or used the services of any professional agencies to produce this work	✔
In addition, I understand that any false claim in respect of this work will result in disciplinary action in accordance with University regulations	✔

DECLARATION:

I am aware of and understand the University's policy on plagiarism and I certify that this assignment is my own work**,** except where indicated by referencing, and that I have followed the good academic practices noted above

Signature...........**Sarah Dickson**..Date.......**03/11/14**....................

Acknowledgements

TABLE OF CONTENTS

Pages

Chapter One: Systematic Review

Psychopathy and post-traumatic stress: a systematic literature review

Sarah J Dickson

Submitted in partial fulfilment of the requirements for the degree of

Doctorate in Clinical Psychology (DClinPsy)

Address for correspondence:

Sarah Dickson

Mental Health & Wellbeing

Administration Building

Gartnavel Royal Hospital

1055 Great Western Road

Glasgow

G12 0XH

Prepared in accordance with submission guidelines for The Journal of Forensic Psychiatry and Psychology (Appendix 1.1)

Psychopathy and post-traumatic stress: a systematic literature review

This review provides a synthesis and critical appraisal of the literature investigating the relationship between psychopathy and posttraumatic stress/acute stress symptoms. A secondary question addressed whether there are differential relationships between psychopathy subtypes/factors and posttraumatic stress. A comprehensive search strategy applied to MEDLINE, EMBASE, Web of Science, PsychINFO and PILOTS yielded 607 papers. Evaluation against the inclusion criteria resulted in 9 papers: 7 papers with a further 2 identified from reference lists. Studies varied from adequate to high quality, with the majority rated as adequate. There was evidence of a relationship between psychopathy and posttraumatic stress. Findings were conflicting regarding the direction of this relationship. Differential relationships were found for psychopathy factors/subtypes with posttraumatic stress. The conclusions must be interpreted with caution given the small number of studies and methodological limitations. Preliminary gender and age differences are discussed.

Keywords: psychopathy; trauma; posttraumatic stress; posttraumatic stress disorder (PTSD)

Introduction

Rationale for review

Traditional conceptualisations of psychopathy proposed that psychopaths were unable to experience deep emotions including anxiety (Cleckley, 1941)[1]. When referring to psychopathy and anxiety, Cleckley (1976) claimed "It is doubtful if in the whole of medicine any other two reactions stand out in clear contrast" (p. 259). This would imply that psychopaths are unable to experience conditions characterised by fear and negative alterations in mood including PTSD (Davidson & Foa, 1991). Furthermore, Karpman (1941, 1948) proposed there are two types of psychopaths; primary and secondary psychopaths, both characterised by antisocial and criminal behaviour but with different etiological underpinnings. Primary psychopaths were thought to have an affective deficit from birth, whereas secondary psychopaths were thought to have the capacity to experience anxiety, as a result of a stressful environment and traumatic life events. When considered at this subtype level, secondary psychopaths may be considered more vulnerable to PTSD.

Researchers have proposed that exposure to trauma plays a role in the etiology of psychopathy (Poythress et al., 2006) and some studies have found a positive association between exposure to traumatic events and psychopathy (e.g. Dembo et al., 2007; Krischer & Sevecke, 2008; Moeller & Hell, 2003). Others have hypothesised that this link may be due to the psychopath's impulsive and irresponsible behaviour predisposing them to dangerous situations (Frick et al., 1999). Given that exposure to trauma is a prerequisite for the development of PTSD, psychopaths may be at increased risk of PTSD. Individually psychopathy and PTSD have been found to be more prevalent in prison populations (Goff et al., 2007; Hare, 2003). This may potentially suggest a co-occurrence between the two. In addition to comorbidity, some have highlighted an overlap in symptomatology between these conditions, for example constricted affect and detachment from others may resemble the callous and unemotional traits associated with psychopathy (Sharf et al., 2014). Thus, it may be difficult to distinguish between these clinical presentations.

---

[1] Different definitions of psychopathy are used throughout the literature and the author notes that there are clear ethical and clinical challenges of labelling an individual 'a psychopath'. Where the term 'psychopath' is used throughout this review, this refers to individuals displaying psychopathic traits as assessed using psychological measures.

Whilst there has been a focus on the link between psychopathy and anxiety for some time, more recent research has explored the relationship between psychopathy and posttraumatic stress, either as their primary research question or as part of wider studies. There have been conflicting findings with some studies showing a positive association, some a negative association and others a differential relationship between the different factors of psychopathy. Thus the interactions between these complex conditions are not well understood. Increased knowledge of the link between these conditions may facilitate psychological and risk formulations, differential diagnosis and the development of tailored interventions. The purpose of this review is to synthesise and critically appraise the available empirical literature examining this relationship, thus informing future research.

Psychopathy

Cleckley in his monograph "The Mask of Sanity" (1941) proposed sixteen criteria which he believed defined the construct of psychopathy. These criteria can be categorised under the labels of positive psychological adjustment, behavioural pathology, impaired social relatedness and emotional unresponsiveness (Patrick, 2006). Hare later built upon Cleckley's description and developed the Psychopathy Checklist (PCL-R; Hare, 1991) in an attempt to operationalise and assess the construct of psychopathy in adults. It is generally accepted within the literature that psychopathy is a multifaceted construct comprised of interpersonal (i.e. arrogant and deceitful), affective (i.e. deficient affective experience) and behavioural (i.e. impulsive and irresponsible) features (Cooke & Michie, 2001; Hare & Neuman, 2005). There has been considerable debate regarding the inclusion of antisocial behaviour as a fourth factor, as proposed by Hare & Neuman (2005) with some arguing that antisocial behaviour is a consequence of psychopathy and not a central component (Skeem & Cooke, 2010).

It has been proposed that the primary and secondary subtypes may parallel these factors, with primary psychopaths reflecting the interpersonal and affective features and secondary psychopaths reflecting the antisocial and lifestyle features of psychopathy (Hicks et al., 2004). Consistent with this, some have found that the interpersonal and affective facets were associated with less anxiety whilst the behavioural facets were associated with heightened anxiety (Blonigen et al., 2012).

Post-Traumatic Stress Disorder (PTSD)

According to the Diagnostic and Statistical Manual of Mental Disorders (DSM-V), a diagnosis of PTSD must include exposure to a traumatic event, either directly, indirectly or as a witness. Furthermore, symptoms of intrusion, avoidance and alterations in arousal must be present in addition to persistent negative alterations in cognitions and mood (APA, 2013). Historically PTSD has been categorised as an anxiety disorder, however with the recent introduction of DSM-V, it has been categorised under "trauma and stressor-related disorders". The majority of symptoms are retained from DSM-IV and PTSD can still be considered as being characterised by fear and avoidance (Davidson & Foa, 1991). Acute Stress Disorder is characterised by similar symptoms to PTSD, however is marked by a more immediate, short term presentation (DSM-V; American Psychiatric Association [APA], 2013). As ASD and PTSD capture similar symptoms and are closely related, both classifications are considered relevant to the systematic review.

Fear conditioning

Fear conditioning is thought to play a central role in the aetiology of PTSD. This involves classical conditioning, whereby a once neutral stimulus triggers a fear response as a result of its association with a traumatic event (Foa et al., 1989). The individual then avoids this stimulus, thereby reducing their anxiety. Consequently, this avoidance is negatively reinforced, preventing extinction of the fear response (Mowrer, 1960).

Research has supported the role of fear conditioning in studies where, relative to traumatised individuals without PTSD and healthy controls, individuals with PTSD exhibit significantly greater physiological responses (e.g. increased heart rate) in response to reminders of a traumatic event (e.g. Blanchard et al., 1994; Ehlers et al., 2010). Conversely, studies have found that psychopaths exhibit lower levels of physiological responses during exposure to aversive stimuli (e.g. electric shock) relative to controls during classical conditioning (e.g. Lykken, 1957) and aversive delay conditioning paradigms (e.g. Birbaumer et al., 2005). Lykken (1957) found that psychopaths responded similarly to controls on self-report anxiety measures, suggesting an underlying fear deficit at the autonomic arousal as opposed to at a cognitive level. This is commonly referred to as the low-fear hypothesis (Lykken, 1957).

Attentional bias

An attentional bias to threat-related stimuli is also thought to perpetuate PTSD (Foa & Riggs, 1993) and studies have shown that individuals with PTSD display an attention bias towards threatening stimuli such as angry faces (Fani et al., 2012) and threatening words (Pineles et al., 2007) relative to trauma exposed controls without PTSD. This is in contrast to individuals with psychopathy who did not take longer to respond when positive or negative emotional stimuli were present, suggesting that they were not distracted by these stimuli (Mitchell et al., 2006). This is consistent with Neuman's (1997) response modulation hypothesis which claims that psychopaths are less capable of shifting their attention from one domain to another, thus are less likely to process peripheral information not central to the task at hand.

Neurobiological research

Studies have found that PTSD is associated with increased activity of the amygdala; a brain structure involved in emotional processing and fear conditioning (Shin et al., 2006). Conversely, studies have revealed reduced activity in the amygdala, amongst other structures in psychopaths relative to controls during an aversive delay conditioning task (e.g. Birbaumer et al., 2005).

Collectively the above findings suggest that psychopaths may be less vulnerable to developing posttraumatic stress symptoms.

Research questions

1) What is the relationship between psychopathy and PTSD (and acute stress) symptoms?

2) Are there differential relationships between the psychopathy subtypes or factors and PTSD symptoms or acute stress symptoms?

Methods

Search strategy

Several electronic databases were searched. These included Ovid MEDLINE ® (1946 – Apr 2014), Ovid EMBASE (1947 – Apr 2014), Web of Science (1900 – Apr 2014), PsychINFO (1991-Apr 2014) and the PILOTS database which covers the published international

literature on traumatic stress. Where possible, searches were limited to publications in English. All possible combinations of the following psychopathy and posttraumatic stress terms were included, with the truncation command (*) utilised to identify all possible endings to the specified term.

Psychopath OR psychopathy OR psychopathic OR callous* OR unemotional OR sociopath*

AND

PTSD OR post trauma* OR post-trauma* OR posttrauma* OR acute stress* OR traumatic stress OR stress reaction* OR stress disorder* OR traumatic neuros*

This yielded 607 results. Following the removal of duplicates, 496 papers were identified as potentially relevant from the electronic search. Figure 1 outlines the screening process which identified 7 papers eligible for inclusion. The reference lists of these papers were searched. This identified a further 2 eligible papers yielding a total of 9 papers.

Inclusion criteria:

1) Includes a validated measure of psychopathy.

2) Includes a measure of PTSD symptoms or acute stress symptoms.

3) Must report on the relationship between psychopathy and PTSD symptoms or acute stress symptoms.

4) Published in a peer-reviewed journal.

5) Published in English.

Exclusion criteria:

1) Studies that do not include a validated measure of psychopathy.

2) Studies that do not include a measure of posttraumatic stress or acute stress.

3) Studies not published in a peer-reviewed journal.

4) Studies not published in English.

5) Reviews, discussion articles, case studies, book chapters or qualitative studies.

Studies identified from Electronic Searches (n = 607)

Ovid MEDLINE (1946-Apr 2014) = 30
Ovid EMBASE (1947-Apr 2014) = 275
PILOTS = 53
PsychINFO (1991- Apr 2014) = 69
Web of Science (1900- Apr 2014) = 180

Excluded Duplicates (n = 111)

Titles Screened (n = 496)

Excluded by Title (n = 440)

Abstracts Screened (n = 56)

Excluded by Abstract (n = 31)
Reasons for exclusion:
No measure of psychopathy (n = 12)
No measure of post-traumatic stress (n=10)
No measure of both (n =2)
Book publication (n=1)
Discussion article (n=4)
Review article = (n=2)

Full Text Screened (n = 25)

Excluded by Full Text (n =18)
Reasons for exclusion:
No measure of psychopathy (n=3)
No measure of posttraumatic stress (n=4)
Psychopathy measure not valid (n=5)
Not available in English (n=1)
Did not report on the relationship between
psychopathy and posttraumatic stress (n=2)
Included only subdomains of PTSD (n=2)
Included only subdomains of psychopathy (n=1)

Papers identified from Electronic Search (n=7)
Papers identified from Reference Lists (n=2)
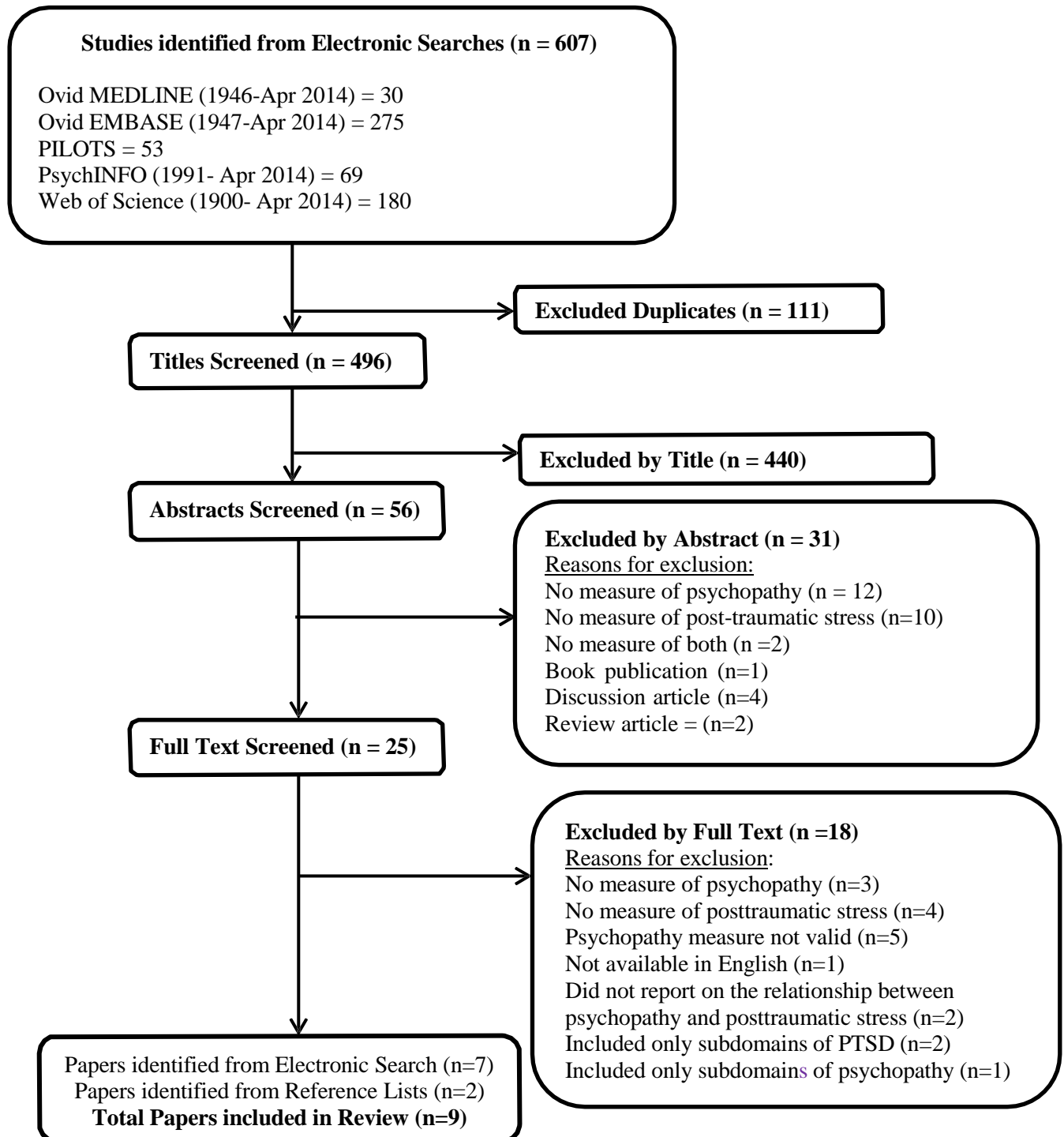**Total Papers included in Review (n=9)**

Figure 1. Flowchart of the screening process.

Quality rating criteria

A recent systematic review concluded that the majority of quality assessment tools for observational studies have not been rigorously developed, that there is a lack of consensus on what domains should be considered and there was "no single obvious choice among the most comprehensive tools we have reviewed" (Jarde et al., 2012).

In the absence of a recommended tool for observational studies, a quality assessment tool was developed for this review (Appendix 1.2). This tool was based on the tools published by DuRant (1994) and Downs and Black (1998) as they were considered among the best tools in another review (Deeks et al., 2003), and included most of the domains identified by Jarde et al., (2012) as important for assessing the methodological quality of observational studies. The tool was designed to extract the relevant data for the review questions with the aim of reviewing the quality of the evidence. There is inevitably an overlap between methodological quality and the quality of reporting and this is reflected in some of the items included in the quality assessment tool.

Some of the items were only applicable to case-control designs. Thus a score of 37 was possible for cross-sectional studies, whilst a score of 42 was possible for case-control designs. Scores were converted to percentages. For the purpose of this review, less than 50% was considered low quality, 50-60% adequate, 61-70% moderate and above 70% as high quality.

To determine inter-rater reliability, 6 of the papers were rated by an independent reviewer (DM) who was blinded to the ratings provided by the principal assessor. The overall level of agreement was 86%. Inconsistencies were resolved via discussion, increasing the level of agreement to 97%.

Results

Table 1 provides a summary of the papers reviewed. Based on the information provided, six of the papers included adults between 17 and 73 years (Blackburn et al., 2003; Blonigen et al., 2012; Hicks et al., 2010; Moeller & Hell, 2003; Pham, 2012; and Willemsen et al., 2012). Two of the papers included youths between 9 and 18 years (Kubak & Salekin, 2009; Salekin et al., 2004) and one included individuals between 14 and 21 years (Myers et al., 2012).

Several populations were studied including male forensic psychiatric patients (Blackburn et al., 2003; Pham, 2012), male prisoners (Moeller & Hell, 2003; Willemsen, 2012), female prisoners (Blonigen et al., 2012; Hicks et al., 2010) and juvenile offenders (Kubak & Salekin, 2009; Myers et al., 2012; Salekin et al., 2004).

Five of the papers addressed the primary review question as the main focus of their research (Blonigen et al., 2012; Kubak & Salekin, 2009; Moeller & Hell, 2003; Pham, 2012 and Willemsen et al., 2012). The remaining studies examined the relationship between psychopathy and general psychopathology and one addressed an unrelated question (Myers et al., 2012). This study was included as psychopathy and PTSD were assessed; therefore the study met the inclusion criteria. Four of the papers addressed the secondary review question (Blonigen et al., 2012; Hicks et al., 2010; Moeller & Hell, 2003; Willemsen et al., 2012).

The papers were varied in quality with one considered high quality, three considered moderate and five considered adequate (Table 2).

Critical appraisal

Studies rated high quality

Blackburn et al., (2003) investigated the overlap between DSM-III Axis I (Mental Disorders) and Axis II (Personality Disorders) including psychopathy in male „mentally disordered offenders" from high-security hospitals. This study is considered cross-sectional with regards to the review questions as groups were formed based on legal classifications and were not distinguishable on the basis of psychopathy as rated using the PCL-R. The CIDI was used to assess for diagnoses including PTSD. The study found that those scoring above 25 on the PCL-R were 2.65 times more likely to be diagnosed with PTSD relative to those scoring below 25. After base rates were accounted for, psychopathy was only significantly associated with PTSD and drug abuse. Furthermore, PTSD was related to personality disorders which are considered more strongly linked with violence (Blackburn & Coid, 1998). This study is commended for its random and systematic sampling, use of clinician measures of psychopathy and PTSD by trained individuals, its high inter-rater reliability for psychopathy assessments and the structured assessment of potential confounding variables including personality disorders. Limitations include insufficient detail regarding the administration of the CIDI, absence of analyses of psychopathy factors and the sample bias

towards more stabilized, non-psychotic patients. The exclusion of females and learning disabled individuals means the findings can only be generalised to these populations with caution.

Studies rated moderate quality

Blonigen et al., (2012) and Hicks et al., (2010) appear to have overlapping samples. Whilst this is not explicitly stated, Blonigen et al., (2012) reported that they expanded on the study by Hicks et al., (2010). Blonigen et al, (2012) included 226 female inmates from a Federal Correctional Institution in Florida recruited via random sampling. Hicks et al., (2010) reported that participants (n=140) were from a larger sample (n=226) of inmate *volunteers* with the same location and identical demographic information reported. Thus, the sampling is ambiguous and it is unclear whether the assessments were administered on more than one occasion in which case practice effects may introduce bias. The similarities and distinguishing features of the studies are discussed separately.

Both studies used the PCL-R and PCL-C to assess psychopathy and PTSD respectively. Identical inter- rater reliability was obtained for the PCL-R, suggesting that psychopathy was assessed once for the purpose of both studies. Strengths of these studies include their detailed exploration of psychopathy at the factor and facet level, the administration of the PCL-R with high-inter rater reliability and assessment of potentially confounding variables including trauma. However, the measure of trauma included only abuse and direct experiences, thus may not be considered a comprehensive measure of trauma. Limitations include the use of a self-report measure to assess PTSD. Furthermore, as noted by Blonigen et al., (2012), the PCL-C does not require symptoms to be linked to a specific traumatic event, thus it may be tapping into related conditions (e.g. depression). It also assesses symptoms over the past month and the   prisoners may not have had the same exposure to traumatic experiences in this time given that they were incarcerated.

Blonigen et al., (2012) investigated the cross-sectional relationship between psychopathy, PTSD and Borderline Personality Disorder (BPD). They assessed BPD given its high comorbidity with PTSD (Pagura et al., 2010) and high prevalence in incarcerated females (Warren et al., 2002). Those with higher psychopathy scores were found to have more PTSD symptoms. This was due to the moderate association between factor 2 scores (lifestyle and antisocial psychopathy traits) and PTSD, with the antisocial traits uniquely associated with

PTSD. However, the link between the antisocial traits and PTSD lost significance when BPD was accounted for. Thus, based on the measures used, BPD explained this relationship. Conversely factor 1, including the interpersonal and affective traits was unrelated to PTSD. This study highlights the differential relationship between psychopathy factors in females, with factor 2 more closely linked to PTSD. Strengths of this study include the random sampling, increasing the samples representativeness in terms of the correctional institution.

Hicks et al., (2010) investigated whether psychopathy subtypes would be found in female prisoners and whether these subtypes would differ on variables including PTSD. Using a case-control design, inmates were divided into a psychopathy (n=70) and control group (n=70) based on PCL-R scores > 25 and ≤ 17 respectively. Cluster analysis was used to divide the psychopathy group into primary and secondary psychopaths based on scores on a well-validated, self-report measure of personality. The secondary psychopaths had significantly higher PCL-R factor 2 scores due to significantly higher scores on the antisocial facet relative to primary psychopaths. Furthermore, secondary psychopaths had significantly more PTSD symptoms relative to primary psychopaths and controls, whilst the primary psychopaths did not differ significantly from controls in terms of PTSD. Thus, the link with PTSD may be due to factor 2 and in particular the antisocial psychopathy traits.

This study highlights the heterogeneous nature of psychopathy, with primary psychopaths considered psychologically resilient and secondary psychopaths less so, thus requiring more mental health care. These subtypes appear to parallel factors 1 and 2 of the PCL-R and may be indicative of different causal pathways. Strengths of this study include the use of separate PCL-R cut-offs to create distinct psychopathy and non-psychopathy groups from the same population, with baseline group comparisons conducted. Thus any differences can more confidently be attributed to psychopathy. Limitations include the voluntary sampling and lack of detail regarding the final sample, precluding an evaluation of the samples representativeness.

Willemsen et al., (2012) explored the cross-sectional relationship between psychopathy, exposure to trauma and posttraumatic stress. Male prisoners were assessed for psychopathy and DSM-IV Axis 1 disorders including PTSD using the PCL-R and SCID-1. This study revealed that the more highly an individual scored for psychopathy, including the interpersonal and affective traits, the less posttraumatic stress was experienced. However, the lifestyle and antisocial traits were unrelated to posttraumatic stress. Where high levels of the affective deficit were present, the impact of the versatility of traumatic events was reduced.

Thus, the affective traits moderated the link between traumatic exposure and posttraumatic stress. The authors conclude that these affective traits may protect against posttraumatic stress and are marked by reduced fear conditioning. Strengths of this study include its detailed theoretical underpinnings and exploration of psychopathy, consideration of potentially confounding variables including the number and versatility of traumatic events and the use of clinician rated measures for psychopathy and posttraumatic stress, with high inter-rater reliability obtained for both. However, only traumatic events from adulthood were considered and therefore may be viewed as a less comprehensive assessment of trauma. Limitations include the self-report measure of posttraumatic stress, the voluntary sample and the limitations to these approaches as discussed above. The all-male prison sample and exclusion of psychotic prisoners limits the ability to generalize the findings to these populations.

Studies rated adequate quality

Kubak & Salekin (2009) and Salekin et al., (2004) appear to include overlapping samples. Although this is not explicitly stated, the same location, demographic information and measures were used to assess psychopathy and PTSD. However, the findings for the relationship between these variables are not identical. It is unclear whether these measures were administered on one occasion or repeatedly in which case practice effects may introduce bias. These studies are evaluated collectively then independently.

Both studies included youth offenders from a court evaluation unit. The PCL: YV, APSD and SRP-II were administered to assess psychopathy and the APS to assess psychopathology including PTSD. Whilst, Salekin et al., (2004) modified the SRP-II to make it "developmentally appropriate", there is no evidence of this by Kubak & Salekin (2009). However, no details were provided regarding the modifications and the measure has not been validated in youths, thus overall this remains a limitation. Collective strengths include the combination of clinician and self-report measures of psychopathy thereby increasing the reliability of this assessment, the focus on youths and inclusion of females, thus addressing gaps in the research. Limitations included insufficient detail regarding recruitment methods, inclusion criteria, and administration of the PCL: YV (e.g. whether or not the rater received training), the use a self-report measure for PTSD in isolation and failure to measure potential confounding variables (e.g. traumatic experiences). Furthermore, neither study explores the

link between PTSD and psychopathy at the factor level. Thus the second review question was not addressed.

Salekin et al., (2004) investigated the validity of youth psychopathy, including its link with psychopathology. They found that higher psychopathy scores, assessed using the APSD were associated with more PTSD symptoms. Conversely, the relationship between psychopathy, measured using the PCL: YV, SRP-II and PTSD were not significant. Preliminary analyses on gender revealed that associations between psychopathy and other measures were stronger for males than females. However, these analyses were not reported as similar patterns were evident and there were an insufficient number of females. Thus, it is unclear whether this referred to the psychopathy – PTSD relationship. Overall there was high comorbidity between psychopathy and other conditions and the authors concluded that youths developing psychopathy may also present with internalizing psychopathology. Strengths of the study include the high inter-rater reliability for the PCL: YV and the measurement of potentially confounding variables including Disruptive Behaviour Disorders which were also associated with PTSD.

Kubak and Salekin (2009) explored the relationship between psychopathy and anxiety with a particular interest in PTSD. They found that higher levels of psychopathy, measured using the PCL: YV and APSD were associated with higher levels of PTSD. The relationship between psychopathy (as assessed by the SRP-II) and PTSD were non-significant; however the association was negative in direction. Strengths of this study include the analyses across age. This revealed that the strength of the relationship between factor 1 of psychopathy and "virtually all DSM-IV anxiety disorders" reduced with age. It is unclear whether this refers to PTSD as the data is not reported. Limitations include the failure to measure potentially confounding variables including Disruptive Behaviour Disorders, given that only the anxiety scales of the APS were administered.

Pham (2012) assessed the relationship between psychopathy and traumatic stress in male forensic psychiatric patients in a high security hospital using the PCL-R and the SASRQ respectively. Part of the study compared "psychopaths" versus "non-psychopaths" based on PCL-R scores > 27 and < 15 respectively. The study found that higher levels of psychopathy were associated with less traumatic stress symptoms including re-experiencing, dissociation and inadaptation. Only the affective facet of the PCL-R was significantly negatively correlated and predictive of all traumatic stress symptoms. Therefore, the authors concluded

that this affective deficit may protect against traumatic stress. When the groups were compared, 77% of "non-psychopaths" compared with 31% of "psychopaths" met diagnosis for Acute Stress Disorder. Strengths of this study include the detailed exploration of psychopathy and acute stress factors, the use of dimensional and categorical methods, the use of the PCL-R by trained professionals with high inter-rater reliability and the application of separate PCL-R cut-offs to create distinct "psychopathy" versus "non-psychopathy" groups for comparison. The assessment of potential confounding variables including trauma and major mental disorders are additional strengths. However, personality disorders were not assessed. This might have been useful to determine whether the findings were specific to psychopathy, particularly given the high prevalence of childhood conduct disorder in the sample which may indicate the presence of antisocial personality disorder. Limitations include insufficient detail regarding sampling, the small sample and reliance on self-report measures of traumatic stress. Recall bias may have been particularly problematic given that the SASRQ assesses symptoms in the 30 days following the traumatic event and this event had often occurred over 10 years ago.

Moeller and Hell (2003) investigated the prevalence of affective disorder, trauma, PTSD and their relationship to psychopathy in male prisoners. Based on a PCL-R cut-off score of 25, a "psychopath" and "non-psychopath" group were formed. The SCID-1 for DSM-IV was administered to measure PTSD. They found that none of the "psychopath" group met diagnostic criteria for PTSD versus three in the "non-psychopath" group. Given that psychopaths reported more traumatic events, the authors concluded that those with psychopathy may possess adaptive coping strategies to prevent them developing PTSD following trauma. This study is commended for using clinician administered measures to assess psychopathy and PTSD, the consistent administration of the SCID-I by the same author and measurement of potential confounding variables (e.g. trauma, drug abuse). However, there is no evidence of blinding to group allocation and insufficient information regarding whether the measures were administered by trained individuals. These factors may have introduced rater bias. Furthermore, the small sample, particularly in the "psychopath" group and absence of baseline group comparisons make it difficult to ascertain the extent to which group differences are due to psychopathy. Whilst it is a strength that inmates were "screened unselected" shortly after admission, the sample may only be representative of this time period as opposed to longer term prisoners, community, psychiatric or female populations.

Myers et al., (2012) conducted a descriptive study investigating the role of psychopathy in adolescent parricide offenders. Psychopathy was assessed using the PCL-R or PCL: YV dependent on age, whilst diagnoses of psychopathology including PTSD were based on clinical interviews, psychological testing including the TSCC and a review of collateral and file information. The findings revealed that only two youths scored above 10 on the PCL. Six youths were diagnosed with PTSD; however they had PCL: YV scores below 10. Thus, those with PTSD did not present with psychopathy. Conversely, those with elevated psychopathy scores did not meet diagnosis for PTSD. Strengths of this study include the comprehensive assessment procedure and administration of PCL measures by trained and experienced professionals. Whilst no conclusions are drawn regarding the psychopathy – PTSD relationship, this study is suggestive of a negative relationship. These conclusions are extremely tentative and must be interpreted with caution given the lack of statistical analyses, small sample, absence of psychopathy factor level scores and scores on measures of psychopathology including PTSD. Furthermore, as recognised by the authors, their familiarity with the cases and studies hypotheses may have compromised the reliability of their assessments.

Table 1.Description of sample characteristics, measures utilised and relevant findings.

Study	Sample Size	Sample Characteristics	Measure of PTSD Symptoms	Measure of Psychopathy	Relevant Findings
Blackburn et al., 2003	175	**Population:** Mentally disordered offenders from high-security hospitals: Ashworth Hospital, England (n=115) & The State Hospital, Scotland (n=60). **Gender:** All Male. **Ethnicity**: Not reported. **Recruitment:** **Ashworth Hospital** –From the personality disorder unit, 55 (79%) of the 70 approached, agreed to participate. From the mental health directorate 60 (65%) of the eligible 93 patients agreed to participate. Of those who did not participate, 33 were excluded on the basis of nursing advice and 15 refused. Non-participants did not differ from participants on age or duration of admission. Non-participants were more psychotic. **The State Hospital –** Excluding females and those with a Learning Disability, every second	**CIDI - version 2.1 [Structured Interview]** Description Determines whether DSM-IV & ICD-10 diagnoses satisfied from self-report information. This includes a category on PTSD. This study assessed lifetime & 12 month prevalence of these disorders.	**PCL-R [Semi-Structured Interview & File Review]** Description 20-item rating scale assessing psychopathic traits in adults. Administered by trained professionals. High inter-rater reliability obtained.	Psychopathy (PCL-R ≥ 25) co-occurred significantly with PTSD (OR = 2.65, p<0.01).

		patient was identified. Non-participants were older, had longer admissions and were more psychotic than participants. **Groups:** Sample divided into 3 groups based on mental health legislation (not according to PCL-R score). Mental Health Act (1983) – Ashworth Hospital Psychopathy (n=54) – Age (M=40.94, SD=9.86) Mental Illness (n=61) – Age (M=36.64, SD=9.50) Mental Health (Scotland) Act (1984) – The State Hospital Mental Disorder (n=60) – Age (M=34.13, SD=9.35)			
Blonigen et al., 2012	226	**Population:** Prison inmates from a Federal Correctional Institution in Tallahassee, Florida. **Gender:** All Female. **Age:** M = 31.9, SD = 6.8, range = 19-53. **Ethnicity:** African American (57.1%, n = 129), Caucasian (29.6%, n =67), Latino (10.6%, n =24), Asian (0.4%, n =1), Other (2.2%, n=5).	**PCL-C** **[Self-Report]** Description 17-item measure that	**PCL-R** **[Semi-Structured Interview & File Review]** Description 20-item rating scale	Correlations revealed a significant positive correlation between the PCL-R Total Score and PTSD (r = .20, p<.01). **Factor Level**

| | | **Recruitment:** Before recruitment, participants were randomly selected from the prison roster & invited to pre-participation screening. Those meeting inclusion criteria were recruited (i.e. English-language proficiency, no imminent release date & based on file review no evidence of psychosis, bipolar disorder or cognitive impairment). | asks individuals to rate the severity with which they have been bothered by the 17 DSM-IV PTSD symptoms over the past month. | assessing psychopathic traits in adults. Administered by trained students. High inter-rater reliability obtained. | Factor 1 of the PCL-R was not significantly correlated with PTSD (r=.04, ns). A significant positive correlation between Factor 2 of the PCL-R & PTSD (r=.28, p<.01). **Facet Level** Interpersonal facet of the PCL-R was unrelated to PTSD (r=.02, ns) Affective facet of the PCL-R was unrelated to PTSD (r=.05, ns). Lifestyle facet was positively correlated with PTSD (r=.22, p<.01). Antisocial facet was positively correlated with PTSD (r=.31, p<.01) **Regression Analyses** |

					Only the antisocial facet of the PCL-R was uniquely associated with PTSD. Borderline Personality Disorder (assessed by MBPD), mediated this relationship.
Hicks et al., 2010	140	Participants were members from a larger population (n=226) of female prison inmate *volunteers* from a Federal Correctional Institution. Larger sample described by Blonigen et al., (2012) – **see above.** **Recruitment:** Inclusion criteria - no imminent release date, no evidence of severe or persistent mental illness as determined by file evidence & competence in English. **Groups:** PCL-R ≥ 25 = Psychopathic group (n=70) PCL ≤ 17 =Non-Psychopathic Controls (n=70)	**PCL-C** **[Self-Report]** Description 17-item measure that asks individuals to rate the severity with which they have been bothered by the 17 DSM-IV PTSD symptoms over the past month.	**PCL-R** **[Semi-Structured Interview & File Review]** Description 20-item rating scale assessing psychopathic traits in adults. Administered by trained psychology students. High inter-rater reliability	Post hoc tests using Turkey's procedure revealed: Primary (M=28.6, SD=3.4) and Secondary psychopaths (M=29.3, SD=2.7) had significantly higher PCL-R scores compared with controls (M=11.2, SD = 4.2). Primary (M=12.1, SD =2.3) and Secondary

		The Psychopathic group was divided based on scores on the 11 primary scales of the MPQ-BF: Primary Psychopathy group (n=31) Secondary Psychopathy groups (n=39)		obtained.	psychopaths (M=11.5, SD=2.1) did not differ significantly in PCL-R Factor 1 scores.
		Final groups: Primary Psychopaths (n=31) Secondary Psychopaths (n=39) Non-Psychopathic Controls (n=70)			Secondary psychopaths had significantly higher PCL-R Factor 2 scores (M=13.9, SD=1.9) relative to primary psychopaths (M=12.7, SD=1.7).
					Secondary psychopaths had significantly higher PCL-R Antisocial facet scores (M=5.4, SD=1.9) relative to primary psychopaths (M=4.2, SD=1.5).
					Secondary psychopaths reported significantly

					more PTSD symptoms (M=2.7, SD=0.8) relative to primary psychopaths (M=1.9, SD=0.8) and Controls (M=2.0, SD=0.9). No significant difference in PTSD symptoms between controls (M=2.0, SD=0.9) and primary psychopaths (M=1.9, SD=0.8).
Kubak & Salekin 2009	130	**P**opulation: Juvenile Offenders at a Court Assessment Unit in a Southeastern state. **Gender:** 92 (70.8%) Male, 38 (29.2%) Female. **Age:** M=14.86, SD=1.64, range 9-18 yrs. **Ethnicity:** 51 (39.2%) African American, 9 (6.9%) Caucasian Americans, 62 (47.7%) Hispanic Americans, 5 (3.8%) Haitian Americans & 3 (2.3%) mixed ethnicity. **Education:** M=8.6 yrs., SD =1.46 **Offence:** theft, armed robbery, battery, throwing projectiles, other violent offences.	**APS - Anxiety Disorder Scales. [Self-Report]** Description Based on DSM-IV criteria. A 346-item self-report measure which assesses symptoms of clinical and personality	**PCL:YV [Semi-Structured Interview &File Review]** Description A 20-item scale to assess psychopathy in youth.	The PCL: YV was significantly positively correlated with PTSD (r=.20, p<.05). The APSD was significantly positively correlated with PTSD (r=.37, p<.001). The SRP-II was

25

		Recruitment: Sampling method not explicitly reported. Inclusion criteria not reported. The sample was reduced to 103 for statistical analyses. This was due to one of the measures being discontinued after 103 youths had been assessed.	disorders & distress in adolescents. This includes a scale for PTSD.	**APSD** **[Self-Report]** <u>Description</u> A 20-item self-report measure to screen for psychopathy in youth. **SRP-II** **[Self-Report]** <u>Description</u> A 60-item self-report version of the PCL-R. Assess to what extent an individual is judged to be a prototypical psychopath.	negatively correlated with PTSD, however this was not statistically significant (r = -.07, ns).
Moeller & Hell 2003	102	**Population:** Prison inmates sentenced to a Swiss Correctional Institution. **Gender:** All Male **Age:** M = 21.3, range 17-27 yrs. **Ethnicity:** Not reported	**SCID-I for DSM-IV** **[Structured Clinical Interview]** Administered by author. No information reported	**PCL-R** **[Semi-Structured Interview &File Review]** <u>Description</u> 20-item rating scale assessing psychopathic	None of the "psychopaths" met diagnostic criteria for PTSD. Three of the "non-psychopaths" met diagnostic criteria for PTSD.

		Recruitment: Data collected from Jul 1999 to Feb 2002. Individuals were screened unselected during their first 3 months following admission. Of the 109 prisoners who were admitted during this time, 7 were unavailable. Gr**oups:** PCL-R > 25 = Psychopath group (n=16) PCL ≤ 25 = Non-psychopathic group (n=86)	on whether they were trained in the SCID-1. Inter-rater reliability not assessed.	traits in adults. Administered by a senior physician. No information reported on whether they were trained in the PCL-R. Inter-rater reliability not assessed.	
Myers et al., 2012	10	**Population:** juvenile parricide offenders who were subsequently assessed during pre-trial forensic psychiatric evaluations. **Gender:** 9 Males / 1 Female. **Age:** At the time of the crimes, M=17.7, SD=2.3, range = 14-21 yrs. **Ethnicity:** 8 White, 2 Black. **IQ** (measure not reported): M=106, SD=10.4, Range = 88-116. **Recruitment:** These participants were seen in several states over a 15 year period. Six were referred for evaluation	A Comprehensive assessment - clinical interviews, neuropsychiatric and psychological testing (e.g. MMPI-Adolescent, **TSCC**, IQ Assessment) & review of collateral information. Diagnoses were made by the authors based on all	**PCL-R & PCL: YV (dependent on age) [Semi-Structured Interview & File Review]** Administered by trained and experienced professionals (the authors).	20% (n=2) had PCL scores above 10. 60% (n=6) of the sample had a diagnosis of PTSD. Of the 6 participants that met diagnostic criteria for PTSD, all 6 had PCL: YV scores of <10.

		by defence and four by prosecution. All participants had been referred to an adult court for prosecution.	information. **TSCC** **[Self-Report]** Measures posttraumatic stress in children.	No formal assessment of inter-rater reliability obtained.	
Pham 2012	48	**Population:** Forensic patients admitted to a high-security psychiatric hospital under the Belgian Social Defence Act. **Gender:** All Male **Age:** M=35.59, SD=9.38 **Ethnicity:** Not reported. **Language**: French-speaking **IQ (Assessed by WAIS-R):** M=82.11, SD=12.29 **Duration of Confinement:** M=69.78 months, SD=46.74 **Crimes:** All had committed a criminal offence but were deemed unable to control their actions. **Recruitment:** Method of sampling not reported. Inclusion criteria not reported. For Part 1 all participants were included (n=48). For Part 2, a sub-sample (n=26) was selected and divided into groups based	**SASRQ** **[Self-report]** Description Two-part self-report instrument. In part 1, participants are asked to describe all traumatic events experienced. In part 2, they identify the most traumatic event and assess its impact on a 6-point Likert scale. The items relate to DSM-IV	**PCL-R** **[Semi-structured interview & file review]** Description 20-item rating scale assessing psychopathic traits in adults. Administered by clinical psychologists trained in the PCL-R. Inter-rater reliability not assessed.	The SASRQ Total Score was negatively correlated with PCL-R Total Score (r= -.30, p<.05). A significant negative correlation between PCL-R Total Score and dissociation (r = -.36, p<.05), re-experiencing (r= -.30, p<.05) and inadaptation (r= -.29, p<.05) symptoms. The Affective facet was the only facet significantly negatively correlated with all traumatic stress factors.

		on PCL-R scores. **Groups:** PCL-R ≥ 27 = Psychopath group (n=13) PCL-R≤15 = Non-Psychopath group (n=13)	symptoms in the 30 days following the event.		Linear regression analyses confirmed that only the affective facet of psychopathy was a negative predictor of SASRQ total score and dissociation, re-experiencing, avoidance, & inadaptation subscales. Psychopaths had significantly lower scores on the SASRQ, reflecting less traumatic stress symptoms. Therefore, the prevalence of Acute Stress Disorder was significantly lower among psychopaths (31%) relative to non-psychopaths (77%).
Salekin et al., 2004	130	The same sample as used in Kubak & Salekin (2009). See above for the demographics of the	**APS – All Scales** **[Self-Report]**	**PCL:YV** **[Semi-Structured**	A significant positive correlation between

				APSD scores and PTSD (r=.37, p<.01).
		sample.	**Interview & File Review]**	
		As with Kubak & Salekin (2009), the sample was reduced to 103 for statistical analyses as the APS data was not available for all participants. No reason is reported for this.	Description	A positive correlation
			A 20-item scale to assess psychopathy in youth. High inter-rater reliability obtained.	between PCL: YV scores and PTSD (r=.20), however not significant
		Description		at a level of p<.01. A
		A 346-item, multiscale inventory that addresses DSM-IV childhood and adolescent disorders. This includes a PTSD scale within the clinical scales.	**APSD** **[Self-Report]** Description A 20-item self-report measure to screen for psychopathy in youth.	positive correlation between SRP-II scores and PTSD (r=.20), however not significant at a level of p<.01.
			SRP-II **[Self-Report]** Description A 60-item self-report version of the PCL-R. Assesses to what extent an individual is judged to be a	

				prototypical psychopath. The SRP-II was modified to make it developmentally appropriate for youths.	
Willemsen et al., 2012	81	**Population:** Prison inmates recruited from two prisons in Flanders, Belgium. **Gender:** All Male **Age**: M=39.8 yrs., SD=12.17, range=20-73 yrs. **Ethnicity:** 84% White, 14% North Africans, and 2% other. **Offence:** 27% attempted manslaughter or murder, 25% violent crime (robbery, assault or battery), 41% sexual crime (indecent assault or rape of a minor or adult), 7% other (drugs, fraud, burglary). **Recruitment:** Individuals participated on a voluntary basis. No incentive was provided. Inclusion criteria included competency in Dutch, not on remand, declared fully responsible for their own actions and not psychotic.	**SCID-I [Structured Clinical Interview]** Description The presence of re-experiencing, hyperarousal & avoidance scored on a 3-point Likert scale. Posttraumatic stress scale was calculated by adding up the scores on the 17 symptoms. High inter-rater reliability obtained.	**PCL-R [Semi-Structured Interview & File Review]** Description 20-item rating scale assessing psychopathic traits in adults. Administered by a clinical psychologist (author) trained in the PCL-R. High inter-rater reliability obtained.	Negative bivariate association between PCL-R total, interpersonal & affective facet scores with posttraumatic stress. An interaction between the affective facet & versatility of traumatic exposure had a significant negative effect on posttraumatic stress. The lifestyle & antisocial facets were not significantly associated with posttraumatic stress.

Abbreviations

APS   = Adolescent Psychopathology Scale - Anxiety Disorder Scales (Reynolds, 1998)

APSD   = Antisocial Process Screening Device (Frick & Hare, 2001)

CIDI   = Composite International Diagnostic Interview (version 2.1; World Health Organization, 1997)

MPQ-BF   = Multidimensional Personality Questionnaire – Brief Form (Patrick et al., 2002)

MBPD   = Minnesota Borderline Personality Disorder (Bornovalova et al., 2011)

PCL-C   = PTSD Checklist – Civilian Version (Weathers et al., 1993)

PCL-R   = Psychopathy Checklist – Revised (Hare, 1991)

PCL: YV   = Psychopathy Checklist – Youth Version (Forth et al., 2003)

SASRQ   = Stanford Acute Stress Reactions Questionnaire (Cardena et al., 1996)

SCID-I for DSM-IV = Structured Clinical Interview for DSM-IV (Wittchen et al., 1997)

SRP-II   = Self-report psychopathy-II (Hare, 1991)

TSCC   = Trauma Symptom Checklist for Children (Brierre, 1996)

WAIS-R   = Wechsler Adult Intelligence Scale – Revised (Wechsler, 1981)

Table 2. Quality ratings for each of the papers reviewed.

Study	Quality rating (%)	Quality category
Blackburn et al., (2003)	73.0%	high
Blonigen et al., (2012)	67.6%	moderate
Willemsen et al., (2012)	62.2%	moderate
Hicks et al., (2010)	61.9%	moderate
Kubak & Salekin (2009)	56.8%	adequate
Myers et al., (2012)	56.8%	adequate
Pham (2012)	54.8%	adequate
Moeller & Hell (2003)	52.3%	adequate
Salekin et al., (2004)	51.4%	adequate

Discussion

The purpose of this review was to critically appraise the empirical literature investigating the relationship between psychopathy and posttraumatic stress or acute stress disorder symptoms. A secondary question addressed whether there are differential relationships between psychopathy subtypes or factors with these symptoms.

Of the nine papers identified, two sets of papers had overlapping samples, providing seven distinct samples. Five of the papers addressed the primary review question as their main research question, four of which were published in the last five years. This may be suggestive of increased recognition of this question in its own right. Four of the papers provided sufficient detail to address the second review question. The papers reviewed varied in quality from adequate to high quality; however the majority were rated as adequate suggesting potential areas for improvement.

What is the relationship between psychopathy and PTSD (and acute stress) symptoms?

All nine papers found evidence of a relationship between psychopathy and posttraumatic stress/acute stress symptoms. Five of the papers, comprised of three distinct samples found a positive relationship between psychopathy and posttraumatic stress (Blackburn et al., 2003; Blonigen et al., 2012, Hicks et al., 2010, Kubak & Salekin, 2009; Salekin et al., 2004). Thus, psychopathy was associated with higher levels of posttraumatic stress symptoms. Conversely, four papers with distinct samples found a negative relationship between psychopathy and posttraumatic stress (Moeller & Hell, 2003; Myers et al., 2012; Pham, 2012 and Willemsen et al., 2012). Thus, psychopathy was associated with less posttraumatic stress.

Overall, the papers revealing a positive relationship were of a slightly higher quality and had a larger sample size collectively (n = 504) than those which found a negative relationship (n = 241), with one paper considered high quality, two considered moderate and two considered adequate. Only one of the papers reporting a negative relationship was considered moderate quality with the remaining three considered adequate. Thus, the strongest evidence is suggestive of a positive relationship between these conditions. Although these studies were of a slightly higher quality, this study is extremely tentative and must be interpreted with caution given the small number of studies and lack of consensus regarding the methods used to assess posttraumatic stress. Furthermore, the relationships found may be confined to the particular populations studied.

An evaluation of the evidence

Collectively, after accounting for the overlapping samples, the studies reporting a positive relationship were comprised of adults and youths (approximately 20%) with approximately equal numbers of males and females. Conversely, those finding a negative relationship were primarily male adults (with the exception of a single female). This may be suggestive of gender differences with female psychopaths more at risk and male psychopaths less at risk of developing posttraumatic stress. This is perhaps not surprising given that females have been shown to be more vulnerable to PTSD following traumatic experiences (Breslau, 2002). However, it may be that the symptoms manifest differently in male and female psychopaths, with females presenting with more conventional symptoms. The findings may also suggest

age differences with more comorbidity in youths, with this effect reducing or potentially reversing with age. Thus, as psychopathy traits become more fixed, the association with posttraumatic stress may weaken.

Whilst both positive and negative relationships were found across prisoners, forensic psychiatric patients and juvenile offenders, a greater proportion (approximately 76%) of those finding a negative relationship were from prison populations.

The majority of studies finding a positive relationship were cross-sectional with the exception of one (Hicks et al., 2010). Furthermore, all except one (Blackburn et al., 2003) assessed posttraumatic stress using symptom severity and not diagnostic cut-offs. Conversely, an equal proportion of cross-sectional and case-control methods were used by those reporting a negative relationship and the majority assessed PTSD in terms of those meeting diagnosis.

Whilst, PCL measures were used to assess psychopathy across all studies, these assessments may be considered more reliable in those studies which found a positive relationship. All were administered by trained individuals and high inter-rater reliability was obtained in four of the papers. However, there was greater variability in its administration in those studies reporting a negative relationship. Only two reported that the assessor was trained and assessed inter- rater reliability. For those reporting a positive relationship, all except one (Blackburn et al., 2003) used self-reports to assess PTSD, whereas a combination of self-report and clinical structured interviews were evident in those finding a negative relationship. Thus, the PTSD assessment in those reporting a positive relationship may be considered less reliable given the biases inherent with self-reports stated earlier. This may be particularly problematic in psychopathic populations given their propensity for impression management (Kubak & Salekin, 2009). It may be argued that traits such as grandiosity may act as a barrier towards the individual sharing their weaknesses or areas of difficulty.

The observed relationships are dependent on the conceptualisation of psychopathy and may not extend to other measures. For example PTSD was no more prevalent within the legal category of psychopathic disorder yet it was when defined using PCL-R scores (Blackburn et al., 2003). Also, two of the studies found a relationship between psychopathy and PTSD, only when using particular measures (Kubak & Salekin, 2009; Salekin et al., 2004).

Are there differential relationships between the psychopathy subtypes or factors and PTSD/ acute stress symptoms?

Four studies including three distinct adult samples addressed the second review question (Blonigen et al., 2012; Hicks et al., 2010; Pham, 2012; Willemsen et al., 2012). These provide evidence of differential relationships between the psychopathy subtypes and factors with PTSD. Generally, factor 2 including the lifestyle and in particular the antisocial facets were found to be associated with an increase in posttraumatic stress in female prisoners. However, the interpersonal and affective traits were unrelated to posttraumatic stress. These studies support the position that individuals with these traits are likely to behave in a way that places them in dangerous situations where they are at increased risk of experiencing traumatic events.

Conversely, in male prisoners and forensic psychiatric patients, interpersonal and affective facets were associated with less posttraumatic stress, with the affective facet reducing the impact of exposure to traumatic experiences (Willemsen et al., 2012). Similarly, Pham (2012) found that the affective facet was the only facet which was significantly associated with, and predicted posttraumatic stress. The lifestyle and antisocial traits however, were not related to posttraumatic stress in these samples. These studies suggest that the affective deficit commonly seen in psychopaths may protect them from developing posttraumatic stress. Whilst this may be considered advantageous, it may mean that they are not deterred by situations which may normally be perceived as stressful or traumatic. Thus, they may continue to place themselves into situations which may have adverse consequences for themselves or others. These differential relationships may partly explain the mixed findings in the research, given that many studies have not examined psychopathy in this level of detail.

Limitations of the research

Collective limitations included the reliance on self-report measures of PTSD, insufficient reporting of sampling and insufficient detail regarding the administration of clinician rated measures. None of the studies indicated that the raters were blinded to group allocation. Whilst this is less of an issue for cross-sectional methods, interviewer bias may present if the interviewer is aware of how the participant scored for psychopathy. The majority of the

studies were cross-sectional, thus causation cannot be determined. None of the case-control studies matched cases to controls, not all conducted baseline group comparisons and the sample sizes were small, limiting the reliability of these findings. The justification of sample size was not reported. Although potential confounding variables such as exposure to trauma were sometimes measured, this was never controlled for within the analyses and not all studies measured this. Furthermore, comorbidity within the sample was rarely addressed within the analyses and it is unclear whether related psychiatric conditions or other variables were impacting on the relationship between psychopathy and posttraumatic stress.

Recommendations for future research

Future research addressing the relationship between psychopathy and PTSD should include clinician assessments and various other measures for both psychopathy and PTSD. This is crucial to determine whether differences in the relationship between these conditions are due to the measures used or a reflection of the heterogeneity of individuals presenting with psychopathic traits. It may be argued that to provide an accurate assessment of such conditions, extensive knowledge and experience is required. This is important when considering the potential overlap in symptomatology (Sharf et al., 2014), thus requiring expertise to make a differential diagnosis.

Future studies should measure psychopathy at the factor and facet level in larger samples with sufficient numbers of males and females to test for gender differences. Similarly, only one study investigated PTSD at the level of individual symptoms, highlighting an avenue for future research. This may determine whether there are differential relationships between psychopathy and the various symptoms of PTSD. If a consistent relationship is found between psychopathy and PTSD, studies should determine whether this is upheld using case-control designs with cases and controls matched on relevant variables, or at a minimum compared for baseline differences. Where professionals are involved in the assessments, they should be blinded to group allocation to reduce rater bias. Regardless of design, future studies should seek to measure and control for confounding variables either by exclusion or using statistical methods. Relevant variables may include exposure to traumatic events and personality disorders, particularly antisocial personality disorder or disruptive behaviour disorders in youth. This would help to ascertain whether the relationship is specific to psychopathy. It is recommended that studies follow reporting guidelines, e.g. the STROBE (Strengthening the Reporting of Observational studies in Epidemiology) to ensure detailed

and transparent reporting (Von Elm et al., 2007). This will allow for an accurate interpretation of the relationship between psychopathy and posttraumatic stress.

Strengths & limitations of review

Whilst attempts were made to ensure the search strategy was comprehensive, the use of specific terminology and bias towards peer reviewed journal publications in English may have excluded potentially relevant papers. The lack of transparency regarding overlapping samples and the consequent reduction in distinct samples included in this review reduces the strength of conclusions that can be drawn. Whilst the quality assessment tool achieved high inter-rater reliability, there is the potential for subjectivity in relation to ratings. Furthermore, this is not a standardized tool and does not provide comparability across reviews.

Conclusions

This review found limited evidence of a relationship between psychopathy and PTSD, with mixed evidence regarding the direction of this relationship. The studies reporting a positive relationship were found to be of a higher quality overall. Furthermore, differences in gender and age between studies reporting a positive relationship and studies reporting a negative relationship, suggest that gender and age may influence the relationship between psychopathy and posttraumatic stress.

There was evidence of a differential relationship between the psychopathy subtypes and factors with posttraumatic stress. In general, factor 1 was associated with a reduction and factor 2 with an increase in posttraumatic stress. Further research is required to investigate psychopathy at this level to determine whether this finding is replicated, and whether gender effects are evident.

The findings from this review are extremely tentative and should be interpreted with caution given the small number of studies and the methodological quality of these studies. Furthermore, the heterogeneous nature of these studies including the designs, measures used (particularly to assess PTSD), and the populations studied make it difficult to synthesize the findings and draw firm conclusions.

Overall, this review highlights a dearth of research exploring the link between psychopathy and PTSD, particularly in females and youths. The quality of existing evidence is variable and future studies should follow the recommendations above to improve the methodological rigour and reporting of their studies. Only with further studies can firmer conclusions regarding the relationship between psychopathy and PTSD be drawn.

References

(* = reviewed studies)

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Birbaumer, N., Veit, R., Lotze, M., Erb, M., Hermann, C., Grodd, W., & Flor, H. (2005). Deficient fear conditioning in psychopathy: A functional magnetic resonance imaging study. *Archives of general psychiatry*, 62(7), 799-805.

Blackburn, R. & Coid, J. W. (1998). Psychopathy and the Dimensions of Personality Disorder in Violent Offenders. *Personality and Individual Differences, 25,* 129-145.

*Blackburn, R., Logan, C., Donnelly, J., & Renwick, S. (2003). Personality disorders, psychopathy and other mental disorders: Co-morbidity among patients at English and Scottish high-security hospitals. *The Journal of Forensic Psychiatry & Psychology*, *14*(1), 111-137.

Blanchard, E. B., Hickling, E. J., Taylor, A. E., Loos, W. R., & Gerardi, R. J. (1994). The psychophysiology of motor vehicle accident related posttraumatic stress disorder. *Behavior Therapy, 25*(3), 453-467.

*Blonigen, D. M., Sullivan, E. A., Hicks, B. M., & Patrick, C. J. (2012). Facets of psychopathy in relation to potentially traumatic events and posttraumatic stress disorder among female prisoners: The mediating role of borderline personality disorder traits. *Personality Disorders: Theory, Research, and Treatment*, *3*(4), 406-414.

Bornovalova, M. A., Hicks, B. M., Patrick, C. J., Iacono, W. G., & McGue, M. (2011). Development and validation of the Minnesota Borderline Personality Disorder scale. *Assessment, 18,* 234–252.

Briere, J. (1996). *Trauma Symptom Checklist for Children: Professional manual.* Florida: Psychological Assessment Resources Inc.

Cardena, E., Classen, C., Koopman, C., & Spiegel, D. (1996). Review of the Stanford Acute Stress Reaction Questionnaire (SASRQ). In B. H. Stamm (Ed.), *Measurement of stress, trauma, and adaptation* (pp. 293–297). Lutherville, MD: Sidran Press.

Cleckley, H. (1941). *The mask of sanity; an attempt to reinterpret the so-called psychopathic personality.*

Cleckley, H. (1976). *The mask of sanity* (5th ed.). St. Louis, MO: Mosby.

Cooke, D. J., & Michie, C. (2001). Refining the construct of psychopathy: towards a hierarchical model. *Psychological assessment, 13*(2), 171-188.

Davidson, J. R., & Foa, E. B. (1991). Diagnostic issues in posttraumatic stress disorder: Considerations for the DSM-IV. *Journal of Abnormal Psychology*, *100*(3), 346-355.

Deeks, J. J., Dinnes, J., D'Amico, R., Sowden, A. J., Sakarovitch, C., Song, F., Petticrew, M., & Altman, D.G. (2003). Evaluating non-randomized intervention studies. *Health Technology Assessment, 7*(27), 1-173.

Dembo, R., Jainchill, N., Turner, C., Fong, C., Farkas, S., & Childs, K. (2007). Levels of psychopathy and its correlates: A study of incarcerated youth in three states. *Behavioral Sciences & the Law, 25*, 717-738.

Downs, S. H., & Black, N. (1998). The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *Journal of epidemiology and community health, 52*(6), 377-384.

Durant, R. H. (1994). Checklist for the evaluation of research articles. *Journal of adolescent health, 15*(1), 4-8.

Ehlers, A., Suendermann, O., Boellinghaus, I., Vossbeck-Elsebusch, A., Gamer, M., Briddon, E., ... & Glucksman, E. (2010). Heart rate responses to standardized trauma-related pictures in acute posttraumatic stress disorder. *International Journal of Psychophysiology*, *78*(1), 27-34.

Fani, N., Tone, E. B., Phifer, J., Norrholm, S. D., Bradley, B., Ressler, K. J., ... & Jovanovic, T. (2012). Attention bias toward threat is associated with exaggerated fear expression and impaired extinction in PTSD. *Psychological medicine*, *42*(3), 533-543.

Foa, E. B., & Riggs, D. S. (1993). Post-traumatic stress disorder in rape victims. In J. Oldham, M. B. Riba, & A. Tasman (Eds.), *American Psychiatric Press Review of Psychiatry, 12*, 273-303. Washington, DC: American Psychiatric Press.

Foa, E. B., Steketee, G., & Rothbaum, B. O. (1989). Behavioral/cognitive conceptualizations of post-traumatic stress disorder. *Behavior therapy*, *20*(2), 155-176.

Forth, A. E., Kosson, D. S., & Hare, R. D. (2003). *The psychopathy checklist: Youth version.* Toronto, ON, Canada: Multi-Health Systems.

Frick, P. J., Lilienfeld, S. O., Ellis, M., Loney, B., & Silverthorn, P. (1999). The association between anxiety and psychopathy dimensions in children. *Journal of Abnormal Child Psychology, 27,* 383-392.

Frick, P. J., & Hare, R. D. (2001). *The antisocial processing device (APSD).* Toronto. ON, Canada: Multi-Health Systems.

Goff, A., Rose, E., Rose, S., & Purves, D. (2007). Does PTSD occur in sentenced prison populations? A systematic literature review. *Criminal behaviour and mental health, 17(3),* 152-162.

Hare, R. D. (1991). *The Hare Psychopathy Checklist-Revised.* Toronto: Multi-Health Systems.

Hare, R. D. (1991). *The Self-Report Psychopathy Scale—II.* Vancouver, BC, Canada: Unpublished test, University of British Columbia.

Hare, R. D. (2003). *The Hare Psychopathy Checklist Revised manual* (2nd ed.). Toronto: Multi-Health Systems.

Hare, R. D., & Neumann, C. S. (2005). Structural models of psychopathy. *Current Psychiatry Reports, 7*, 57–64.

Hicks, B. M., Markon, K. E., Patrick, C. J., Krueger, R. F., & Newman, J. P. (2004). Identifying psychopathy subtypes on the basis of personality structure. *Psychological assessment*, *16*(3), 276-288.

*Hicks, B. M., Vaidyanathan, U., & Patrick, C. J. (2010). Validating female psychopathy subtypes: Differences in personality, antisocial and violent behavior, substance abuse, trauma, and mental health. *Personality Disorders: Theory, Research, and Treatment*, *1*(1), 38-57.

Jarde, A., Losilla, J. M., & Vives, J. (2012). Methodological quality assessment tools of non-experimental studies: a systematic review. *Anales de Psicología, 28(2),* 617-628.

Karpman, B. (1941). On the need of separating psychopathy into two distinct clinical types: the symptomatic and the idiopathic. *Journal of Criminal Psychopathology, 3,* 112-137.

Karpman, B. (1948). Conscience in the psychopath: another version. *American Journal of Orthopsychiatry*, *18*(3), 455-491.

Krischer, M. K., & Sevecke, K. (2008). Early traumatization and psychopathy in female and male juvenile offenders. *International journal of law and psychiatry*, *31*(3), 253-262.

*Kubak, F. A., & Salekin, R. T. (2009). Psychopathy and anxiety in children and adolescents: New insights on developmental pathways to offending. *Journal of Psychopathology and Behavioral Assessment*, *31*(4), 271-284.

Lykken, D. T. (1957). A study of anxiety in the sociopathic personality. *The Journal of Abnormal and Social Psychology*, *55*(1), 6-10.

Mitchell, D. G., Richell, R. A., Leonard, A., & Blair, R. J. R. (2006). Emotion at the expense of cognition: Psychopathic individuals outperform controls on an operant response task. *Journal of Abnormal Psychology*, *115*(3), 559.

*Moeller, A. A., & Hell, D. (2003). Affective disorder and „psychopathy" in a sample of younger male delinquents. *Acta Psychiatrica Scandinavica*, *107*(3), 203-207.

Mowrer, O. H. (1960). *Learning theory and behaviour*. New York, NY: Wiley.

*Myers, W. C., & Vo, E. J. (2012). Adolescent parricide and psychopathy. *International Journal of Offender Therapy and Comparative Criminology*, *56*(5), 715-729.

Newman, J. P., Schmitt, W. A., & Voss, W. D. (1997). The impact of motivationally neutral cues on psychopathic individuals: Assessing the generality of the response modulation hypothesis. *Journal of Abnormal Psychology, 106*(4), 563-575.

Pagura, J., Stein, M. B., Bolton, J. M., Cox, B. J., Grant, B., & Sareen, J. (2010). Comorbidity of borderline personality disorder and posttraumatic stress disorder in the U.S. population. *Journal of Psychiatric Research, 44,* 1190 –1198.

Patrick, C. J., Curtin, J. J., & Tellegen, A. (2002). Development and validation of a brief form of the Multidimensional Personality Questionnaire. *Psychological Assessment, 14,* 150–163.

Patrick, C. J. (2006). Back to the future: Cleckley as a guide to the next generation of psychopathy research. In C.J. Patrick (Ed.), *Handbook of Psychopathy*, p 605-617). New York: Guilded Press.

*Pham, T. H. (2012). Psychopathy and traumatic stress. *Journal of Personality Disorders*, *26*(2), 213-225.

Pineles, S. L., Shipherd, J. C., Welch, L. P., & Yovel, I. (2007). The role of attentional biases in PTSD: Is it interference or facilitation? *Behaviour Research and Therapy, 45*(8), 1903-1913.

Patrick, C. J. (2006). Back to the future: Cleckley as a guide to the next generation of psychopathy research. In C.J. Patrick (Ed.), *Handbook of Psychopathy*, p 605-617). New York: Guilded Press.

Poythress, N. G., Skeem, J. L., & Lilienfeld, S. O. (2006). Associations among early abuse, dissociation, and psychopathy in an offender sample. *Journal of Abnormal Psychology, 115,* 288-297.

Reynolds, W. M. (1998). *Adolescent psychopathology scale: Administration and interpretation manual.* Odessa, FL: Psychological Assessment Resources.

*Salekin, R. T., Leistico, A. M. R., Neumann, C. S., DiCicco, T. M., & Duros, R. L. (2004). Psychopathy and comorbidity in a young offender sample: Taking a closer look at psychopathy's potential importance over disruptive behavior disorders. *Journal of Abnormal Psychology*, *113*(3), 416-427.

Sharf, A., Kimonis, E. R., & Howard, A. (2014). Negative life events and posttraumatic stress disorder among incarcerated boys with callous-unemotional traits. *Journal of Psychopathology and Behavioral Assessment*, 1-14.

Shin, L. M., Rauch, S. L., & Pitman, R. K. (2006). Amygdala, medial prefrontal cortex, and hippocampal function in PTSD. *Annals of the New York Academy of Sciences*, *1071*(1), 67-79.

Skeem, J. L., & Cooke, D. J. (2010). Is criminal behavior a central component of psychopathy? Conceptual directions for resolving the debate. *Psychological Assessment*, *22*(2), 433-445.

Von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gøtzsche, P. C., & Vandenbroucke, J. P. (2007). The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Preventive medicine, 45*(4), 247-251.

Warren, J., Hurt, S., Loper, A., Bale, R., Friend, R., & Chauhan, P. (2002). Psychiatric symptoms, history of victimization, and violent behavior among incarcerated female felons: An American perspective. *International Journal of Law & Psychiatry, 25,* 129 –149.

Weathers, F., Litz, B., Herman, D., Huska, J., & Keane, T. (1993). The PTSD Checklist (PCL): Reliability, validity, and diagnostic utility, Paper presented at the Annual Convention of the International Society for Traumatic Stress Studies, San Antonio, TX.

Wechsler, D. (1981). *Manual for the Wechsler Adult Intelligence Scale-Revised*. New York: Psychological Corporation.

*Willemsen, J., De Ganck, J., & Verhaeghe, P. (2012). Psychopathy, traumatic exposure, and lifetime posttraumatic stress. *International Journal of Offender Therapy and Comparative Criminology*, *56*(4), 505-524.

Wittchen, H.-U., Wunderlich, U., Gruschitz, S., & Zaudig, M. (1997). *StrukturiertesKlinisches Interview fuˇr DSM-IV, Achse I (SKID-I).* Gottingen: Hogrefe.

World Health Organization (1997). *Composite International Diagnostic Interview (Automated Version)*. Geneva: WHO.

Chapter Two: Major Research Project

The psychopathy checklist youth version (PCL: YV):
an investigation into its inter-rater reliability.

Sarah J Dickson

Submitted in partial fulfilment of the requirements for the degree of
Doctorate in Clinical Psychology (DClinPsy)

Address for correspondence:

Sarah Dickson

Mental Health & Wellbeing

Administration Building

Gartnavel Royal Hospital

1055 Great Western Road

Glasgow

G12 0XH

Plain English Summary

The psychopathy checklist youth version (PCL: YV): an investigation into its inter-rater reliability.

Background

Psychopathy is a type of personality disorder with three main characteristics; an arrogant and deceitful interpersonal style, deficient affective experience and an impulsive behavioural style (e.g. Cooke & Michie, 2001). Research has consistently shown psychopathy to be linked with negative outcomes such as violence, antisocial behaviour and criminality (e.g. Eden's et al., 2007). A diagnosis may make it difficult to access treatment as psychopathy has a long-standing reputation as being 'untreatable'.

The PCL: YV (Forth et al., 2003) is a clinician measure used to assess psychopathic traits in adolescents. This study focused on the inter-rater reliability (IRR) of this measure. IRR is the level of agreement between two or more raters when using a measure.

Previous studies have assessed the Inter-rater reliability (IRR) of the PCL: YV by having a few raters rate a small sample of cases. In clinical practice, however, clinicians have various levels of experience, and the young people themselves present with varying levels of psychopathy. Factors like these have been shown to effect ratings on other measures. No study has explored these characteristics in relation to the PCL: YV.

Research Questions

1. What level of IRR do professionals achieve when using the PCL: YV?

2. To what extent do professionals agree with experts when using the PCL: YV?

3. Is IRR of judgements associated with the rater's professional characteristics or the severity of psychopathic traits that the young person presents with?

Methods

Participants

Health professionals from relevant services were invited by email to participate in the study. Participants were required to meet the 'User Qualifications' outlined in the PCL: YV manual (with the exception of two criteria which were met by attending training offered by this research). Nineteen professionals participated.

Design & Data Collection

Six male case studies (two with low, two with medium and two with high levels of psychopathy) were developed. Experts (professionals with extensive experience in the assessment of psychopathy) rated these cases using the PCL: YV.

Participants attended a free, two-day training event on the PCL: YV offered as part of this research. Afterwards, they rated the six case studies using the PCL: YV and completed a Staff Information questionnaire which gathered professional information.

Main Findings

This study found a high level of IRR overall for the PCL: YV. However, there was lower IRR for particular traits (impulsivity, poor anger control & early behaviour problems). Also, the cases with medium levels of psychopathy had lower IRR. Professional characteristics (e.g. length of experience) did not affect IRR. However, professionals who had experience using structured assessments of personality were more confident in their PCL: YV ratings. They also said they would feel more comfortable using the PCL: YV to assess psychopathy in adolescents.

Recommendations

Professionals should seek supervision and consultation when using the PCL: YV, paying particular attention to the traits with lower IRR. Training should cover these traits more thoroughly. Future research should focus specifically on cases with medium levels of psychopathy, to determine whether lower IRR is replicated.

References

Cooke, D. J., & Michie, C. (2001). Refining the construct of psychopathy: towards a hierarchical model. *Psychological Assessment, 13*(2), 171-188.

Edens, J. F., Campbell, J. S., & Weir, J. M. (2007). Youth psychopathy and criminal recidivism: a meta-analysis of the psychopathy checklist measures. *Law and Human Behavior, 31*(1), 53-75.

Forth, A. E., Kosson, D.S., & Hare, R.D. (2003). *Psychopathy Checklist: Youth Version (PCL: YV)*. Toronto: Multi-Health Systems.

The psychopathy checklist youth version (PCL: YV): an investigation into its inter-rater reliability.

Research suggests that the roots of psychopathy are evident in youth. Studies have consistently shown psychopathy to be associated with a myriad of adverse outcomes including violence and antisocial behaviour. Thus, the early identification of psychopathic traits may facilitate early intervention and risk management strategies. A diagnosis of psychopathy has important clinical and legal implications. Thus, it is crucial that measures used to assess psychopathy are valid and reliable. This study investigated the inter-rater reliability (IRR) of the Psychopathy Checklist Youth Version (PCL: YV). Several studies have highlighted sources of rater bias with adult measures of psychopathy, amongst other measures. Therefore, this study also addressed whether IRR was associated with the rater's professional characteristics or the severity of cases presented. Six case vignettes were developed (two with low, two with moderate and two with high levels of psychopathic traits). These were rated by experts in the field. Nineteen multidisciplinary health professionals recruited from relevant services participated in this study. All participants attended training on the PCL: YV and then rated the case vignettes using the PCL: YV and completed a Staff Information Questionnaire. This study found high IRR overall for the PCL: YV. However, lower IRR was obtained for particular items ("Impulsivity", "Poor anger control", "Early behaviour problems") and the moderate cases. The rater's professional characteristics were not found to influence IRR. Potential explanations for these findings are discussed followed by recommendations for future research.

Keywords: PCL: YV; inter-rater reliability; psychopathy checklist; youth psychopathy

Introduction

Psychopathy

Clinical descriptions of psychopathy have existed from the 1800's; however Cleckley (1941) was the first to outline criteria to define the construct. Central to this, he proposed that psychopathy was a severe condition masked by an outward appearance of robust mental health. Patrick (2006) classified these criteria under the labels of: positive psychological adjustment, behavioural pathology and impaired social relatedness and emotional unresponsiveness. Modern conceptualisations of psychopathy tend to be based loosely on Cleckley's prototype, though there has been extensive debate regarding the factor structure of psychopathy. It is now generally accepted that "psychopathy" as a construct includes interpersonal (arrogant and deceitful), affective (deficient affective experience) and behavioural (impulsive behavioural style) features (Cooke & Michie, 2001; Hare &Neuman, 2005). The inclusion of antisocial behaviour as a fourth factor (Hare & Neuman, 2005) remains under debate and some have argued that this should be considered a behavioural consequence of psychopathy as opposed to a central component (Skeem & Cooke, 2010). Therefore, there is no universal definition of psychopathy and readers should remain mindful of the definitions used when interpreting the literature.

Psychopathy in youth

It has been suggested that the roots of psychopathy are both present and evident in youth (Hare, 1991), though the developmental pathways to this disorder are not yet fully understood. Over the last decade considerable attention has been paid to the exploration of psychopathy in youth, however, this has been controversial. Concerns include the stability of psychopathic traits into adulthood, the danger of pathologising developmentally appropriate behaviour, the stigmatizing nature of the term (Boccacini et al., 2008) and its impact on clinical and legal decisions (Edens & Vincent, 2008). Psychopathy may manifest differently across the lifespan, therefore it is necessary that developmental markers for the condition in adulthood are identified and that the measures used are developmentally informed (Johnstone & Cooke, 2004).

As has been demonstrated in the adult literature, there is evidence to suggest that youth psychopathy is statistically associated with a myriad of adverse outcomes for the sufferer and society. A series of meta-analyses, for example, found that psychopathy predicted

institutional misconduct (Edens & Campbell, 2007) as well as general, violent and sexual recidivism (Edens et al., 2007). Furthermore, callous and unemotional traits are considered a central component of psychopathy (Hare, 1993) and have been found to delineate a subgroup of youth with more severe, persistent and pervasive antisocial behaviour (Frick & White, 2008). Consistent with this, the DSM-V (Diagnostic and Statistical Manual of Mental Disorders, 5th edition) has introduced a callous-unemotional specifier to further classify youths meeting criteria for conduct disorder (DSM-V; APA, 2013). If psychopathic traits can be identified early on, this may facilitate early intervention and risk management strategies (Johnstone & Cooke, 2004).

Assessments of psychopathy using structured clinical tools are used in most western countries to inform legal decisions including whether youths should be transferred to the adult court (Skeem et al., 2011). The language used when describing young people with these traits has been shown to have a powerful influence on the attributions and expectations of professionals and people in society. Therefore, great caution and sensitivity is required when discussing and describing this group. In one study using case vignettes, jurors judged youths described as „a psychopath" at increased risk of future offending and deserving of harsher punishment relative to youths described as meeting criteria for psychopathy (but not described as a psychopath) or conduct disorder (Boccaccini et al., 2008). In terms of clinical applicability, psychopathy is often thought to be synonymous with untreatability (Olver & Stockdale, 2010) and anecdotal evidence highlights the difficulty securing treatment for youths diagnosed with psychopathy (Johnstone, 2011).

Assessment of psychopathy in youth

The most commonly used measures of psychopathic traits in youth are derived from the widely used Psychopathy Checklist-Revised (Hare, 1991, 2003) and include the Antisocial Process Screening Device (APSD; Frick & Hare, 2001) and the Psychopathy Checklist: Youth Version (PCL: YV; Forth et al., 2003). Whilst other measures are available, these are primarily self-report or parent report measures. An exception to this is the Comprehensive Assessment of Psychopathic Personality (CAPP) and its youth derivative; however this currently remains a research instrument (Cooke et al., 2004). Thus, the PCL: YV is the only clinician rated measure which is manualised and widely available for use. In the development of the PCL: YV, the item names, descriptors and information sources were modified from the adult PCL-R to make them developmentally appropriate (Forth et al., 2003). Despite this,

concerns have been raised with regards to particular items, which have been argued to be more appropriate for adults, for example those involving relationship histories (Edens et al., 2001). The PCL:YV is designed to assess psychopathic traits in individuals aged 12-18 years, is administered and rated in the same format as the PCL-R with the 20 items rated from 0 (item definitely does not apply) to 2 (item definitely applies). No cut-off score is provided for clinical purposes, albeit the authors acknowledge that cut-offs may be used in research to establish subgroups. Throughout the research cut-offs from the adult literature are often used, with no theoretical or statistical rationale for doing so. Consistent with the PCL-R, the PCL: YV is structured using the four factor model and some have criticised its over-reliance on behavioural features (Dawson et al., 2012).

Inter-rater reliability

Inter-rater reliability (IRR) is the level of agreement between two or more raters using the same measure to assess the same construct independently (Rosenfeld & Penrod, 2011). Given the potential consequences of a diagnosis of psychopathy, is it crucial that the measures used are valid and reliable. As highlighted by Eden's et al., (2010) "Although the validity of scores from any type of test or rating scale is certainly a critical topic to consider, validity is predicated on these scores being sufficiently reliable to allow for meaningful assessments of the variable of interest".

The authors of the PCL: YV report "excellent" IRR for the total score based on nineteen youth samples, across various settings. It is acknowledged however that double ratings were available for only a very small number of young people (Forth et al., 2003). There was also variability across the samples and individual items and although most were described as having "acceptable" IRR, there were exceptions to this. Poor IRR was found for the impulsivity item in the probation sample in addition to Items number 4, 8 and 13 within the clinic/community sample (Forth et al., 2003).

A recent meta-analysis explored the IRR of the PCL: YV and found "very strong" IRR (Olver & Stockdale, 2010). The authors however noted considerable differences in IRR across the total and factor scores. Therefore "while the research suggests that the tool and its components have strong reliability overall, there are cases where weaker and possibly unacceptably low levels of reliability are obtained" (Olver & Stockdale, 2010).

Although numerous studies have assessed the IRR of the PCL: YV as part of their methodologies, to our knowledge none have focused exclusively on this. The methods used to date have involved a few professionals rating a random subset of cases. This may not be representative of the heterogeneous characteristics of the raters and young people seen in clinical practice. In reality young people present with varying levels of psychopathic traits and those conducting the assessments present with varying levels of training and experience. However, the characteristics of the raters and the cases are often not reported. Thus, it is unclear whether rater and/or case characteristics may impact on the level of IRR. Furthermore, IRR is often only reported for the total score and occasionally the factor scores and is not known at the item level.

With regards to the adult PCL-R (Hare, 1991; 2003), some studies have found lower levels of IRR than is commonly reported in the literature. For example, Edens et al., (2010) found lower levels of IRR for the PCL-R in an applied setting where independent ratings were made on archival data on sex offenders. This was particularly true for the factor 1 "personality scores" and it may be that these traits are less observable and thus more dependent on clinical opinion (Edens et al., 2010). Another study found that PCL-R ratings were related to the raters own personality as assessed by the NEO personality inventory (Miller et al., 2011).

Lastly, studies have found evidence of "adversarial allegiance" using the PCL-R (Murrie et al., 2008; 2009; 2013), that is the "tendency for experts to reach conclusions that supports the party who retains them" (Murrie et al., 2013). Consistent with this, Eden's and Vincent (2008) describe a real-life example of adversarial allegiance in a US murder trial with the defence assigning a score of 11 and the prosecution a score of 19 using the PCL-R.

Additional factors including the training and experience of the raters have been investigated for other measures. In one study higher IRR was found when the Risk of Sexual Violence Protocol was administered by professionals highly trained in risk assessment, however those with less training agreed less with experts and overestimated risk (Sutherland et al., 2012). Furthermore, this study also found evidence of lower agreement for cases presenting with moderate levels of risk. This may suggest that these cases are less clear cut and hence more dependent on clinical judgement. Another study found that experience and training did not predict the IRR for a CBT formulation task; however whether or not the rater had PhD training predicted the accuracy of their judgements (Persons & Bertagnolli, 1999).

Collectively, these findings highlight sources of bias influencing the rating and IRR of the PCL- R amongst other measures. It is unclear whether these extend to the PCL: YV. This study therefore seeks to investigate the IRR of the PCL: YV, the effects of the rater's professional characteristics and the severity of psychopathic traits presented by the young person.

Research questions

1) What level of IRR do multidisciplinary professionals achieve when using the PCL: YV?

2) To what degree do multidisciplinary professionals agree with experts (professionals with expertise in the assessment of psychopathy), when using the PCL: YV?

3) Is IRR associated with case characteristics (i.e. the severity of psychopathic traits)?

4) Is the IRR associated with individual rater's professional characteristics?

Method

Ethical approval

Ethical approval was granted by the Medical, Veterinary and Life Sciences College Ethics Committee for the University of Glasgow (Appendix 2.1). Management approval was obtained from Dr Julie Metcalfe (Clinical Director for Child and Adolescent Mental Health) and Dr Michael Smith (Lead Associate Medical Director for Mental Health) from NHS Greater Glasgow and Clyde. Management permission was provided from NHS Lothian, Dumfries and Galloway, Fife and Ayrshire and Arran.

Justification of sample size

PASS (Power and Sample Size for Windows; Hintze, 2008) was used to calculate the sample size based on the formula outlined by Walter et al., (1998). A minimum of 6 cases and 18 raters were required based on the power fixed at .80, significance level of .05, null hypothesis of ICC=.30 and alternative hypothesis of ICC=.70. Sensitivity analyses revealed that

significantly increasing the number of raters above 18 had a minor effect on power.

Case vignettes

Six factitious male case vignettes (see Appendix 2.4 for an example) were developed based on the theoretical and clinical experience of the research team. Where elements of real cases were included, personally identifiable information was significantly altered to ensure anonymity. These were designed to capture a broad range of severity of psychopathic traits. Two were designed to include very low, two with moderate and two with very high levels of psychopathy. The recommended research cut-offs outlined in the PCL: YV were used with < 20 representing low, 20-30 representing moderate and >30 representing high severity cases. The case vignettes were between 4 and 5 pages long and were structured in accordance with the PCL: YV Interview Guide. The case vignettes were structured under the following headings: Reason for Referral and Information Sources, Presentation at Interview, Family Life, Developmental History, School History and Adjustment, Work History and Money, Long-term goals, Peer/Sexual Relationships and Interpersonal Traits, Substance Use and Lifestyle, Other Antisocial Behaviour and Mental Health. References were made throughout the vignettes to collateral information, for example from parents.

Expert review process

Four professionals (three Consultant Forensic Clinical Psychologists and one Clinical Psychologist) with expertise in the assessment of psychopathy and experience working with adolescents and adults in forensic settings were recruited. These formed the expert panel. To reduce the demands on the expert's time, the cases were pre-rated by the trainee and Dr Lorraine Johnstone (Consultant Clinical Forensic Psychologist). Each expert was randomly allocated a low, moderate and high severity case. Thus, each case was rated by two experts plus the pre-rating providing three ratings per case. The experts were asked to verify whether they agreed with the proposed ratings and complete a feedback questionnaire (Appendix 2.6) to comment on the authenticity of the cases and determine whether they represented the intended severity.

Experts provided detailed feedback and were in agreement with the majority of pre-ratings. Where there was disagreement, the largest discrepancy was by a single point. The modal score from the three ratings provided an expert score. The experts were in agreement regarding the authenticity of the cases. The majority were in agreement with the

categorisation of the cases into low, moderate or high severity. One expert disagreed with the classification of two of the cases however was in agreement with the scoring. This expert classified a moderate scoring case in the high category as this case scored highly on interpersonal and affective traits, which the expert considered more representative of psychopathy. They also classified a high scoring case in the moderate category as there was evidence of abuse resulting in the expert conceptualising the traits within a complex trauma presentation. Following discussion, it was agreed that the cases would be categorised on the basis of their score as opposed to the interpretation of the score.

Recruitment

A recruitment email was distributed to Specialist Children's Services, Forensic Mental Health Services and relevant university departments inviting staff to participate. These services were selected in the hope that staff would be more likely to meet the PCL: YV user criteria and that the training would be clinically relevant. The email included the Participant Information Sheet (Appendix 2.2) and details regarding the training. Participants signed up for the study by email.

Inclusion criteria

Participants were required to meet the PCL: YV user criteria for clinical settings, with the exception of criteria 3 and 5 which were met by attending the training event.

1) Possess an advanced graduate degree in the social, medical or behavioural sciences;

2) Have appropriate professional credentials (e.g. be registered, licensed, or legally entitled to conduct psychological assessments) or be working under the direct supervision of a registered professional;

3) Be familiar with the clinical and research literature pertaining to psychopathy, adult and adolescent;

4) Experience working with adolescents or completion of graduate courses in adolescent development;

5) Adequate training and experience in the PCL: YV.

Training event

Participants attended a two-day training event on the PCL: YV, delivered by a Consultant Clinical Forensic Psychologist with relevant experience. This included didactic teaching and discussion to familiarise participants with the PCL: YV. Consistent with the PCL: YV training recommendations, the training covered the nature and assessment of psychopathy and the PCL: YV assessment procedure and scoring (Forth et al., 2003). During the last day participants provided written consent to participate in the study (Appendix 2.3). A fully crossed design was used whereby all participants rated all six case vignettes using the PCL: YV. Participants also completed a purpose designed Staff Information Questionnaire (Appendix 2.5) developed to gather professional information. The order of completion of the case vignettes was randomised. Participants took approximately four hours to rate the case vignettes.

Participants

Twenty-one participants volunteered to participate in this study. Two participants withdrew due to clinical commitments. Nineteen (3 male and 16 female) professionals participated in this study with the majority recruited from NHS Greater Glasgow & Clyde[2]. Participants worked in an Adolescent Inpatient Service (n=6), Looked After and Accommodated Children's Service (n=3), Forensic Child and Adolescent Mental Health Service (n=3), relevant university departments such as the Centre for Youth and Criminal Justice (n=2) and adult Forensic Mental Health (community and secure inpatient) Services (n=5). Participants included Psychiatrists (n=9), Clinical Psychologists (n=7), Forensic Psychologist (n=1), Social Worker (n=1) and Child Psychotherapist (n=1). The sample included fully qualified staff (n=13) and staff in training (n=6). All participants had experience working with adolescents (M = 5.84, SD = 6.17, range = 0.5-25 years). Twelve had experience working in forensic mental health (M= 1.85, SD=2.58, range= 0.25– 9 years). None of the participants had received training in the PCL: YV or used it in clinical practice. All but one met the PCL: YV user qualifications, however this individual met the research qualifications and this was considered sufficient for participation.

---

[2] To retain confidentiality, the health board of each participant could not be reported.

Statistical analyses

Decisions regarding statistical analyses were based on the relevant statistical literature and advice from independent statisticians from the Robertson Centre for Biostatistics at the University of Glasgow. The analyses were conducted using SPSS (Statistical Package for the Social Sciences) for Windows, version 21 and Microsoft Excel, 2010.

A combination of inferential and descriptive statistics was used to address the research questions. The primary research questions (1 and 2) were addressed using Intraclass Correlation Coefficients (ICC) and percentage agreement statistics. These were calculated for individual items, total and factor scores across all cases. Question 3 was addressed by comparing percentage agreement statistics across each of the six case vignettes. Percentage agreement and inferential statistics including Correlations, Independent-Samples t-tests and Mann-Whitney U tests were conducted for question 4.

The total and factor scores are provided separately throughout the analyses as the total score ranges from 0 to 40, whereas factors 1 and 2 scores range from 0 to 8 and factors 3 and 4 scores range from 0 to 10. It should be noted that factor 1 (Interpersonal) includes items 1, 2, 4 and 5 and factor 2 (Affective) includes items 6, 7, 8 and 16. Factor 3 (Behavioural) includes items 3, 9, 13, 14 and 15 and factor 4 (Antisocial) includes items 10, 12, 18, 19 and 20. Due to copyright restrictions, only six PCL: YV items could be cited. The three items with the highest agreement and three items with the lowest agreement are discussed. The remaining items are identified by their number only.

PCL: YV scoring & prorating procedure

The PCL: YV total score was calculated by totalling the raw scores from the twenty items. The four factor scores were calculated by totalling the relevant items noted above. The prorating procedure outlined in the PCL: YV manual (Forth et al., 2003) was used to adjust the total and factor scores for missing items. This involved identifying the number of omitted items for each and consulting the appropriate table to identify the adjusted score. Thus, there was no missing data for total or factor scores as this was managed using this procedure. The methods used to manage missing data at the item level are described below.

Missing data

Overall, 13 out of 2280 scores were omitted, resulting in 0.57% of missing data. Data were missing for eight of the twenty items. This included one or two missing scores for these items, with the exception of Item 10 (Poor Anger Control) which had four missing scores. There was a relatively even spread of missing data across most cases. However, Case 2 (low severity) had no missing data and Case 3 (moderate severity) had more missing data.

When computing ICC's, a single missing score resulted in the deletion of the whole case vignette from that items analysis. Such a reduction in the number of cases would have a substantial impact on power. Thus, consistent with the PCL: YV manual, a sensitivity analyses was conducted using several methods to manage missing data. This included omitting raters with missing data on an item-by-item basis, replacing missing data with the value of 1 or with the mean score for that item across the whole sample. When these different methods were used, at most there was a highly negligible difference in the results. To remain consistent with the PCL: YV manual, missing values were replaced by the mean value for that item.

Intraclass correlations

The Intraclass correlation coefficient (ICC) measures the level of agreement, corrected for chance by comparing the variability of different raters rating the same item with the total variation across all raters and items (Hallgren, 2012). Whilst alternative analytic methods were considered (Light (1971) generalized kappa coefficient, Fleiss (1971) generalized kappa coefficient, weighted Kappa and Krippendorff's alpha), the ICC was considered the more appropriate form of analyses. The ICC is the recommended statistic when there are multiple raters and the data is ordinal (Uebersax, 2014). The ICC is commonly used in the PCL: YV literature, thus allowing comparability across studies.

Different ICC's are available for different methodological designs. The two-way random effects model (ICC2) was deemed the most appropriate as every rater assessed all cases and we were interested in the effects of both the raters and cases (Gwet, 2012). The ICC"s were calculated for absolute agreement for single measures. Predetermined benchmarks recommended by Fleiss (1981) were used to provide a qualitative evaluation of the level of agreement. These were: ICC < .40 = "poor", ICC .40 - .75 = "intermediate to good" and ICC > .75 = "excellent". Furthermore, it is recommended that both ICC and raw agreement are

reported (Uebersax, 2014), thus percentage agreement was also calculated.

Percentage agreement

Percentage agreement statistics were calculated using the formula outlined in Araujo and Born (1985) in relation to the mean (rounded to the nearest whole score), modal and expert score. These were adjusted to account for missing values. Thus, the percentage agreement values represented the proportion of scores in agreement with the mean, mode and expert score respectively. Whilst percentage agreement statistics have received criticism for their lack of control for chance-agreement, this measure is intuitive and provides important information regarding raw agreement at a practical level (Uebersax, 2014). Furthermore, the values are considered relative to one another in terms of individual items, cases and raters.

The mean and modal scores might be considered as indicating different assessment methods used in clinical practice. The mean value represents the score which may be achieved through discussion and negotiation within a team whereas the modal value represents the most common score assigned by the team.

Results

Research question 1

What level of inter-rater reliability do multidisciplinary professionals achieve using the PCL: YV?

Tables 1 to 3 report the range of percentage agreements with the mean and mode scores (rounded to the nearest percentage) and the ICC's for individual items, factors and total scores across all cases and raters. For ICC's the 95% confidence interval and number of missing scores replaced by the mean score for each item are provided.

As shown in Table 1, the average percentage agreements with the mean and mode were both 80% and ranged from 63% to 93%. The ICC"s ranged from .41 to .94 indicating "intermediate" to "excellent" levels of agreement, with these values positively correlated with the percentage agreement with the mean (Spearman's Rho = .82, $p < .01$), mode (Spearman's Rho = .91, $p < .01$) and expert (Spearman's Rho = .77, $p < .01$).

The overall level of agreement calculated using the ICC was conducted by stacking the raw scores for all twenty items for each case as separate items across all raters. This resulted in an ICC2 of .78 indicating "excellent" reliability overall.

Items 2 (Grandiose sense of self-worth), 6 (Lack of remorse) and 8 (Callous/lack of empathy) achieved the highest percentage agreement with the mean and mode, with each scoring above 85% agreement. The same items achieved the highest reliability using ICC's, with each achieving an ICC2 above .90. However, Items 10 (Poor anger control), 12 (Early behaviour problems) and 14 (Impulsivity) had the lowest percentage agreement, achieving around 70% or less agreement with the mean and mode. These items also had the lowest reliability using ICC's, producing ICC's less than .65.

As shown in Table 2, the percentage agreement with the mean and mode for the total score was 13% and 24% respectively. The percentage agreements with the mean and mode were lower for factors 3 (Behavioural) and 4 (Antisocial) relative to factors 1 (Interpersonal) and 2 (Affective). As illustrated in Table 3, the ICC"s for the total score was .96 with factors scores obtaining ICC's ranging from .86 to .95. The same pattern was evident with factors 3 and 4 achieving lower ICC's relative to factors 1 and 2; however these were still within the "excellent" range.

Table 1. Percentage agreement and ICC2 for individual PCL: YV items.

Item No.	Percentage Agreement				ICC2	
	Mean	Mode	Expert	Missing Score (n)	ICC2	95% CI
1	84%	84%	84%	0	.81	.60 - .96
2	88%	88%	87%	0	.92	.80 - .99
3	76%	76%	69%	1	.73	.49 - .94
4	86%	86%	65%	1	.83	.64 - .97
5	86%	86%	71%	1	.87	.72 - .98
6	90%	90%	90%	0	.92	.82 - .99
7	78%	78%	77%	0	.82	.62 - .97
8	93%	93%	93%	0	.94	.86 - .99
9	77%	77%	67%	0	.81	.60 - .96
10	70%	70%	36%	4	.41	.19 - .82
11	79%	79%	77%	2	.78	.57 - .96
12	63%	63%	47%	0	.56	.30 - .89
13	74%	74%	64%	1	.75	.52 - .95
14	66%	66%	50%	2	.62	.37 - .91
15	76%	76%	66%	0	.74	.51 - .95
16	84%	85%	84%	0	.87	.71 - .98
17	76%	79%	79%	0	.81	.62 - .96
18	83%	83%	76%	0	.80	.59 - .96
19	84%	84%	84%	0	.79	.57 - .96
20	84%	84%	68%	1	.79	.58 - .96
Mean	**80%**	**80%**	**72%**	**1**	**.78**	**.73 - .83**

Table 2. Percentage agreement for PCL: YV total and factor scores.

	Percentage Agreement		
	Mean	**Mode**	**Expert**
Total Score	13%	24%	15%
Factor 1 – Interpersonal	51%	64%	51%
Factor 2 – Affective	68%	69%	68%
Factor 3 – Behavioural	27%	38%	21%
Factor 4 - Antisocial	28%	36%	22%
Mean	**37%**	**46%**	**35%**

Table 3. ICC2 for PCL: YV total and factor scores.

	ICC2	95% Confidence Interval
Total Score	.96	.90 - .99
Factor 1 – Interpersonal	.94	. 85 - .99
Factor 2 – Affective	.95	.89 - .99
Factor 3 – Behavioural	.89	.74 - .98
Factor 4 - Antisocial	.86	.69 - .97
Mean	**.92**	**.90 - .94**

Research question 2

To what degree do multidisciplinary professionals agree with expert ratings using the PCL: YV?

The average percentage agreement with the expert score was 72% with individual items achieving between 36% and 93% agreement (Table 1). Thus, overall participants agreed less with experts than with one another. Items 2 (Grandiose sense of self-worth), 6 (Lack of remorse) and 8 (Callous/lack of empathy) also achieved the highest percentage agreement with the expert, obtaining over 85% agreement.

Consistent with the findings for the mean and mode, Items 10 (Poor anger control), 12 (Early behaviour problems) and 14 (Impulsivity) had the lowest agreement, achieving approximately 50% or less agreement with the expert. Visual inspection of these items revealed that the experts and participants different mostly on the high severity cases, with participants assigning higher scores than experts.

As reported in Table 2, 15% were in agreement with the expert for the total score, with percentage agreement ranging between 21% and 68% for the factor scores. Again, factors 3 and 4 achieved lower agreement with the expert, relative to factors 1 and 2.

Secondary research questions

Research question 3

Is the inter-rater reliability of the PCL: YV associated with the severity of psychopathic traits?

Table 4 compares the average percentage agreement with the mean, mode and expert score, across all twenty items and raters for each case vignette. The highest agreement was found for the low severity cases, closely followed by the high severity cases, with all achieving over 70% agreement. The moderate severity cases (3 and 4) had the lowest level of agreement.

As shown in Table 5, the same pattern is evident at the level of total and factor scores with cases 3 and 4 obtaining lower agreement overall. However, this pattern is less clear cut at the individual case level, particularly for factors 3 and 4, where the level of agreement is more comparable across cases. The pattern of greater agreement for factors 1 and 2 relative to factors 3 and 4 was still evident when investigating the cases individually. However, there were exceptions to this in the moderate and high severity cases.

Table 4. Average % agreement (with mean, mode, expert) for the six cases.

	Case 1 Chris	Case 2 Steven	Case 3 Adam	Case 4 Frazer	Case 5 John	Case 6 Danny
Severity of Psychopathic Traits	Low	Low	Mod	Mod	High	High
% agreement - mean	93%	87%	66%	68%	87%	78%
% agreement - mode	93%	87%	67%	68%	87%	78%
% agreement - expert	89%	82%	60%	59%	73%	71%

Table 5. Average % agreement (with mean, mode, expert) for total and factor scores for the six cases.

	case 1 Chris	case 2 Steven	case 3 Adam	case 4 Frazer	case 5 John	case 6 Danny
Severity of psychopathic traits	**low**	**low**	**Mod**	**Mod**	**High**	**High**
PCL:YV total						
% agreement -mean	11%	11%	11%	5%	26%	11%
% agreement - mode	32%	21%	11%	21%	26%	21%
% agreement - expert	26%	11%	11%	21%	16%	11%
factor 1 - interpersonal						
% agreement – mean	100%	95%	42%	26%	21%	21%
% agreement – mode	100%	95%	42%	32%	58%	58%
% agreement - expert	100%	95%	5%	26%	58%	21%
factor 2 - affective						
% agreement – mean	84%	100%	26%	37%	79%	84%
% agreement – mode	84%	100%	32%	37%	79%	84%
% agreement - expert	84%	100%	26%	37%	79%	84%
factor 3 - behavioural						
% agreement – mean	21%	26%	26%	37%	26%	26%
% agreement – mode	58%	32%	26%	37%	47%	26%
% agreement – expert	21%	26%	0%	37%	26%	16%
factor 4 antisocial						
% agreement – mean	47%	21%	16%	21%	53%	11%
% agreement – mode	47%	26%	26%	32%	53%	32%
% agreement – expert	42%	21%	26%	21%	11%	11%
Mean	**57%**	**52%**	**22%**	**28%**	**44%**	**34%**

Research question 4

Is the inter-rater reliability of the PCL: YV associated with the raters professional characteristics?

To obtain a variable representing the level of agreement at an individual level, percentage agreement statistics were conducted for individual raters. Each individual's percentage agreement with the mean, mode and expert scores across all items and cases were calculated.

These continuous variables were correlated with the continuous professional variables to determine whether there was a significant relationship between them.

All continuous variables were tested for normality, with normality judged by the Shapiro-Wilk's test (indicated by a significance value of $\geq$ .05) and visual inspection of histograms, Q-Q plots and boxplots. Variables including average percentage agreement with the mean, mode and expert and confidence and comfort variables (in relation to their use of the PCL: YV) were normally distributed. The remaining continuous variables were not normally distributed. Thus, the non-parametric Spearman's Rho correlation was selected.

Where outliers were present, visual inspection indicated that these were true scores. None of their z-scores exceeded 3.29, and the mean and 5% trimmed mean values were similar indicating that these extreme values were not strongly influencing the mean (Field, 2012; Pallant, 2010). Thus, these data points were retained. An exception to this was the length qualified variable, therefore a sensitivity analysis was conducted.

As shown in Table 6, none of these relationships were statistically significant. However, when the analyses were re-run excluding the outliers for the length qualified variable, this revealed a negative correlation between length qualified and perceived difficulty administering the PCL: YV, $r = -.57$, $n = 15$, $p = .03$.

Table 6. Spearman's rho correlation coefficients (two-tailed): associations between professional variables and average % agreement with the mean, mode, expert, perceived comfort, confidence and difficulty using the PCL: YV.

	% Agree mean	% Agree mode	% Agree expert	Comfort	Confidence	Difficulty
Length of experience with adolescents	-.18 $p = .47$	-.07 $p = .79$	-.17 $p = .50$	-.03 $p = .89$	-.24 $p = .34$	-.06 $p = .81$
Length of experience in forensic settings	.05 $p = .85$	-.00 $p = .99$	.09 $p = .72$	.32 $p = .18$	.35 $p = .15$	-.06 $p = .81$
Length of time qualified	-.03 $p = .91$	.04 $p = .88$	-.12 $p = .63$	-.36 $p = .14$	-.23 $p = .37$	-.45 $p = .07$

** denotes a correlation significant at 0.01 level.

For categorical variables, where one or more group was not normally distributed, the non-parametric Mann-Whitney U Test was used. Otherwise, Independent-Samples t-tests were conducted, referring to Levene's test of equal variances to determine the appropriate statistic. As detailed below, the relationship between the professional variables and IRR were mostly non-significant (see Appendix 2.7 for the full reporting of non-significant findings).

Qualified vs. in training

Independent-samples t-tests found no statistically significant difference between fully qualified staff and staff in training in terms of their average percentage agreement with the mean, mode or expert score or confidence in the accuracy of their PCL: YV ratings. A Mann-Whitney U Test revealed that there were no significant differences between fully qualified staff and staff in training in terms of perceived comfort or difficulty using the PCL: YV.

Forensic experience vs. no forensic experience

Independent-samples t-tests revealed no significant difference between staff with and without experience working in forensic mental health in terms of their percentage agreement with the mean, mode and expert score. Furthermore, these groups were not significantly different in terms of perceived comfort, confidence or difficulty when using the PCL: YV.

Completion of formal training in the structured assessment of personality or risk

Independent-samples t-tests revealed no significant difference between those who had and had not received training in the structured assessment of personality or risk in terms of their percentage agreement with the mean, mode or expert. Further independent-samples t-tests revealed no significant difference in perceived comfort, confidence and difficulty administering the PCL: YV between these two groups.

Experience conducting structured assessments of personality

Independent samples t-test found no significant difference between those with experience administering structured assessments of personality, versus those without this experience in terms of their percentage agreement with the mean, mode or expert. However, those who had conducted structured assessments of personality (M=6.68, SD = .75) were more confident in the accuracy of their PCL: YV ratings relative to those without this experience (M=5.17, SD = 1.34), t (16) = 2.81, p = .01, d = .33. Those who conducted structured assessments as part of

their past or current role (*M*=7.33, SD=1.03) reported that they would feel more comfortable using the PCL: YV to assess psychopathy in adolescents, relative to those without this experience (*M*=5.54, SD=1.71), t (17) = 2.35, *p* = .03, *d* = .25. A Mann- Whitney U Test revealed no significant difference between these two groups in terms of perceived difficulty using the PCL: YV.

Discussion

To our knowledge, this is the first study to assess the inter-rater reliability (IRR) of the PCL: YV in such depth, particularly at the level of individual items. The secondary research questions investigated the effects of professional characteristics and severity of psychopathy on IRR. These are discussed in turn followed by an evaluation of the strengths and limitations of the study and directions for future research.

Primary research questions

This study found "excellent" IRR overall when using the PCL: YV. This is reflected by an ICC of .78 for the items collectively, and an ICC of .96 for the PCL: YV total score. This is consistent with the PCL: YV manual which reported ICC's ranging from .90 to .93 for the total score (Forth et al., 2003) and a meta-analysis reporting "very strong" IRR for the total score (Olver & Stockdale, 2010). Thus, this study extended these findings for the total score, when rated by several multidisciplinary professionals. However, all participants received PCL: YV training immediately prior to rating the case vignettes and this may have improved IRR.

It may be argued that information at the factor and individual item level is more clinically meaningful than the total score as these may inform tailored interventions and risk management strategies. All four factors obtained strong IRR, with the interpersonal and affective factors obtaining higher IRR than the behavioural and antisocial factors. This is in contrast to meta-analytic findings which found stronger reliability for the latter factors (Olver & Stockdale, 2010) and research which found stronger agreement for the latter factors on the PCL-R (Edens et al., 2010). Other studies, however found a similar pattern to that found in this study for the PCL: YV (e.g. Fink et al., 2012). These conflicting findings indicate that further research is warranted.

At an item level, there was considerable variation in IRR. The items, "Grandiose sense of self-worth", "Lack of remorse" and "Callous/lack of empathy" achieved the highest agreement. These traits are included within the interpersonal and affective facets and it may be argued that these are more clearly defined constructs which are considered more fixed personality traits. Thus, the raters may have found it easier to identify whether these traits were present or not. However, these findings are in contrast to the PCL: YV manual (Forth et al., 2003) which reports lower agreement for "Lack of Remorse" in their probation sample and lower agreement for "Callous/Lack of Empathy" in their community sample.

The items with least agreement in this study fell under the behavioural and antisocial factors and included "poor anger control", "impulsivity" and "early behaviour problems". It could be argued that poor anger control and impulsivity are more common in adolescents given that this developmental period is characterised by emotional lability. Thus, these traits may be viewed as more acceptable in this population and the raters may have found it difficult to distinguish between developmentally-appropriate behaviour and psychopathology. These traits may be considered more transient and are inferred from behaviours. Thus, judgement may play a greater role, making them more difficult to rate. Finally, the professional groups included may have had less exposure to these behaviours relative to other disciplines (e.g. nursing), that have more direct contact with patients. Thus, their perception of what is considered normal in terms of these traits may be more dependent on other factors. For the item "early behaviour problems" raters should consider behavioural problems before the age of 10. This caveat is not highlighted on the rating forms and perhaps not all raters recalled this instruction, leading to inconsistencies. It may be beneficial to include this caveat on the rating form to remind professionals when rating this item. The finding regarding lower IRR for the "impulsivity" item was also found in the PCL: YV manual (Forth et al., 2003), whereas "poor anger control" and "early behaviour problems" did not obtain lower levels of agreement within their samples.

This study found that on average the raters agreed more with one another than they did with the experts. However, the same pattern was evident with the items noted above also obtaining the most and least agreement with the expert. Inspection of these items revealed that this was primarily due to the high severity cases where the expert assigned lower scores than the raters. It may be that the expert's extensive experience in forensic services resulted in a higher normative level of what they would consider severe traits.

Secondary research questions

Due to the small number of cases, inferential statistics could not be used to address the effect of the severity of psychopathy on the level of agreement. Instead a comparison of percentage agreement statistics across cases was used. This revealed less agreement for the moderate severity cases, relative to the low and high severity cases. This makes intuitive sense as these cases may be less clear cut than those presenting with extreme traits. It should be noted that one of the experts commented that these cases were most representative of cases seen in clinical practice. Whilst, this finding is tentative and based on descriptive statistics, similar findings were shown for the Risk of Sexual Violence Protocol (Sutherland et al., 2012) and is therefore of interest.

In terms of rater characteristics, professional variables (length of time qualified, working with adolescents or working within forensic services) were not shown to be significantly associated with IRR. Whether the rater was qualified or not, had experience in forensic services, had received training (or had experience) in the structured assessment of personality or risk did not significantly influence IRR. However, those with experience in the administration of structured assessments of personality reported feeling more confidence in the accuracy of their ratings and more comfortable using the PCL: YV to assess psychopathy in youth. Thus, the applied use of such assessments in general influenced professionals' opinions regarding their ability to use the PCL: YV. Also, the longer professionals had been qualified, the easier they found it to administer the PCL: YV. However, confidence and comfort and ease of use do not necessarily equate with accuracy. Therefore, professionals with such experience should remain cautious when using the PCL: YV.

Strengths & limitations

Strengths of this study include the use of comprehensive case vignettes developed in collaboration with experts in the field, including the expert rating procedure used to enhance their authenticity. A large proportion of the sample were recruited from Specialist Children's Services that work with children and adolescents presenting with severe emotional and behavioural difficulties. Therefore, the sample can be considered reasonably representative

as youths displaying psychopathic traits may be most likely to be seen within such services. This is particularly true for the Forensic Child and Adolescent Mental Health Service.

The use of case vignettes, however, may be considered a limitation. Particular information is difficult to portray through written methods and raters may have felt emotionally detached from the cases in comparison to a clinical interview. This may be particularly true for individuals with psychopathic traits which may evoke strong emotional reactions in the rater and potentially influencing their ratings. Furthermore, these cases were pre-rated to reduce the demands on the expert raters. Whilst the experts provided feedback regarding these cases including any ratings which they disagreed with, it would have been more reliable to have had the experts rate the case vignettes from scratch.

The cases may have portrayed the psychopathic traits more clearly than what is seen in clinical practice, which may have increased IRR. In response to this criticism, the cases included extreme versions of the traits to provide a clear distinction between the different severities. As acknowledged by one of the experts the high severity case vignettes rarely present in clinical practice. However, it was noted that when they do, they cause significant problems for the system. Furthermore, this study was focused on IRR and did not aim to assess construct validity.

An increased number of cases would have been more representative of clinical practice and provided greater power. A weakness of the study was that the sample size calculation was based on the primary research questions and was underpowered for the secondary questions. Furthermore, one participant did not meet the full PCL: YV criteria for clinical use and this may be considered a weakness in terms of ecological validity. However, all met the requirements for research purposes. The short timeframe between attending the training and rating the case vignettes may have improved reliability with participants more likely to recall the newly acquired information. Conversely, they may not have had time to process and reflect on the information.

Future directions

Firstly, future research could benefit from increasing the authenticity of the case information using retrospective file reviews of real cases or video recordings of clinical interviews. It is unclear how these would influence reliability, plausibly in either direction. Secondly,

including a greater number of cases of moderate severity would allow IRR to be tested using inferential statistics and provide further investigation of this finding. Also, recruiting larger, more heterogeneous samples would allow for a more comprehensive evaluation of the effect of professional variables on IRR, including training and experience using the PCL: YV which was not possible in this study. Thirdly, research could investigate additional rater and case characteristics found to influence PCL-R ratings including adversarial allegiance (Murrie et al., 2011) and the raters" personality (Miller et al., 2011). The gender of the case may also be relevant as there is research to suggest psychopathy may present differently in females (Salekin et al., 2001). This in addition to our relative lack of research and hence our understanding of the construct in females may make them more difficult to assess, thus impacting on IRR. Whilst it was beyond the scope of this study, the inclusion of a qualitative element investigating the interpretation of PCL: YV scores would be useful as this is likely to play a role in clinical decision making. Finally, studies utilising the PCL: YV should seek to obtain a measure of IRR as part of their methodologies, provide transparent reporting of the rater and case characteristics involved in the IRR process and where possible calculate IRR at the level of individual items.

Conclusions

This study found evidence that the PCL: YV can obtain "excellent" IRR when used by multidisciplinary professionals. Whilst, the total and factor scores all achieved "excellent" IRR, there was less agreement for the behaviour and antisocial factors. There was variable agreement for individual items, with "Impulsivity", "Poor anger control" and "Early behaviour problems" obtaining lower levels of agreement. It is recommended that professionals using the PCL: YV pay particular attention to these items and that those providing training on the PCL: YV cover these items more thoroughly. Similarly, the lower agreement obtained for the moderate severity cases, highlights the need for close supervision or consultation from experts in this area when using the PCL: YV. Consistent with the recommendations by Olver & Stockdale (2010), it may be beneficial to conduct joint sessions with two clinicians rating the youth independently and resolving inconsistencies through discussion. Whilst this may be considered resource intensive, the potential implications of labelling a young person with psychopathy are significant. Thus, these precautions are important in ensuring that accurate assessments are made based upon the individual's developmental stage, and that youths are not labelled unnecessarily with the stigmatising label of psychopathy and the resulting clinical and social implications.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Araujo, J., & Born, D. (1985). Calculating percentage agreement correctly but writing its formula incorrectly. *The Behavior Analyst*, *8*(2), 207-208.

Boccaccini, M. T., Murrie, D. C., Clark, J. W., & Cornell, D. G. (2008). Describing, diagnosing, and naming psychopathy: How do youth psychopathy labels influence jurors? *Behavioral Sciences & the Law*, *26*(4), 487-510.

Cleckley, H. (1941). *The mask of sanity*. St. Louis, MO: Mosby.

Cooke, D. J., & Michie, C. (1999). Psychopathy across cultures: North America and Scotland compared. *Journal of Abnormal Psychology*, *108*(1), 58-68.

Cooke, D. J., & Michie, C. (2001). Refining the construct of psychopathy: Towards a hierarchical model. *Psychological Assessment, 13*(2), 171-188.

Cooke, D. J., Hart, S. D., Logan, C., & Michie, C. (2004). Comprehensive assessment of psychopathic personality–Institutional Rating Scale (CAPP-IRS). Unpublished manuscript.

Dawson, S., McCuish, E., Hart, S. D., & Corrado, R. R. (2012). Critical Issues in the Assessment of Adolescent Psychopathy: An Illustration Using Two Case Studies. *International Journal of Forensic Mental Health, 11*(2), 63-79.

Edens, J. F., Boccaccini, M. T., & Johnson, D. W. (2010). Interrater reliability of the PCL-R total and factor scores among psychopathic sex offenders: Are personality features more prone to disagreement than behavioral features? *Behavioral Sciences & the Law, 28*(1), 106-119.

Edens, J. F., Campbell, J. S., & Weir, J. M. (2007). Youth psychopathy and criminal recidivism: A meta-analysis of the psychopathy checklist measures. *Law and Human Behavior*, *31*(1), 53.

Edens, J.F., & Campbell, J.S. (2007). Identifying youths at risk for institutional misconduct: A meta-analytic investigation of the psychopathy checklist measures. *Psychological Services, 4*(1), 13-27.

Edens, J. F., Skeem, J. L., Cruise, K. R., & Cauffman, E. (2001). Assessment of "juvenile psychopathy" and its association with violence: a critical review*. *Behavioral Sciences & the Law, 19*(1), 53-80.

Edens, J. F., & Vincent, G. M. (2008). Juvenile psychopathy: A clinical construct in need of restraint? *Journal of Forensic Psychology Practice*, *8*(2), 186-197.

Field, A. (2013). Discovering statistics using IBM SPSS statistics. Sage.

Fink, B. C., Tant, A. S., Tremba, K., & Kiehl, K. A. (2012). Assessment of Psychopathic Traits in an Incarcerated Adolescent Sample: A Methodological Comparison. *Journal of Abnormal Child Psychology, 40*(6), 971-986.

Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological bulletin*, *76*(5), 378-382.

Fleiss, J. L. (1981). *Statistical methods for raters and proportions* (2nd ed). New York: John Wiley & Sons.

Forth, A. E., Kosson, D.S., & Hare, R.D. (2003). *Psychopathy Checklist: Youth Version (PCL: YV).* Toronto: Multi-Health Systems.

Frick, P. J., & White, S. F. (2008). Research review: The importance of callous unemotional traits for developmental models of aggressive and antisocial behavior. *Journal of Child Psychology and Psychiatry, 49*(4), 359-375.

Frick, P., & Hare, R.D. (2001). *The Antisocial Process Screening Device*. Toronto: Multi-Health Systems.

Gwet, K.L. (3rd ed). *Handbook of Inter-Rater Reliability: The Definitive Guide to Measuring the Extent of Agreement Among Multiple Raters.* LLC: Advanced Analytics.

Hallgren, K. A. (2012). Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial. *Tutorials in quantitative methods for psychology, 8*(1), 23-34.

Hare, R.D. (1993).*Without conscience: The disturbing world of the psychopaths among us.* New York: Guilford Press.

Hare, R. D. (1991).*Manual for the Revised Psychopathy Checklist*. Toronto, Multi-Health Systems.

Hare, R. D. (2003).*Manual for the Revised Psychopathy Checklist* (2nd ed.). Toronto, Multi-Health Systems.

Hare, R. D., & Neumann, C. S. (2005). Structural models of psychopathy. *Current Psychiatry Reports, 7*, 57–64.

Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures, 1*(1), 77-89.

Hintze, J. (2008). PASS (Power and Sample Size for Windows). NCSS, LLC. Kaysville, Utah, USA. Retrieved August 2013, from www.ncss.com.

Johnstone, L & Cooke, D.J. (2004). Psychopathic-like traits in childhood: conceptual and measurement concerns. *Behavioral Sciences & the Law 22* (1), 103-125.

Johnstone, L. (2011). Bergen Conference on the Treatment of Psychopathy. Retrieved December 2013, from http://vimeopro.com/bergenconference/bctp2011/video/32542222

Light, R. J. (1971). Measures of response agreement for qualitative data: Some generalizations and alternatives. *Psychological Bulletin*, *76*(5), 365-377.

Miller, A. K., Rufino, K. A., Boccaccini, M. T., Jackson, R. L., & Murrie, D. C. (2011). On individual differences in person perception: Raters" personality traits relate to their Psychopathy Checklist-Revised scoring tendencies. *Assessment*, *18*(2), 253-260.

Murrie, D. C., & Balusek, K. (2008). Forensic assessment of violence risk in adversarial proceedings: Pursuing objectivity and avoiding bias. *Journal of Forensic Psychology Practice, 7*(4), 141-153.

Murrie, D. C., Boccaccini, M. T., Turner, D. B., Meeks, M., Woods, C., & Tussey, C. (2009). Rater (dis) agreement on risk assessment measures in sexually violent predator proceedings: Evidence of adversarial allegiance in forensic evaluation? *Psychology, Public Policy, and Law, 15*(1), 19.

Murrie, D. C., Boccaccini, M. T., Guarnera, L. A., & Rufino, K. A. (2013). Are forensic experts biased by the side that retained them? *Psychological science, XX*(X), 1-9.

Olver, M., & Stockdale, K. (2010). Psychopathy and youth violence: research, controversies and clinical utility. *The British Journal of Forensic Practice*, *12*(2), 3-13.

Pallant, J. (2010). *SPSS survival manual: A step by step guide to data analysis using SPSS.* McGraw-Hill International.

Patrick, C. J. (2006). Back to the future: Cleckley as a guide to the next generation of psychopathy research. In C.J. Patrick (Ed.), *Handbook of Psychopathy*, p 605-617). New York: Guilded Press.

Persons, J. B., & Bertagnolli, A. (1999). Inter-rater reliability of cognitive-behavioral case formulations of depression: A replication. *Cognitive Therapy and Research*, *23*(3), 271-283.

Rosenfeld, B., & Penrod, S. D. (2011). *Research Methods in Forensic Psychology.* John Wiley & Sons.

Salekin, R. T., Rogers, R., & Machin, D. (2001). Psychopathy in youth: Pursuing diagnostic clarity. *Journal of Youth and Adolescence, 30*(2), 173-195.

Skeem, J. L., & Cooke, D. J. (2010). Is criminal behavior a central component of psychopathy? Conceptual directions for resolving the debate. *Psychological assessment*, *22*(2), 433.

Skeem, J. L., Polaschek, D. L., Patrick, C. J., & Lilienfeld, S. O. (2011). Psychopathic personality bridging the gap between scientific evidence and public policy. *Psychological Science in the Public Interest, 12*(3), 95-162.

Sutherland, A. A., Johnstone, L., Davidson, K. M., Hart, S. D., Cooke, D. J., Kropp, P. R., ... & Stocks, R. (2012). Sexual violence risk assessment: An investigation of the interrater reliability of professional judgments made using the risk for sexual violence protocol. *International Journal of Forensic Mental Health, 11*(2), 119-133.

Uebersax, J. (2009). *Statistical Methods for Rater and Diagnostic Agreement.* Retrieved May, 2014, from http://www.john-uebersax.com/stat/agree.htm

Walter, S. D., Eliasziw, M., & Donner, A. (1998). Sample size and optimal designs for reliability studies. *Statistics in medicine, 17*(1), 101-110.

Chapter 3

Developing the Therapeutic Alliance: A reflective account

Abstract

Introduction – This reflective account focuses on the development of communication skills and more specifically the therapeutic alliance. Gibbs' (1988) model of reflective practice is used to provide a structured framework to this account. Reference is also made to Stoltenberg and Delworth's (1987) integrated developmental model of supervision, to conceptualise key learning experiences from a developmental perspective. **Reflections –** This account reflects on experiences spanned across my clinical training. In particular, my experience of working with a patient during my first year forensic mental health placement is discussed, in addition to my experience of working with a young person with significant communication difficulties. This is followed by a review of personal strengths and areas for development and finally a meta-reflection of the process of writing this account.

Chapter 4

Experiences of Consultation: A reflective account

Abstract

Introduction – This reflective account focuses on my experiences of providing consultation, what I have learned from those experiences and the ways in which they have influenced my professional development. Atkins and Murphy's (1994) cycle of reflection is used to provide a structured framework to this account. However, reference is also made to Stoltenberg and Delworth's (1987) Integrated Developmental Model, where appropriate. **Reflections** – Throughout this account, I reflect upon two experiences of providing consultation during my second year, learning disability placement. The first experience involved working with a large care team; however I had no direct involvement with the patient. The second experience involved working primarily with key members of staff in addition to observations of the patient. To conclude, a meta-reflection of the process of writing this account is provided.

Appendix 1.1 Journal submission guidelines

Advice to authors on preparing a
manuscript for *The Journal of Forensic
Psychiatry and Psychology*

NB: Please follow any specific instructions for authors provided by the Editor of the journal

Font: Times New Roman, 12 point. Use margins of at least 2.5 cm (1 inch).

Title: Bold, type the first word and proper nouns only in capital letters. Any sub-title should follow a colon and every word should be lower case (except proper nouns).

Authors' names: Give the names of all contributing authors on the title page exactly as you wish them to appear in the published article.

Affiliations: List the affiliation of each author (department, university, city, country).

Correspondence details: Please provide an institutional email address for the corresponding author. Full postal details are also needed by the publisher, but will not necessarily be published.

Anonymity for peer review: Ensure your identity and that of your co-authors is not revealed in the text of your article or in your manuscript files when submitting the manuscript for review. Advice on anonymizing your manuscript is available here.

Abstract: Indicate the abstract paragraph with a heading or by reducing the font size. Advice on writing abstracts is available here.

Keywords: Please provide three to six keywords to help readers find your article. Advice on selecting suitable keywords is available here.

Headings: Please indicate the level of the section headings in your article:
• First-level headings (e.g. Introduction, Conclusion) should be in bold, with an initial capital for each main word.
• Second-level headings should be in bold italics, with an initial capital letter for each main word.
• Third-level headings should be in italics, with an initial capital letter for each main word.
• Fourth-level headings should also be in italics, at the beginning of a paragraph, with only the first word capitalized (except proper nouns). The text follows immediately after a full stop (full point) or other punctuation mark.

Tables and figures: Indicate in the text where the tables and figures should appear, for example by inserting [Table 1 near here]. The actual tables and figures should be supplied either at the end of the text or in a separate file as requested by the Editor. Ensure you have permission to use any figures you are reproducing from another source. Advice on artwork is available here.

Running heads and **received dates** are not required when submitting a manuscript for review.
If your article is accepted for publication, it will be copy-edited and typeset in the correct style for the journal.
If you have any queries, please contact us at authorqueries@tandf.co.ukmentioning the full title of the journal you are interested in, or see our Author Services homepage. ,

Taylor & Francis Standard Reference Style: APA

APA (American Psychological Association) references are widely used in the social sciences, education, engineering and business. For detailed information, please see the Publication Manual of the American Psychological Association, 6th edition, http://www.apastyle.org/ and http://blog.apastyle.org/.

Appendix 1.2 Quality assessment tool

Quality Assessment Tool

(based on Downs & Black, 1998 and DuRant, 1994)

Introduction	
1. Have the authors established a theoretical framework/ rationale for their study?	2 = comprehensive rationale 1 = adequate rationale 0 = rationale unclear
2. Do the research questions/hypotheses logically flow from the theoretical model/rationale?	2 = all hypotheses/questions logically flow 1 = majority of hypotheses/questions logically flow 0 = little or no logical flow evident
Methods	
3. Are the design/methods selected appropriate to adequately test the research questions?	1 = Yes 0 = No 0 = Unable to determine
4. Is there evidence of protection of subjects – i.e. approved by an institutional review board?	2 = Reviewed by ≥ 2 1 = Reviewed by 1 0 = Not reported / Unable to determine
5. Is the study sample(s) appropriate to the problem being studied or the hypothesis being tested?	1 = sample clearly related to rationale 0 = sample appears unrelated
6. Was there a comparison (i.e. a control group or comparative data)	1 = Yes 0 = No
7. How were the comparison (i.e. control) subjects selected?	1 = Comparison subjects pre-selected prior to assessment of variables 0 = Comparison subjects defined after assessment of variables 0 = Not reported / Unable to determine
8. Were the subjects approached to participate in the study representative of the entire population from which they were recruited? (i.e. considered more representative if they include the entire source population, an unselected sample of consecutive patients, or a random sample)	4 = geographical cohort 3 = random sample 2 = convenience sample 1 = volunteer sample 0 = unable to determine
9. Were those subjects who were prepared to participate representative of the entire population from which they were recruited? (i.e. the proportion who agreed to participate should be stated. Validation that the sample was representative would include demonstrating that the distribution of the main confounding factors was the same in the study sample and the source population)	1 = representativeness of sample clearly demonstrated 0 = representativeness of sample not demonstrated
10. Were the cases and controls recruited from the same population?	1 = Yes 0 = No 0 = Unable to determine
11. Were the experimental conditions the same for cases and controls? (e.g. were they recruited over the same time period, was the time period between administration of measures the same etc)	1 = Yes 0 = No 0 = Unable to determine
12. Were controls matched to cases?	2 = matched on ≥ 2 variables 1 = matched on 1 variable 0 = not matched 0 = Unable to determine
13. What kind of population do cases represent?	1 = appropriate heterogeneity 0 = restrictively homogenous 0 = Unable to determine

(i.e. are they a heterogeneous representation of the disease/outcome in question or are they a highly selective population with limited generalizability)	
14. Have the number of non-respondents/refusals/subjects lost to follow-up been kept small (<10%)	1 = Yes 0 = No 0 = Unable to determine

Measurement

15. What method of assessment was employed for psychopathy?	3 = Both Self-report and clinician rated assessment/interview 2 = Clinician rated assessment / interview 1 = Self-report 0 = Unable to determine
16. If clinician measure used for psychopathy, was a measure of inter-rater reliability employed?	1=Yes 0=No 0=Unable to determine or N/A
17. If clinician measures were used for psychopathy, is there evidence that they were trained in the measure?	1=Yes 0=No 0=Unable to determine or N/A
18. Are the tools/assessments used to measure psychopathy valid and reliable? *(i.e. have they undergone validity and reliability testing? If more than one measure used, 1=majority of measures valid & reliable)*	1=Yes 0 = No 0 = Unable to determine
19. What method of assessment was employed for posttraumatic stress/acute stress symptoms?	3 = Both Self-report and clinician rated assessment/interview 2 = Clinician rated assessment / interview 1 = Self-report 0 = Unable to determine
20. If clinician measures used for posttraumatic stress, was a measure of inter-rater reliability employed?	1= Yes 0= No 0 = Unable to determine or N/A
21. If clinician measures were used for posttraumatic stress, is there evidence that they were trained in the measure?	1= Yes 0 = No 0 = Unable to determine or N/A
22. Are the tools/assessments used to measure posttraumatic stress / acute stress symptoms valid and reliable? *(i.e. have they undergone validity and reliability testing? If more than one measure used, 1=majority of measures valid & reliable).*	1 = Yes 0 = No 0 = Unable to determine
23. Were the outcome variables measured using appropriate 'blinded' methods?	1 = Yes 0 = No 0= Not Reported / Unable to determine
24. Were any efforts made to control for recall bias? *(i.e. are multiple methods used, file review information or collateral information used to measure important variables that could be influenced by recall bias?)*	1 = efforts to control for recall bias are evident. 0 = efforts to control for recall bias are not evident 0 = Unable to determine or N/A

Results

25. Were the statistical tests chosen to analyse the data appropriate?	1 = tests selected clearly consistent with research q's 0 = consistency with research q's
26. Is there sufficient analysis to determine whether significant	1= confounding variables addressed 0 = no evidence of confounding variables being

differences may be due to some other variable (e.g. lack of comparability of the groups in sex, age or clinical characteristics or in other relevant variables)	addressed
27. Are adequate summary data presented (i.e. are continuous level data presented as means & standard deviations)? (including controls and/or normative scores if comparisons made)	1= Yes 0 = No
28. Is the study sample large enough to test the hypotheses – i.e. did it have sufficient power to detect a clinically important effect where the probability value for a difference being due to chance is less than 5%	1= Yes 0 = No 0 = Unable to Determine
Discussion	
29. Are conclusions substantiated by the data that are presented in the results section?	1 = conclusions clearly drawn from results presented 0 = conclusions drawn from sources other than data generated
30. Are generalizations confined to the population from which the sample was drawn?	1= generalisations not confined to population from which the sample was drawn. 0 = generalisations confined to the population from which the sample was drawn. 0 = Unable to determine.
31. Are the limitations of the study considered and are they taken into consideration when conclusions are drawn?	1 = potential influence of limitations on results acknowledged 0 = potential influence of limitations on results not acknowledged

Appendix 2.1 Ethics approval letter

University of Glasgow | College of Medical, Veterinary & Life Sciences

16th December 2013

Dear Professor Davidson

MVLS College Ethics Committee

Project Title: An investigation of the Inter-Rater Reliability of the Psychopathy Checklist: Youth Version (PCL: YV): effects of rater and case characteristics.

Project No: 200130027

The College Ethics Committee has reviewed your application and has agreed that there is no objection on ethical grounds to the proposed study. It is happy therefore to approve the project, subject to the following conditions:

- Project end date: **31 August 2014**

- The research should be carried out only on the sites, and/or with the groups defined in the application.

- Any proposed changes in the protocol should be submitted for reassessment, except when it is necessary to change the protocol to eliminate hazard to the subjects or where the change involves only the administrative aspects of the project. The Ethics Committee should be informed of any such changes.

- If the study does not start within three years of the date of this letter, the project should be resubmitted.

- You should submit a short end of study report to the Ethics Committee within 3 months of completion.

Yours sincerely

Dorothy McKeegan

Dr Dorothy McKeegan
College Ethics Officer

Dr Dorothy McKeegan

Senior Lecturer

R303 Level 3
Institute of Biodiversity Animal Health and Comparative Medicine
Jarrett Building
Glasgow G61 1QH Tel: 0141 330 5712
E-mail: Dorothy.McKeegan@glasgow.ac.uk

Appendix 2.2 Participant information sheet

PARTICIPANT INFORMATION SHEET

An investigation of the Inter-Rater Reliability of the Psychopathy Checklist: Youth Version (PCL: YV): effects of rater and case characteristics.

1. What is this study about?

The concept of callous-unemotional traits is one that is becoming embedded in practice with conduct disordered children and young offenders. It is soon to be included in the DSM-V as a diagnostic specifier. It is therefore essential that clinicians that rate youth as having these difficulties do so with reliable and valid measures. We are therefore inviting you to take part in our study which aims to determine the inter-rater reliability of the PCL: YV – a recommended procedure for identifying callous-unemotional traits.

Inter-rater reliability is the degree of agreement among raters when using a measure independently. For any part of a diagnosis or construct to have use, it must be reliably assessed. The PCL: YV is an expert rated measure used to assess a range of traits in young people which correspond to the model of callous-unemotional traits and psychopathy. Although many studies have assessed inter-rater reliability as part of their research, no study has focused exclusively on this. Furthermore, the methods used to assess inter-rater reliability in many of the existing studies are often not representative of the heterogeneous characteristics of clinicians and young people seen in clinical practice. In addition, despite being examined in the adult field, no existing research using youth protocols has considered the effects of individual characteristics on rating.

As such, this study is broad in its scope and will explore the extent to which raters agree with expert consensus judgements, the effects of rater and patient characteristics on PCL: YV ratings and inter-rater reliability.

2. What will happen if I agree to take part?

If you agree to take part you will be asked to attend one of the two day training events provided by this study. These will be held on [Insert Date-Date] and [Insert Date-Date] at [Insert Location]. The training will focus on the PCL: YV. On the final afternoon of the training, you will be asked to complete a shortened version of the PCL: YV for six fictitious case vignettes. You will also be asked to complete a short

Staff Information Questionnaire gathering demographic and professional information. Overall, it should take you approximately three hours for you to complete this study.

3. Do I have to take part?

No, it is up to you whether you want to take part. If you agree to take part, we will ask you to sign a consent form. You are free to withdraw from this study at any time, without giving a reason.

4. Will my participation be anonymous?

Yes. Throughout the study your data will be identified by a randomly assigned unique number. All information you provide will be stored in a locked filing cabinet within NHS premises or the Section of Psychological Medicine, Gartnavel Royal Hospital. An electronic copy of this information will be stored on a password protected NHS computer and University of Glasgow encrypted and password protected laptop for data analyses. After five years the data will be destroyed.

5. Potential disadvantages of taking part?

Taking part in the study requires a commitment of your time and effort. It is possible that you may find some of the details in the case vignettes distressing. However, as NHS employees working within mental health services, you are likely to be familiar with such material. We really appreciate you taking the time to participate in this study and understand that this would be a considerable addition to your workload.

6. Potential benefits of taking part?

By taking part in this study, you will be given the opportunity to attend a training event where you can develop your knowledge, understanding and skills in using the PCL: YV as well as the use (and misuse) of the construct of psychopathy. You will be able to list your attendance at this course as formal CPD. There are no immediate personal benefits from participating in this study,

7. What will happen to the results from this study?

Upon completion of this study, a report will be produced to describe the findings. No personally identifiable information will be used in any report or publication. You will also receive a summary of the results from which you will be able to extract your own results by searching for your unique number. You may request to see a copy of the final report.

8. Will I receive payment or expenses?

Unfortunately we are unable to provide participants with expenses or

payment for taking part in this study. However, training on the PCL: YV is being provided as part of this study and this contributes to savings to organisations.

9. Who has reviewed this study?

This study has been reviewed and given favourable opinion by the MVLS (Medical, Veterinary & Life Sciences) Ethics Committee for Glasgow University. It has also been approved by the University of Glasgow Doctorate of Clinical Psychology, Major Research Project submissions process.

10. Who is funding this research?

This study is funded by the University of Glasgow.

11. What do I do if I have any questions or complaints about this study?

If you have any questions or concerns about any aspect of this study, please contact Sarah Dickson. Alternatively, you can contact other members of the research team, Professor Kate Davidson or Dr Lorraine Johnstone. If you would like to speak to an independent person about this project then please contact Dr Hamish McLeod on 0141 211 3922. If you still have concerns and wish to complain formally, you can do this through the NHS Complaints Procedure.

12. Contact Details

Research Team

Principal Investigator:
Sarah Dickson
Trainee Clinical Psychologist
Tel: 0141 211 3552
Email: sarahdickson@nhs.net

Project Supervised by:
Dr Lorraine Johnstone
Consultant Clinical Forensic
Psychologist & Honorary
Research Fellow
Tel: 0141 276 3850
Email: Lorraine.Johnstone@ggc.scot.nhs.uk

Professor Kate Davidson
Consultant Clinical Psychologist &
Honorary Professor of Clinical Psychology
Tel: 0141 211 3900
Email: Kate.Davidson@glasgow.ac.uk

THANK YOU FOR READING THIS INFORMATION SHEET.

Appendix 2.3 Participant consent form

PARTICIPANT CONSENT FORM

An investigation of the Inter-Rater Reliability of the Psychopathy Checklist: Youth Version (PCL: YV): effects of rater and case characteristics.

Participant ID: _____

Please read each statement carefully and write your initials in the box if you agree with it.

1. I confirm that I have read and understand the Participant Information Sheet dated 21/10/2013 (version number 3) for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.

2. I understand that my participation is voluntary and that I am free to withdraw at any time without having to give a reason.

3. I understand that the information I provide will be confidential and that personally identifiable information will not be included in any report or publication of this study.

4. I agree to take part in the above study.

Name of Participant: **Date:**   **Signature:**

_____   _____   _____

Name of Principal Investigator: **Date:**   **Signature:**

_____   _____   _____

Please note: 1 copy to be retained by the participant and 1 to be returned to the research team.

Appendix 2.4 Example of case vignette (moderate severity)

1. REASON FOR REFERRAL & INFORMATION SOURCES

Adam was referred for a psychological assessment to try to gain some insight into his antisocial behaviour which most recently culminated in the assault of a pregnant member of staff.

The opinion presented below is based on the following information sources:

1) Two interviews with Adam
2) Interview with Mrs Bell, Adams maternal grandmother
3) Review of Social Work records
4) Conversations with Mr & Mrs Johnston, Adam's former foster carers
5) Conversations with staff from the residential unit
6) Review of psychiatric files, volume 1
7) Review of school reports

2. PRESENTATION AT INTERVIEW

At interview, Adam appeared slightly uncomfortable and was noted to be sitting forward in his seat with his hands clasped. As the interview progressed he appeared to relax, sat back in his seat and unzipped his jacket. Adam provided all information asked of him and seemed to be upfront and to the point regarding both positive and negative aspects of his behaviour. However, it was noted that he had a tendency to justify his antisocial behaviour. Adam appeared relatively flat in mood throughout both interviews, however displayed signs of anxiety when discussing particular topics including the recent assault on the pregnant woman. Adam provided information which was generally consistent with collateral information. He was oriented to place, time and person and there was no evidence suggestive of acute mental illness.

3. FAMILY LIFE

Adam is 15 years old and is currently accommodated in a local children's residential unit. Adam's mother, Miss Cara Bell fell pregnant with Adam when she was 18 and was living with Adams father, who was aged 27 years. Reports indicate that Adam's father left Miss Bell a few months after Adam was born and has never reinstated contact. Social work reports document Miss Bell's substance misuse, financial difficulties and promiscuous sexual behaviour during the first two years of Adam's life. Available health records highlight a series of appointments where Miss Bell failed to attend with Adam to receive his physical health check-ups. Following increasing concerns regarding Miss Bell's ability to care for Adam, a Children's Hearing was held when Adam was 3 years old and the decision was made for Adam to be placed in a kinship care with his maternal grandmother, Mrs Jean Bell.

Mrs Bell described Adam as a "hyper" child and said "he was always up to mischief". She recalled the unpredictable nature of Adam's behaviour and said that he had a tendency to get himself "hyped up" and was frequently verbally abusive towards her. According to Mrs Bell she struggled to cope with Adam's "outbursts" and claimed "he would be fine one minute then screaming the next". She admitted that she had purposefully avoided taking Adam out during busy times in case he would have one of his "meltdowns" in public. She said that she found this difficult to cope with and recalled Adam being referred to a Child Mental Health Service at the age of 6 regarding his behaviour. At this time, he was noted to have some attachment problems. Mrs Bell reported that her daughter had been "trying to get off the drugs" at this time and admitted to allowing her to have contact with Adam. She claimed that she now realises that this was unfair on Adam and said that her daughter repeatedly presented in an intoxicated state to see him. Consequently, a Children's Hearing concluded that Adam should be accommodated by the local

authority at the age of 7. It is noted that this was due to Mrs Bell's inability to safeguard Adam. Mrs Bell maintained that she had continued to have frequent contact with Adam since he was taken from her care. She reported that recently she has been in and out of hospital due to her physical health and had been unable to visit Adam as frequently. This has correlated with a change in Adam's behaviour. He frequently asks about his grandmother and her wellbeing. It is noted that Adam had intermittent telephone contact and supervised contact with his mother since the age of 12. He also asks about her on a regular basis and admits to worrying that something might happen to her.

From the ages of 7 to 14, Adam stayed with foster carers.   He found it difficult to develop relationships as he felt disloyal to his mother and grandmother. Adam's former foster carers Mr and Mrs Johnston described Adam as a "quiet". They claimed that he never asked for much and appeared to enjoy spending time by himself in his bedroom. However, they said he would thwart their attempts to get him to help out within the home. They also spoke of a spell when he had a short temper and frequently instigated heated arguments with them. It was noted that there was no specific trigger. Mrs Johnston recalled Adam running away on several occasions resulted in him being returned by the police several hours later. Adam was unable to provide an explanation as to why he had ran away other than that he "felt like it" at the time. He claimed to have no plan of where he was going and said that he never had much money to go any distance. He said he now realised he had been immature. Mrs Johnston maintained that they had tried to stay in contact with Adam since moving to the residential unit, however claimed that Adam usually refused to speak with them. She said that Adam would occasionally phone her and was good at making her feel guilty, usually by complaining about his care at the residential unit and claiming that he missed her.

When asked about his period in care, Adam was upfront about his verbal aggression and bad temper. He reported that he felt "guilty" for his behaviour. However, he claimed that this did not warrant him being moved into residential care and claimed "it wasn't as bad as they made out". Adam claimed that he was "close" to his grandmother and reported that he had been worried about her recently due to her hospital admissions. He said that he had had dreams about her regularly and he sometimes feels sick when he thinks that she might not get better. It was noted that Adam's demeanour changed when discussing this and he appeared anxious. Residential staff provided information consistent with Adam's account. They said that Adam frequently asked to phone his grandmother and had expressed his anxiety about his grandmother finding out about his recent assault, claiming that this would "stress her out". Adam claimed to have a positive relationship with his mum; however staff reported that Adam was often dismissive and confrontational during supervised contact with his mum. They said that Adam frequently changed his mind regarding contact with his mum. They said that he had a tendency to change his mind shortly before she was due to visit and would ask staff to phone her and tell her not to come. It is noted, that Adam sometimes appeared to regret his decision within a matter of hours.

4. DEVELOPMENTAL HISTORY

Social work records document Miss Bell's long-standing history of substance misuse and she admitted to drinking alcohol and taking drugs throughout her pregnancy. Adam was born four weeks premature by caesarean section, weighing 5lb and 7oz and after 10 days in hospital was discharged with Miss Bell. Mrs Bell was unable to recall precisely when Adam met his developmental milestones however she claimed that these were delayed. She estimated that he started talking when he was around one and a half years old.

She described Adam as a "good" baby and said "he'd go to anyone". However, she said that Adam was "a handful" during his toddler years and recalled his frequent tantrums and short temper. However, she said that she was not sure he was atypical of a "terrible two". She claimed that Adam also suffered from night terrors and nocturnal enuresis when he was fist accommodated with her and said that this continued until he started school.

5. SCHOOL HISTORY & ADJUSTMENT

It is noted that Adam was held back at nursery for a year prior to starting primary school, as it was felt that he was not ready for the transition to school. Due to the frequent changes in his foster placements, Adam attended three mainstream primary schools and a mainstream secondary school prior to attending the residential educational facility.

Whilst it would appear that Adam was well behaved at times, school reports document his oppositional nature and tendency to challenge his teachers. His teachers made reference to the "unpredictable" nature of his behaviour and claimed that he had a short temper. It would appear that this escalated into verbal abuse directed at his teachers on numerous occasions. It was also noted that Adam had thrown objects at his peers on a few occasions, when having one of his "tantrums". From primary 2 onwards, there were reports of Adam running out of the school on several occasions requiring the police to be called. When asked about this, Adam was unable to provide an explanation for this, other than that he "felt like it" at the time. Reports from secondary school document a similar pattern of behaviour however it was noted that there were more frequent reports of Adam threatening his peers with physical violence and making reference to being in a gang. However, it would appear that this did not escalate into any physical violence. In terms of his attendance rates, school reports state that Adam was frequently caught truanting during his secondary school education.

Adam was described as of "average" ability. His teachers said that he was capable of concentrating on his work but that he often failed to apply himself in several of his classes and frequently failed to submit homework assignments. When asked about school, Adam claimed that he could definitely have done better but just did not realise that school was so important. He said that he had always found most of his subjects "dull" however said that he enjoyed computing and claimed to know everything about them. Reports from the residential educational facility highlight Adam's lack of interest in his classes and his tendency to disengage from the work. School records indicate that Adam has now passed his standard grade exams with general and credit grades. Adam spoke about his academic achievements, however claimed that he wished he had worked harder.

6. WORK HISTORY & MONEY

Adam reported that he has never had a job but that he would like to gain employment in the future. He reported that he had enquired about working in the local shop near his former foster carer's home and had been informed that they would consider giving him a "Saturday shift". Adam claimed that he had been disappointed about his move to the residential unit and this lost opportunity. Mrs Johnston provided confirmation that Adam had enquired about this job.

7. LONG-TERM GOALS

When thinking about his future career prospects, Adam said that he would maybe like to pursue a vocational course and then perhaps consider a job such as building or plumbing. He commented that he had also thought about working with computers, but was unsure of whether he would be suited to an "office job".

When asked what qualifications he would require for this Adam shrugged his shoulders. He commented, "I've not looked into it properly yet" and said that he hadn't decided whether or not he would want to stay on at school for his final year. He commented, "I keep changing my mind".

When asked where he would like to live or whether or not he wanted a family, he laughed and said "I'm only 15, I'm not thinking about wains yet!" He then commented "honestly, I don't know".

8. PEER/SEXUAL RELATIONSHIPS & INTERPERSONAL TRAITS

It would appear that Adam has developed a heterosexual orientation. He reported that he first had sexual intercourse when he was 15 years old and said he had had sexual intercourse with 3 girls. He said that this had been within the context of intimate relationships however claimed that these were "nothing serious" with the longest lasting three weeks. He admitted to being unfaithful to all three girlfriends and said that he "couldn't help" himself when he had the opportunity to sleep with other girls. He recalled his former girlfriend becoming suspicious after looking through his phone. Adam reported "I just turned it on her and threatened to finish it if she couldn't trust me". It was noted that Adam appeared smug about this and went on to describe how his girlfriend had come "crawling back" to him and how he had slept with her later that evening. He appeared proud of his ability to lie to his girlfriends and claimed that they had never found out the full extent of his cheating.

When asked how his relationships ended, Adam claimed that he usually got "bored" and would stop contacting them. He said that on one occasion, he had ended a relationship on the basis that she had been "frigid" and was "shit in bed". He recalled this individual contacting him since and claimed to be "dingying" her phone calls. He said that he had not maintained contact with any of his ex-girlfriends but that he had kept their numbers for when he needed "a shag". Adam spoke about his peers at school and how he had managed to break up a young couple by telling the female that he had seen her boyfriend cheating on her. He admitted that this had not been the case, however claimed that he wanted to "bang" her and knew that was the only way to do it. There was no further information to suggest that Adam engaged in any unusual or excessive sexual behaviour.

In terms of peer relationships, Adam reported that he had made friends with a group of males during his first year at secondary school and claimed that they were "close". He spoke of his friends Sean, Marky and Stuart and said that they continued to text each other and meet up. He said that his mate Stuart was possibly moving down South and said that this would be "shit" if that were to happen. Residential staff corroborated Adam's account and said that he often spoke about these friends and that they sometimes visited him in the unit. The staff claimed that Adam also appeared to have made some genuine friendships with other young people in the unit and he was often seen socialising with the same peers. However, they noted that he had befriended a younger male and there was evidence to suggest that he had been using this individual to store stolen goods for him in the unit. They said that on one occasion, this individual had been found in possession of a staff member's wallet and they believed that Adam had been behind this.

9. SUBSTANCE USE & LIFESTYLE

Adam reported that he first started drinking alcohol when he was 14 years old. He said that he used to drink "most weekends" before being taken into the residential unit. He said that he also smoked cannabis around this time and that this had usually been when he skipped school. Adam denied taking any other types of drugs and there was no information available to dispute this.

With regards to his free time, Adam claimed that he enjoyed watching TV and playing his PlayStation. In particular he spoke about his love for comedy shows and named several comedians that he admired. Adam claimed that prior to being accommodated in the residential unit; he met up with his friends every weekend. Adam spoke about his time in the residential unit and expressed his annoyance at the excessive number of activities and outings which were planning, commenting "sometimes you just want to chill".

10. OTHER ANTISOCIAL BEHAVIOUR

It was noted that there were several referrals to the Children's Hearing System, the first when Adam was 11 years old. This involved referrals for shoplifting and substance use.

In terms of the index offence, Adam was alleged to have got into an argument with a residential staff member over his school timetable. It was noted that the situation escalated quickly and Adam starting shouting prior to lifting a nearby chair and throwing it towards her. This resulted in the staff member sustaining injuries to her upper body and head and she was off work. Adam claimed that "it all happened so quickly" and said that he regretted his actions and was feeling worried about the staff member returning to work. He said that he kept thinking about the incident for days after and upon discovering she was pregnant had "freaked out" in case she had lost her baby. Adam did not know that she was pregnant at the time as she had not told anyone. It was noted that Adam's demeanour changed when discussing this and he appeared anxious. According to residential staff Adam became withdrawn following the incident and was noted to ask staff if the victim was okay and when she would be back at work. The staff claimed that Adam appeared to get "a real fright" and they believed that Adam was sorry for his behaviour.

With regards to Adams general presentation on the unit, staff reported that he appeared flat and rarely appeared happy or excited by anything. They claimed that Adam could be oppositional and usually refused to complete the chores asked of him. They said that he could be "lazy" and said that he would often try and get other young people to do his homework. It would appear that this had happened on several occasions without staff knowing about it at the time. The staff claimed that they are now wise to this and ensure that a member of staff observes him doing this.

The staff reported that Adam had developed good relationships with certain staff members and appeared to take a dislike to others. This was apparent in staffs conflicting views, with some describing him as "manipulative" and others stating that he was a "nice" boy. It was noted that the staff had to provide comprehensive handovers to each other at the end of their shifts, as Adam had a tendency to lie about things to gain privileges such as additional TV time.

11. MENTAL HEALTH

A review of Adam's psychiatric files document involvement from a Child and Adolescent Mental Health Service (CAMHS) at the age of six and it was noted that his symptoms may have been indicative of Attention Deficit Hyperactivity Disorder (ADHD). However, it was noted that the lack of structure and neglect experienced in his early life may have impacted on his ability to regulate his emotions. More recently, Adam has been receiving clinical psychology input regarding his anger and currently appears to be engaging with this. At the beginning of this a Becks Youth Inventory (BYI) was completed and Adam scored highly on the anger, disruptive behaviour and self-concept domains. When asked whether he was happy or not, Adam claimed "I don't know, but I'm not unhappy". He said that he always felt "pissed off" however was unable to elaborate on this or describe his feelings any further.

Appendix 2.5 Staff information questionnaire

Staff Information Questionnaire

1. **What is your gender?**   Male ☐   Female ☐

2. **To which profession do you belong?**

Clinical Psychology ☐   Forensic Psychology ☐   Psychiatry ☐   Social Work ☐

Other, please state _____

   3. **Are you training in that profession?**   Yes ☐   No ☐

   If so, what is your training grade? _____

4. **Please list the professional qualifications you hold (e.g. BSc, MA, DClinPsy etc.)**

5. **If applicable, approximately how many years have you been fully qualified in that profession? (for medics, years since qualification?** _____

6. **Do you have experience working with adolescents?**   Yes ☐   No ☐

7. **If applicable, how many years experience do you have working with adolescents?** _____

8. **Do you have experience working in forensic mental health?** Yes ☐   No ☐

9. **If applicable, how many years experience do you have working in forensic mental health?**

10. **Have you completed any formal training in the structured assessment of personality or risk?**
                                                                    Yes ☐   No ☐
11. **If so, what formal training have your received?** _____

12. **Does your current or previous roles involve conducting structured assessments of personality?**   Yes ☐   No ☐

13. Work Experience (Please tick the appropriate boxes to indicate your experience)

|	What is your <u>current</u> post?	What services have you previously worked in?
Child Services		
CAMHS		
Child Inpatient Unit		
Adolescent Inpatient Unit		
Adult Services		
Adult Mental Health Service – Community		
Adult Mental Health Service – Inpatient		
Older Adult Mental Health - Community		
Older Adult Mental Health - Inpatient		
Forensic Services		
Forensic CAMHS		
Adult Forensic – Community		
Adult Forensic – Low secure		
Adult Forensic – Medium secure		
Adult Forensic Service – High secure		
Other Relevant Services		
Please state: _____		

For the following questions, please do **not** include the training provided as part of this study when providing your responses for the PCL: YV.
For those questions where a yes or no is required, please circle Y or N to indicate your answer.

|	Are you familiar with this measure?	Have you attended formal training in this measure?	No. of days formal training?	Have you used this measure in your clinical practice?	If yes, roughly how many times?	How useful was it? 1 → 5 Not Useful At all / Very Useful
PCL-R (Hare Psychopathy Checklist Revised)	Y / N	Y / N		Y / N		
PCL:SV (Hare Psychopathy Checklist, Screening Version)	Y / N	Y / N		Y / N		
PCL:YV (Psychopathy Checklist, Youth Version)	Y / N	Y / N		Y / N		
APSD (Antisocial Process Screening Device)	Y / N	Y / N		Y / N		
CPS (Child Psychopathy Scale)	Y / N	Y / N		Y / N		
YPI (Youth Psychopathic Traits Inventory)	Y / N	Y / N		Y / N		
PCS (Psychopathy Content Scale)	Y / N	Y / N		Y / N		
PPI (Psychopathic Personality Inventory)	Y / N	Y / N		Y / N		
CBCL (Child Behaviour Checklists)	Y / N	Y / N		Y / N		
CRS-R (Connors Rating Scales Revised)	Y / N	Y / N		Y / N		

14. <u>Psychopathy</u>

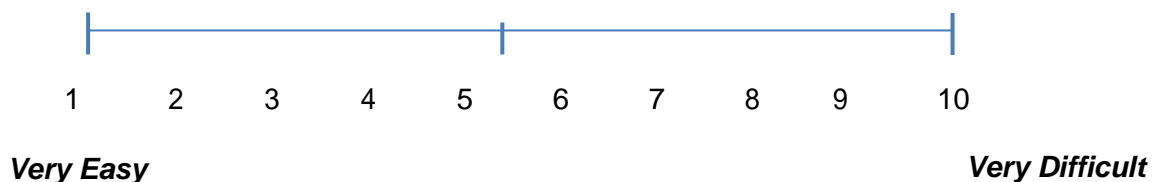**Please circle a number on the scales below to indicate your responses.**

How comfortable would you feel about using the PCL: YV to assess psychopathy in adolescents?

```
   1    2    3    4    5    6    7    8    9    10
```

I would not feel   *I would feel*
comfortable at all   *comfortable*
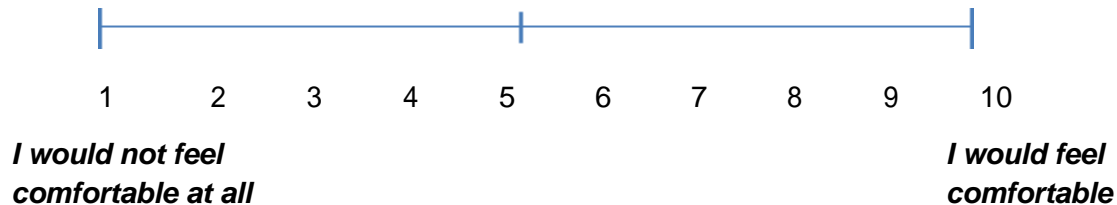
How confident were you overall in the accuracy of your ratings in this study?

```
   1    2    3    4    5    6    7    8    9    10
```

Not confident at all   *Very Confident*
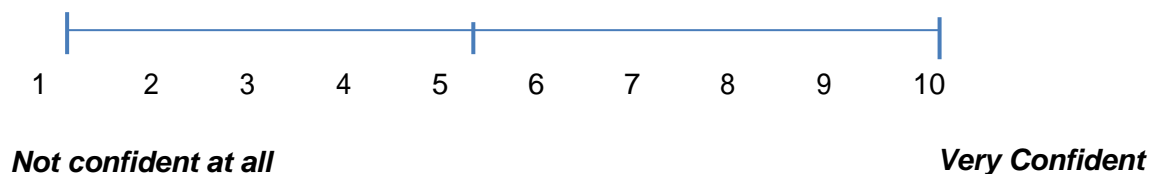
Did you find the PCL: YV easy or difficult to administer when assessing the level of psychopathic traits for the case vignettes?

```
   1    2    3    4    5    6    7    8    9    10
```

Very Easy   *Very Difficult*

<u>**Thank you for your participation.**</u>

Appendix 2.6 Expert rater feedback form

THE PSYCHOPATHY CHECKLIST: YOUTH VERSION (PCL: YV)

INTER-RATER RELIABILITY STUDY

EXPERT PANEL REVIEW

CASE BEING REVIEWED: ADAM

I would like to take this opportunity to thank you for participating in the expert review process for this study. The participants in this study will be required to rate six fictitious case vignettes using a shortened version of the PCL: YV. As you are aware it is not possible for us to develop case vignettes which are representative of the range of young people seen in clinical practice using only six cases. Instead we have sought to develop case vignettes varying in severity of psychopathic traits and introduced this expert review process as a way of maximising the validity of these cases. These case vignettes were developed to represent young people where it was felt that an assessment of psychopathic traits may be warranted.

Your expert judgement will allow us to ensure the authenticity of this case and will be used within our data analysis.

---

1. <u>In your own opinion</u>, please state whether you think this case presents with Low, Moderate or High levels of psychopathy?

LOW☐

MEDIUM☐

HIGH☐

---

2. Does this case vignette 'feel' authentic?

   YES☐   NO☐

   If not, please suggest what information should be included, altered or removed to ensure that this case feels authentic?

3. **Are there any items where you feel there is insufficient information provided to rate the item?** *(If so, please provide the name of the item and suggest what information should be included).*

3. **Are there any suggestions or improvements you would make for this case*?* *(If so, please state what information should be included, altered or removed to improve the case)*

Thank you for your participation.

Appendix 2.7 Reporting of non-significant findings with categorical variables

In training vs. fully qualified

An independent-samples t-test found no significant difference between staff in training (M=71.35, SD=3.86) and qualified staff (M=71.44, SD=2.74) in their agreement with the mean, t (17) = -.06, p=.96, d= .00018.

An independent-samples t-test found no significant difference between staff in training (M=73.68, SD=6.28) and qualified staff (M=73.06, SD=3.99) in their agreement with the mode, t (17) = .26, p=.79, d=.0040.

An independent-samples t-test found no significant difference between staff in training (M=64.33, SD=1.29) and qualified staff (M=64.57, SD=1.29) in their agreement with the experts, t (17) =-.16, p=.88, d=.0014.

An independent-samples t-test found no significant difference between staff in training (M=5.83, SD=1.94) and those qualified (M=5.67, SD=1.16) in terms of their confidence in the accuracy of their ratings, t (16) =.23, p=.82, d=.00329.

A Mann-Whitney U Test revealed no significant difference between staff in training (Md=7.00, n=6) and qualified staff (Md=6.00, n=13) in terms of reported comfort using the PCL: YV, U = 32, z=-.63, p=.53, r= -.14.

A Mann-Whitney U Test revealed no significant difference between staff in training (Md=7.50, n=6) and qualified staff (Md=6.50, n=12) in terms of difficulty using the PCL: YV, U=18, z=-1.76, p=.08, r = -.41.

Forensic experience vs. no forensic experience

An independent-samples t-test found no significant difference between those with (M=71.75, SD=3.07) and without (M= 70.95, SD=3.12) forensic experience, in terms of agreement with the mean, t (17) =.56, p=.59, d=.017.

An independent-samples t-test found no significant difference between those with (M=73.63, SD=4.95) and without (M=72.75, SD=4.50) forensic experience, in terms of agreement with the mode, t (17) =.40, p=.70, d=.0091.

An independent-samples t-test found no significant difference between those with (M=64.85, SD=3.39) and without (M=64.00, SD=2.50) forensic experience, in terms of agreement with the expert, t (17) =.60, p-.56, d=.0208.

An independent-samples t-test found no significant difference between those with (M=6.64, SD=1.36) and without (M=5.38, SD=1.99) forensic experience, in terms of reported comfort using the PCL: YV, t (17) =1.64, p=.12, d=.137.

An independent-samples t-test found no significant difference between those with (M=6.10, SD=1.19) and without (M=5.25, SD =1.58) forensic experience in terms of their confidence in the accuracy of their ratings, t (16) =1.30, p=.21, d=.0955.

An independent-samples t-test found no significant difference between those with (M=6.90, SD=.99) and without (M=6.75, SD=1.28) forensic experience in terms of difficulty using the PCL: YV, t (16) =.28, p=.78, d=.00487.

Completion of formal training in the structured assessment of personality or risk

An independent-samples t-test found no significant difference between staff who had (M=71.28, SD=3.16) and had not (M=71.53, SD=3.08) received such training in their agreement with the mean, t (17) =-.17, p=.86, d=.00177.

An independent-samples t-test found no significant difference between staff who had (M=72.61, SD=4.56) and had not (M=73.84, SD=4.91) received such training in their agreement with the mode, t (17) =-.56, p=.58, d=.0183.

An independent-samples t-test found no significant difference between staff that had (M=64.73, SD=3.74) and had not (M=64.28, SD=2.33) received such training in their agreement with the expert, t (17) =.32, p=.75, d=.0060.

An independent-samples t-test found no significant difference between staff that had (M=6.67, SD=1.50) and had not (M=5.60, SD=1.84) received such training in their comfort using the PCL: YV, t (17) =1.38, p=.19, d=.100.

An independent-samples t-test found no significant difference between staff that had (M=6.25, SD=1.282) and those who had not (M=5.30, SD=1.418) received such training in their confidence in the accuracy of their ratings, t (16) =1.47, p=.16, d=.11927.

An independent-samples t-test found no significant difference between staff that had (M=6.88, SD=1.13) and those who had not (M=6.80, SD=1.14) received such training in their perceived difficulty when using the PCL: YV, t (16) =.14, p=.89, d=.00122.

Experience conducting structured assessments of personality (current or previous role)

An independent-samples t-test found no significant difference between staff who had (M=70.18, SD=3.25) and had not (M=71.97, SD=2.87) conducted structured assessments of personality, in terms of their agreement with the mean, t (17) = -1.21, p=.24, d=.079.

An independent-samples t-test found no significant difference between staff who had (M=70.73, SD=4.08) and had not (M=74.42, SD=4.57) conducted structured assessments of personality, in terms of their agreement with the mode, t (17) =-1.69, p=.11, d=.14.

An independent-samples t-test found no significant difference between staff who had (M=65.63, SD=3.96) and had not (M=63.97, SD=2.45) conducted structured assessments of personality, in terms of their agreement with the expert, t (17) =1.133, p=.273, d=.07.

A Mann-Whitney U Test found no significant difference between staff who had (Md=7.00, n=6) and had not (Md=7.00, n=12) conducted structured assessments of personality, in terms of perceived difficulty using the PCL: YV, U=33, z=-.29, p=.77, r=-.07.

Appendix 2.8 Research proposal

An investigation of the Inter-Rater Reliability of the Psychopathy Checklist: Youth Version (PCL: YV): effects of rater and case characteristics.

<u>**Psychopathy**</u>

Early descriptions of psychopathy were outlined by Cleckley (1941) in his monograph, "The Mask of Sanity". He describes interviews with fifteen incarcerated psychopaths and outlined criteria which he believed captured the construct of psychopathy. These can be described under: positive psychological adjustment, behavioural pathology and impaired social relatedness and emotional unresponsiveness (Patrick, 2006, p 611-12). Positive psychological adjustment included intelligence, social adeptness, absence of irrationality and nervousness and low incidence of suicide. Behavioural pathology included irresponsibility, impulsive antisocial behaviour, failure to learn from experience, lacked life plans and promiscuity. Impaired social relatedness and emotional unresponsiveness included a lack of remorse and shame, reduced affective reactions, an inability to love, egocentricity, absence of loyalty, deceitfulness, insincerity and lacking insight. Current conceptualisations of psychopathy generally refer to a Personality Disorder, characterised by three symptom groups; an arrogant and deceitful interpersonal style, deficient affective experience and an impulsive behavioural style (Cooke & Michie, 2001).

Psychopathy is a serious condition. It is statistically associated with a myriad of adverse outcomes for the sufferer and society. Studies have consistently found that psychopathy is linked with poor response to treatment, substance misuse, criminality and violence within adult populations (Salekin *et al.,* 2006; Leistico *et al.,* 2008). It is therefore crucial that attempts are made to identify psychopathic traits early on to prevent or intervene to mitigate against the development or continuation of this condition. This has justified the extension of the construct to youth. Consistent with this, research has found that youth showing similar symptoms to those observed in adult psychopaths are also elevated on measures of antisocial behaviour and aggression (Forth & Book, 2010).

Links with Conduct Disorder

Over the last decade, the link between psychopathic traits and conduct disorder has been an area of considerable focus. Many studies have focused on the callous-unemotional (CU) traits model, commonly measured using the Antisocial Process Screening Device (APSD, Frick & Hare, 2001). A study of high risk boys found that higher levels of CU predicted more severe forms of antisocial behaviour, after controlling for other symptoms of conduct disorder (Pardini & Fite, 2012). Similarly, Kahn *et al.,* (2012) found high levels of CU traits were associated with more severe aggression and cruelty. Others have found that adolescent males scoring highly on the PCL:YV display more externalizing behaviours including aggression, more conduct disorder symptoms and substance abuse compared to those with low scores on this measure (Gretton *et al.,* 1994; Forth & Burke, 1998). Similarly, studies using the PCL: YV have found that CU traits (measured by four items of the PCL: YV) were associated with more severe, violent offending (Vincent *et al.,* 2003).

Diagnostic Specifier

It has been proposed that a callous-unemotional specifier should be included as a subtype of conduct disorder within the DSM-V (Kahn et al., 2012). Given that CU traits are directly derived from the adult psychopathy literature, and are likely to be assessed in the process of diagnosis using youth assessment protocols such as PCL:YV and APSD, it is vital that studies are conducted to test their validity and reliability.

Assessment in Adults

The most extensively used and researched measure of adult psychopathy is the Psychopathy Checklist (PCL-R) (Hare, 2003). The PCL-R has demonstrated high levels of validity and reliability (Hare & Neuman, 2006). With regards to inter-rater reliability, Intraclass correlation coefficients have ranged from around .86 to .94 (Hare, 2003). However, not all studies have demonstrated high levels of IRR and some have revealed concerning variations (Murrie et al., 2007; 2009; 2013).

Assessment in Children

Several conceptual models exist which purport to measure psychopathic traits in childhood (Lynam & Gudonis, 2005). However, they are far from a panacea. Johnstone and Cooke, (2007) and others have argued that many conceptual, methodological and developmental

challenges remain to be resolved.

The two most commonly used measures are those derived from the PCL-R: the Psychopathy Checklist, Youth Version (PCL: YV; Forth, Kosson, & Hare, 2003) and Antisocial Process Screening Devise (APSD; Frick & Hare, 2001). The APSD provides a screen of psychopathic traits in 4-18 year olds and self-report, parent and teacher versions are available. The PCL: YV is a clinician rated measure for individuals aged 13 and over. It involves gathering information from multiple sources, a semi-structured interview with the young person, collateral interviews and a file review. Following this, the clinician decides the degree to which each of the 20 symptoms are present and rates them from 0 (item definitely does not apply) to 2 (item definitely applies). There is no established cut-off score, however studies tend to report a cut-off of over 30 (Vincent et al., 2008). Despite the wide use of this measure, some have raised concerns that it may over rely on the behavioural aspect of psychopathy with half the items assessing this domain (Dawson et al., 2012). Others including Johnstone and Cooke (2004) and Johnstone (2010) have argued that considerable care is needed in the process of differential diagnosis.

Other measures include the Child Psychopathy Scale (CPS; Lynam, 1997), Youth Psychopathic Traits Inventory (YPI; Andershed *et al.,* 2002) and more recently measures based on the Comprehensive Assessment of Psychopathic Personality (CAPP) model. However, the CAPP currently remains a research instrument (Cooke et al., 2004).

This research will focus on the PCL: YV as this is the only clinician rated measure which is manualised and widely available for use**.**

Recent studies indicate that psychopathy is increasingly being referred to in court cases in North America (Skeem et al., 2011) where diagnoses of psychopathy are used to decide whether or not the child should be moved to adult criminal justice system. Studies have found that the label "psychopath" can lead jurors to believe the individual is at higher risk of future offending and that they require greater punishment (Boccaccini et al., 2008). Recent changes to the Criminal Justice and Licensing (Scotland) Act (2010) state that if an individual accused or murder, is diagnosed with psychopathy, they can put forward a "plea of diminished responsibility" and may face a lesser charge. Consequently, there is likely to be

an increase in the use of diagnosis for legal decision making in Scotland. In clinical practice, some have highlighted difficulties obtaining treatment for young people diagnosed with psychopathy (Johnstone, Bergen Conference 2011).

Validity and Reliability

The potential consequences of diagnosis and mis-diagnosis are significant. Therefore, it's important that assessments are accurate. This relies on valid and reliable assessment tools. A measure is said to be valid if it measures what it is intended to measure. Reliability refers to the consistency of a measure (Rosenfeld & Penrod, 2011).

Some studies have investigated the construct validity of the PCL: YV using factor analysis, which is the degree to which it accurately identifies the construct of psychopathy. Some have reported that a two-factor model fits the data with factor one representing interpersonal-affective symptoms and factor two representing socially deviant lifestyle symptoms (Forth et al., 1990), however others have failed to support this structure in adolescents (Salekin et al.,2006). Some have found a three-factor model comprising interpersonal, affective and behavioural dimensions to be a better fit for youth psychopathy (Cooke & Michie, 2001) and others have found a four-factor model with the addition of an antisocial domain best fits this concept in adolescents (Salekin et al., 2006). There is considerable debate in the literature as to which model fits best.

Predictive validity, i.e. the extent to which a score on a measure of psychopathy predicts scores on some external criterion, has been studied. However, some have argued that predictive validity is only meaningful at a group level and not precise at the individual level (Hart, 2007). That is, we cannot predict the future with precision. However, we can attempt to ensure that we use measures fairly and consistently across clinicians. This is important for the PCL: YV where clinical judgement plays a role in the scoring.

Inter-rater Reliability

Inter-rater reliability (IRR) is the level of agreement between two or more raters using the same measure independently. Depending on the type of variable, Cohen or Fleiss" Kappa, Pearson, Spearman or Intraclass correlations (ICC) can be used to measure IRR. General guidelines state that ICCs of less than .40 are poor, between .40 and .59 are fair, .60 to .74 is

good and over .74 is excellent (Fleiss, 1981). More generally, ICCs of .70 or above are considered acceptable. Some have argued that "*Although the validity of scores from any type of test or rating scale is certainly a critical topic to consider, validity is predicated on these scores being sufficiently reliable to allow for meaningful assessments of the variable of interest*" (Edens et al., 2010).

Although published studies utilising the PCL: YV (Appendix X) have assessed IRR as part of their research, no study has focused exclusively on this. The studies reported in Appendix X report IRR values ranging from .80 to .98 for Total PCL: YV Scores. For those reporting the IRR of individual factor scores, there appeared to be greater variation in these scores, with some achieving less than desirable levels of IRR (e.g. Spain et al., 2004).The methods used to date have primarily involved randomly selecting a subset of cases for rating by a few raters. This may not be representative of the heterogeneous characteristics of potential raters and young people seen in clinical settings. In reality young people present with varying levels of psychopathy and risk and staff may have varying levels of training and experience. It may be that there is less rater agreement for adolescents presenting with moderate levels of psychopathic traits as opposed to those with very mild or severe traits which may be easier to identify. Of the published studies, the raters characteristics were often not provided (Appendix X). It is unclear whether rater and case characteristics may influence IRR of the PCL: YV.

Despite the high levels of IRR often reported, studies in the adult literature have found that factors including adversarial allegiance, the raters personality and training may influence scoring and IRR of the PCL-R amongst other measures (Miller et al, 2007, 2009, 2013; Murrie et al, 2009; Persons et al, 1999).

<u>Adversarial Allegiance</u>

Adversarial allegiance is "the tendency for forensic evaluators to form opinions in a manner that better supports the party that retains them" (Murrie *et al.,* 2013). A series of studies found a difference in the PCL-R ratings provided by opposing sides (prosecution and defence) where both raters assessed the same individual. In most cases, the scores were in the expected direction, i.e. lower scores by defence and higher scores by prosecution (Murrie et al., 2007, 2009, and 2013). Also, two of these studies included other

measures but found that the difference was strongest for the PCL-R. The authors interpreted this as reflecting the subjective nature of the PCL-R protocol. The Intraclass correlations found in these studies were lower than is often reported (ICC=.39, ICC=.42). In the most recent study, wide variations in scores were noted within the same group.

Training

To our knowledge, the effects of training and experience on PCL ratings have not been studied. However, this has been studied for other tools. Persons et al., (1999) found that length of experience and level of training did not predict IRR for a CBT formulation task. However, whether the individual had PhD training or not predicted the accuracy of their judgements. Another study investigating the IRR of the Risk for Sexual Violence Protocol found higher levels of IRR for those highly trained in risk assessment (Sutherland et al., 2012). Also, those with less training agreed less with expert judgements and assigned higher estimations of risk.

Together, these findings suggest the measures including those used to assess psychopathy may be vulnerable to adversarial allegiance and rater bias due to personality traits or level of experience and training. These variables have not been studied in relation to the PCL: YV; however there is no reason to assume these biases would not extent to this measure. Another factor which could influence IRR is the gender of the young person.

Gender

Cleckley's (1941/1988) descriptions of psychopathy have greatly influenced current conceptualizations of psychopathy. Only two of his case studies included women, therefore our current understanding of psychopathy may be biased towards males (Kreis & Cooke, 2012).

In psychopathy research, females have been described as a "token inclusion in research" where they are measured against male templates (Logan, 2011). For instance, the PCL: YV was derived from the PCL-R, a measure developed and validated primarily on males. Therefore, there may be inherent bias in the measure for males. Youth psychopathy is particularly complex as psychopathy may be expressed differently across age and gender. Given the controversies extending the construct to youth, research has focused primarily on

the validity of the construct in youth overall and few studies have incorporated gender differences (Verona et al., 2010). Furthermore, the samples used to identify the psychometric properties of the PCL: YV were primarily male with only 6 of the 19 samples including males and females. The authors describe the level of IRR as "excellent" however state that independent double ratings were only available for a small number of cases. Furthermore, correlations between gender and total score across all samples found a small but significant effect, with males scoring slightly higher than females. Overall, the authors of the PCL: YV acknowledge the limited research on females and state "PCL-YV total scores do not appear to be unduly influenced by the youth's age, ethnicity or gender" (Forth *et al.,* 2003; p.51). Others argue there is insufficient research to support this claim (Odgers et al., 2005).

Some studies suggest psychopathy may present differently in males and females. In a recent prototypical analyses study, professionals rated adult cases using the CAPP (Kreis & Cooke, 2011). Results revealed similarities in psychopathy across gender. However, "emotional instability", "unstable self-concept" and "manipulative" were more relevant for females (Kreis & Cooke, 2011). Another study comparing the CAPP and Psychopathy Checklist Screening Version (PCL: SV) for two female case studies found similar results (Kreis & Cooke, 2012). Again "emotional instability" was viewed as more prominent in females. If the PCL: SV had been used alone, this information would have been missed. Furthermore, females may express their psychopathic traits within close relationships and domestic settings presenting additional challenges for the assessment of psychopathy in females (Kreis & Cooke, 2012).

A similar study in youth found 23 items and 14 items were rated as prototypical for males and females respectively (Salekin et al., 2001). Also, "sexual promiscuity" and "stays out at night without parental permission" were identified as specific to females. Therefore, psychopathy may be less prevalent in female youth or the traits reflected in current measures may not adequately capture the construct in females.

Also, stereotypes of the rater may influence assessments of psychopathy. Logan (2011) highlighted that "gender stereotypes dictate our expectations of how individual's behave and how we interpret their behaviour" (Logan, 2011). One study found that violence was viewed as inconsistent with the female stereotype and "neutralizing discourses" were used to reframe

female violence as a result of mental illness or trauma (Adshead, 2011).

These findings suggest that psychopathy may present differently across gender and the gender stereotypes may introduce rater bias. Therefore, clinicians may find it more difficult to assess psychopathy in young females and consequently levels of IRR may be lower. However, available studies often do not indicate whether female cases were used in the measure of IRR, therefore it is unclear whether this is the case.

Aims/Hypothesis

To identify:

4. What level of IRR do multidisciplinary professionals achieve when using the PCL: YV?
5. To what extent do professionals agree with expert ratings when using the PCL: YV?
6. Is IRR of judgements associated with rater characteristics (e.g. length of experience, amount of training)?
7. Is IRR of judgements made using the PCL: YV influenced by characteristics of the case studies (level of psychopathy and gender)?

Hypotheses

1) The PCL: YV will demonstrate lower levels of IRR reported in the literature due to the increased number of raters.
2) Those with more experience (forensic experience and experience of using the PCL: YV in clinical practice) and more training will agree more with expert ratings and have higher rates of IRR.
3) Compared to male case studies, lower levels of IRR will be found for female case studies. The levels of IRR will differ across cases with low, medium and high levels of psychopathy. Lower levels of IRR will be found for the medium cases compared with the low and high level cases.

Participants

NHS staff recruited from forensic services and child and adolescent mental health services across the Scottish Health Boards. This will include the Scottish Clinical Forensic group and the Division of Forensic Psychology network.

Participants must meet the PCL: YV manual criteria except criteria 3 and 5 which will be met by attending the training:

1) Possess an advanced graduate degree in either a social, medical or behavioural science (e.g. MA, M.S.W, PhD, DClinPsy, M.D., M.B)?

2) Have appropriate professional credentials (e.g. be registered, licensed, or legally entitled to conduct psychological assessments) OR be working under the direct supervision of a registered professional.

3) Be familiar with clinical and research literature pertaining to psychopathy, adult and adolescent.

4) Have experience working with adolescents or completion of graduate courses in adolescent development.

5) Have adequate training and experience in the use of the PCL: YV.

If the participant does not possess the above criteria, the manual states that a qualified clinician should supervise, and take responsibility for those individuals. Therefore, participants will still be able to participate if they do not have the above qualifications but are supervised by someone who does.

Recruitment Procedures

Circular invitation to relevant NHS employees and through attendance at PCL: YV training.

Measures

➢ PCL: YV.

➢ Purpose designed staff information questionnaire based on previous research (Sutherland et al., 2012). This will gather professional information (e.g. length of time qualified).

Design

3x2 Factorial design. Independent Variables:

1) Level of Psychopathy – low, medium, high.
2) Gender of Case Vignette – male, female.

Research Procedures

➢ Six factitious case studies based loosely on real cases will be developed. These will vary in level of psychopathy (low, medium, high).

➢ These case vignettes will be replicated, changing the gender to create 12 cases.

➢ Recruit 6 „expert‟ raters with expertise in assessment/ diagnosis of psychopathy.

➢ Experts will be randomly allocated 1-2 case vignettes to rate using the PCL: YV. The gender of the case vignette will be concealed, with the case named "X". Amendments may be made to the case vignettes following expert feedback.

➢ An email will be sent to appropriate staff groups inviting them to participate in the study. The information sheet and consent form will be attached. The email will contain details regarding the training and how to sign up if they wish to participate.

➢ Two, two-day, free PCL: YV training events lead by Dr Lorraine Johnstone and myself, to delegates consenting to take part. This will be held in either NHS Greater Glasgow & Clyde or NHS Forth Valley premises. This is the only time that participants will have direct contact with the research team.

➢ Participants will be assigned either the first set of case studies or the second. The ordering of the case vignettes will be counterbalanced.

➢ Participants will rate six cases and complete the Staff Information Questionnaire.

Data analysis

Questions 1 & 2 = Case 2 Intraclass Correlation Coefficients (ICC2) and percentage agreement statistics. Percentage agreement with the mean, mode and expert rating will be calculated. This will be done by taking the ratings from all 12 vignettes.

Question 3 = correlations between the continuous professional characteristics and participant agreement on PCL: YV scores will be conducted.

Question 4 = the ICC and average percentage agreement (with mean, mode and expert) will

be calculated for each of the twelve cases. This will allow comparisons to be made across the cases varying in level of psychopathy and gender. The cases will then be grouped into male and female and an ANOVA conducted between these groups. The same will be done for low, medium and high cases.

Justification of Sample Size

Prior studies have found Intraclass correlations between 0.81 and 0.98 for the PCL: YV Total Scores (see table 2). These studies involve a few raters rating a random sample of the total cases. This study will involve many raters rating case vignettes varying in level of psychopathy and gender, therefore it is anticipated that the level of inter-rater reliability will not be as high as this.

To determine the sample size, PASS sample size software based on the formula outlined in Walter et al., (1998) was used. Based on power being set at 0.80, alpha at 0.05, R0 = 0.3 and R1 = 0.7, a minimum of 6 vignettes and 18 raters are required. This left two options:

1. Develop 3 male and 3 female vignettes to be rated by all staff.
2. Develop 6 vignettes and replicate these changing the gender to provide 12 vignettes, in which case double the sample size would be required.

After consideration, option 2 was favoured: any effect could more confidently be attributed to gender as the content of the vignettes will be identical. Therefore, a minimum of 36 participants will be required.

Ethical Issues

Case vignettes will be based upon the clinical and theoretical experience of the research team. Where details are taken from real cases which the field supervisor for this research has worked with, the information will be significantly altered to ensure anonymity. Participants will receive an information sheet (Appendix X) and given the opportunity to discuss any questions. Written informed consent will be obtained at the start of the training event (Appendix X).

Information in the vignettes could potentially be distressing. As participants are NHS

employees working in mental health services, they are likely to come across similar material in clinical practice. Participants will be informed that Dr Lorraine Johnstone will be available at the end of the training should they have any concerns or questions they wish to discuss. Participant information will be anonymised and handled in line with the Data Protection Act (1998) and NHS procedures.

Participants may achieve low levels of agreement with expert ratings. This could have implications for patients they go on to assess. A disclaimer will be included within the recruitment email and at the beginning of the training. This will clearly state that attendance will not guarantee competence in the use of the PCL: YV. Following submission and approval of the research, participants will receive a summary of the results. This will include all participants" results with each participant represented by a number. Therefore only they will know their results. Participants will be informed that it is their responsibility as a clinician to ensure they seek further support (e.g. training, supervision, experience) prior to administering the PCL:YV in clinical practice should this be necessary.

Practical Applications

This research will ascertain the IRR of the PCL: YV and identify the influences of rater and case characteristics on the IRR of the PCL: YV. This may inform staff training programs in the assessment of psychopathic traits in youth.

References

Adshead, G. (2011). Same but Different: Constructions of Female Violence in Forensic Mental Health. *International Journal of Feminist Approaches to Bioethics, 4*(1), 41-68.

Andershed, H., Hodgins, S., & Tengström, A. (2007). Convergent Validity of the Youth Psychopathic Traits Inventory (YPI) Association with the Psychopathy Checklist: Youth Version (PCL: YV). *Assessment, 14*(2), 144-154.

Book, A. S. (2010).Psychopathic Traits in Children and Adolescents. *Handbook of Child and Adolescent Psychopathy*, 251.

Andershed, Kerr, M., Stattin, H., & Levander, S. (2002). Psychopathic traits in non-referred youths: initial test of a new assessment tool. In S. Blaauw (Ed.), *Psychopaths: Current international perspectives.* The Hague: Elsevier.

Boccaccini, M. T., Murrie, D. C., Clark, J. W., & Cornell, D. G. (2008). Describing, diagnosing, and naming psychopathy: how do youth psychopathy labels influence jurors? *Behavioral Sciences & the Law*, *26*(4), 487-510.

Campbell, M. A., Porter, S., & Santor, D. (2004). Psychopathic traits in adolescent offenders: An evaluation of criminal history, clinical, and psychosocial correlates. *Behavioral Sciences & the Law, 22*(1), 23-47.

Chauhan, P., Ragbeer, S. N., Burnette, M. L., Oudekerk, B., Reppucci, N. D., & Moretti, M. M. (2014). Comparing the Youth Psychopathic Traits Inventory (YPI) and the Psychopathy Checklist–Youth Version (PCL-YV) Among Offending Girls. *Assessment, 21*(2), 181-194.

Cleckley, H. (1941). *The mask of sanity* (1[st]ed). St. Louis, MO: Mosby.

Cooke, D. J., & Michie, C. (2001). Refining the construct of psychopathy: towards a hierarchical model. *Psychological Assessment, 13*(2), 171.

Cooke, D.J., Hart, S.D., Logan, C., & Michie, C. (2004). Comprehensive Assessment of Psychopathic Personality – Institutional Rating Scale (CAPP-IRS). Unpublished manuscript.

Corrado, R. R., Vincent, G. M., Hart, S. D., & Cohen, I. M. (2004). Predictive validity of the Psychopathy Checklist: Youth Version for general and violent recidivism. *Behavioral Sciences & the Law, 22*(1), 5-22.

Das, J., de Ruiter, C., Lodewijks, H., & Doreleijers, T. (2007). Predictive validity of the Dutch PCL: YV for institutional disruptive behavior: Findings from two samples of male adolescents in a juvenile justice treatment institution. *Behavioral Sciences & the Law, 25*(5), 739-755.

Dawson, S., McCuish, E., Hart, S. D., & Corrado, R. R. (2012). Critical Issues in the Assessment of Adolescent Psychopathy: An Illustration Using Two Case Studies. *International Journal of Forensic Mental Health, 11*(2), 63-79.

Dolan, M. C., & Rennie, C. E. (2007). Is juvenile psychopathy associated with low anxiety and fear in conduct-disordered male offenders? *Journal of Anxiety Disorders, 21*(8), 1028-1038.

Dolan, M. C., & Rennie, C. E. (2006). Reliability and validity of the psychopathy checklist: Youth version in a UK sample of conduct disordered boys. *Personality and Individual Differences, 40*(1), 65-75.

Douglas, K. S., Epstein, M. E., & Poythress, N. G. (2008). Criminal recidivism among juvenile offenders: Testing the incremental and predictive validity of three measures of psychopathic features. *Law and Human Behavior, 32*(5), 423-438.

Edens, J. F., Boccaccini, M. T., & Johnson, D. W. (2010). Inter-rater reliability of the PCL-R total and factor scores among psychopathic sex offenders: Are personality features more prone to disagreement than behavioral features? *Behavioral Sciences & the Law, 28*(1), 106-119.

Fink, B. C., Tant, A. S., Tremba, K., & Kiehl, K. A. (2012). Assessment of Psychopathic Traits in an Incarcerated Adolescent Sample: A Methodological Comparison. *Journal of Abnormal Child Psychology, 40*(6), 971-986.

Fleiss, J. L. (1982). *Statistical Methods for Rates and Proportions*. (2nded). John Wiley. *New York*.

Forth, A. E., Kosson, D.S., & Hare, R.D. (2003). *Psychopathy Checklist: Youth Version (PCL: YV).* Toronto: Multi-Health Systems.

Forth, A. E., & Book, A. S. (2010). Psychopathic traits in children and adolescents: The relationship with antisocial behaviors and aggression. *Handbook of Child and Adolescent Psychopathy*, 251-283.

Forth, A. E., & Burke, H. C. (1998). Psychopathy in adolescence: Assessment, violence, and develop- mental precursors. In D. Cooke, A. E. Forth, & R. D. Hare (Eds.), *Psychopathy: Theory, research, and implications for society* (pp. 205–229). Dordrecht: Kluwer.

Frick, P., & Hare, R.D. (2001). *The Antisocial Process Screening Device*. Toronto: Multi-Health Systems.

Hare, R. D. (2003). Manual for the Revised Psychopathy Checklist (2nded). Toronto, ON: Multi-Health Systems.

Hare, R. D., & Neumann, C. S. (2006). The PCL-R assessment of psychopathy. *Handbook of Psychopathy*, 58-88.

Hart, S. D., Michie, C., & Cooke, D. J. (2007). Precision of actuarial risk assessment instruments evaluating the margins of error of group v. individual predictions of violence. *The British Journal of Psychiatry*, *190*(49), s60-s65.

Hemphälä, M., & Tengström, A. (2010). Associations between psychopathic traits and mental disorders among adolescents with substance use problems. *British Journal of Clinical Psychology, 49*(1), 109-122.

Johnstone, L., & Cooke, D. J. (2007). Psychopathy and young offenders. *Psychiatry, 6*(10), 429-432.

Johnstone, L., & Cooke, J. (2004). "Psychopathic-like traits in childhood: Conceptual and measurement concerns." *Behavioral Sciences & the Law, 22*(1), 103-125.

Kahn, R. E., Frick, P. J., Youngstrom, E., Findling, R. L., &Youngstrom, J. K. (2012). The effects of including a callous–unemotional specifier for the diagnosis of conduct disorder. *Journal of Child Psychology and Psychiatry, 53*(3), 271-282.

Kreis, M. K., & Cooke, D. J. (2011). Capturing the psychopathic female: A prototypicality analysis of the Comprehensive Assessment of Psychopathic Personality (CAPP) across gender. *Behavioral Sciences & the Law, 29*(5), 634-648.

Kreis, M. K., Cooke, D. J., Michie, C., Hoff, H. A., & Logan, C. (2012). The Comprehensive Assessment of Psychopathic Personality (CAPP): Content validation using prototypical analysis. *Journal of Personality Disorders, 26*(3), 402-413.

Lee, Z., Vincent, G. M., Hart, S. D., & Corrado, R. R. (2003). The validity of the Antisocial Process Screening Device as a self-report measure of psychopathy in adolescent offenders. *Behavioral Sciences & the Law, 21*(6), 771-786.

Lynam, D.R. (1997). Pursuing the psychopath: capturing the fledgling psychopath in a nomological net. *Journal of Abnormal Psychology, 106*, 425-438.

Lynam, D. R., & Gudonis, L. (2005). The development of psychopathy. *Annual Review of Clinical Psychology, 1,* 381-407.

Marczyk, G. R., Heilbrun, K., Lander, T., & Dematteo, D. (2003). Predicting juvenile recidivism with the PCL: YV, MAYSI, and YLS/CMI. *International Journal of Forensic Mental Health, 2*(1), 7-18.

McCrory, E., Hickey, N., Farmer, E., & Vizard, E. (2008). Early-onset sexually harmful behaviour in childhood: a marker for life-course persistent antisocial behaviour? T*he Journal of Forensic Psychiatry & Psychology, 19*(3), 382-395.

Miller, A. K., Rufino, K. A., Boccaccini, M. T., Jackson, R. L., & Murrie, D. C. (2011). On individual differences in person perception: Raters˝ personality traits relate to their Psychopathy Checklist-Revised scoring tendencies. *Assessment, 18*(2), 253-260.

Murrie, D. C., & Balusek, K. (2008). Forensic assessment of violence risk in adversarial proceedings: Pursuing objectivity and avoiding bias. *Journal of Forensic Psychology Practice, 7*(4), 141-153.

Murrie, D. C., Boccaccini, M. T., Turner, D. B., Meeks, M., Woods, C., & Tussey, C. (2009). Rater (dis) agreement on risk assessment measures in sexually violent predator proceedings: Evidence of adversarial allegiance in forensic evaluation? *Psychology, Public Policy, and Law, 15*(1), 19.

Odgers, C. L., Reppucci, N. D., & Moretti, M. M. (2005).Nipping psychopathy in the bud: an examination of the convergent, predictive, and theoretical utility of the PCL-YV among adolescent girls. *Behavioral Sciences & the Law, 23*(6), 743-763.

O'Neill, M. L., Lidz, V., & Heilbrun, K. (2003). Predictors and correlates of psychopathic characteristics in substance abusing adolescents. *International Journal of Forensic Mental Health, 2*(1), 35-45.

Pardini, D. A., & Fite, P. J. (2010). Symptoms of Conduct Disorder, Oppositional Defiant Disorder, Attention-Deficit/Hyperactivity Disorder, and Callous-Unemotional Traits as Unique Predictors of Psychosocial Maladjustment in Boys: Advancing an Evidence Base for

DSM-V. *Journal of the American Academy of Child & Adolescent Psychiatry, 49*(11), 1134-1144.

Penney, S. R., & Moretti, M. M. (2007). The relation of psychopathy to concurrent aggression and antisocial behavior in high-risk adolescent girls and boys. *Behavioral   Sciences & the Law, 25*(1), 21-41.

Rosenfeld, B., & Penrod, S. D. (2011). *Research Methods in Forensic Psychology*. John Wiley & Sons.

Persons, J. B., & Bertagnolli, A. (1999). Inter-Rater Reliability of Cognitive-Behavioral Case Formulations of Depression: A Replication. *Cognitive Therapy and Research, 23*(3), 271-283.

Salekin, R. T., Brannen, D. N., Zalot, A. A., Leistico, A. M., & Neumann, C. S. (2006). Factor Structure of Psychopathy in Youth Testing the Applicability of the New Four-Factor Model. Criminal Justice and Behavior, 33(2), 135-157.

Salekin, R. T., Rogers, R., & Machin, D. (2001). Psychopathy in youth: Pursuing diagnostic clarity. *Journal of Youth and Adolescence, 30*(2), 173-195.

Schmidt, F., Campbell, M. A., & Houlding, C. (2011). Comparative analyses of the YLS/CMI, SAVRY, and PCL: YV in adolescent offenders: A 10-year follow-up into adulthood. *Youth Violence and Juvenile Justice, 9*(1), 23-42.

Schrum, C. L., & Salekin, R. T. (2006). Psychopathy in adolescent female offenders: An item response theory analysis of the Psychopathy Checklist: Youth Version. *Behavioral Sciences & the Law, 24*(1), 39-63.

Skeem, J. L., & Cauffman, E. (2003). Views of the downward extension: Comparing the youth version of the Psychopathy Checklist with the Youth Psychopathic Traits Inventory. *Behavioral Sciences & the Law, 21*(6), 737-770.

Skeem, J. L., Polaschek, D. L., Patrick, C. J., & Lilienfeld, S. O. (2011). Psychopathic personality bridging the gap between scientific evidence and public policy. *Psychological Science in the Public Interest, 12*(3), 95-162.

Spain, S. E., Douglas, K. S., Poythress, N. G., & Epstein, M. (2004). The relationship between psychopathic features, violence and treatment outcome: The comparison of three youth measures of psychopathic features. *Behavioral Sciences & the Law, 22*(1), 85-102.

Sutherland, A. A., Johnstone, L., Davidson, K. M., Hart, S. D., Cooke, D. J., Kropp, P. R., ... & Stocks, R. (2012). Sexual violence risk assessment: An investigation of the interrater reliability of professional judgments made using the risk for sexual violence protocol. *International Journal of Forensic Mental Health, 11*(2), 119-133.

Vincent, G. M., Odgers, C. L., McCormick, A. V., & Corrado, R. R. (2008). The PCL: YV and recidivism in male and female juveniles: A follow-up into young adulthood. *International Journal of Law and Psychiatry, 31*(3), 287-296.

Vitacco, M. J., Caldwell, M. F., Van Rybroek, G. J., & Gabel, J. (2007).Psychopathy and behavioral correlates of victim injury in serious juvenile offenders. *Aggressive Behavior, 33*(6), 537-544.

Vitacco, M. J., Neumann, C. S., Caldwell, M. F., Leistico, A. M., & Van Rybroek, G. J. (2006). Testing factor models of the Psychopathy Checklist: Youth Version and their association with instrumental aggression. *Journal of Personality Assessment, 87*(1), 74-83.

Walter, S. D., Eliasziw, M., & Donner, A. (1998). Sample size and optimal designs for reliability studies. *Statistics in medicine*, *17*(1), 101-110.

Welsh, J. L., Schmidt, F., McKinnon, L., Chattha, H. K., & Meyers, J. R. (2008). A Comparative Study of Adolescent Risk Assessment Instruments Predictive and Incremental Validity. *Assessment, 15*(1), 104-115.